Millimeter Wave Systems for Wireless Cellular Communications

Lou Zhao

A thesis submitted to the Graduate Research School of
The University of New South Wales
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

School of Electrical Engineering and Telecommunications
Faculty of Engineering
The University of New South Wales

August 2018

Thesis/Dissertation Sheet

UNSW
SYDNEY | Australia's Global University

Surname/Family Name	:	**ZHAO**
Given Name/s	:	**LOU**
Abbreviation for degree as give in the University calendar	:	**PhD**
Faculty	:	**Engineering**
School	:	**School of Electrical Engineering & Telecommunications**
Thesis Title	:	**Millimeter Wave Systems for Wireless Cellular Communications**

Abstract 350 words maximum: (PLEASE TYPE)

This thesis considers channel estimation and multiuser (MU) data transmission for time division duplex (TDD) massive multiple-input multiple-output (MIMO) systems with fully digital/hybrid structures in millimeter wave (mmWave) channels. The work reported in this thesis contains novel mmWave transmission schemes and performance analysis, which provides insights into the design of hybrid mmWave networks. It contains three main contributions.

In this thesis, we first propose a tone-based linear search algorithm to facilitate the estimation of angle-of-arrivals (AoAs) of the strongest line-of-sight channel components as well as scattering components of the users at the base station (BS) with fully digital structure. Our results show that the proposed maximum-ratio transmission (MRT) based on the strongest components can achieve a higher data rate than that of the conventional MRT, under the same mean squared errors (MSE) of channel estimation. In addition, we quantify the achievable rate degradation due to phase quantization errors and propose an angular domain user scheduling algorithm for mmWave systems to improve the users' receive signal-to-interference-plus-noise ratio (SINR).

Second, we develop a low-complexity channel estimation and beamformer/precoder design scheme for hybrid mmWave systems, which utilizes tone signals and orthogonal pilots for the design of analog beamforming matrices and digital precoding matrices, respectively. In addition, the proposed scheme applies to both non-sparse and sparse mmWave channel environments. We then leverage the proposed scheme to investigate the downlink achievable rate performance. The results show that the proposed scheme obtains a considerable achievable rate of fully digital systems. Taking into account the effect of various types of errors, we investigate the achievable rate performance degradation of the considered scheme.

Third, we extend our proposed scheme to a multi-cell MU mmWave MIMO network. We derive the closed-form approximation of the normalized MSE of channel estimation under pilot contamination over Rician fading channels. Furthermore, we derive a tight closed-form approximation and the scaling law of the average achievable rate. Our results unveil that channel estimation errors, the intra-cell interference, and the inter-cell interference caused by pilot contamination over Rician fading channels can be efficiently mitigated by simply increasing the number of antennas equipped at the desired BS.

Copyright Statement

Authenticity Statement

Originality Statement

I hereby declare that this submission is my own work and to the best of my knowledge it contains no material previously published or written by another person, or substantial portions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institute, except where due acknowledgment is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

Signed ...

Date .....

Dedicated to my parents, my wife, and my son.

Abstract

This thesis considers channel estimation and multiuser (MU) data transmission for time division duplex (TDD) massive multiple-input multiple-output (MIMO) systems with fully digital/hybrid structures in millimeter wave (mmWave) channels. The work reported in this thesis contains novel mmWave transmission schemes and performance analysis, which provides insights into the design of hybrid mmWave networks. It contains three main contributions.

In this thesis, we first propose a tone-based linear search algorithm to facilitate the estimation of angle-of-arrivals (AoAs) of the strongest line-of-sight channel components as well as scattering components of the users at the base station (BS) with fully digital structure. Our results show that the proposed maximum-ratio transmission (MRT) based on the strongest components can achieve a higher data rate than that of the conventional MRT, under the same mean squared errors (MSE) of channel estimation. In addition, we quantify the achievable rate degradation due to phase quantization errors and propose an angular domain user scheduling algorithm for mmWave systems to improve the users' receive signal-to-interference-plus-noise ratio (SINR).

Second, we develop a low-complexity channel estimation and beamformer /precoder design scheme for hybrid mmWave systems, which utilizes tone signals and orthogonal pilots for the design of analog beamforming matrices and digital precoding matrices, respectively. In addition, the proposed scheme applies to

both non-sparse and sparse mmWave channel environments. We then leverage the proposed scheme to investigate the downlink achievable rate performance. The results show that the proposed scheme obtains a considerable achievable rate of fully digital systems. Taking into account the effect of various types of errors, we investigate the achievable rate performance degradation of the considered scheme.

Third, we extend our proposed scheme to a multi-cell MU mmWave MIMO network. We derive the closed-form approximation of the normalized MSE of channel estimation under pilot contamination over Rician fading channels. Furthermore, we derive a tight closed-form approximation and the scaling law of the average achievable rate. Our results unveil that channel estimation errors, the intra-cell interference, and the inter-cell interference caused by pilot contamination over Rician fading channels can be efficiently mitigated by simply increasing the number of antennas equipped at the desired BS.

Acknowledgments

This work would not have been done without the encouragement and the support from some people I have met during my fascinating Ph.D. journey.

First and foremost, I would like to thank my supervisor Professor Jinhong Yuan with my deepest gratitude for offering me a Ph.D. opportunity three and half years ago and having faith in me ever since. He guides me throughout this work with patience, technical and personal support, and encouragement. He tolerates my shortcomings and helps me to overcome my weakness. Besides, he is passionate about new technologies and exciting ideas. I feel fortunate to pursue my Ph.D. degree under his supervision.

Second, I would like to thank my co-supervisor, Dr. Derrick Wing Kwan Ng, for his guidance, constructive suggestions, and help on my research. Derrick is talented, extremely self-disciplined, and very hard-working. I will regard him as the role model of my academic career.

I would also like to thank Dr. Giovanni Geraci and Dr. Tao Yang for the first step of my research. Many thanks to Dr. Lei Yang, Dr. Yixuan Xie, Dr. Tao Huang, Dr. Mengyu Huang, Dr. Chenxi Liu, Dr. Shuang Tian, Dr. Nan Yang, Dr. Shihan Yan, Dr. Yansha Deng, Dr. Yeqing Hu, and others for having a lot of discussions with me. Special thanks to my colleagues and friends at the University of New South Wales, you guys make the Ph.D. journey funny and interesting. I will always hold dear the days and nights spent in the office.

Finally, my deepest appreciation goes to my beloved family for everything they have done for me. I would like to dedicate this thesis to my parents, my wife, and my son. To my son, you make my life brighter.

List of Publications

Journal Articles:

1. **L. Zhao**, D. W. K. Ng, and J. Yuan, "Multi-user precoding and channel estimation for hybrid millimeter wave systems," ***IEEE J. Sel. Areas Commun.***, vol. 35, no. 7, pp. 1576–1590, Jul. 2017.

2. **L. Zhao**, G. Geraci, T. Yang, D. W. K. Ng, and J. Yuan, "A tone-based AoA estimation and multiuser precoding for millimeter wave massive MIMO," ***IEEE Trans. Commun.***, vol. 65, no. 12, pp. 5209–5225, Dec. 2017.

3. **L. Zhao**, Z. Wei, D. W. K. Ng, J. Yuan, and M. C. Reed, "Multi-cell hybrid millimeter wave systems: Pilot contamination and interference mitigation," accepted, ***IEEE Trans. Commun.***, Jun. 2018.

4. Z. Wei, **L. Zhao**, J. Guo, D. W. K. Ng, and J. Yuan, "Multi-Beam NOMA for Hybrid mmWave Systems," under review, ***IEEE Trans. Commun.***, Jun. 2018.

Conference Articles:

1. **L. Zhao**, T. Yang and G. Geraci and J. Yuan "Downlink multiuser massive MIMO in Rician channels under pilot contamination" in ***Proc. IEEE Intern. Commun. Conf. (ICC)***, Kuala Lumpur, May 2016, pp. 1–6.

2. **L. Zhao**, D. W. K. Ng, and J. Yuan, "Multiuser precoding and channel estimation for hybrid millimeter wave MIMO systems," in ***Proc. IEEE Intern. Commun. Conf. (ICC)***, Paris, May 2017, pp. 1–7.

3. **L. Zhao**, Z. Wei, D. W. K. Ng, J. Yuan and M. C. Reed, "Mitigating pilot contamination in multi-cell hybrid millimeter Wave Systems," in ***Proc. IEEE Intern. Commun. Conf. (ICC)***, Kansas City, May 2018.

4. Z. Wei, **L. Zhao**, J. Guo, D. W. K. Ng, and J. Yuan, "A multi-beam NOMA framework for hybrid mmWave systems," in ***Proc. IEEE Intern. Commun. Conf. (ICC)***, Kansas City, May 2018 (ICC best paper awards).

Abbreviations

3-D	Three-dimensional
5G	Fifth-generation
AoA	Angle-of-arrival
AoD	Angle-of-departure
ADC/DAC	Analog-to-digital converter/Digital-to-analog converter
AWGN	Additive white Gaussian noise
BS	Base station
CSI	Channel state information
CW	Continuous wave
CRLB	Cramér Rao lower bound
dB	Decibel
DS	Delay spread
D2D	Device-to-device
DPC	Dirty paper coding
FDD	Frequency division duplex
HPBW	Half-power beam width
i.i.d.	Independent and identically distributed
ISI	Inter-symbol interference
I/Q	In-phase/Quadrature
IoT	Internet-of-things
LOS	Line-of-sight
LO	Local oscillator
LS	Least squares

MU	Multi-user
mmWave	Millimeter wave
MIMO	Multiple-input multiple-output
MMSE	Minimum mean square error
MRC	Maximal-ratio combining
MRT	Maximal-ratio transmission
MSE	Mean squared error
NMSE	Normalized mean squared error
NLOS	Non line-of-sight
NOMA	Non-orthogonal multiple access
OLB	Open-loop beamforming
PA	Power amplifier
PDP	Power delay profile
PAC	Pilot-aided-CSI
PESA	Passive electronically scanned array
QoS	Quality-of-service
RA	Random angle
RF	Radio frequency
RCS	Radar cross-section
RACH	Random access channel
r.m.s.	Root mean square
SS	Synchronization signals
SLOS	Strongest LOS
SLPS	SLOS-plus-scattering components
SNR	Signal-to-noise ratio
SINR	Signal-to-interference-plus-noise ratio
SVD	Singular value decomposition
TDD	Time division duplex
UMi	Urban micro-cell

UPA	Uniform panel array
ULA	Uniform linear array
US	User scheduling
WLAN	Wireless local area networks
ZF	Zero-forcing

List of Notations

Boldface upper-case letters denote matrices, boldface lower-case letters denote vectors, and italics denote scalars.

$\mathbf{X}^T$ Transpose of $\mathbf{X}$

$\mathbf{X}^*$ Complex conjugate of $\mathbf{X}$

$\mathbf{X}^H$ Conjugate transpose of $\mathbf{X}$

$\mathbf{X}^{-1}$ Inverse of $\mathbf{X}$

$\mathbf{X}\{i,j\}$ The element in the row i and the column j of $\mathbf{X}$

$\det(\mathbf{X})$ Determinant of $\mathbf{X}$

$\mathrm{tr}(\mathbf{X})$ Trace of $\mathbf{X}$

$|x|$ Absolute value (modulus) of the a complex scalar x

$\mathbb{C}^{M \times N}$ The space of all $M \times N$ matrices with complex entries

$\|\cdot\|_\mathrm{F}$ Frobenius norm of a vector or a matrix

$\mathbf{0}$ Zero matrix. A subscript can be used to indicate the dimension

$\mathbf{I}_\mathrm{N}$ N dimension identity matrix

j Imaginary unit $j = \sqrt{-1}$

$\mathbb{E}_x[\cdot]$ Statistical expectation with respect to random variable x

$\mathcal{CN}$ Complex Gaussian distribution

$\ln(\cdot)$ Natural logarithm

$\log_2(\cdot)$ Logarithm in base two

$\lambda_i(\cdot)$ The i-th maximum eigenvalue of a matrix

$\mathrm{diag}\{\boldsymbol{a}\}$ A diagonal matrix with the entries of $\boldsymbol{a}$ on its diagonal

$\lim$ Limit

$\mathrm{sinc}(x)$ Sinc function with input x

$\max\{\cdot\}$	Maximization
$\min\{\cdot\}$	Minimization
s.t.	Subject to

Contents

List of Figures

List of Algorithms

Chapter 1

Introduction

The significantly increased data rate requirement and quality-of-service (QoS) triggered by the proliferation of smartphones and tablets cannot be satisfied by current communication networks. In particular, the tremendous number of devices and appliances are expected to connect wirelessly to the Internet, e.g. device-to-device (D2D) communications and internet-of-things (IoT) applications, the required capacity exceeds the limit of existing networks. To address the aforementioned requirements, the fifth generation (5G), the true revolution of technologies in both the radio access network and the mobile core network, is expected to support $100 - 1000$ times higher system capacity than current the fourth generation (4G) systems and have attracted tremendous interests from both academia and industry, e.g. [1–30]. For example, 5G should support 10 Gb/s peak data rate, 50 Mb/s guaranteed data rate, 10 Tb/s/km^2 mobile data volume, less than 1 ms end-to-end latency, 1 million/km^2 number of devices, higher than 99.999% reliability, and less than 1 meter outdoor terminal location accuracy. As a result, denser network, larger bandwidth, and higher spectral efficiency are necessary to satisfy these demands. To meet the stringent spectral efficiency requirement, massive multiple-input multiple-output (MIMO)

technology is introduced by using hundreds of antennas equipped at the base station (BS) to serve tens of users simultaneously [27, 28, 31, 32]. It is proved that, by adopting low computational complexity linear precoding schemes, such as maximal-ratio transmission (MRT) and zero-forcing (ZF), massive MIMO systems can significantly improve the spectral efficiency and fully explore a large amount of available spatial degrees of freedom [14, 15, 28, 33–35]. Besides, to overcome the bandwidth limitation in licensed band, millimeter wave (mmWave) systems that migrate from sub-6 GHz to higher frequencies in unlicensed band can inevitably offer a huge trunk of a bandwidth of the order of gigahertz, e.g. an unlicensed spectrum ranging from 30 GHz to 300 GHz, to achieve ultra-high data rate communication [2, 3, 5, 17–21, 33, 36–39]. In practice, mmWave and massive MIMO can complement each other to overcome some drawbacks in each technology. For example, due to the short wavelength of mmWave frequencies, mmWave technology can shrink the physical size of antenna arrays of massive MIMO systems. Therefore, the antenna array of mmWave massive MIMO is small which enables flexible and practical deployment. On the other hand, massive MIMO, equipping a BS with hundreds of antennas, can provide significant antenna array gains to help mmWave communication systems to compensate the inherent high propagation path loss, low penetration coefficients, and high signal attenuation caused by raindrop absorption [29, 40]. Thus, the combination of mmWave communication systems with massive MIMO systems is considered as one of the promising candidate technologies for 5G communication systems with many potential opportunities for research [3, 18, 33, 37, 41].

1.1   Background

In this section, we will briefly introduce the concepts of massive MIMO and the characteristics of mmWave channels.

1.1.1   The Concepts of Massive MIMO

Massive MIMO technology is one of the key technologies for 5G. Compared with the conventional multi-user (MU) MIMO systems, massive MIMO is a special form of MIMO systems with hundreds of antennas equipped at the BS to simultaneously serve tens of users in their cells [42]. In general, a large amount of experimental and theoretical research shows that several favorable properties can be obtained when the number of antennas equipped at the BS is significantly larger than the number of users.

There are several main advantages when the number of antennas is sufficiently large: (1) the effects of small-scaling fading vanish (also known as channel hardening); (2) simple linear signal processing algorithms can achieve considerable rate performance compared to optimal algorithms, e.g. dirty paper coding (DPC); (3) higher spectral efficiency and higher energy efficiency [28, 43–47]. The more antennas the transmitter/receiver are equipped with, the better performance in terms of data rate and link reliability the system can achieve. However, the cost, the energy consumption, and the complexity of hardware (power amplifier (PA) and analog-to-digital converter/digital-to-analog converter (ADC/DAC)) increase with the increasing number of antennas [14, 21, 27, 37, 48].

Key Aspects of Massive MIMO:

A lot of efforts have been dedicated to the investigation of massive MIMO from different aspects, e.g. architecture aspects [16, 43, 49–51], capacity and fundamental

aspects [28, 44–46, 52–54], channel state information (CSI) acquisition [2, 55–63], detection algorithms [36, 64–67], downlink precoding design [3, 20, 68–73], channel measurements and modeling [51, 74, 75], resource allocation [15, 42, 76–80], hardware impairments [1, 80–83], beamforming and precoding [84–89] and performance analysis [42, 76, 90–97]. These aspects of massive MIMO had been well studied and interested readers may refer to these works for details.

1.1.2 MmWave Channels Statistics

The study of mmWave channels can be dated back to decades ago [98, 99]. Due to the high propagation path loss, mmWave technology has not been widely adopted for cellular application in the past.

Nowadays, the spectral resource below 6 GHz becomes tense while sufficiently large bandwidth is still available at the mmWave band. Thus, large efforts have been devoted to the research of mmWave communication and there are some practical communication systems utilizing the mmWave band, e.g. mmWave wireless local area networks (WLAN) IEEE 802.11 ad. However, these frequencies have not been well-explored for cellular applications. As mmWave technology is considered as a key feature for 5G, the features of mmWave channels must be considered carefully for system design.

Path Loss:

According to the free space propagation path loss model provided by the Friis transmission formula, the received power can be expressed as [100]

$$P_{\mathrm{r}}(d_{\mathrm{r}}) = P_{\mathrm{t}} G_{\mathrm{t}} G_{\mathrm{r}} \left(\frac{\lambda}{4\pi} \right)^2 d_{\mathrm{r}}^{-n}, \tag{1.1}$$

where λ is wavelength, d is the transmission distance between the desired BS and the user, n is the path loss exponent, and P_t is the transmit power. In addition, G_t and G_r are the antenna gains of the transmitter and receiver, respectively.

We can rewrite Equation (1.1) in the form of path loss. The commonly adopted empirical propagation path loss model is considered as a function of distance and carrier frequency, which is given by

$$\varpi \text{ [dB]} = 10\alpha \log_{10} d_r + \xi \log_{10}\left(4\pi\frac{1}{\lambda}\right), \tag{1.2}$$

where in a log-log plot of the path loss curve, slope α and intercept ξ are estimated by least squares (LS) linear regression over the measured data [101, 102]. For different scenarios, α and ξ may have different values, e.g. line-of-sight (LOS) environment and non line-of-sight (NLOS) environment. In addition, the propagation path loss model should also take into account the impacts of rain attenuation, atmospheric absorption, and shadowing [18, 29, 101–103].

Generally speaking, the severe path loss of mmWave channels is one of the key technical challenges in preventing mmWave technology to be implemented in conventional cellular communication systems [29, 101, 102, 104]. Nowadays, there are two methods to overcome such a challenge, e.g. increasing antenna array gain and decreasing the propagation distance. In fact, it is expected that small cell will serve as a core structure of future cellular systems [49, 105]. Hence, applying mmWave technology to small-cell urban mobile networks becomes a new trend [106]. In particular, it can significantly reduce the requirement of transmit power of mmWave systems. In addition, the high antenna array gain can be enabled by adopting the massive numbers of antennas.

Power Delay Profile and Delay Spread:

The power delay profile gives the intensity of a signal received through a multi-path channel as a function of time delay. The shape of power delay profile (PDP) in mmWave channel measurements is a superposition of multiple exponentially decaying spectrums. The delay spread (DS) is usually quantified through root mean square (r.m.s.) and considered as a collection of the multi-path containing significant energy which spread over the PDP. Besides, the DS of the channel is an important metric to understand the required system overhead to facilitate communication [107]. To provide sufficient channel characteristics information for the system design (concerning the length of cyclic prefix for a multi-carrier design), it is necessary to proceed the PDP estimation in mmWave channel measurements, which is called channel sounding.

Channel sounding for the estimation of PDP and calculating DS can be performed with both omni-directional antenna as well as directional antenna array/horn antennas. The basic procedure of channel sounding with directional horn antennas is listed as follows [107–109]:

- Perform an azimuthal scan and produce $\dfrac{360}{T}$ number of slices (the beamwidth of main lobe of directional horn antenna can cover T degrees)

- Mark the absolute propagation times for different sounding slices

- Synthesize the equivalent omni-directional PDP over an absolute propagation time axis

The PDP sounding and DS calculation for the indoor environment can be conducted via omni-directional antennas due to the short propagation distance. The results illustrate that for most indoor scenarios the DS estimated via directional antennas as well as omni-directional antennas are small, e.g. $30 - 50$ ns

indoor office and $50-90$ ns indoor shopping mall [107–109]. However, the outdoor PDP environment sounding should be conducted via directional horn antennas with large antenna-gain to compensate the severe propagation path loss. Then, by synthesizing the estimated PDP via directional antennas over an absolute propagation time axis, we can obtain the equivalent omni-directional PDP. For the equivalent omni-directional PDP, there are different outdoor scenarios where a significantly large DS have been observed, e.g. $150-300$ ns for street canyon settings and even up to $800-1000$ ns for open square settings [107–109]. In addition, the simulation results conducted by adopting ray tracing demonstrate that scattering and reflections from small objects affect the mmWave propagation channel [102, 108, 110]. These simulation and experiment results for outdoor scenarios can be explained with radar cross-section (RCS) effect, where small objects that do not participate in electromagnetic propagation at low frequencies show up at higher frequencies [102, 107, 108]. According to the simulation and experiment results as mentioned earlier, the sparse scattering assumption of mmWave channels holds for most indoor scenarios. The omni-directional delay spread is comparable with after-beamforming delay spread. However, for non-sparse mmWave channels, the omni-directional delay spread is significantly larger than that of after-beamforming delay spread. Thus, how to support these extremes without incurring a high amount of signaling overhead is an important unsolved problem. Fortunately, field test results showed that the after-beamforming delay spread could be significantly smaller than the omni-directional delay spread [107–109]. Furthermore, the blocking capability improves with the increasing number of antennas equipped at the BS/horn antenna gains [1, 111]. Specifically, the beamwidth of the main lobe becomes narrower and the magnitude of sidelobes becomes lower when the number of antennas becomes larger [1, 102, 111]. Thus,

the non-sparse scattering environment and the analog beamforming technology should be taken into account in the design of future mmWave systems, e.g. channel estimation and frequency selectivity [29, 101, 102, 107–109].

1.2   Challenges and Motivations

There are plenty of implementation challenges for mmWave massive MIMO communication systems, such as efficient hardware structures, low-cost hardware elements, hybrid precoding strategies, and channel estimation algorithms. In this thesis, channel estimation for mmWave hybrid systems as well as the impact of hardware impairments on rate performance are the primary focus of this thesis.

1.2.1   Hardware Constraint: Fully Digital vs Hybrid Structure

From the literature, it is certain that conventional fully digital MIMO systems, in which each antenna connects with a dedicated radio frequency (RF) chain, are impractical for mmWave systems due to the prohibitively high cost, e.g. tremendous energy consumption of high-resolution ADC/DAC and PAs [17, 112–114]. As a result, several mmWave hybrid systems were proposed as compromised solutions which strike a balance between hardware complexity and system performance [3, 18, 20, 33, 37, 115–118], such as fully access hybrid structure, subarray hybrid structure and lens-based hybrid structure [17, 36, 112, 113], as shown in Figure 1.1. In particular, the trade-offs between system performance, hardware complexity[1], and energy consumption are still unclear [17, 112].

In the hybrid structure, N_S data streams are first through the baseband digital

---

[1]The hardware includes PA, ADC/DAC, phase shifters, and antenna array.

Fully access structure   Subarray structure   Lens antenna structure

(a)   (b)   (c)
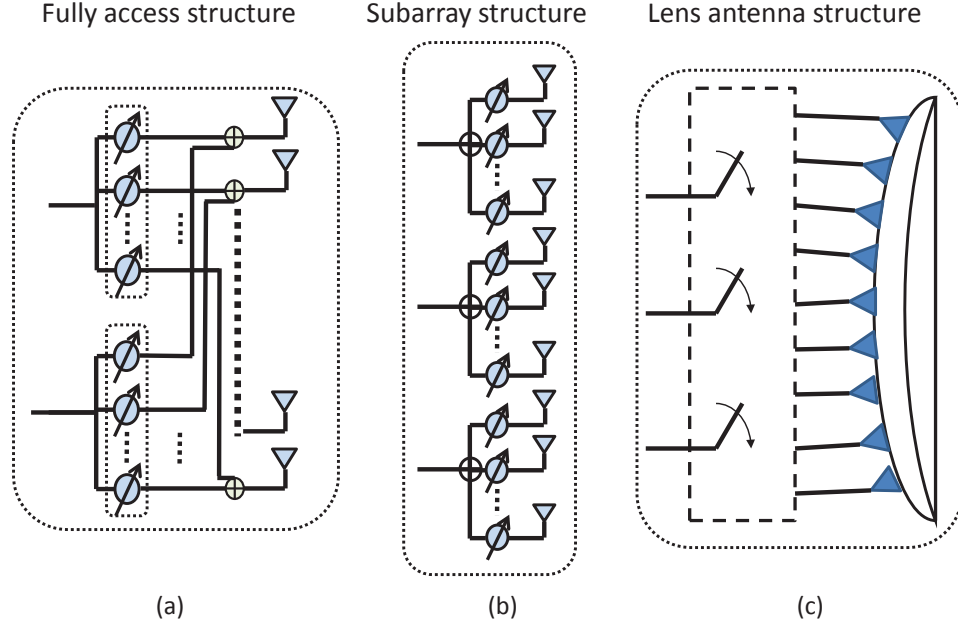
Figure 1.1: Illustration of different hybrid structures.

precoding, and then the output is through the N RF chains. After the RF analog beamforming, the RF signals are outputted to the M antennas. With $M \gg N$, the number of RF chains can be significantly reduced for practical implementation. Specifically, the use of a large number of antennas, connected with only a small number of independent RF chains at transceivers, is adopted to exploit the large array gain to compensate the inherent high path loss in mmWave channels [29, 102]. In the fully access structure shown in Figure 1.1(a), an RF chain is connected to all the antennas through the network of analog phase shifters. In the subarray structure shown in Figure 1.1(b), the array is divided into sub-arrays, and each sub-array is fed by its RF chain. The lens antenna array structure shown in Figure 1.1(c), which is formed by an electromagnetic lens with energy focusing capability and a matching antenna array with elements located on the focal surface, can simultaneously realize signal-emitting and phase-shifting. In

general, the hybrid system imposes a restriction on the number of RF chains which introduces a paradigm shift in the design of both resource allocation algorithms, transceiver signal processing, and channel estimation algorithms.

1.2.2   Hybrid Precoding: Basic Concepts

For a hybrid mmWave system, the utmost important issues should be developing channel estimation algorithms and hybrid precoding algorithms [18].

To address the importance in designing hybrid precoding algorithms, we take a MU hybrid mmWave system with the fully access structure as an example, which consists of one M-antenna BS with N RF chains and K single-RF-chain users with P-antenna.

During the downlink transmission, the signal received at user k is given by

$$y_k = \boldsymbol{f}_k^H \mathbf{H}_k \mathbf{F}_{\mathrm{BS}} \mathbf{W}_{\mathrm{BS}} \mathbf{s} + \boldsymbol{f}_k \mathbf{n}_k, \tag{1.3}$$

where $\mathbf{H}_k \in \mathbb{C}^{P \times M}$ is the downlink mmWave channel between the BS and user k, $\mathbf{F}_{\mathrm{BS}} \in \mathbb{C}^{M \times N}$ is the analog beamformer at the BS, $\mathbf{W}_{\mathrm{BS}} \in \mathbb{C}^{N \times N}$ is the digital precoder at the BS, $\boldsymbol{f}_k^H \in \mathbb{C}^{1 \times P}$ is the analog beamformer at user k, $k \in \{1, \cdots, K\}$, $\mathbf{s} = [\ s_1, \ldots, s_N\ ] \in \mathbb{C}^{N \times 1}$ is the transmitted symbols, $\mathbb{E}[\mathbf{ss}^H] = \frac{E_{\mathrm{s}}}{N} \mathbf{I}_{\mathrm{P}}$, E_{s} is the total transmit power, $\mathbf{n}_k \in \mathbb{C}^{P \times 1}$ is an additive white Gaussian noise (AWGN) vector, $\mathbf{n}_k \sim \mathcal{CN}\left(\mathbf{0}, \sigma_n^2 \mathbf{I}_{\mathrm{P}}\right)$, and σ_n^2 is the noise variance at each antenna equipped at a user. The achievable system sum rate is given by [3, 18]

$$R_{\mathrm{SUM}} = \sum_{k=1}^{K} R_k. \tag{1.4}$$

In Equation (1.4), R_k is the rate achieved by user k and can be expressed as

$$R_k = \log_2 \left[1 + \frac{\frac{E_\text{s}}{N} |\boldsymbol{f}_k^H \mathbf{H}_k \mathbf{F}_\text{BS} \boldsymbol{w}_{\text{BS},k}|^2}{\frac{E_\text{s}}{N} |\boldsymbol{f}_k^H \sum\limits_{j \neq k}^{N} \mathbf{H}_k \mathbf{F}_\text{BS} \boldsymbol{w}_{\text{BS},j}|^2 + \sigma_n^2} \right], \tag{1.5}$$

where $\boldsymbol{w}_{\text{BS},k}$ is the k-th column of $\mathbf{W}_\text{BS}$.

The optimal joint digital precoder and analog beamformer design is to maximize the achievable sum rate. Thus, the optimal precoders design can be found by solving the following optimization problem [3, 18, 20, 36, 47, 103]:

$$\left\{ \mathbf{W}_\text{BS}^\text{opt}, \mathbf{F}_\text{BS}^\text{opt}, \{\boldsymbol{f}_k^\text{opt}\}_{k=1}^K \right\} = \arg\max \sum_{k=1}^{K} \log_2 \left[1 + \frac{\frac{E_\text{s}}{N} \left| \boldsymbol{f}_k^H \mathbf{H}_k \mathbf{F}_\text{BS} \boldsymbol{w}_{\text{BS},k} \right|^2}{\frac{E_\text{s}}{N} \left| \boldsymbol{f}_k^H \sum\limits_{j \neq k}^{N} \mathbf{H}_k \mathbf{F}_\text{BS} \boldsymbol{w}_{\text{BS},j} \right|^2 + \sigma_n^2} \right]$$

$$\tag{1.6}$$

$$\text{s.t.} \quad \text{C1}: |\boldsymbol{f}_k\{i\}|^2 = 1, \ \forall k, i,$$

$$\text{C2}: |\boldsymbol{F}_\text{BS}\{i,j\}|^2 = 1, \ \forall i, j,$$

$$\text{C3}: \|\mathbf{F}_\text{BS}\mathbf{W}_\text{BS}\|_\text{F}^2 = N,$$

where C1 and C2 are the hardware constraint for the design of analog beamformers and the total power constraint is enforced by normalizing $\mathbf{F}_\text{BS}$ in C3.

The problem in Equation (1.6) is a non-convex optimization problem. In particular, the digital precoder design is under the constraint of the number of RF chains and the designed analog beamformers. Thus, the optimal solution to the hybrid precoding design is not known in general and only suboptimal iterative solutions exist [3, 70, 112, 116, 119].

In order to avoid the exhaustive search for the optimal analog beamforming

and digital precoding design, Ref. [3] proposed a two-stage algorithm which designs analog beamformer and digital precoder separately. In the first-stage, for each user k, $k \in \{1, \cdots, K\}$, the BS and user k select their analog beamformers respectively to solve

$$\left\{ \{ \boldsymbol{f}_k^{\text{opt}} \}_{k=1}^K, \{ \boldsymbol{f}_{\text{BS},k}^{\text{opt}} \}_{k=1}^K \right\} = \arg \max \left\| \boldsymbol{f}_k^H \mathbf{H}_k \boldsymbol{f}_{\text{BS},k} \right\| \tag{1.7}$$

$$\text{s.t.} \quad \text{C1}: \ |\boldsymbol{f}_k\{i\}|^2 = 1, \ \forall k, i,$$

$$\text{C2}: \ |\boldsymbol{f}_{\text{BS},k}\{j\}|^2 = 1, \forall k, j.$$

In fact, the main idea of the first-stage is to jointly design analog beamformers at the user and the BS to maximize the desired signal power of each user. For the analog beamformer design, Equation (1.7) is widely adopted in many works [112, 120]. It is true that the power of the received signal before analog beamforming is typically low due to high mmWave path losses, additional blockage, and penetration losses. Thus, it is reasonable that the proposed analog beamformer design tries to maximize the received power of the desired signal. However, it neglects the resulting interference among different users [1, 47, 107]. Actually, the MU inter-user interference due to the analog beamformers may cause severe impacts and should be taken into account. Thus, we should consider the trade offs between the received power of the desired signal and the MU interference for the joint analog beamformer design [106–109, 118, 121]. Then, a novel low-complexity practical algorithm for the design of analog beamformers is highly desirable.

In the second-stage of the hybrid precoding design, all the users feed back the downlink equivalent CSI to the BS based on the jointly designed analog beamformers:

$$\mathbf{H}_{\text{eq}} = \left[\ \mathbf{h}_{\text{eq},1}, \ldots, \mathbf{h}_{\text{eq},K} \ \right], \tag{1.8}$$

where $\mathbf{h}_{\mathrm{eq},k} = \left(\boldsymbol{f}_k^{\mathrm{opt}}\right)^H \mathbf{H}_k \boldsymbol{f}_{\mathrm{BS},k}^{\mathrm{opt}}, \forall k \in \{1, \cdots, K\}$. The digital precoder of the BS based on the feedback of equivalent channel to manage the MU interference

$$\mathbf{W}_{\mathrm{BS}} = \mathbf{H}_{\mathrm{eq}}^H \left[\mathbf{H}_{\mathrm{eq}}\mathbf{H}_{\mathrm{eq}}^H\right]^{-1}. \tag{1.9}$$

However, explicit CSI feedback from users is still required for estimating these channel. In practice, CSI feedback may cause high complexity and extra signallings overhead. In addition, there will be a system rate performance degradation due to the limited amount of the feedback and the limited resolution of CSI quantization.

Therefore, a low computational complexity mmWave channel estimation algorithm, which does not require explicit CSI feedback, is necessary to unlock the potential of hybrid mmWave systems.

1.2.3 Combination of MmWave Systems and A Small-cell Scenario

To further improve the network spectral efficiency, the small-cell scenario has been proposed to cooperate with mmWave systems [22, 37, 119]. In fact, the combination of hybrid mmWave systems and a small-cell scenario is still unexplored and possesses many exciting research opportunities [37]. For example, cell shrinking/small-cell scenario can bring numerous benefits for hybrid mmWave systems. In particular, it facilitates the reuse of the same piece of spectrum across a large geographic area for achieving a high network spectral efficiency [122] and reducing of the severe large-scale propagation path loss by shortening the distances between transceivers [22, 119]. For the multi-small-cell mmWave network, the interference received at the desired user originates from two sources, as il-

lustrated in Figure 3 of [119]: interference among different BSs and interference within the desired cell. It is mentioned in [22] that, mmWave beams are highly directional, which completely changes the interference behavior as well as the sensitivity to beams misalignment. In particular, the interference adopts an on/off behavior where strong interference occurs intermittently [22]. With a shrinking cell radius, distances from neighbouring BSs to the desired user decrease, which may lead to severe inter-cell interference to the desired user. Recently, it is also mentioned in work [123] that the received interference significantly increases when directional transmission is simultaneously adopted in both transceivers for the same amount of total emitted energy. Besides, due to the impact of imperfect CSI on the design of downlink precoder, the desired BS will cause severe intra-cell interference on the desired user. Thus, for the multi-small-cell mmWave downlink transmission with imperfect CSI and a large number of antennas, the desired user may suffer severe intra-cell and inter-cell interference. Hence, it is necessary to study the performance in the multi-small-cell hybrid mmWave network.

1.2.4   MmWave Channel Estimation

Several channel estimation algorithms are widely adopted in different works, such as multi-resolution hierarchical codebook algorithms, compress sensing algorithms, open-loop beamforming algorithms, and pilot-aided algorithms, cf. [5, 24, 26, 37, 57, 59, 60, 112, 124–136]. Some of the algorithms mentioned above are based on the assumption of sparse mmWave channels, which may not hold true for some scenarios as shown recent field measurements. The field test results show that both strong LOS components and non-negligible scattering components may exist in mmWave propagation channels, especially in urban areas, e.g. building valley environment [29, 101, 102, 107–109]. Therefore, we introduce two widely adopted

algorithms which do not rely on the assumption of sparse mmWave channels.

Pilot-aided Channel Estimation and Pilot Contamination:

Pilot-aided channel estimation, which relies on orthogonal pilot symbols, is widely adopted by MIMO systems and is suitable for sparse or non-sparse environments. The majority of contributions in the literature considered that massive MIMO systems rely on pilot sequences to estimate CSI [21, 28, 44, 53, 96]. Due to the limited orthogonal pilot resources, they must be reused across different cells. Thus, the reuse of orthogonal pilot sequences causes interference during channel estimation, which is known as pilot contamination [28, 53]. As a result, the downlink transmission based on the CSI obtained via contaminated pilots causes severe intra-cell and inter-cell interference in the desired cell. In fact, pilot contamination is considered as a fundamental performance bottleneck of the conventional multi-cell MU massive MIMO systems, since the resulting channel estimation errors do not vanish even if the number of antennas is sufficiently large, cf. [28, 35, 46, 53, 86, 137–139].

Recently, various algorithms [53, 58, 60, 63, 140–144] have been proposed to alleviate the impact of pilot contamination, e.g. data-aided iterative channel estimation algorithms, pilot design algorithms, multi-cell minimum mean square error (MMSE) based precoding algorithm, and so forth. However, some algorithms, e.g. multi-cell MMSE algorithm, are mostly based on the assumption that the desired BS can have perfect knowledge of covariance matrices of pilot-sharing users in neighbouring cells, which is overly optimistic. Besides, the condition that the desired BS has the perfect knowledge of covariance matrices is a necessary but not sufficient condition for pilot contamination mitigation, cf. [58,63,144]. The algorithm proposed in [58] requires that covariance matrices of pilot-sharing users in

neighbouring cells are orthogonal, which is unlikely in practice [144]. Also, the requirement of [63] for completely eliminating pilot contamination is that the number of antennas equipped at the BS and the size of a coherence time block jointly go to infinity. In the literature, most of existing multi-cell massive MU-MIMO works for pilot contamination [22, 28, 53, 60, 86] assumed that cross-cell channels from pilot-sharing users in neighbouring cells to the desired BS are Rayleigh fading channels with zero means. However, as discussed in Section IV of [145], it is probably not accurate for modeling the inter-cell interference as a Gaussian random variable with a small-cell setting. In fact, recent field measurements have confirmed that the strongest angle-of-arrival/angle-of-departure (AoA/AoD) components always exist in the inter-cell mmWave channels in small-cell systems [29, 101, 102, 146]. Besides, the mean values of cross-cell channels are not zero and different from each other. In other words, the distribution of cross-cell mmWave channels is different from that of the sub-6 GHz channels. Thus, the results obtained in works mentioned above for pilot contamination mitigation and performance analysis, e.g. [22, 28, 37, 53, 60, 144], cannot be applied directly. Furthermore, a thorough study on the impact of pilot contamination in such a practical network system with small cell radius has not been reported yet.

Open-Loop Beamforming Channel Estimation:

Channel estimation of practical mmWave systems may rely on AoA estimation in mmWave channels [32, 33, 37]. As a result, for mmWave channel estimation, the open-loop beamforming (OLB) channel estimation has been widely adopted. Generally, the OLB algorithm is widely used in analog beamforming and hybrid precoding. In particular, the BS transmits L beamforming vectors to its users. Then, each user reports the index of a beam with the largest gain in

the L beams via a feedback link with $\log_2 L$ numbers of bits. However, due to the limited amount of feedback bits, the OLB algorithm does not allow the BS to accurately estimate the channel responses of a large number of users. Interestingly, fully digital systems, which can be considered as a subcase of hybrid systems, can also adopt the OLB algorithm for AoAs estimation [37, 147]. Thus, the fully digital system, which contains more RF chains than analog and hybrid structures, can simultaneously transmit predesigned beams in all estimation directions., e.g. [29, 114, 135, 148, 149]. For the OLB channel estimation, the required number of beams depends on the required resolutions and is predesigned by the codebook. The number of beams for AoA estimation does not necessarily scale up with the number of transmit antennas. However, to serve an area with an increasing user density, it is expected that the BS should increase the spatial resolution. For example, if the system requires an half-power beam width (HPBW) coverage for the OLB training, that the required number of beams should be $\frac{2M}{1.782}$ [111], where M is the number of antennas equipped at the BS. In this case, the required number of beams for the OLB training scales with an increasing number of antennas, which may not be suitable for mmWave systems employing massive MIMO technology. In fact, it is expected that the required coherence time resources of channel estimation for mmWave massive MIMO systems shall neither scale up with the number of users nor with the number of antennas [33, 37, 41]. Therefore, a simple AoA channel estimation algorithm with low overhead is highly desirable for hybrid mmWave system design.

1.2.5 TDD or FDD Feedback?

In practice, the conventional fully digital systems and the hybrid systems are designed based on different hardware architectures. Hence, intuitively, the channel

estimation algorithms for hybrid mmWave systems are different from that for fully digital systems. Currently, the majority of contributions in the literature focus on the development of CSI feedback based channel estimation methods for frequency division duplex (FDD) hybrid mmWave systems, e.g. [3, 10, 32, 96, 118, 150]. This is motivated by the assumption of the sparsity of mmWave channels that the numbers of resolvable AoA/AoD paths are finite and limited. Thus, the CSI acquisition via feedbacks only leads to a small amount of signaling overhead compared to non-sparse CSI acquisition. Generally, due to the high propagation path loss, the sparsity may only exist in outdoor long distance propagation mmWave channels [102]. In some scenarios, the assumption of the sparsity of mmWave channel may not hold anymore. For example, for practical urban micro-cell (UMi) scenarios, such as the city center, the number of scattering clusters increases significantly and the channels are expected to be non-sparse. In [29, 101, 102], recent field test results, as well as ray-tracing simulation results, have shown that reflections from street signs, lamp posts, parked vehicles, passing people, etc., could reach a receiver from all possible directions in urban micro-cell (UMi) scenarios. In other words, the AoAs of scattering components between the users and the BS are uniformly distributed between $[0, \pi]$. In [151], the authors revealed that the presence of significant amount of scatterers in propagation environment has a non-negligible impact on the achievable rate performance of FDD systems when the conventional CSI feedback algorithms are adopted. Besides, the required amount of feedback signalling overhead increases tremendously with the number of scattering components, consuming a significant portion of system resources [3]. Moreover, recent field test results have also verified that, for both fully digital and hybrid massive MIMO systems in different channel environments, time division duplex (TDD) based beamforming exploiting channel reciprocity for channel es-

timation can outperform the FDD based beamforming scheme utilizing CSI feedback in terms of achievable sum-rate [107, 151]. To overcome the aforementioned common drawbacks of conventional CSI feedback based FDD mmWave channel estimation algorithms, a TDD-based beamforming channel estimation algorithm for mmWave channels is strongly welcomed. However, time synchronization and calibration issues between the forward and reverse links need to be taken into account [18].

1.2.6 Hardware Impairments

In practice, considering the high cost and the high complexity of hardware for high frequency bands, it is likely to build mmWave systems with low-cost components that are prone to hardware imperfection. Thus, constraints due to non-ideal hardware should be taken account in the designs of mmWave MIMO systems. In practice, hardware components may have various types of impairments that may degrade the achievable rate performance. For example, practical transceiver hardware is impaired by phase noise, limited phase shifters resolution, non-linear power amplifiers, in-phase/quadrature (I/Q) imbalance, and limited ADC/DAC resolution [83, 88, 114, 136, 152–154].

Generally, three different hardware impairments are considered in the literatures, e.g. additive distortions, incorporate multiplicative phase-drifts, and quantization loss caused by low-resolution ADC/DAC. The authors of [152] proved that for a fully digital massive MIMO system, the additive distortion caused by hardware impairments creates finite ceilings on the channel estimation accuracy and on the uplink/downlink capacity. In other words, the system performance cannot be improved by increasing the number of antennas equipped at the BS as well as the signal-to-noise ratio (SNR). In addition, the work in [153] concluded

that the impact of phase error on hybrid beamforming has a further reduction on the potential gain brought by mmWave systems. In fact, the achievable rate degradation caused by phase errors can be compensated by simply employing more transmit antennas, e.g. phase errors in phase shifters induced by thermal noise, transceiver RF beamforming errors caused by AoA estimation errors, and channel estimation errors affected by independent additive distortion noises [83, 152, 153]. However, the rate performance degradation caused by joint hardware imperfections, i.e., random phase errors, transceiver analog beamforming errors, and channel estimation errors, has not been thoroughly discussed.

1.3   Thesis Outline and Contributions

The main contributions and the outline of the thesis are summarized as follows.

In Chapter 2, we investigate channel estimation and MU downlink transmission of a TDD massive MIMO system in mmWave channels. We propose a tone-based linear search algorithm to facilitate the estimation of AoAs of the strongest LOS (SLOS) channel component as well as scattering components of the users at the BS. Based on the estimated AoAs, we reconstruct the SLOS component and scattering components of the users for downlink transmission. We then derive the achievable rates of MRT and ZF precoding based on the SLOS component and the SLOS-plus-scattering components (SLPS), respectively. Taking into account the impact of pilot contamination, our analysis and simulation results show that the SLOS-based MRT can achieve a higher data rate than that of the traditional Pilot-aided-CSI based (PAC-based) MRT, under the same mean squared error (MSE) of channel estimation. As for ZF precoding, the achievable rates of the SLPS-based and the PAC-based are identical. Furthermore, we quantify the achievable rate degradation of the SLOS-based MRT precoding caused

by phase quantization errors in the large number of antennas regime. We show that the impact of phase quantization errors on the considered systems cannot be mitigated by increasing the number of antennas and therefore the resolutions of RF phase shifters is critical for the design of efficient mmWave massive MIMO systems.

The results in Chapter 2 have been presented in the following publications:

- **L. Zhao**, G. Geraci, T. Yang, D. W. K. Ng, and J. Yuan, "A tone-based AoA estimation and multiuser precoding for millimeter wave massive MIMO," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5209–5225, Dec. 2017.

- **L. Zhao**, T. Yang and G. Geraci and J. Yuan "Downlink multiuser massive MIMO in Rician channels under pilot contamination" in *Proc. IEEE Intern. Commun. Conf. (ICC)*, Kuala Lumpur, May 2016, pp. 1–6.

In Chapter 3, we develop a low-complexity channel estimation for hybrid mmWave systems, where the number of RF chains is much lower than the number of antennas equipped at each transceiver. The proposed mmWave channel estimation algorithm first exploits multiple frequency tones to estimate the strongest AoAs at both BS and user sides for the design of analog beamforming matrices. Then all the users transmit orthogonal pilot symbols to the BS along the directions of the estimated strongest AoAs in order to estimate the channel. The estimated channel will be adopted to design the digital ZF precoder at the BS for the multi-user downlink transmission. The proposed channel estimation algorithm is applicable to both non-sparse and sparse mmWave channel environments. Furthermore, we derive a tight achievable rate upper bound of the digital ZF precoding with the proposed channel estimation algorithm scheme. Our analytical and simulation results show that the proposed scheme obtains a considerable achievable rate of fully digital systems, where the number of RF chains equipped

at each transceiver is equal to the number of antennas. Besides, by taking into account the effect of various types of errors, i.e., random phase errors, transceiver analog beamforming errors, and equivalent channel estimation errors, we derive a closed-form approximation for the achievable rate of the considered scheme. We illustrate the robustness of the proposed channel estimation and multi-user downlink precoding scheme against the system imperfection.

The results in Chapter 3 have been presented in the following publications:

- **L. Zhao**, D. W. K. Ng, and J. Yuan, "Multi-user precoding and channel estimation for hybrid millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1576–1590, Jul. 2017.

- **L. Zhao**, D. W. K. Ng, and J. Yuan, "Multiuser precoding and channel estimation for hybrid millimeter wave MIMO systems," in *Proc. IEEE Intern. Commun. Conf. (ICC)*, Paris, May 2017, pp. 1–7.

In Chapter 4, we investigate the system performance of a multi-cell MU hybrid mmWave MIMO network. Due to the simultaneous reuse of pilot symbols among different cells, the performance of channel estimation is expected to be degraded by pilot contamination, which is considered as a fundamental performance bottleneck of conventional multi-cell MU massive MIMO networks. To analyze the impact of pilot contamination to the system performance, we first derive the closed-form approximation expression of the normalized MSE of the channel estimation algorithm proposed in [1] over Rician fading channels. Our analytical and simulation results show that the channel estimation error incurred by the impact of pilot contamination and noise vanishes asymptotically with an increasing number of antennas equipped at each RF chain at the desired BS. Furthermore, by adopting ZF precoding in each cell for downlink transmission, we derive a tight closed-form approximation of the average achievable rate per user. Our results

unveil that the intra-cell interference and inter-cell interference caused by pilot contamination over Rician fading channels can be mitigated effectively by simply increasing the number of antennas equipped at the desired BS.

The results in Chapter 4 have been accepted to appear in the following publications:

- **L. Zhao**, Z. Wei, D. W. K. Ng, J. Yuan, and M. C. Reed, "Multi-cell hybrid millimeter wave systems: Pilot contamination and interference mitigation," accepted, *IEEE Trans. Commun.*, Jun. 2018.

- **L. Zhao**, Z. Wei, D. W. K. Ng, J. Yuan and M. C. Reed, "Mitigating pilot contamination in multi-cell hybrid millimeter Wave Systems," in *Proc. IEEE Intern. Commun. Conf. (ICC)*, Kansas City, May 2018.

In Chapter 5, future works and conclusions are presented.

Chapter 2

Fully Digital mmWave MIMO Systems: A Tone-based Channel Estimation Strategy

2.1   Introduction

In this chapter, we consider a mmWave MU massive MIMO system. First, we propose a novel and simple channel estimation algorithm, which is inspired by the signal processing of radar and sonar systems. We illustrate the proposed channel estimation for mmWave channels via analysis and simulation results. In addition, we analyze the rates achieved by utilizing MRT and ZF precoding, which are based on the estimated components of mmWave channels. Furthermore, we discuss some hardware constraints that may affect the achievable rate performance, e.g. the limited resolutions of digital RF phase shifters. Our main contributions are summarized as follows:

- We propose a novel MU channel estimation scheme via tone-based AoA

estimation. By introducing multiple frequency tones, the proposed scheme can simultaneously and efficiently estimate all users' AoAs. More importantly the accuracy of the proposed channel estimation algorithm increases with the number of antennas available at the BS. Compared to existing the pilot-aided channel estimation and the OLB channel estimation methods, our proposed AoA estimation over Rician fading channles neither causes pilot contamination nor requires a channel estimation overhead which increases with the numbers of users and antennas equipped at the BS.

- Based on various components of the estimated CSI, we derive the achievable rates of the MRT and the ZF precodings. In particular, the SLOS is adopted for the SLOS-based MRT. In addition, the combination of multiple AoAs from the users to the BS are utilized for the SLPS-based ZF downlink transmissions. We also derive closed-form approximations for the achievable rates of the MRT and the ZF precoding strategies in mmWave channels when the traditional PAC estimation is adopted. We analytically show that the achievable rates per user for the SLPS-based and the PAC-based ZF precoding strategies are identical, if the MSEs of CSI are the same. Also, it is interesting to note that the achievable rate per user of the SLOS-based MRT is better than that of the PAC-based MRT.

- To obtain system design insights, we perform an asymptotic analysis on achievable rates of the MRT and ZF precoding strategies. The analysis and simulation results verify that the mmWave massive MIMO can achieve a high sum rate performance even with dense users population. On the other hand, we show that the resolutions of RF phase shifters is critical for mmWave massive MIMO design. In fact, the negative impact of limited resolutions of shifters on the system rate performance cannot be mitigated
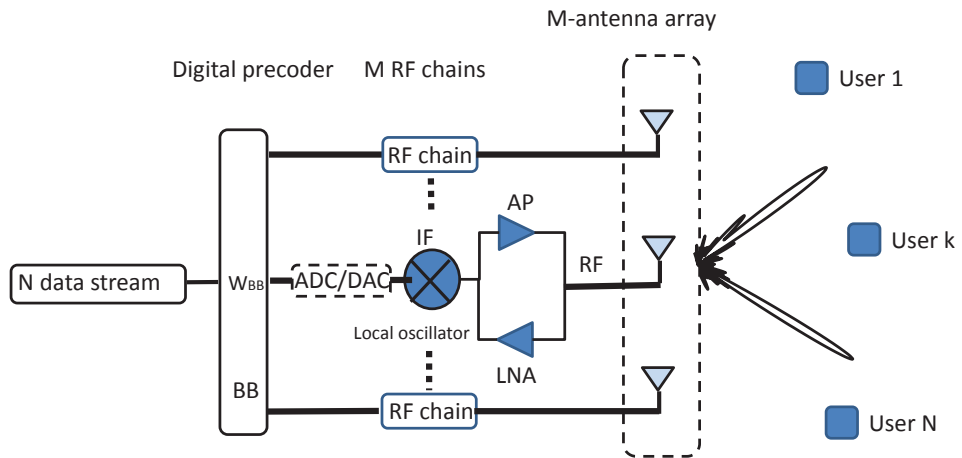
Figure 2.1: Fully digital massive MIMO architecture.

by increasing the number of antennas.

The rest of the chapter is organized as follows. Section 2.2 describes the system model considered in the chapter. In Section 2.3, we detail the proposed channel estimation algorithm. In Section 2.4, we derive the downlink achievable rate performance based on the CSI estimated by the proposed algorithm. In Section 2.5, we compare rate performance between the conventional algorithms and that investigated in Section 2.4. The impact of imperfect hardware are discussed in Section 2.6. Finally, Section 2.7 summarizes the chapter.

2.2 System Model

In this chapter, an mmWave massive MU MIMO system is considered, which is shown in Figure 2.1. The system consists of T neighboring cells and there are N users per cell. The BS in each cell is equipped with $M \geq 1$ antennas to serve the N single-antenna users simultaneously, $M \gg N$. Due to the significant propagation attenuation at mmWave, the system is dedicated to cover a small area, e.g. with a cell radius of approximately 150 m. The channels from the BS to

the users in the same cell are modeled by Rician fading and with large K-factors
($7 \sim 17$ dB) [29, 155, 156]. We assume that the users and the BSs in all cells are
fully synchronized in time and that the uplink and downlink in each cell adopt
TDD [28, 53].

Let $\mathbf{H} \in \mathbb{C}^{M \times N}$ be the uplink channel matrix between the users and the BS
in the desired cell. We assume that $\mathbf{H}$ is a narrowband slow time-varying block
fading Rician fading channel, i.e., the channel is constant in a time slot but varies
slowly from one time slot to another. Each time slot is divided into three phases:
channel estimation phase, uplink transmission phase, and downlink transmission
phase. We denote the k-th column vector of $\mathbf{H}$ as $\mathbf{h}_k \in \mathbb{C}^{M \times 1}$, representing the
channel vector between the BS and user k in the desired cell. According to the
field measurement of mmWave channels [29], a strong LOS component is expected
and other propagation paths can be considered as scattering components. We
assume that the large scale path loss can be compensated by using automatic
gain control (AGC). Thus, the channel vector of user k can be expressed as a
combination of a deterministic strongest LOS channel vector $\mathbf{h}_{\mathrm{L},k} \in \mathbb{C}^{M \times 1}$ and a
multiple-path scattered channel vector $\mathbf{h}_{\mathrm{S},k} \in \mathbb{C}^{M \times 1}$, i.e.,

$$\mathbf{h}_k = \left(\sqrt{\frac{\upsilon_k}{\upsilon_k + 1}} \right) \mathbf{h}_{\mathrm{L},k} + \left(\sqrt{\frac{1}{\upsilon_k + 1}} \right) \mathbf{h}_{\mathrm{S},k}, \tag{2.1}$$

where $\upsilon_k > 0$ is the Rician K-factor of user k, which denotes the ratio between
the power of the SLOS component and the power of the scattering components.
In general, we can re-express Equation (2.1) as

$$\mathbf{H} = \mathbf{H}_{\mathrm{L}}\mathbf{G}_{\mathrm{L}} + \mathbf{H}_{\mathrm{S}}\mathbf{G}_{\mathrm{S}}, \tag{2.2}$$

where $\mathbf{G}_{\mathrm{L}} \in \mathbb{C}^{N \times N} = \mathrm{diag}\left\{\overline{\boldsymbol{\upsilon}}_{\mathrm{L}}\right\}$, $\mathbf{G}_{\mathrm{S}} \in \mathbb{C}^{N \times N} = \mathrm{diag}\left\{\overline{\boldsymbol{\upsilon}}_{\mathrm{S}}\right\}$. In addition, we

have $\overline{v}_L = \left[\sqrt{\frac{v_1}{v_1+1}}, \cdots, \sqrt{\frac{v_N}{v_N+1}} \right]^T$, and $\overline{v}_S = \left[\sqrt{\frac{1}{v_1+1}} \cdots \sqrt{\frac{1}{v_N+1}} \right]^T$. We adopt uniform linear array (ULA) as in [135, 157]. Here, $\mathbf{h}_{L,k} \in \mathbb{C}^{M \times 1}$ is the k-th column vector of the SLOS matrix $\mathbf{H}_L$ and it can be expressed as

$$\mathbf{h}_{L,k} = \left[1, \quad \cdots, \quad e^{-j2\pi(M-1)\frac{d}{\lambda}\cos(\theta_k)} \right]^T, \tag{2.3}$$

where d is the distance between the neighboring antennas at the BS, λ is the wavelength of the carrier frequency, and $\theta_k \in [0, +\pi]$ is the AoA of the SLOS component from user k to the BS in the desired cell. For convenience, we set $d = \dfrac{\lambda}{2}$ for the rest of the chapter which is an assumption commonly adopted in the literature [158, 159]. The scattering component $\mathbf{h}_{S,k}$ is the k-th column vector of the scattering matrix $\mathbf{H}_S$ and can be expressed as

$$\mathbf{h}_{S,k} = \sqrt{\frac{1}{L}} \sum_{l=1}^{L} \alpha_{k,l} \mathbf{a}_{k,l}, \ k = 1, \cdots, N, \tag{2.4}$$

where L denotes the number of propagation paths, $\alpha_{k,l} \sim \mathcal{CN}(0, 1)$ is the complex path gain and $\theta_{k,l} \in [0, +\pi]$ is the AoA associated to the (k, l)-th propagation path [117], and $\mathbf{a}_{k,l} \in \mathbb{C}^{M \times 1}$ is the l-th propagation path of user k given by

$$\mathbf{a}_{k,l} = \left[1, \quad \cdots, \quad e^{-j2\pi(M-1)\frac{d}{\lambda}\cos(\theta_{k,l})} \right]^T. \tag{2.5}$$

We assume that all the users have various AoAs which are uniformly distributed over $[0, +\pi]$ and they are separated by at least hundreds of wavelengths [28, 160]. To facilitate the investigation of channel estimation and downlink transmission, we assume that perfect long-term power control is performed to compensate for path loss and shadowing between the desired BS and the desired users and that equal power allocation is used among different data streams of the users [3, 46, 115].

2.3   Channel Estimation and Downlink Transmission

In this section, we first propose a novel MU mmWave channel estimation algorithm based on various single carrier frequency tones, which is inspired by signal processing of radar and sonar systems [147, 161]. Based on the estimated SLOS component and scattering components, we propose the SLOS-based MRT and the SLPS-based ZF precoders for MU downlink precoding. In addition, the achievable rate performance of these two precoders are derived and used for comparison in the following section.

2.3.1   MU MmWave Channel Estimation

We note that the channel estimation of mmWave channels is equivalent to the AoA estimation of the SLOS and scattering components. In this section, we propose to use the frequency tone resources to simultaneously estimate the AoAs of the users without causing collision among different cells. The details are discussed in the following paragraphs.

At the beginning of a coherence time block, all the users register at the desired BS and transmit unique tones separated in frequency. This initialization can be done during the time synchronization in the handshaking between the desired BS and the users. The frequency tone of user k for the AoA estimation is a single carrier continuous wave (CW) signal, $s_k = \sin(2\pi f_k t)$, where f_k is the carrier frequency for user k and t is the time variable of the tone signal. We also denote that f_c is the system carrier frequency, BW is the total available bandwidth for all tone frequencies, and f_c is at the very center of the total bandwidth.

For the channel estimation, we choose to transmit tone signals from the users

to the BS via one of the omni-directional antennas equipped at the users.

The reasons are three:

- The utilizing of only one omni-directional antenna at the user to transmit the tone signal is to sound mmWave channels in every direction;

- In practice, the number of antennas equipped at the users is much smaller than the number of antennas equipped at the BS, M. As long as a sufficiently large number of antennas, M, is equipped at the BS, the power gain loss can be compensated;

- For the AoA estimation, tone signal is robust to the background noise. It is known that the thermal noise power is determined by the signal bandwidth and the noise spectral density level. Due to the extreme narrow bandwidth of the tone signal, the receive thermal noise power is low. Thus, it is easy to detect the tone signal even with a slightly low receive signal power [147].

For any two different frequency tones f_i and f_j, $\forall i \neq j$, in the BW, if the frequency tones satisfy the following condition,

$$\frac{f_i - f_j}{f_\mathrm{c}} < 2 \times 10^{-4}, \tag{2.6}$$

the AoA estimation performance by using ULA with any tones in BW is identical[1]. The condition also can be re-expressed as $\frac{BW}{f_\mathrm{c}} < 2 \times 10^{-4}$ [111]. For example, if the system operates at a carrier frequency of 30 GHz, then BW is 6 MHz. Any frequency tones in the frequency range of $[30\ \mathrm{GHz} - 3\ \mathrm{MHz},\ 30\ \mathrm{GHz} + 3\ \mathrm{MHz}]$ have an identical performance in the AoA estimation. However, the frequency tones will be affected due to users' doppler frequency shift f_d.

---

[1]The difference of different carrier wavelengths is less than 2 micrometers for a carrier frequency of 30 GHz.
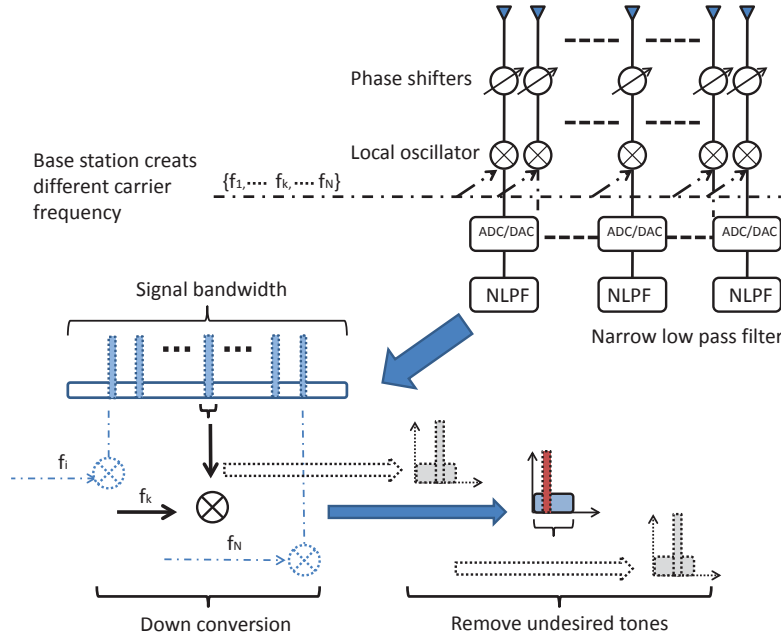
Figure 2.2: Tone-based AoA channel estimation.

In addition, to overcome the collisions caused by users' doppler frequency shift f_d between neighboring frequency tones, a frequency protectional gap $B_k = |f_k - f_{k+1}|$ need to satisfy

$$B_k > 2f_d. \tag{2.7}$$

For example, if the system operates at 30 GHz and the maximum speed of users is 72 km/h, then the minimum frequency protection gap is $B_k = 4$ kHz. In practice, a single carrier frequency CW signal can be easily generated by a commercial 30 GHz signal generator [162], which has a narrow bandwidth of $100 \sim 200$ Hz. Therefore, for a bandwidth of 6 MHz we can generate 1400 single carrier CW signals, which can estimate 1400 users' AoAs simultaneously. For mmWave massive MIMO systems, according to the 3GPP medium user density standard [163], the number of users in each cell ranges from 40 to 400. Therefore, the frequency resources can support at least three cells without causing any collision in the AoA estimation among different cells.

The basic idea of utilizing frequency tone signal for the AoA estimation is borrowed from single-carrier monopulse radar systems [147]. The key idea is that the narrow-band single-carrier frequency CW signal can be performed for the AoA estimation. Therefore, orthogonal pilot symbols are no longer required for the AoA estimation of mmWave channels. In addition, different users can be distinguished at the BS via different carrier frequencies. At the beginning of a coherence time block, all the users register at the desired BS. At the same time, the desired BS instructs all the users' unique tone carrier frequencies. At the user's side, according to the allocated tone carrier frequencies, a single-carrier frequency CW pass-band signal is created in the user's RF chain and broadcasted via one of its antennas. For example, the frequency tone of user k for the AoA estimation is $s_k = \sin(2\pi f_k t)$, where f_k is the carrier frequency for user k and t is the time variable of the tone signal. The pass-band received signal of user k at the BS is given by

$$\boldsymbol{\varphi}_k = \left(\sqrt{\frac{\upsilon_k}{\upsilon_k + 1}} \mathbf{h}_{\mathrm{L},k} + \sqrt{\frac{1}{\upsilon_k + 1}} \mathbf{h}_{\mathrm{S},k} \right) s_k + \widehat{\mathbf{z}}, \qquad (2.8)$$

where $\widehat{\mathbf{z}} \in \mathbb{C}^{M \times 1}$ is the AWGN vector at the BS antennas, whose entries follow independent and identically distributed (i.i.d.) $\mathcal{CN}(0, \sigma^2)$ and σ^2 is the noise variance at each antenna of the BS. At the BS, different frequency tones will be generated as the reference signals for the down-conversion of received signals, as shown in Figure 2.2. Take user k for example, the reference signal is $s_k = \sin(2\pi f_k t)$ at the BS and the received signal after down-conversion is given by

$$\begin{aligned}
\boldsymbol{\varphi}_k \sin(2\pi f_k t) &= \left(\sqrt{\frac{\upsilon_k}{\upsilon_k + 1}} \mathbf{h}_{\mathrm{L},k} + \sqrt{\frac{1}{\upsilon_k + 1}} \mathbf{h}_{\mathrm{S},k} \right) s_k \sin(2\pi f_k t) + \widehat{\mathbf{z}} \sin(2\pi f_k t) \\
&= \left(\sqrt{\frac{\upsilon_k}{\upsilon_k + 1}} \mathbf{h}_{\mathrm{L},k} + \sqrt{\frac{1}{\upsilon_k + 1}} \mathbf{h}_{\mathrm{S},k} \right) \frac{1}{2} \left[1 - \cos(4\pi f_k t) \right] + \widehat{\mathbf{z}} s_k. \quad (2.9)
\end{aligned}$$

Then the mixed signal will pass through different narrow low pass filters with an appropriate filter bandwidth to remove high frequency components [147]. After the high frequency components are filtered and removed, the received base-band signal which contains the AoA information can be expressed as

$$\mathbf{r}_k = \left(\sqrt{\frac{\upsilon_k}{\upsilon_k + 1}} \mathbf{h}_{\mathrm{L},k} + \sqrt{\frac{1}{\upsilon_k + 1}} \mathbf{h}_{\mathrm{S},k} \right) + \mathbf{z}. \tag{2.10}$$

To facilitate the estimation of AoAs, we perform a linear search in the angular domain ranged from $0°$ to $180°$ with an angle search step size of $\frac{180}{J}$, where J is the maximum number of search steps. The AoA detection matrix for all users is $\mathbf{\Gamma} \in \mathbb{C}^{M \times J} = \left[\; \overline{\boldsymbol{\gamma}}_1 \ldots \overline{\boldsymbol{\gamma}}_i \ldots \overline{\boldsymbol{\gamma}}_J \; \right]$, where $\overline{\boldsymbol{\gamma}}_i \in \mathbb{C}^{M \times 1}$ is a column vector of matrix $\mathbf{\Gamma}$ given by

$$\overline{\boldsymbol{\gamma}}_i = \frac{1}{\sqrt{M}} \left[\; 1, \quad \cdots \quad , \quad e^{j2\pi(M-1)\frac{d}{\lambda}\cos\left(\widehat{\theta}_i\right)} \; \right]^T, \tag{2.11}$$

where $\widehat{\theta}_i = (i-1)\frac{180}{J}$, $i \in \{1, \cdots, J\}$, stands for a potential AoA of user at the BS. We note here, the coefficient $\frac{1}{\sqrt{M}}$ appears in Equations (2.11) is the normalization factor. The size of matrix $\boldsymbol{\gamma}$ represents the computational complexity of AoA estimation. We then utilize different AoA detection vectors, e.g. $\overline{\boldsymbol{\gamma}}_i^T$, to obtain detection output of user k. Specifically, for $i \in \{1, \cdots, J\}$,

$$\boldsymbol{\omega}_{k,i} = \overline{\boldsymbol{\gamma}}_i^T \mathbf{r}_k = \sqrt{\frac{\upsilon_k}{\upsilon_k + 1}} \overline{\boldsymbol{\gamma}}_i^T \mathbf{h}_{\mathrm{L},k} + \sqrt{\frac{1}{\upsilon_k + 1}} \overline{\boldsymbol{\gamma}}_i^T \mathbf{h}_{\mathrm{S},k} + \overline{\boldsymbol{\gamma}}_i^T \mathbf{z}_k. \tag{2.12}$$

Now we obtain the detection outputs $[\boldsymbol{\omega}_{k,1} \ldots \boldsymbol{\omega}_{k,i} \ldots \boldsymbol{\omega}_{k,J}]$. The potential AoA of the SLOS, which leads to the maximum value among the J observation directions, i.e.,

$$\widehat{\mathbf{h}}_{\mathrm{L},k} = \widetilde{\boldsymbol{\gamma}}_k^* = \underset{\forall \overline{\gamma}_i, \; i \in \{1, \cdots, J\}}{\arg \max} |\boldsymbol{\omega}_{k,i}|, \tag{2.13}$$

is considered as the SLOS component of user k. If the number of search steps J is

sufficiently large, we accurately approximate the channel of user k, $\mathbf{h}_k$, by using AoAs from the D strongest detection outputs. We first sort the detection outputs $[\boldsymbol{\omega}_{k,1} \ldots \boldsymbol{\omega}_{k,i} \ldots \boldsymbol{\omega}_{k,J}]$ in descending order and obtain the ordered detection output vector as $[\boldsymbol{\omega}_{k,(1)} \ldots \boldsymbol{\omega}_{k,(i)} \ldots \boldsymbol{\omega}_{k,(J)}]$, where $|\boldsymbol{\omega}_{k,(i)}| \geqslant |\boldsymbol{\omega}_{k,(i+1)}|$. Then, we can obtain the estimated channel $\widehat{\mathbf{h}}_k$ as

$$\widehat{\mathbf{h}}_k = \sqrt{\frac{\widehat{\upsilon}_k}{\widehat{\upsilon}_k + 1}}\widetilde{\boldsymbol{\gamma}}_k^* + \sqrt{\frac{1}{\widehat{\upsilon}_k + 1}}\sum_{p=1}^{D-1}\frac{\boldsymbol{\omega}_{k,(p)}}{\boldsymbol{\omega}_{k,(1)}}\widetilde{\boldsymbol{\gamma}}_{k,p}^*, \tag{2.14}$$

where $\widetilde{\boldsymbol{\gamma}}_{k,p}^*$ is the p-th strong AoA given by

$$\widetilde{\boldsymbol{\gamma}}_{k,p}^* = \frac{1}{\sqrt{M}}\left[\begin{array}{cccc} 1, & \cdots, & e^{j2\pi(M-1)\frac{d}{\lambda}\cos\left(\widehat{\theta}_{k,p}\right)} \end{array}\right]^T, \tag{2.15}$$

and the Rician K-factor is given by

$$\frac{1}{\widehat{\upsilon}_k} = \sum_{p=1}^{D-1}\left|\frac{\boldsymbol{\omega}_{k,(p)}}{\boldsymbol{\omega}_{k,(1)}}\right|^2. \tag{2.16}$$

We note that the exact number of combined AoAs D is hard to obtain in practice. As a result, the number of combined AoAs is set as a reasonable value for practical implementation. In the sequel, we set as $D = 15$ for simplicity [102]. In fact, the tone-based AoA channel estimation error is mainly caused by the uncertainty of D.

Utilizing the estimated SLOS component as the estimated channel, the MSE of the proposed algorithm in the high SNR regime is given by

$$\text{MSE}_{\text{L},k} \overset{\text{SNR}\rightarrow\infty}{=} \frac{1}{M}\mathbb{E}\left[\left(\mathbf{h}_k - \widehat{\mathbf{h}}_{\text{L},k}\right)^H\left(\mathbf{h}_k - \widehat{\mathbf{h}}_{\text{L},k}\right)\right] = 2\left(1 - \sqrt{\frac{\upsilon_k}{\upsilon_k + 1}}\right). \tag{2.17}$$

We demonstrate and validate the accuracy of the proposed AoA estimation

via simulation in Figures 2.3(a), 2.3(b), 2.3(c), and 2.3(d). The accuracy of the AoA estimation is determined by the number of search steps J as well as the number of strongest AoAs D.
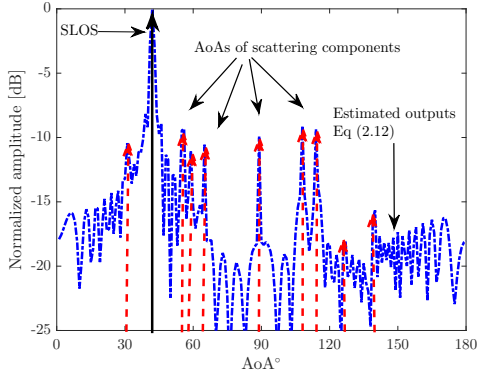
(a) Illustration of AoAs estimation.
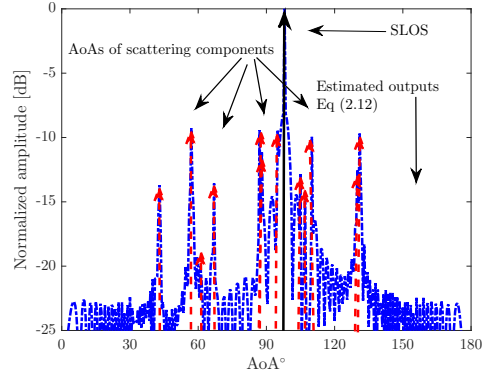
(b) Illustration of AoAs estimation.
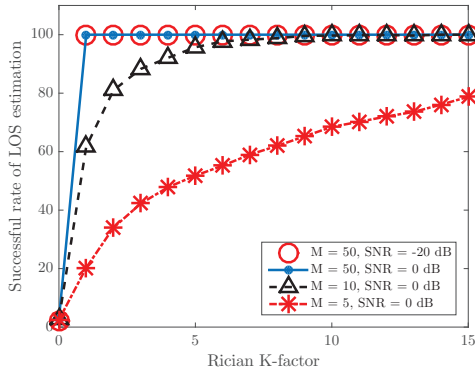
(c) Successful rate of SLOS estimation.
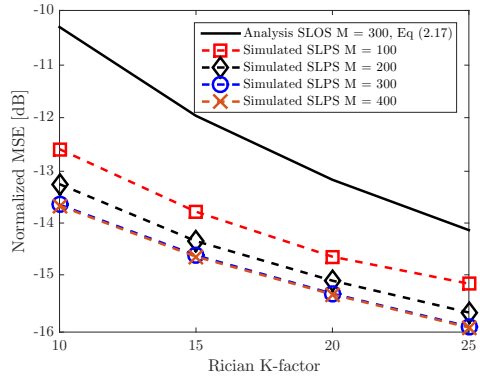
(d) Normalized MSE versus Rician K-factor.

Figure 2.3: (a) Illustration of AoA estimation for the SLOS component and scattering components for $M = 100$ and $\upsilon_k = 10$; (b) Illustration of AoA estimation for the SLOS component and scattering components for $M = 200$ and $\upsilon_k = 10$; (c) Successful rate of the SLOS estimation versus the Rician K-factor for various number of antennas; (d) Normalized MSEs of channel estimation versus the Rician K-factor for various number of antennas. For (a), (b), (c), and (d), we set the number of combination AoAs as $D = 15$ and the number of search steps as $J = 360$.

Searching for $\widehat{\theta}_i$ can be processed via baseband digital signal processing. More details on the Cramér-Rao lower-bound of the ULA AoA estimation can be found in [161]. In practice, larger number of antennas M means higher spatial resolution

for AoAs estimation.

In Figures 2.3(a) and 2.3(b), we illustrate the proposed AoA estimation for different M. The SNR of tone signal adopted in the simulations for Figures 2.3(a) and 2.3(b) is 5 dB.

With a large M, the impact of noise σ^2 on the AoA estimation is mitigated, as shown in Figures 2.3(a) and 2.3(b). These figures indicate that the proposed algorithm is robust to the AWGN when the BS is equipped with a sufficiently large number of antennas. If the norm of the angle difference between the estimated SLOS component and the actual SLOS component is less than 0.1%, we denote the estimation of the SLOS component as successful.

Figure 2.3(c) shows the successful rate of the AoA estimation achieved by the proposed algorithm versus the Rician K-factor for different number of antennas M. It is clear that the successful rate increases with the number of antennas as well as with the Rician K-factor. In addition, we also see that the successful rates of the SLOS estimation for SNR = 0 dB with $M = 50$ are almost the same as those for SNR = -20 dB with $M = 50$. Figure 2.3(d) shows the normalized MSEs of channel estimation versus the Rician K-factor for different number of antennas M. It shows that the SLPS algorithm has better MSE performance than the SLOS algorithm. In Figure 2.3(d), the correctness of Equation (2.17) is also verified. In Figures 2.3(c) and 2.3(d), we also observe that the improvement in MSE is saturated for a large number of antennas. Compared to conventional orthogonal pilot-aided and the OLB algorithm approaches for channel estimation [28, 37, 53], the proposed tone-based AoA estimation scheme is simpler and the required coherence time resources will neither increase with the number of users nor the number of antennas. In addition, the proposed channel estimation relies on the antenna array gain for AoA estimation and the tone-based AoA estimation can be
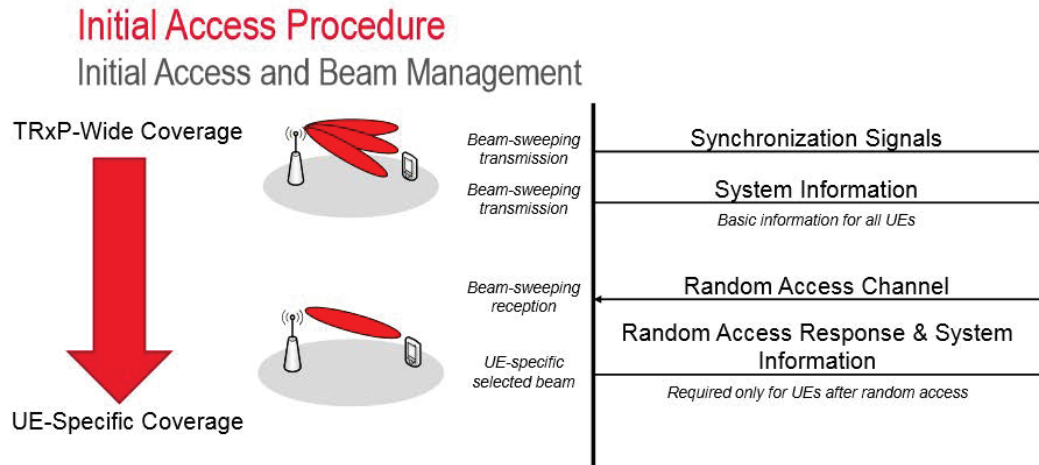
Figure 2.4: Illustration of the practical 5G NR CSI acquisition/initial access procedure.

extended to hybrid systems by following a similar approach as in [1]. Literally, the more antennas used, the more antenna array gain can be obtained. Therefore, it is expected that the required power of the transmitted tone is low in the considered massive MIMO mmWave system. More importantly, the proposed tone-based AoA estimation algorithm can avoid the system performance degradation due to pilot contamination. With a sufficiently large number of antennas, the proposed algorithm can accurately estimate the SLOS component as well as scattering components.

In mmWave systems, channel estimation is important for unlocking the potential of the system performance. Majority of contributions in the literature focus on the development of CSI feedback based channel estimation methods for hybrid mmWave systems, e.g. [3, 10, 32, 96, 118, 150]. In practice, the 5G NR specification includes a set of basic beam-related procedures for the above 6 GHz CSI acquisition [164]. The beam management of 5G NR mmWave systems consists of four different operations which are shown in Figure 2.4, i.e., beam sweeping, beam measurement, beam determination, and beam reporting. Taking a single
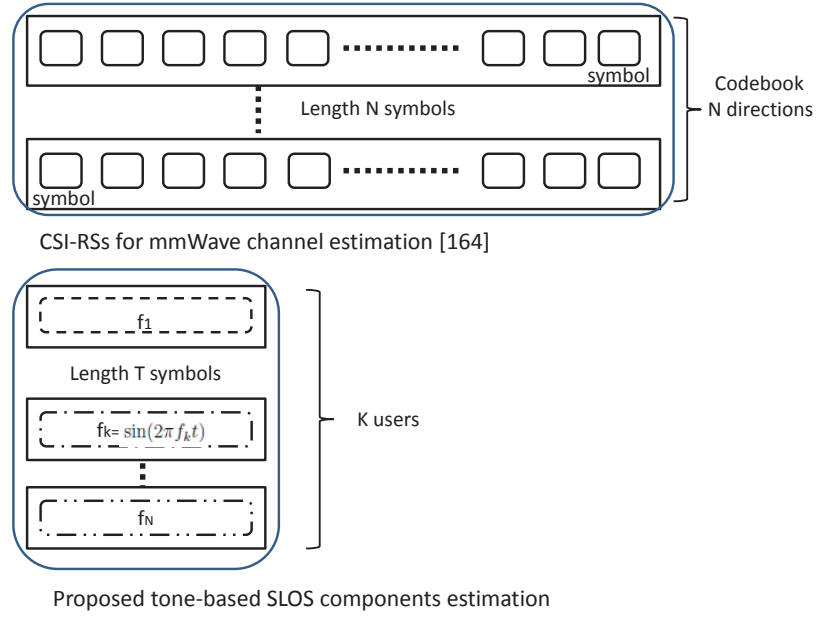
Figure 2.5: Illustration of the overhead signalling difference between the practical 5G NR CSI acquisition procedure and the proposed algorithm.

antenna user in 5G NR stand-alone (SA) downlink scheme for example, the BS sequentially transmits synchronization signals (SS) and CSI reference signals (CSI-RSs) to the users by using a predefined codebook of directions. Then, the users search and track their optimal beams by measuring the collected CSI-RSs. At the end of CSI acquisition phase, beam reporting, the users feed back the determined beam information (beam quality and beam decision) and the random access channel (RACH) preamble to the BS.

Mainly, the difference between the channel estimation algorithm adopted in practical mmWave systems and the proposed tone-based algorithm is the overhead signalling consumed for the channel estimation phase. Figure 2.5 illustrates the overhead signalling difference between our proposed tone-based algorithm and the practical 5G NR mmWave channel estimation protocol. It shows that utilizing tone signals can significantly save the overhead signalling resource. In addition, the MSE performance of aforementioned algorithms are the same. The

reason is simple and straightforward. The antenna array adopted in these two algorithms to facilitate the channel estimation is same and has the same Cramér Rao lower bound (CRLB).

In the following sections, we analyze the system rate performance of downlink transmission based on the estimated CSI via the proposed tone-based AoA estimation scheme.

2.4   Performance Analysis

For the downlink transmission performance analysis, we first discuss the adopted assumptions. In reference [101], which is based on recent field test results, the authors suggested that the angles between the users and the BS are uniformly distributed between $(0, \pi)$. In addition, reference [102] adopted ray-tracing model to characterize the AoAs of users at different locations. The ray-tracing simulation results are verified and supported by field test results (Figure 5 of [102]). It is shown that reflections from street signs, lamp posts, parked cars, passing people, etc., could reach the receiver from all directions. Based on these field test results, it is clear that the AoAs between the BS and the users are caused by many clusters in the propagation directions. In fact, the clusters around different separated users are different with high probability in the urban areas as shown in Figure 2 of [102]. Although some users may share part of the clusters [150], to simplify the average achievable rate performance analysis, we assume that the set of AoAs of multi-path between different users will be different with high probability [165] in the large number of antennas regime. In other words, the scattering components of different users are assumed not highly correlated due to the random clusters in the urban scenario in the considered system. Specifically, in the large number of antennas regime, for $\theta_{k,l} \neq \theta_{j,i}$, it is expected that

$$\mathbf{a}_{k,l}^H \mathbf{a}_{j,i} \overset{M\to\infty}{\to} 0, \ k \neq j. \tag{2.18}$$

We assume that given the random location of the scatters, the set of AoAs of multi-path between different users will be different with high probability [165]. Therefore, for $k \neq j$, we have the following result:

$$\mathbf{h}_{S,k}^H \mathbf{h}_{S,j} = \frac{1}{L} \left(\sum_{l=1}^{L} \alpha_{k,l} \mathbf{a}_{k,l}^H \right) \left(\sum_{l=1}^{L} \alpha_{j,l} \mathbf{a}_{j,l} \right) \overset{M\to\infty}{\to} 0. \tag{2.19}$$

2.4.1   SLOS-based MRT Precoding

MRT is considered as the simplest precoding strategy, due to its low computational complexity. In [159], the authors proved that single-user massive MIMO systems are more suitable to exploit strong LOS components of Rician fading channels by comparison with pure Rayleigh fading environments. In this section, we study the rate performance of the considered MU mmWave systems when M-RT downlink transmission is employed and designed based on the estimated SLOS component. To facilitate the following study, we assume that the perfect SLOS component is available[2] and the downlink MRT precoder is given by $\widetilde{\mathbf{W}}_L = \mathbf{H}_L^*$. The rate performance achieved by using the MRT precoder $\widetilde{\mathbf{W}}_L$ can be considered as the system performance benchmark. We then express the received signal of user k as

$$y_{LMRT}^k = \widetilde{\beta} \mathbf{h}_k^T \mathbf{h}_{L,k}^* x_k + \widetilde{\beta} \mathbf{h}_k^T \sum_{j=1,j\neq k}^{N} \mathbf{h}_{L,j}^* x_j + z_k, \tag{2.20}$$

---

[2]It was verified by simulation that, for the proposed SLOS estimation with a sufficiently large number of antennas at the BS, e.g. 50 antennas, our proposed SLOS estimation algorithm is highly accurate. The simulation results are shown in Figure 2.3(c).

where $\widetilde{\beta} = \sqrt{\frac{1}{\mathrm{tr}(\widetilde{\mathbf{W}}_{\mathrm{L}}\widetilde{\mathbf{W}}_{\mathrm{L}}^{H})}}$ is the power normalization factor, $x_k \in \mathbb{C}^{1 \times 1}$ is the transmitted signal from the BS to user k in the desired cell, $\mathbb{E}(x_k) = 0$, and $\mathbb{E}(x_k^2) = E_s$ is the transmit power. We then present the signal-to-interference-plus-noise ratio (SINR) expression of user k as

$$\mathrm{SINR}_{\mathrm{LMRT}}^{k} = M^2 \left[\frac{1}{\upsilon_k} \sum_{j=1}^{N} \mathbb{E}_{\mathbf{h}_{\mathrm{S}}} \left[\mathbf{h}_{\mathrm{S},k}^{H}\mathbf{h}_{\mathrm{L},j}\mathbf{h}_{\mathrm{L},j}^{H}\mathbf{h}_{\mathrm{S},k} \right] \right.$$

$$\left. + \sum_{j=1,j\neq k}^{N} \mathbf{h}_{\mathrm{L},k}^{H}\mathbf{h}_{\mathrm{L},j}\mathbf{h}_{\mathrm{L},j}^{H}\mathbf{h}_{\mathrm{L},k} + \frac{\upsilon_k+1}{\upsilon_k}\frac{\sigma^2}{E_s\widetilde{\beta}^2} \right]^{-1}, \qquad (2.21)$$

where $\frac{E_s}{\sigma^2}$ represents the transmit SNR. In the large number of antennas and high receive SNR regime, the approximated average achievable rate of the MRT precoding can be expressed as

$$R_{\mathrm{LMRT}}^{k} \underset{\substack{M \to \infty \\ \mathrm{SNR} \to \infty}}{\approx} \log_2 \left[1 + \left[\underbrace{\frac{1}{M^2}\sum_{j=1,j\neq k}^{N}\mathbf{h}_{\mathrm{L},k}^{H}\mathbf{h}_{\mathrm{L},j}\mathbf{h}_{\mathrm{L},j}^{H}\mathbf{h}_{\mathrm{L},k}}_{\text{Part 1}} + \underbrace{\frac{N}{M\upsilon_k}}_{\text{Part 2}} \right]^{-1} \right], \qquad (2.22)$$

where Part 1 in Equation (2.22) is considered as the interference to the desired user k caused by SLOS components of different users. In addition, Part 1 decreases significantly with an increasing number of antennas M. Part 2 in Equation (2.22) is considered as the interference caused by the scattering parts, which is determined by the Rician K-factor and the number of users to the number of antennas ratio.

2.4.2 SLPS-based ZF Precoding

ZF precoding is widely adopted for MU systems due to its capability in interference suppression. It was verified by simulation that the estimated SLPS can achieve a better MSE performance in channel estimation than the estimated S-

LOS, which is shown in Figure 2.3(d). While the estimated SLPS is adopted for the design of ZF precoder, the achievable rate of the SLPS-based ZF can be considered as the rate performance benchmark. Based on the estimated SLPS-based uplink channel $\widehat{\mathbf{H}}$, the downlink precoder is expressed as

$$\widehat{\mathbf{W}} = \left(\widehat{\mathbf{H}}\right)^* \left[\left(\widehat{\mathbf{H}}\right)^T \left(\widehat{\mathbf{H}}\right)^*\right]^{-1}. \tag{2.23}$$

The received signal at user k in the desired cell is given by

$$y_{\mathrm{ZF}}^k = \widehat{\beta}\mathbf{h}_k^T\widehat{\mathbf{w}}_k x_k + \widehat{\beta}\mathbf{h}_k^T \sum_{j=1,j\neq k}^{N} \widehat{\mathbf{w}}_j x_j + \widehat{\beta}\sum_{p=1}^{T}\rho_{p,k}\mathbf{h}_{p,k}^T\widehat{\mathbf{w}}_{p,k}x_p + z_k, \tag{2.24}$$

where the variable p is the number of neighbouring cells, $\widehat{\mathbf{w}}_k \in \mathbb{C}^{M\times 1}$ is the k-th column vector of $\widehat{\mathbf{W}}$ and $\widehat{\beta} = \sqrt{\frac{1}{\mathrm{tr}(\widehat{\mathbf{W}}\widehat{\mathbf{W}}^H)}}$ is the normalization factor. For the ZF precoder under the perfect CSI, the precoder matrix is given by $\mathbf{W} = \mathbf{H}^* \left(\mathbf{H}^T\mathbf{H}^*\right)^{-1}$. Let $\mathbf{w}_k \in \mathbb{C}^{M\times 1}$ denote the k-th column vector of $\mathbf{W}, \forall i \neq k, \mathbf{h}_i^T\mathbf{w}_k = 0$, and $\mathbf{h}_k^T\mathbf{w}_k = 1$. We re-express Equation (2.24) as

$$y_{\mathrm{ZF}}^k = \underbrace{\widehat{\beta}x_k}_{\text{desired signal}} + \underbrace{\widehat{\beta}\mathbf{h}_k^T\Delta\mathbf{W}\mathbf{x}}_{\text{intra-cell interference}} + z_k, \tag{2.25}$$

where $\Delta\mathbf{w}_j \in \mathbb{C}^{M\times 1}$ denotes the j-th column vector of the ZF precoder error matrix $\Delta\mathbf{W} = \widehat{\mathbf{W}} - \mathbf{W}$, and $\mathbf{x} = [x_1,\ldots,x_N]^T$ denotes the transmitted signal vector. We then express the SINR of user k as

$$\mathrm{SINR}_{\mathrm{ZF}}^k = \frac{\widehat{\beta}^2 E_s}{\widehat{\beta}^2 E_s\mathbf{h}_k^T\mathbb{E}_{\Delta\mathbf{H}}\left[\Delta\mathbf{W}\Delta\mathbf{W}^H\right]\mathbf{h}_k^* + \sigma^2}. \tag{2.26}$$

Now we summarize the achievable rate per user in the large number of antennas and high SNR regime in the following theorem.

Theorem 2.1. *For a large receive SNR and large number of antennas, the average achievable rate of the ZF precoding based on the estimated CSI can be approximated by*

$$
R_{\text{ZF}}^k \underset{\substack{M \to \infty \\ \text{SNR} \to \infty}}{\approx} \log_2 \left[1 + \left[\left(\sqrt{1+\delta^2} - 1 \right)^2 + \frac{\left(\sqrt{1+\delta^2} \right) \left(2 - \sqrt{1+\delta^2} \right) \delta^2 \text{tr}\left(\mathbf{K}^{-1}\right)}{1 + \delta^2 \text{tr}\left(\mathbf{K}^{-1}\right)} \right. \right.
$$
$$
\left. \left. - \frac{2\sqrt{1+\delta^2} \left(\sqrt{1+\delta^2} - 1 \right) \delta^2 M \eta_{kk}}{1 + \delta^2 \text{tr}\left(\mathbf{K}^{-1}\right)} \right]^{-1} \right], \tag{2.27}
$$

where $\mathbf{K} = (\mathbf{H}^T \mathbf{H}^*)$, η_{kk} *represents the k-th diagonal element of* $\mathbf{K}^{-1}$, *and* δ^2 *is the normalized MSE of proposed channel estimation algorithm.*

Proof: Please refer to Appendix A. ∎

Observing Equation (2.27), the approximated average achievable rate of the ZF precoding is bounded by above. The finite achievable rate ceiling is mainly determined by the normalized MSE of the estimated SLPS.

2.5　Comparison with PAC-based Precoding

In this section, we analyze the achievable rates performance of the MRT and the ZF precoders in multi-cell mmWave systems via pilot-aided estimated CSI. The adopted CSI for the design of these precoders is under the impact of pilot contamination, which is caused by the reuse of orthogonal pilot symbols among different cells. The analytical results obtained in this section will be used as a performance reference in the subsequent sections. In particular, our analysis quantifies the connection between rate performance and the MSE of channel estimation.

2.5.1 PAC-MRT and PAC-ZF Precoding under Pilot Contamination

Here we study the performance of the MRT precoding based on pilot-aided CSI in the presence of pilot contamination in mmWave channels. The pilot contamination leads to two types of interference in the considered system: intra-cell interference and inter-cell interference [28,53]. For the uplink channel estimation, we denote the received pilot sequences at the BS in the desired cell as

$$\mathbf{Y}_{\mathrm{u}} = \mathbf{H}\boldsymbol{\Psi} + \sum_{p=1}^{T} \mathrm{diag}\left(\rho_{p,k}\right) \mathbf{H}_p \boldsymbol{\Psi} + \mathbf{Z}, \qquad (2.28)$$

where $\boldsymbol{\Psi} \in \mathbb{C}^{N \times N}$ stands for the reused orthogonal pilot sequences matrix and $\rho_{p,k}$ is the cross-cell path loss attenuation coefficient. In Equation (2.28), $\mathbf{H}_p$ is the channel matrix from the users in the p-th neighboring cell to the BS in the desired cell. According to the field test, the probability of existing LOS component decreases significantly with increasing propagation distance [29]. In this case, we model the entries of cross-cell channels $\mathbf{H}_p$ as i.i.d. random variable $\mathcal{CN}(0,1)$.

Conventionally, the MMSE channel estimation, which requires the prior channel information, can achieve a better channel estimation performance than LS channel estimation. However, the prior information of mmWave channels, e.g. the Rician K-factor, is hard to obtain. Therefore, we adopt the LS algorithm for channel estimation instead of the MMSE algorithm, the estimated channel can be expressed as

$$\widehat{\mathbf{H}} = \mathbf{Y}_{\mathrm{u}}\boldsymbol{\Psi}^H \left[\boldsymbol{\Psi}\boldsymbol{\Psi}^H\right]^{-1} = \mathbf{H} + \underbrace{\sum_{p=1}^{T} \mathrm{diag}\left(\rho_{p,k}\right) \mathbf{H}_p + \widehat{\mathbf{Z}}}_{\text{pilot contamination}} = \mathbf{H} + \Delta\mathbf{H} + \widehat{\mathbf{Z}}, \quad (2.29)$$

where $\widehat{\mathbf{Z}} = \mathbf{Z}\boldsymbol{\Psi}^H \left[\boldsymbol{\Psi}\boldsymbol{\Psi}^H\right]^{-1}$ and $\Delta\mathbf{H} \in \mathbb{C}^{M \times N}$ is the equivalent channel estimation error matrix due to the impact of pilot contamination. Let $\Delta\mathbf{h}_k \in \mathbb{C}^{M \times 1}$ be the k-th column of the channel estimation error matrix $\Delta\mathbf{H}$, and the expression of $\Delta\mathbf{h}_k$ is given by

$$\Delta\mathbf{h}_k = \sum_{p=1}^{T} \rho_{p,k}\mathbf{h}_{p,k}, \tag{2.30}$$

where $\mathbf{h}_{p,k}$, with entries follow $\mathcal{CN}(0,1)$, is the channel vector from the p-th neighboring cell.

The relationships between $\Delta\mathbf{h}_k$, $\mathbf{h}_{p,k}$, and δ_{P}^2 can be expressed as

$$\frac{1}{M}\mathbb{E}\left[\Delta\mathbf{h}_k^H \Delta\mathbf{h}_k\right] = \frac{1}{M}\sum_{p=1}^{T}\rho_{p,k}^2\mathbb{E}\left[\mathbf{h}_{p,k}^H \mathbf{h}_{p,k}\right] = \sum_{p=1}^{T}\rho_{p,k}^2 = \delta_{\mathrm{P}}^2, \tag{2.31}$$

where δ_{P}^2 is the normalized MSE of channel estimation [166].

In addition, we have $\mathcal{CN}(\varsigma, 1)$ for the desired signal in the desired cell. The reason is that the perfect long-term power control is performed to compensate for path loss and shadowing between the desired BS and the desired users to facilitate the investigation of channel estimation and downlink transmission [3, 46, 115]. Due to the cross-cell propagation path loss, the received energy of desired pilot symbols at the desired BS is significantly larger than the energy of pilot symbols from neighboring cells to the desired BS.

For example, the 200 meters mmWave path loss is approximately above 140 dB and the 100 meters propagation path loss is about 120 dB [29, 102]. The received energy of desired pilot symbols is about 20 dB larger than that of the energy of contaminated pilot symbols (Table IV of [102]). Thus, the impact of pilot contamination on the channel estimation is assumed as $0 < \delta_{\mathrm{P}}^2 < 1$.

The received signal at user k with the MRT precoder $\widetilde{\mathbf{W}}_{\mathrm{PAC-MRT}} = \mathbf{H}^* +$

$(\Delta \mathbf{H})^*$ is given by

$$y_{\text{PAC-MRT}}^k = \underbrace{\beta \mathbf{h}_k^H \mathbf{h}_k x_k}_{\text{desired signal}} + \underbrace{\beta \sum_{p=1}^{T} \rho_{p,k} \mathbf{h}_{p,k}^H \widehat{\mathbf{w}}_{p,k}^* x_{p,k}}_{\text{inter-cell interference}}$$

$$+ \underbrace{\beta \sum_{j=1,j\neq k}^{N} \mathbf{h}_j^H \mathbf{h}_k x_j + \beta \sum_{j=1}^{N} \Delta \mathbf{h}_j^H \mathbf{h}_k x_j}_{\text{intra-cell interference}} + z_k, \qquad (2.32)$$

where $\beta = \sqrt{\frac{1}{\text{tr}(\widetilde{\mathbf{W}}_{\text{PAC-MRT}} \widetilde{\mathbf{W}}_{\text{PAC-MRT}}^H)}}$ is the normalization factor, $\widehat{\mathbf{w}}_{p,k}^*$ is the down-link precoder of user k in p-th cell, and $z_k \sim \mathcal{CN}(0, \sigma^2)$ accounts for AWGN. Substituting Equation (2.1) into Equation (2.32), the received SINR at user k can be expressed as

$$\text{SINR}_{\text{PAC-MRT}}^k = \left|\mathbf{h}_k^H \mathbf{h}_k\right|^2 \left[\left(\sum_{j=1,j\neq k}^{N} \mathbf{h}_j^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{h}_j \right) \right.$$

$$\left. + \sum_{j=1}^{N} \mathbb{E}_{\Delta \mathbf{h}} \left[\Delta \mathbf{h}_j^H \mathbf{h}_k \mathbf{h}_k^H \Delta \mathbf{h}_j \right] + M \sum_{p=1}^{T} \rho_{p,k}^2 + \frac{E_s}{\beta^2 \sigma^2} \right]^{-1}. \qquad (2.33)$$

We next present an asymptotic SINR of Equation (2.33) in the following corollary.

Corollary 2.1. *As SNR and the number of antennas approach infinity, the average achievable rate of the MRT precoding under pilot contamination can be approximated by*

$$R_{\text{PAC-MRT}}^k \underset{\substack{M\to\infty \\ \text{SNR}\to\infty}}{\approx}$$

$$\log_2 \left[1 + \left[\underbrace{\sum_{j=1,j\neq k}^{N} \frac{\mathbf{h}_{\text{L},j}^H \mathbf{h}_{\text{L},k} \mathbf{h}_{\text{L},k}^H \mathbf{h}_{\text{L},j}}{M^2}}_{\text{Part 1}} + \underbrace{\frac{(N-1)(2v_k+1)}{M(v_k+1)^2}}_{\text{Part 2}} + \underbrace{\frac{N\delta_{\text{P}}^2}{M}}_{\text{Part 3}} + \underbrace{\frac{\delta_{\text{P}}^2}{M}}_{\text{Part 4}} \right]^{-1} \right].$$

$$(2.34)$$

Proof: Please refer to Appendix B.   ■

Part 1 and Part 2 of Equation (2.34) represent the interference caused by the corresponding SLOS components and scattering components of different users, respectively. In addition, Part 3 and Part 4 of Equation (2.34) denote the intra-cell and inter-cell interference caused only by channel estimation error due to pilot contamination, respectively. It is interesting to note that, the intra-cell interference in Part 3 is much larger than that of the inter-cell interference in Part 4.

It is clear that with an increasing number of antennas M, the achievable rate increases. Besides, with increasing the Rician K-factor υ_k, the interference caused by scattering components in Part 2 of Equation (2.34) vanishes.

For the downlink ZF precoding, similar to the signal processing in Equations $(2.24)-(2.25)$, $\widetilde{\mathbf{w}}_{\text{PAC-ZF}}$ is the downlink ZF precoder and $\overline{\beta} = \sqrt{\frac{1}{\text{tr}(\widetilde{\mathbf{W}}_{\text{PAC-ZF}}\widetilde{\mathbf{W}}_{\text{PAC-ZF}}^H)}}$ is the normalization factor.

We then express the SINR of user k under the PAC-based ZF precoding as

$$
\text{SINR}_{\text{PAC-ZF}}^k =
$$

$$
\left[\underbrace{\sum_{p=1}^{T} \rho_{p,k}^2 \mathbb{E}_{\mathbf{h}_{p,k}} \left[\mathbf{h}_{p,k}^T \widehat{\mathbf{w}}_{p,k} \widehat{\mathbf{w}}_{p,k}^H \mathbf{h}_{p,k}^* \right]}_{\text{inter-cell interference}} + \underbrace{\mathbf{h}_k^T \mathbb{E}_{\Delta \mathbf{H}} \left[\Delta \mathbf{W} \Delta \mathbf{W}^H \right] \mathbf{h}_k^*}_{\text{intra-cell interference}} + \frac{\sigma^2}{E_s \overline{\beta}^2} \right]^{-1} . \quad (2.35)
$$

Now we summarize the average achievable rate per user in the large number of antennas and high SNR regime in the following corollary.

Corollary 2.2. *For a large receive SNR and a large number of antennas, the average achievable rate of the ZF precoding under pilot contamination can be approximated by*

$$R_{\mathrm{PAC-ZF}}^k \underset{\substack{\mathrm{SNR}\to\infty}}{\overset{M\to\infty}{\approx}} \log_2\left[1 + \left[\left(\sqrt{1+\delta_{\mathrm{P}}^2}-1\right)^2 - \frac{2\sqrt{1+\delta_{\mathrm{P}}^2}\left(\sqrt{1+\delta_{\mathrm{P}}^2}-1\right)\delta_{\mathrm{P}}^2 M\eta_{kk}}{1+\delta_{\mathrm{P}}^2\mathrm{tr}\left(\mathbf{K}^{-1}\right)}\right.\right.$$

$$\left.\left.+\frac{\delta_{\mathrm{P}}^2}{M} + \frac{\left(\sqrt{1+\delta_{\mathrm{P}}^2}\right)\left(2-\sqrt{1+\delta_{\mathrm{P}}^2}\right)\delta_{\mathrm{P}}^2\mathrm{tr}\left(\mathbf{K}^{-1}\right)}{1+\delta_{\mathrm{P}}^2\mathrm{tr}\left(\mathbf{K}^{-1}\right)}\right]^{-1}\right], \qquad (2.36)$$

where $\mathbf{K} = (\mathbf{H}^T\mathbf{H}^*)$, η_{kk} *represents the k-th diagonal element of* $\mathbf{K}^{-1}$, *and* δ_{P}^2 *is the normalized MSE of LS channel estimation given by Equation (2.31).*

Proof: Following a similar approach as in Appendix A.   ■

Equations (2.34) and (2.36) represent the ceiling of the achievable rate of MU mmWave massive MIMO systems when the conventional PAC is adopted. The performance gap between Equations (2.34) and (2.36) will be quantized via simulation in Figure 2.6 in the following sections.

2.5.2   Achievable Rate Comparison

In this section, we compare the achievable rates of the PAC-based MRT and the PAC-based ZF precoders to the achievable rates of the SLOS-based MRT and the SLPS-based ZF precoders, respectively. For mmWave cellular systems, the deployment of small cell is necessary to facilitate high data rate communications [102]. In our work, the radius of the cell is assumed approximately less than 150 meters. Under such a scenario, for a reasonably large transmit power and the use of large number of antennas, high SNR is expected to the system operating regime[3]. In fact, the derived downlink achievable rate analysis is based on the high SNR assumption with practical interests for high data rate communication. For the comparison between the achievable rates of the PAC-based ZF precoding and the SLPS-based ZF precoding, we observe the mathematical similarity between

---

[3]By calculating the link budget [102], it can be shown that the considered system operates at high SNR regimes.

Equations (2.36) and (2.27). While the MSEs of two different channel estimation are identical, i.e., $\delta^2 = \delta_P^2$, the difference between R_{ZF}^k and $R_{\text{PAC-ZF}}^k$ is negligible.

This observation will also be verified via simulation in Figure 2.6. Then, we compare the achievable rates between the SLOS-based MRT and the PAC-based MRT and summarize the performance comparison in the following theorem.

Theorem 2.2. *If the Rician factor satisfies $\upsilon_k > 2$, the rate performance gap between the proposed SLOS-based MRT and the PAC-based MRT is always larger than 0, even if the MSE of pilot-aided channel estimation is $\delta_P^2 = 0$, i.e.,*

$$\Delta R = \log_2 \left[\frac{1 + \text{SINR}_{\text{LMRT}}^k}{1 + \text{SINR}_{\text{PAC-MRT}}^k} \right] > 0. \tag{2.37}$$

Proof: The result follows by substituting (2.34) and (2.22) into (2.37). ∎

Intuitively, the required coherence time resource for the conventional PAC estimation, increases with an increasing number of users N. Interestingly, the time resource consumption of the proposed AoA channel estimation algorithm, T_{CE}, does not increase with the number of antennas nor the number of users. In this case, we note that the coherence time saved by adopting the proposed algorithm is $N - T_{\text{CE}}$ symbols, which can be used for downlink data transmission. In Figure 2.6, we present a comparison between the achievable sum-rates of the SLOS-based MRT, the PAC-based MRT, the SLPS-based ZF, and the PAC-based ZF for the same MSEs of channel estimation $\delta^2 = \delta_P^2 = 0.02$. This result verifies the accuracy of our derived analytical expressions in Equations (2.22), (2.27), (2.34), and (2.36). The achievable rates of the SLPS-based and the PAC-based ZF precoders are almost identical. As expected, the ZF precoding can achieve a higher rate performance than the MRT precoding due to its interference suppression capability, especially in the high SNR regime. It is interesting that in the high SNR
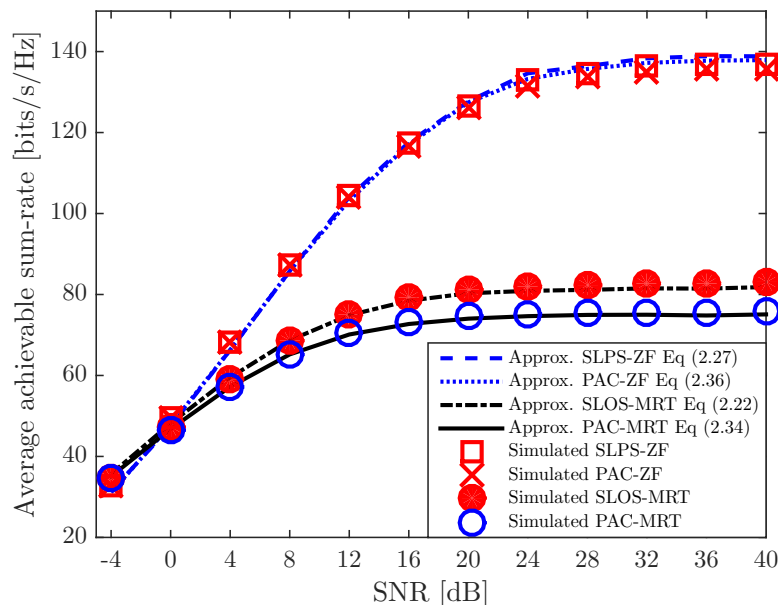
Figure 2.6: Average achievable sum-rate [bits/s/Hz] versus SNR [dB] for different precoding schemes with $M = 100$ and $N = 20$. We set the Rician K-factor as 15 and the normalized MSE = 0.02 for the SLPS-based ZF, the PAC-based ZF, and the PAC-based MRT.

regime, the SLOS-based MRT can achieve a significant performance gain in terms of achievable rate than the PAC-based MRT. The reasons are two-fold. First, for the SLOS-based MRT precoding, the signal energy focuses on the strong LOS propagation path leading to a smaller amount of signal energy leakage into the scattering paths. In fact, the less signal energy leaks into the scattering part, the smaller inter-user interference it will cause. In contrast, pilot-aided CSI naturally contains scattering components, which leads to a inter-user interference. Second, the PAC-based channel estimation error may cause extra MU interference. However, the proposed tone-based AoA channel estimation error, which is due to not taking into account scattering components, does not cause extra MU interference as shown in Equation (2.22). Thus, the SLOS-based MRT can achieve a better SINR performance than the PAC-based MRT.
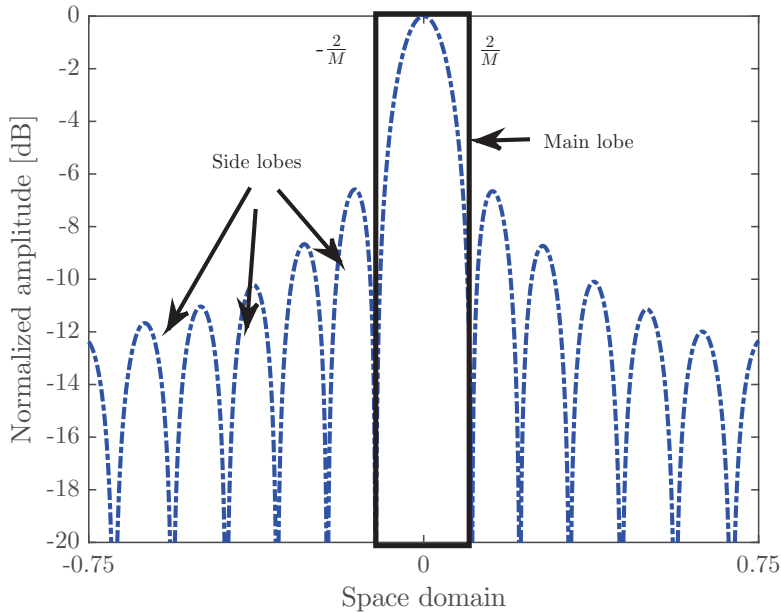
Figure 2.7: Illustration of the main lobe and side lobes of ULA in the space domain with respect to the BS.

2.5.3 Large-system Analysis

In this section, we study the system performance while the numbers of users and antennas are large. In fact, the achievable rate approximation expressions of the MRT and the ZF in Equations (2.22), (2.27), (2.34), and (2.36) still rely on instantaneous channel realization of the SLOS components of all the users. These channel dependent expressions are inconvenient to provide useful guidelines for system design.

Fortunately, similar to conventional massive MIMO systems while the numbers of users and antennas are large, the law of large number can bring us some interesting results [96, 167]. Now, we focus on a scenario where the number of users and the number of antennas are large while their ratio is bounded above by a finite constant. To facilitate the analysis, we assume that all the users are uniformly distributed [168] with certain degrees separation in angular domain.

To illustrate the concepts of angular separation, the main lobe and side lobes of ULA is shown in Figure 2.7.

The range of the main lobe in the space domain[4] for user k [58], Θ_k, is given by

$$\cos(\theta_k) - \frac{2}{M} < \Theta_k < \cos(\theta_k) + \frac{2}{M}, \tag{2.38}$$

where θ_k is the AoA of the SLOS of user k. If the AoA of the SLOS of user i is not in the main lobe of user k, we consider that user i is in the side lobes of user k from the BS point of view. The minimum angular separation between different users from the BS point of view in a far-field region can be expressed as

$$|\cos(\theta_k) - \cos(\theta_i)| \geqslant \frac{4}{M}, \ \forall i \neq k. \tag{2.39}$$

Let us consider the extreme case that all the users are separated one by one and with the minimal angular separation $|\cos(\theta_k) - \cos(\theta_i)| = \frac{4}{M}, \ \forall i \neq k$. The maximum number of users $N_{\max}$, which can be supported by the BS equipped with M antennas, should satisfy the following constraint:

$$\frac{4}{M}(N_{\max} - 1) \leqslant 2. \tag{2.40}$$

Therefore, the maximum number of users $N_{\max}$ is approximately $N_{\max} \approx \frac{M}{2}$. Here, we justify that the density of users due to the angular separation assumption in Equation (2.39) is reasonable with the following example.

For typical mmWave massive MIMO systems operating at 30 GHz, the cell radius is roughly 150 meters. For a BS equipped with a 200-antenna ULA array serving 100 users simultaneously, the user density is 3500 devices/(km)^2, which is the medium user density as stated in 3GPP [163]. Besides, the 200-antenna

---

[4]The spatial domain constraint is given by $-1 \leqslant \cos(\theta_k) \leqslant 1$ [111].

ULA array, which is deployed for half wavelength separation, is only 0.5 meter long. Under such an assumption, we note that the channels between any two users are not orthogonal. Such non-orthogonality has already been verified via field experiments [74, 160].

Based on the angular separation assumption in Equation (2.39), we can then make a further approximation [111] for the achievable rates of the MRT and the ZF precoding in Equations (2.22), (2.27), and (2.34). Part 1 of Equation (2.22) and Part 1 of Equation (2.34) can be rewritten as

$$\sum_{j=1,j\neq k}^{N} \mathbf{h}_{\mathrm{L},k}^{H}\mathbf{h}_{\mathrm{L},j}\mathbf{h}_{\mathrm{L},j}^{H}\mathbf{h}_{\mathrm{L},k} \overset{M,N\rightarrow\infty}{\approx} \frac{N^2}{4}. \tag{2.41}$$

While the number of antennas M and the number of users N are large with their ratio $B = \dfrac{N}{M}$ fixed, the asymptotic SINR expressions of the SLOS-based MRT and the PAC-based MRT in the high SNR regime can be expressed as

$$\mathrm{SINR}_{\mathrm{LMRT}}^{k} \overset{M,N\rightarrow\infty}{\underset{\mathrm{SNR}\rightarrow\infty}{\approx}} \left[\frac{1}{4}\left(\frac{N}{M}\right)^2 + \frac{1}{\upsilon_k}\frac{N}{M}\right]^{-1}, \text{ and} \tag{2.42}$$

$$\mathrm{SINR}_{\mathrm{PAC-MRT}}^{k} \overset{M,N\rightarrow\infty}{\underset{\mathrm{SNR}\rightarrow\infty}{\approx}} \left[\frac{1}{4}\left(\frac{\upsilon_k}{\upsilon_k+1}\right)^2\left(\frac{N}{M}\right)^2 + \frac{N}{M}\frac{(2\upsilon_k+1)}{(\upsilon_k+1)^2} + \frac{N}{M}\delta^2\right]^{-1}, \tag{2.43}$$

respectively.

Interestingly, if Rician K-factor υ_k does not approach infinity, the result shows that Equation (2.42) is always larger than Equation (2.43) for a fixed ratio $B = \dfrac{N}{M}$.

Now we focus on the asymptotic rate performance of the PAC-based ZF precoding and the SLPS-based ZF precoding. To this end, we have

$$\mathbb{E}[\mathbf{H}^{T}\mathbf{H}^{*}] = (\mathbf{G}_{\mathrm{L}})^2\,\mathbf{H}_{\mathrm{L}}^{T}\mathbf{H}_{\mathrm{L}}^{*} + (\mathbf{G}_{\mathrm{S}})^2\,\mathbf{H}_{\mathrm{S}}^{T}\mathbf{H}_{\mathrm{S}}^{*}. \tag{2.44}$$

We can obtain several properties of symmetric matrix $\mathbf{H}_{\mathrm{L}}^T \mathbf{H}_{\mathrm{L}}^*$ via singular value decomposition (SVD) as

$$\mathrm{tr}\left[\mathbf{H}_{\mathrm{L}}^T \mathbf{H}_{\mathrm{L}}^*\right] = \mathrm{tr}\left[\mathbf{H}_{\mathrm{L}}^H \mathbf{H}_{\mathrm{L}}\right] = \mathrm{tr}\left[\mathbf{U}\mathbf{\Sigma}\mathbf{U}^H\right] = \sum_{i=1}^{N}\lambda_i\left(\mathbf{\Sigma}\right) = MN, \qquad (2.45)$$

where $\lambda_i\left(\mathbf{\Sigma}\right)$ is the non-negative eigenvalues of $\mathbf{\Sigma}$ and $\mathbf{U}\mathbf{U}^H = \mathbf{I}_{\mathrm{N}} = \mathbf{U}^H\mathbf{U}$. We can obtain

$$\mathrm{tr}\left[\mathbf{h}_{\mathrm{L},i}^H \mathbf{h}_{\mathrm{L},i}\right] = M = \lambda_i(\mathbf{\Sigma}), \ \forall i \in [1, N]. \qquad (2.46)$$

where $\mathbf{h}_{\mathrm{L},i}$ is the i-th column vector of $\mathbf{H}_{\mathrm{L}}$. As we assume that all the users are distinguishable in the angular domain, we have the following expression:

$$\mathrm{tr}\left\{\left[\mathbf{H}_{\mathrm{L}}^H \mathbf{H}_{\mathrm{L}}\right]^{-1}\right\} = \mathrm{tr}\left[\left(\mathbf{U}^H\right)^{-1}\mathbf{\Sigma}^{-1}\mathbf{U}^{-1}\right] = \mathrm{tr}\left(\mathbf{\Sigma}^{-1}\right) = \sum_{i=1}^{N}\frac{1}{\lambda_i\left(\mathbf{\Sigma}\right)} \approx \frac{N}{M}. \quad (2.47)$$

Based on the previous results Equations (2.27) and (2.47), the SINR expression of user k under the SLPS-based ZF precoding in the high SNR regime is given by

$$\mathrm{SINR}_{\mathrm{ZF}}^k \overset{\underset{M,N\to\infty}{\approx}}{\underset{\mathrm{SNR}\to\infty}{}} \left[\left(\sqrt{1+\delta^2}-1\right)^2 - \frac{2\delta^2\sqrt{1+\delta^2}\left(\sqrt{1+\delta^2}-1\right)\upsilon_k}{\left(1+\delta^2\dfrac{N}{M}\right)\left(\upsilon_k+1\right)}\right.$$
$$\left. + \frac{\left(\sqrt{1+\delta^2}\right)\left(2-\sqrt{1+\delta^2}\right)\delta^2}{\left(1+\delta^2\dfrac{N}{M}\right)}\frac{N}{M}\right]^{-1}. \qquad (2.48)$$

With the fixed system load, the number of users to the number of antennas ratio $B = \dfrac{N}{M}$, we have the observations: a) Equation (2.48) is determined by the MSE of channel estimation δ^2; b) $\mathrm{SINR}_{\mathrm{ZF}}^k$ is saturated with respect to $\upsilon_k \to \infty$.

In Figures 2.8 and 2.9, we validate Theorem 2.2 and the asymptotic analysis in Equations (2.42), (2.43), and (2.48). In Figure 2.8, user scheduling (US) stands
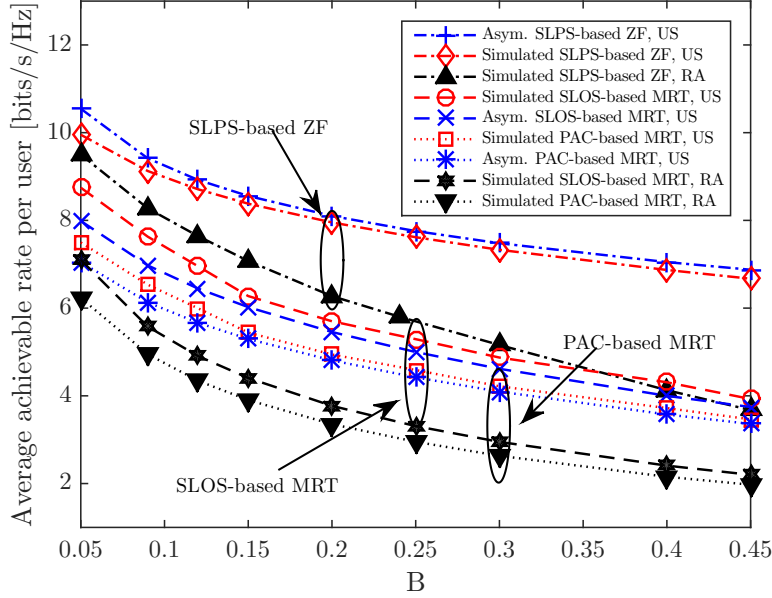
Figure 2.8: Average rate per user [bits/s/Hz] based on deterministic channel versus different $B = \frac{N}{M}$ for $M = 200$, the Rician K-factor 15, SNR = 40 dB, and the normalized MSE = 0.02.

for that the AoAs of all the users' SLOS components satisfy the assumption in Equation (2.39) and random angle (RA) stands for that the AoAs of all the users' SLOS components are uniformly distributed from 0° to 180°. Figure 2.8 shows that the proposed SLOS-based precoder outperforms the conventional imperfect PAC-based precoder for a wide range of scenarios.

In Figure 2.9, interestingly, for a wide range of Rician K-factor, the proposed SLOS-based MRT can achieve a significant higher rate performance than that of the PAC-based MRT, even the pilot-aided estimated CSI is perfect. As the MSE of channel estimation increases, the sum-rate gap between the SLOS-based MRT and the PAC-based MRT increases significantly. The results shown in Figure 2.9 indicates that the SLOS-based MRT is a better choice than the conventional PAC-based MRT for mmWave massive MIMO systems. It is interesting to note that, while the assumption in Equation (2.39) is satisfied, the achievable rate per-
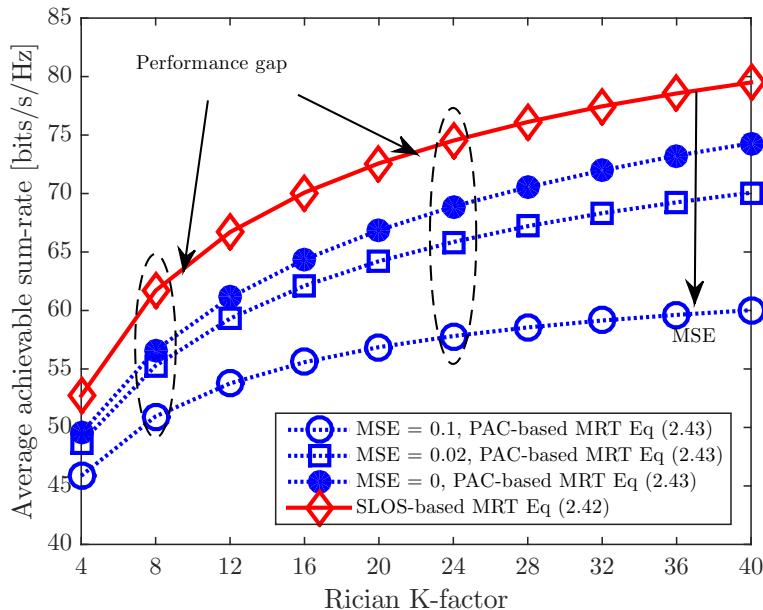
Figure 2.9: Average achievable sum-rate [bits/s/Hz] versus Rician K-factor comparison with different MSEs, $N = 10$, and $M = 100$.

formance under the assumption of US can significantly outperforms the achievable rate performance under RA assumption. These simulation results imply that, an appropriate US algorithm selecting a set of users satisfying Equation (2.39) can improve the achievable rate performance for mmWave massive MIMO communication systems. Therefore, Equation (2.39) can be considered as a guideline of US.

2.6 Impact of Imperfect Hardware: Phase Quantization Error

Hardware impairments in massive MIMO systems causing performance degradation have been widely discussed [38, 80]. The hardware structure is widely adopted as a fully digital conventional MIMO scheme that M RF chains with

M phase shifters are connected to M antennas. In particular, most of the works in the literature modeled that the impact of phase and amplitude errors are additive distortions and hardware imperfections are uncorrelated with the desired signals. Some works, e.g. [80], proved that the additive distortion created by hardware impairments at the BS will not vanish even the number of antennas goes large. In this section, we analyze the impact of phase quantization errors on the average achievable rate performance in the high SNR and large number of antennas regime. The large-system analysis results provide important design guidelines for mmWave massive MIMO systems. Amplitude quantization errors are critical important for the performance of MIMO systems. However, amplitude quantization errors, which are partially contributed by the limited resolution of ADC/DAC and hardware imperfection attenuation, are nonlinear and very hard to model. Currently, to our best knowledge, no existing mathematical model is precise enough for capturing the characteristics of amplitude quantization errors. Thus, we leave this amplitude quantization errors problem as a future work.

Transceiver RF chains need to provide accurate phase information for different antennas at the BS. Yet, digital/analog RF chains have limited resolutions in practical systems. Thus, phase quantization error may cause strong impacts on mmWave massive MIMO systems. For example, if the digital phase shifters of each RF chain can only provide a 2-bit of resolution for quantization, the quantized phase responses are $[0, 90, 180, 270]$ degrees. If the exact phase responses in different antennas are $[82, 345, 248, 151, 53]$ degrees, the obtained phase responses after quantization in different antennas will be $[90, 0, 270, 180, 90]$. The differences between the actual and the quantized phase responses may lead to serious phase mismatch in transmission.

We now discuss the rate performance degradation of the SLOS-based MRT

precoding due to the phase quantization error. Then we compare the obtained result to the conventional pilot-aided CSI based MRT. We derive the relationship between the quantization resolutions of phase shifters and the associated rate performance degradation. We first denote the equivalent channel of user k after phase quantization as[5]

$$\widehat{\mathbf{h}}_{\mathrm{L},k} = \mathbf{\Lambda}_{\Delta k}\mathbf{h}_{\mathrm{L},k}, \tag{2.49}$$

where the quantization error diagonal matrix is expressed as [25, 153]

$$\mathbf{\Lambda}_{\Delta_k} = \mathrm{diag}\left\{\Delta\phi_k\right\}. \tag{2.50}$$

The phase quantization error of user k is $\Delta\phi_k \in \mathbb{C}^{M\times 1} = \left[\ e^{j\overline{\phi}_{k,1}}, \ldots, e^{j\overline{\phi}_{k,M}}\ \right]^T$, where the exact phase quantization error $\overline{\phi}_{k,i}$ on the i-th phase shifter ($\forall i \in [1, M]$) is uniformly distributed over $[-a, +a]$, and a is the maximal phase quantization error [134]. The exact uniformly distributed phase quantization errors on different phase shifters have following property that

$$\mathbb{E}_{\Delta\phi_k}\left[\mathbf{\Lambda}_{\Delta k}\right] = \begin{bmatrix} \int\limits_{-a}^{a}\dfrac{1}{2a}e^{j\Delta\phi_{k,1}}d\Delta\phi_{k,1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \int\limits_{-a}^{a}\dfrac{1}{2a}e^{j\Delta\phi_{k,M}}d\Delta\phi_{k,M} \end{bmatrix}$$

$$= \frac{\sin\left(a\right)}{a}\mathbf{I}. \tag{2.51}$$

Then the received signal of user k under limited phase quantization resolutions mismatch can be expressed as

---

[5]We note that the phase quantization error cannot be modeled as an additive distortion [25, 153].

$$\widehat{y}^k_{\text{LMRT}} = \underbrace{\widetilde{\beta}\sqrt{\frac{\upsilon_k}{\upsilon_k+1}}\mathbf{h}^H_{\text{L},k}\mathbf{\Lambda}_{\Delta k}\mathbf{h}_{\text{L},k}x_k}_{\text{desired signal}} + \underbrace{\widetilde{\beta}\sqrt{\frac{\upsilon_k}{\upsilon_k+1}}\sum_{j=1,j\neq k}^N \mathbf{h}^H_{\text{L},j}\mathbf{\Lambda}_{\Delta j}\mathbf{h}_{\text{L},k}x_j}_{\text{inter-user interference}}$$

$$+\underbrace{\widetilde{\beta}\sqrt{\frac{1}{\upsilon_k+1}}\sum_{j=1}^N \mathbf{h}^H_{\text{L},j}\mathbf{\Lambda}_{\Delta j}\mathbf{h}_{\text{S},k}x_j}_{\text{scattering interference}} + z_k. \tag{2.52}$$

We denote the received signal power of user k as

$$\text{S}_k = \frac{\upsilon_k\widetilde{\beta}^2 E_s}{\upsilon_k+1}\mathbf{h}^H_{\text{L},k}\mathbb{E}_{\Delta\phi_k}\left[\mathbf{\Lambda}_{\Delta k}\right]\mathbf{h}_{\text{L},k}\mathbf{h}^H_{\text{L},k}\mathbb{E}_{\Delta\phi_k}\left[\mathbf{\Lambda}^H_{\Delta k}\right]\mathbf{h}_{\text{L},k}. \tag{2.53}$$

Then we can approximate the inter-user interference by ignoring small random part as:

$$\text{I}^k_{\text{LMRT}} \approx \frac{\widetilde{\beta}^2 E_s MN}{\upsilon_k+1}$$

$$+ \frac{\upsilon_k\widetilde{\beta}^2 E_s}{\upsilon_k+1}\sum_{j=1,j\neq k}^N \mathbf{h}^H_{\text{L},k}\mathbb{E}_{\Delta\phi_j}\left[\mathbf{\Lambda}_{\Delta j}\right]\mathbf{h}_{\text{L},j}\mathbf{h}^H_{\text{L},j}\mathbb{E}_{\Delta\phi_j}\left[\mathbf{\Lambda}^H_{\Delta j}\right]\mathbf{h}_{\text{L},k}. \tag{2.54}$$

For different AoAs, the phase quantization errors are different. We then present the approximated average SINR expression in the high SNR regime as

$$\widehat{\text{SINR}}^k_{\text{LMRT}} \overset{\text{SNR}\to\infty}{\approx} \mathbb{E}\left[\frac{\text{S}_k}{\text{I}^k_{\text{LMRT}}}\right]$$

$$\approx \frac{\dfrac{\upsilon_k\beta^2 E_s}{\upsilon_k+1}\mathbf{h}^H_{\text{L},k}\mathbb{E}_{\Delta\phi_k}\left[\mathbf{\Lambda}_{\Delta k}\right]\mathbf{h}_{\text{L},k}\mathbf{h}^H_{\text{L},k}\mathbb{E}_{\Delta\phi_k}\left[\mathbf{\Lambda}^H_{\Delta k}\right]\mathbf{h}_{\text{L},k}}{\displaystyle\sum_{j=1,j\neq k}^N \mathbf{h}^H_{\text{L},j}\mathbb{E}_{\Delta\phi_j}\left[\mathbf{\Lambda}_{\Delta j}\right]\mathbf{h}_{\text{L},k}\mathbf{h}^H_{\text{L},k}\mathbb{E}_{\Delta\phi_j}\left[\mathbf{\Lambda}^H_{\Delta j}\right]\mathbf{h}_{\text{L},j} + \dfrac{MN}{\upsilon_k}}. \tag{2.55}$$

By exploiting the property that

$$\mathbb{E}_{\Delta\phi_k}\left[\mathbf{\Lambda}_{\Delta k,n}\right] = \int_{-a}^{a}\frac{1}{2a}e^{j\Delta\phi_k}d\Delta\phi_k = \frac{\sin(a)}{a} = \mathbb{E}_{\Delta\phi_k}\left[\mathbf{\Lambda}^H_{\Delta k,n}\right], \tag{2.56}$$

and the SINR expression (2.55) can be rewritten as

$$\widehat{\text{SINR}}_{\text{LMRT}}^{k} \overset{\text{SNR}\to\infty}{\approx}$$

$$\frac{\left[M\,\text{sinc}\,(a)\right]^2}{\dfrac{MN}{\upsilon_k} + \displaystyle\sum_{j=1,j\neq k}^{N} \mathbf{h}_{\text{L},j}^{H}\mathbb{E}_{\Delta\phi_j}\left[\mathbf{\Lambda}_{\Delta j}\right]\mathbf{h}_{\text{L},k}\mathbf{h}_{\text{L},k}^{H}\mathbb{E}_{\Delta\phi_j}\left[\mathbf{\Lambda}_{\Delta j}^{H}\right]\mathbf{h}_{\text{L},j}}. \tag{2.57}$$

For a special case of $a = 0$, the sum of sinc function is $\sum_{n=1}^{M} \dfrac{\sin(a)}{a} = M$. In the considered system, as we assume that all users are separated by hundreds of wavelength and the AoAs angular separation condition in (2.39) is satisfied, the inter-user interferences are from the side lobes. For $M \to \infty$, $N \to \infty$, and $\dfrac{N}{M}$ is fixed, we have the following asymptotic SINR expression for the SLOS-based MRT in the high SNR regime:

$$\widehat{\text{SINR}}_{\text{SLOS}-\text{MRT}}^{k} \overset{\substack{M,N\to\infty \\ \text{SNR}\to\infty}}{\approx} \frac{\left[\text{sinc}^2(a)\right]}{\dfrac{1}{\upsilon_k}\dfrac{N}{M} + \dfrac{1}{4}\left(\dfrac{N}{M}\right)^2}. \tag{2.58}$$

For $M \to \infty$, $N \to \infty$, and $\dfrac{N}{M}$ is fixed, we have the following asymptotic SINR expression for the PAC-based MRT in the high SNR regime:

$$\widehat{\text{SINR}}_{\text{PAC}-\text{MRT}}^{k} \overset{\substack{M,N\to\infty \\ \text{SNR}\to\infty}}{\approx} \frac{\left[\text{sinc}^2(a)\right]}{\dfrac{1}{4}\left(\dfrac{\upsilon_k}{\upsilon_k+1}\right)^2\left(\dfrac{N}{M}\right)^2 + \dfrac{N}{M}\dfrac{(2\upsilon_k+1)}{(\upsilon_k+1)^2} + \dfrac{N}{M}\delta^2}. \tag{2.59}$$

From Equations (2.58) and (2.59), we observe that the SINR expressions are affected by the range of phase quantization error a. With the fixed quantization error, the performance degradation may not vanish, even if more antennas are adopted for transmission.

By comparing Equations (2.58) to (2.59), the SLOS-based MRT still achieves a better rate performance than that of the PAC-based MRT, despite the presence of phase quantization errors.
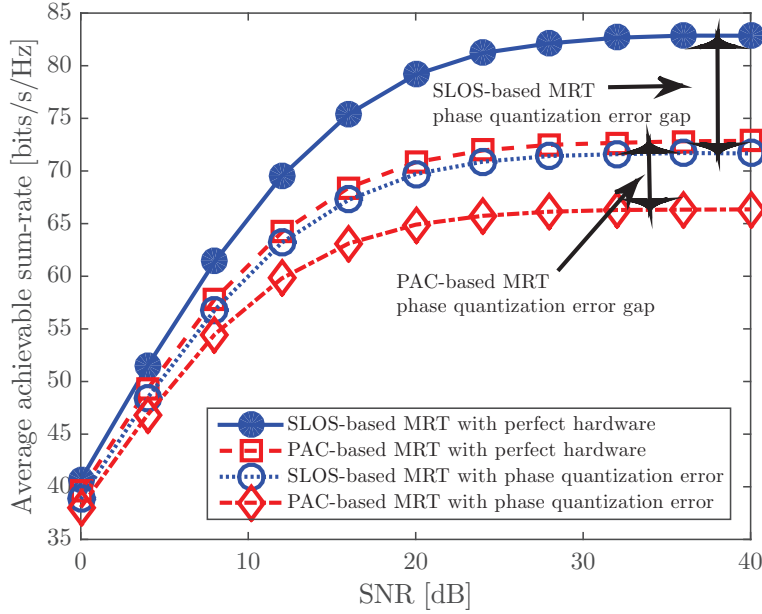
Figure 2.10: Average achievable sum-rate [bits/s/Hz] under phase quantization error versus SNR [dB], for $N = 10$, $M = 100$, the normalized MSE $= 0.02$ and Rician K-factor of 50.

Accordingly, the considered mmWave massive MIMO system requires the support of high phase quantization resolutions RF hardware in mmWave channels.

Remark 2.1. *In order to maintain the system performance (no more than a 3-dB of SINR degradation compared to ideal hardware), the quantization resolutions of phase shifters in each RF chain should satisfy:*

$$\left(\sum_{n=1}^{M} \frac{\sin(a)}{a} \right)^2 \geqslant \frac{M^2}{2}. \tag{2.60}$$

Based on Equations (2.58) and (2.42), we can obtain the rate performance degradation of the SLOS-based MRT under different quantization resolutions in the following simulation. In Figures 2.10 and 2.11, we illustrate the impact of phase quantization error on system rate performance. We assume that the resolutions of the phase quantization per RF chain is 2 bits. In Figure 2.10, we show
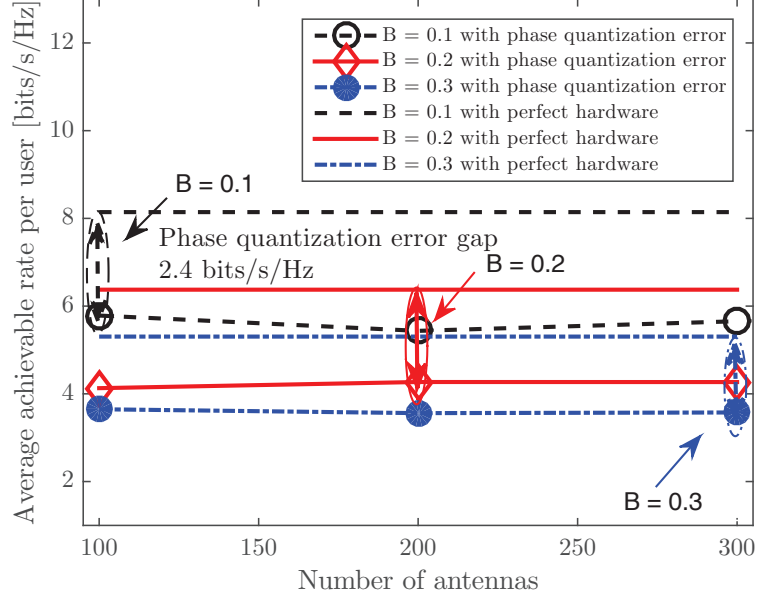
Figure 2.11: The impact of 2-bit phase quantization resolutions on average achievable rate per user of SLOS-based MRT versus the number of antennas with the normalized MSE = 0.02 and Rician K-factor of 50.

the performance degradation for the SLOS-based MRT and the PAC-based MRT for $M = 100$, $N = 10$. In the high SNR regime, the system rate performance degradation due to the limited phase quantization resolutions is significant. However, the SLOS-based MRT can still achieve a better sum rate, which verifies the correctness of our derived results in Equations (2.58) and (2.59).

In Figure 2.11, we demonstrate the derived results for different system loads. In particular, we observe that the impact of phase quantization errors will not vanish with the increase of the number of antennas. The rate performance degradation caused by phase quantization errors is 2.4 bits/s/Hz per user, which is close to the predicted value (2.98 bits/s/Hz per user) via Equation (2.58). It indicates that the hardware requirement of mmWave massive MIMO system is very stringent and should be taken into account in system design.

2.7 Summary

In this chapter, we first proposed a frequency tone-based AoAs channel estimation algorithm in mmWave massive MIMO systems and we analyzed the performance of the associated SLOS-based MRT and the SLPS-based ZF precoder. We also derived the achievable rates of the PAC-based MRT and the PAC-based ZF precoding under the impact of pilot contamination. In particular, we compared the achievable rates of the SLOS-based MRT and the SLPS-based ZF precoding to the achievable rates of the PAC-based MRT and the PAC-based ZF precoding, respectively. We showed that, for the same MSE of channel estimation and large numbers of antennas and users, the achievable rates of the SLPS-based and the PAC-based ZF precoding are identical. Besides, our proposed SLOS-based MRT precoder can achieve a higher rate than that of the traditional PAC-based M-RT precoder. Furthermore, we analyzed the impact of phase quantization error on system rate performance, revealing that such impact cannot be mitigated in mmWave massive MIMO systems by increasing the number of antennas.

Chapter 3

Hybrid mmWave MIMO Systems: Multi-user Precoding and Channel Estimation

3.1   Introduction

In this chapter, we consider a TDD MU hybrid mmWave system. In the previous chapter, we proposed a frequency tone-based channel estimation algorithm for the estimation of strongest AoAs. Besides, we proved that our proposed SLOS-based MRT precoder can achieve a higher rate performance than that of the traditional PAC-based MRT precoder. Based on these prior information, we propose and detail a novel non-feedback non-iterative channel estimation algorithm which is applicable to both non-sparse and sparse mmWave channels.

The proposed three-step channel estimation algorithm is an extension of the tone-based AoA channel estimation algorithm proposed for fully digital mmWave MIMO in Chapter 2. Also, we analyze the achievable rate performance of the

mmWave system using digital ZF precoding based on the estimated channel information. Furthermore, we analyze the performance degradation under some practical hardware imperfections, such as random phase errors, RF transceiver beamforming errors, and channel estimation errors.

Our main contributions are summarized as follows:

- We propose a three-step MU channel estimation scheme for mmWave channels. In the first two steps, we estimate the strongest AoAs at both the BS and the users sides instead of estimating the combination of multiple AoAs. The estimated strongest AoAs will be exploited for the design of analog beamforming matrices at the BS and users. In the third step, all the users transmit orthogonal pilot symbols to the BS along the beamforming paths of the strongest AoA directions to facilitate the equivalent channel estimation, which will be exploited to design the digital ZF precoder at the BS for the downlink transmission. Our proposed hybrid scheme can suppress the downlink MU interference effectively via its analog beamforming and digital precoder. Firstly, the proposed analog beamforming allow signal transmission and reception along the strongest AoA direction, which reduces the interference outside the strongest AoA directions and utilizes the transmission power more efficiently. Secondly, the digital ZF precoder can suppress the MU interference within the strongest AoA directions.

- We analyze the achievable rate performance of the proposed scheme based on the estimated equivalent channel CSI, analog beamforming matrices, and digital ZF precoding. While assuming the equivalent CSI is perfectly known at the BS, we derive a tight performance upper bound on the achievable rate of our proposed scheme. Also, we quantify the performance gap between the proposed hybrid scheme and the fully digital system in terms

of achievable rate per user. It is interesting to note that the performance gap is determined by the ratio between the power of the strongest AoA component and the power of the scattering component, Rician K-factor υ. The performance gap of the average achievable rate per user between the hybrid system and the fully digital system is only $\left| \log_2 \left(\frac{\upsilon}{\upsilon+1} \right) \right|$ bits/s/Hz in the large numbers of antennas regime.

- We further analyze the system performance degradation and derive the closed-form approximation of achievable rate under various types of errors, i.e., random phase errors, transceiver analog beamforming errors, and equivalent channel estimation errors, in the high receiver signal-to-noise ratio (SNR) and the large numbers of antennas regimes. Interestingly, our results confirm that the impact of phase errors and transceiver beamforming errors will not cause a performance ceiling in terms of achievable rate. Besides, the performance gap in terms of the achievable rate between the system under phase errors and transceiver beamforming errors and the system with perfect hardware is approximated.

The rest of the chapter is organized as follows. Section 3.2 describes the system model considered throughout the chapter. In Section 3.3, we detail the proposed channel estimation algorithm for hybrid systems. In Section 3.4, we investigate the downlink transmission rate performance based on the equivalent CSI estimated in Section 3.3. Besides, we further have rate performance analysis under the impact of hardware impairments in Section 3.5. Numerical results and related discussions are presented in Section 3.6. Finally, Section 3.7 summarizes the chapter.

3.2   System Model
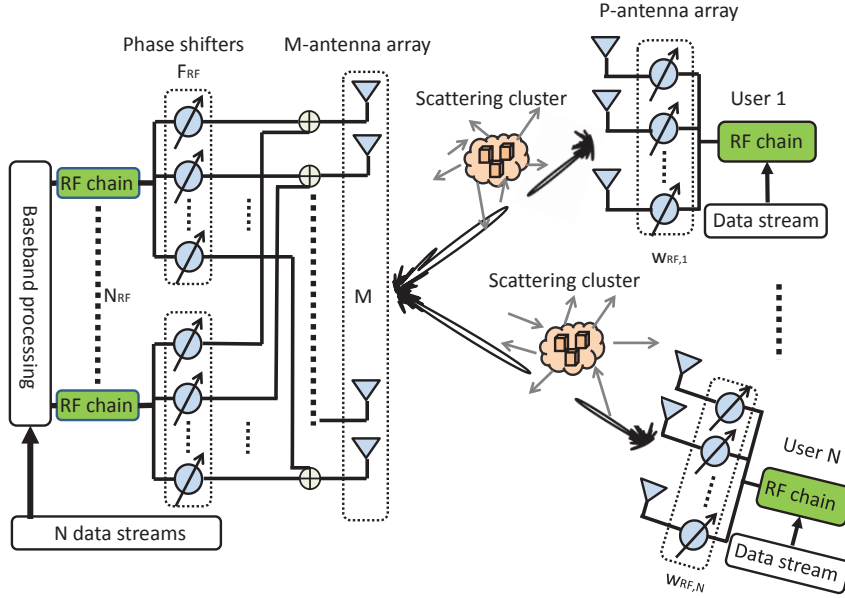
Figure 3.1: A mmWave massive MIMO communication system with a hybrid system of transceivers.

We consider a MU hybrid mmWave system which consists of one BS and N users in a single cell, as shown in Figure 3.1. Generally, there are two kind of hybrid structures which are widely adopted by researchers [3, 17]: the full access hybrid architecture and the subarray hybrid architecture. The full access hybrid architecture, where each RF chain is connected to all the antennas, can provide a higher array gain and a narrower beam width than that of the subarray hybrid architecture, where each RF chain is connected to a part of the antennas. In this chapter, we adopt the full access hybrid architecture since it offers higher flexibility in the design of channel estimation algorithm. We assume that the BS is equipped with $M \geq 1$ antennas and N_{RF} RF chains to serve the N users. Besides, each user is equipped with P antennas and a single RF chain. We also assume that $M \geqslant N_{\mathrm{RF}} \geqslant N$. In the following sections, we set $N = N_{\mathrm{RF}}$ to
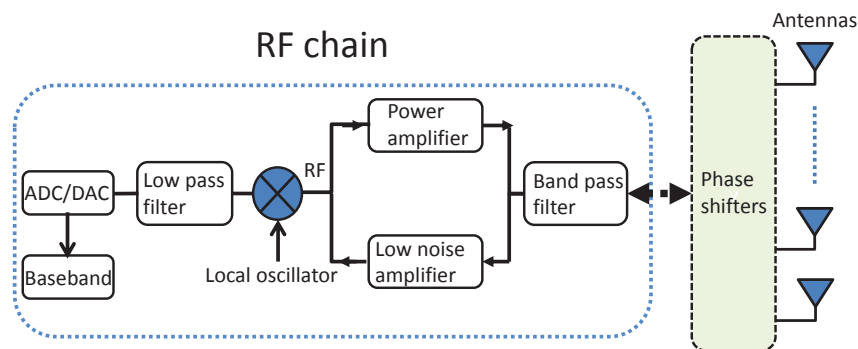
Figure 3.2: A block diagram of an RF chain for an antenna array.

simplify the analysis[1].

Each RF chain at the BS can access to all the antennas by using M phase shifters, as shown in Figure 3.2. At each BS, the number of phase shifters is $M \times N_{\mathrm{RF}}$. Due to significant propagation attenuation at mmWave frequency, the system is dedicated to cover a small area, e.g. cell radius is ~ 150 m. We assume that the users and the BS are fully synchronized and TDD is adopted to facilitate uplink and downlink communications [28].

In previous work [3], mmWave channels were assumed to have sparse propagation paths between the BS and the users. Yet, in recent field tests, especially in the urban microcell environments, both a strong LOS component and non-negligible scattering components may exist in mmWave propagation channels [29, 102, 165]. Therefore, for the urban short-distance propagation environment, mmWave channels are more suitable to be modeled by non-sparse Rician fading and with a large Rician K-factor [29, 102, 155]. On the other hand, for the suburban long-distance propagation environment, mmWave channels can be modeled by a sparse channel model. The reason is that scattering components will vanish during the long-distance propagation because of the high reflection loss and large

---

[1]We note that our proposed channel estimation scheme, precoding scheme, and analysis can be generalized to the case of $N_{\mathrm{RF}} \geq N$, at the expense of a more involved notation.

propagation path loss. In addition, the blockage of LOS component is critical for mmWave systems and widely considered in previous works [113, 169]. In [169], the authors proved that the optimal cell size to achieve the maximum SINR scales with the average size of the area that is LOS to a user. Therefore, with a properly designed radius of cells, one should expect the existence of at least one LOS component from the BSs to any user. When the LOS component from a certain BS to a user is blocked, the user can still exploit other existing LOS components from other BSs for channel estimation and data transmission.

Let $\mathbf{H}_k \in \mathbb{C}^{M \times P}$ be the uplink channel matrix between user k and the BS in the cell. We assume that $\mathbf{H}_k$ is a slow time-varying block Rician fading channel, i.e., the channel is constant in a block but varies slowly from one block to another. Then, in this chapter, we assume that the channel matrix $\mathbf{H}_k$ can be decomposed into a deterministic LOS channel matrix $\mathbf{H}_{\mathrm{L},k} \in \mathbb{C}^{M \times P}$ and a scattered channel matrix $\mathbf{H}_{\mathrm{S},k} \in \mathbb{C}^{M \times P}$ [165], i.e.,

$$\mathbf{H}_k = \underbrace{\mathbf{H}_{\mathrm{L},k}\mathbf{G}_{\mathrm{L},k}}_{\text{LOS component}} + \underbrace{\mathbf{H}_{\mathrm{S},k}\mathbf{G}_{\mathrm{S},k}}_{\text{Scattering component}} , \tag{3.1}$$

where $\mathbf{G}_{\mathrm{L},k} \in \mathbb{C}^{P \times P}$ and $\mathbf{G}_{\mathrm{S},k} \in \mathbb{C}^{P \times P}$ are diagonal matrices with entries

$$\mathbf{G}_{\mathrm{L},k} = \mathrm{diag}\left\{\sqrt{\frac{\upsilon_k}{\upsilon_k + 1}}\right\} \text{ and } \mathbf{G}_{\mathrm{S},k} = \mathrm{diag}\left\{\sqrt{\frac{1}{\upsilon_k + 1}}\right\}, \tag{3.2}$$

respectively, and $\upsilon_k > 0$ is the Rician K-factor of user k. Besides, Equations (3.1) and (3.2) are the generalization of mmWave channel models, which capture both the scattered and non-scattered components. In general, we can adopt different array structures, e.g. ULA and uniform panel array (UPA) for both the BS and the users. Here, we adopt the ULA as it is commonly implemented in practice [3]. We assume that all the users are separated by hundreds of wavelengths or

more [28]. Thus, we can express the deterministic LOS channel matrix $\mathbf{H}_{\mathrm{L},k}$ of user k as [170]

$$\mathbf{H}_{\mathrm{L},k} = \mathbf{h}_{\mathrm{L},k}^{\mathrm{BS}} \mathbf{h}_{\mathrm{L},k}^{H}, \tag{3.3}$$

where $\mathbf{h}_{\mathrm{L},k}^{\mathrm{BS}} \in \mathbb{C}^{M \times 1}$ and $\mathbf{h}_{\mathrm{L},k} \in \mathbb{C}^{P \times 1}$ are the antenna array response vectors of the BS and user k respectively.

In particular, $\mathbf{h}_{\mathrm{L},k}^{\mathrm{BS}}$ and $\mathbf{h}_{\mathrm{L},k}$ can be expressed as [111, 170]

$$\mathbf{h}_{\mathrm{L},k}^{\mathrm{BS}} = \begin{bmatrix} 1, & \ldots, & e^{-j2\pi(M-1)\frac{d}{\lambda}\cos(\theta_k)} \end{bmatrix}^{T} \text{ and} \tag{3.4}$$

$$\mathbf{h}_{\mathrm{L},k} = \begin{bmatrix} 1, & \ldots, & e^{-j2\pi(M-1)\frac{d}{\lambda}\cos(\phi_k)} \end{bmatrix}^{T}, \tag{3.5}$$

respectively, where d is the distance between the neighboring antennas and λ is the wavelength of the carrier frequency. Variables $\theta_k \in [0, +\pi]$ and $\phi_k \in [0, +\pi]$ are the angles of incidence of the LOS path at antenna arrays of the BS and user k, respectively. For convenience, we set $d = \dfrac{\lambda}{2}$ for the rest of the chapter which is an assumption commonly adopted in the literature [111, 170]. Without loss of generality, we assume that the scattering component $\mathbf{H}_{\mathrm{S},k}$ consists N_{cl} clusters and each cluster contributes $N_{\mathrm{l},i}$ propagation paths [165], which can be expressed as

$$\mathbf{H}_{\mathrm{S},k} = \sqrt{\frac{1}{\sum_{i=1}^{N_{\mathrm{cl}}} N_{\mathrm{l},i}}} \sum_{i=1}^{N_{\mathrm{cl}}} \sum_{l=1}^{N_{\mathrm{l},i}} \alpha_{i,l} \mathbf{h}_{i,l}^{\mathrm{BS}} \mathbf{h}_{k,i,l}^{H} = \begin{bmatrix} \mathbf{h}_{\mathrm{S},1}, & \ldots, & \mathbf{h}_{\mathrm{S},k}, & \ldots, & \mathbf{h}_{\mathrm{S},P} \end{bmatrix}, \tag{3.6}$$

where $\mathbf{h}_{i,l}^{\mathrm{BS}} \in \mathbb{C}^{M \times 1}$ and $\mathbf{h}_{k,i,l} \in \mathbb{C}^{P \times 1}$ are the antenna array response vectors of the BS and user k associated to the (i,l)-th propagation path, respectively. Here $\alpha_{i,l} \sim \mathcal{CN}(0,1)$ represents the path attenuation of the (i,l)-th propagation path and $\mathbf{h}_{\mathrm{S},k} \in \mathbb{C}^{M \times 1}$ is the k-th column vector of $\mathbf{H}_{\mathrm{S},k}$. With the increasing number of clusters, the path attenuation coefficients and the AoAs between the users and

the BS become randomly distributed [102, 165]. Therefore, we model the entries of scattering component $\mathbf{H}_{\mathrm{S},k}$ in a general manner as an i.i.d. random variable[2] $\mathcal{CN}(0, 1)$.

3.3 Proposed Channel Estimation for Hybrid System

In this section, we propose and detail our mmWave channel estimation for hybrid mmWave systems. In practice, the hybrid system imposes a fundamental challenge for mmWave channel estimation. Unfortunately, the conventional pilot-aided channel estimation algorithm for fully digital systems, e.g. [3, 5], is not applicable to the considered hybrid mmWave system. The reasons are that the number of RF chains is much lower than the number of antennas equipped at the BS and the transceiver beamforming matrix cannot be acquired.

To address this important issue, we propose a novel channel estimation algorithm, which contains three steps as shown in Figure 3.3 and Algorithm 3.1. In the first and second steps, we introduce unique unmodulated frequency tones to estimate the strongest AoAs at the BS and user sides. The unique frequency tones and linear search algorithm are inspired by signal processing in monopulse passive electronically scanned array (PESA) radar and sonar systems [170]. The unmodulated frequency tone-based AoA beam sweeping search signal processing procedures are different while adopting different hardware architectures, e.g. hybrid architecture and fully digital architecture. These estimated strongest AoAs will be exploited to develop analog transmit and receive beamforming matrices at

---

[2]To facilitate the study of the downlink hybrid precoding, we assume that perfect long-term power control is performed to compensate for path loss and shadowing at the desired users and equal power allocation among different data streams of the users [3, 46, 115]. Thus, the entries of scattering component $\mathbf{H}_{\mathrm{S},k}$ are modeled by i.i.d. random variables.
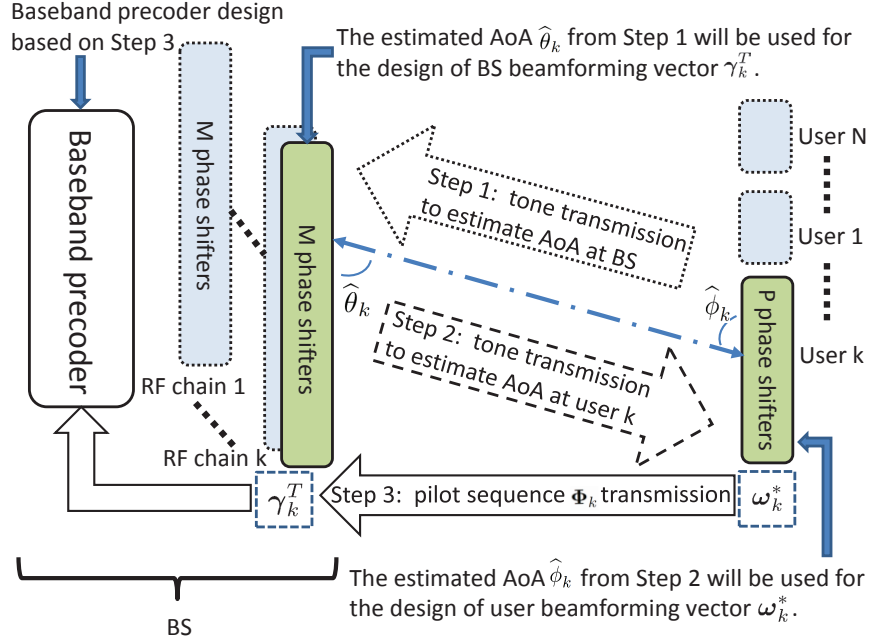
Figure 3.3: An illustration of the proposed channel estimation algorithm for hybrid mmWave systems.

the BS and users. In the third step, the users transmit orthogonal pilot symbols to the BS along the beamforming paths in order to estimate the equivalent channel via the strongest AoA directions. Then, the estimated channel will be used for the design of BS digital baseband precoder for the downlink transmissions by exploiting the reciprocity between the uplink and downlink channels.

3.3.1 Details of Proposed Channel Estimation

Step 1, <u>Line 1 in Algorithm 3.1</u>: Firstly, all the users transmit unique frequency tones to the desired BS in the uplink simultaneously. For user k, an unique unmodulated frequency tone, $x_k = \cos(2\pi f_k t), k \in \{1, \cdots, N\}$, is transmitted from one of the omni-directional antennas in the antenna array to the BS. Here, f_k is the single carrier frequency and t stands for time and $f_k \neq f_j, \forall k \neq j$. For the AoA estimation, if the condition $\dfrac{f_k - f_j}{f_c} < 10^{-4}, \forall k \neq j$, is satisfied, the

---

Algorithm 3.1 Channel Estimation Algorithm for Hybrid Systems

---

Require: Multiple single-carrier frequency tone signals $[f_1, \ldots, f_N]$, pilot sequences matrix $\mathbf{\Psi}$, the detection matrices $[\mathbf{\gamma}_1, \ldots, \mathbf{\gamma}_N]$ of AoA at the BS, and the detection matrices $[\mathbf{\Omega}_1, \ldots, \mathbf{\Omega}_N]$ of AoA at the users

STEP 1: <u>Estimate the AoA at the BS and</u>
<u>design BS analog beamforming matrix</u>

1: The estimation of AoA at the BS: all the users transmit their unique frequency tones by using only one omni-directional antenna

2: The BS calculates $r_{k,i}^{\mathrm{BS}}$, $i \in \{1, \ldots, J\}$, as shown in (3.9) to estimate the uplink AoA of user k and its corresponding beamforming vector: $\widetilde{\mathbf{\gamma}}_k = \underset{\forall \gamma_{k,i},\, i \in \{1,\ldots,J\}}{\arg\max} \left| r_{k,i}^{\mathrm{BS}} \right|$

3: The BS analog beamforming matrix is
$\mathbf{F}_{\mathrm{RF}} = \left[\ \widetilde{\mathbf{\gamma}}_1, \ldots, \widetilde{\mathbf{\gamma}}_N\ \right]$

STEP 2: <u>Estimate the AoA at the users and</u>
<u>design user analog beamforming matrix</u>

4: The estimation of AoA at the users: the BS transmits frequency tones back to all the users using $\mathbf{F}_{\mathrm{RF}}$ as a transmit beamforming matrix

5: Calculate $r_{k,i}^{\mathrm{UE}}$, $i \in \{1, \ldots, J\}$, as shown in (3.13) to estimate the downlink AoA of user k and its corresponding beamforming vector: $\widetilde{\mathbf{\omega}}_k^* = \underset{\forall \mathbf{\omega}_{k,i},\, i \in \{1,\ldots,J\}}{\arg\max} \left| \gamma_{k,i}^{\mathrm{UE}} \right|$

6: The users analog beamforming matrix is
$\mathbf{Q}_{\mathrm{RF}} = \left[\ \widetilde{\mathbf{\omega}}_1^*, \ldots, \widetilde{\mathbf{\omega}}_N^*\ \right]$

STEP 3: <u>Estimate equivalent channel and</u>
<u>design digital ZF precoder</u>

7: All the users transmit orthogonal pilot sequences by using $\mathbf{Q}_{\mathrm{RF}}$ as beamforming matrix and the BS uses $\mathbf{F}_{\mathrm{RF}}^T$ as beamforming matrix to receive pilot sequences

8: The BS obtains and calculates $\widehat{\mathbf{H}}_{\mathrm{eq}}^T$ as shown in (3.16) $\widehat{\mathbf{H}}_{\mathrm{eq}}^T = \mathbf{\Psi}^H \left[\ \mathbf{s}_1, \ldots, \mathbf{s}_N\ \right]$

9: The BS sets the baseband digital ZF precoder as
$\overline{\mathbf{W}}_{\mathrm{eq}} = \widehat{\mathbf{H}}_{\mathrm{eq}}^* (\widehat{\mathbf{H}}_{\mathrm{eq}}^T \widehat{\mathbf{H}}_{\mathrm{eq}}^*)^{-1}$

AoA estimation difference by using different tones is generally negligible [111], where f_{c} is the system carrier frequency. The pass-band received signal of user k at the BS, $\mathbf{y}_k^{\mathrm{BS}}$, is given by

$$\mathbf{y}_k^{\mathrm{BS}} = \left(\sqrt{\frac{\upsilon_k}{\upsilon_k + 1}} \mathbf{h}_{\mathrm{L},k}^{\mathrm{BS}} + \sqrt{\frac{1}{\upsilon_k + 1}} \mathbf{h}_{\mathrm{S},k} \right) x_k + \mathbf{z}_{\mathrm{BS}}, \tag{3.7}$$

where $\mathbf{z}_{\mathrm{BS}}$ denotes the thermal noise at the antenna array of the BS, $\mathbf{z}_{\mathrm{BS}} \sim \mathcal{CN}\left(\mathbf{0}, \sigma_{\mathrm{BS}}^2 \mathbf{I}\right)$, and σ_{BS}^2 is the noise variance at each antenna of the BS. To facilitate the estimation of AoA, we perform a linear search in the angular domain ranged from $0°$ to $180°$ with an angle search step size of $\dfrac{180}{J}$. Therefore, the AoA detection matrix $\mathbf{\Gamma}_k \in \mathbb{C}^{M \times J}$, $\mathbf{\Gamma}_k = \left[\; \boldsymbol{\gamma}_{k,1}, \ldots, \boldsymbol{\gamma}_{k,J}\; \right]$, contains J column vectors. In general, the typical value of the minimum search steps J depends on the number of antennas M used for the AoA search. In particular, $J \approx \frac{2M}{1.782}$ [111]. The i-th vector $\boldsymbol{\gamma}_{k,i} \in \mathbb{C}^{M \times 1}, i \in \{1, \cdots, J\}$, stands for a potential AoA of user k at the BS and is given by

$$\boldsymbol{\gamma}_{k,i} = \frac{1}{\sqrt{M}} \left[\; 1, \quad \ldots, \quad e^{j2\pi(M-1)\frac{d}{\lambda}\cos\left(\widehat{\theta}_i\right)} \;\right]^T, \tag{3.8}$$

where $\widehat{\theta}_i = (i-1)\frac{180}{J}, i \in \{1, \cdots, J\}$, is the assumed AoA and $\boldsymbol{\gamma}_{k,i}^H \boldsymbol{\gamma}_{k,i} = 1$. For the AoA estimation of user k, $\mathbf{\Gamma}_k$ is implemented in the M phase shifters connected by the k-th RF chain. The local oscillator (LO) of the k-th RF chain at the BS generates the same carrier frequency f_k to down convert the received signals to the baseband, as shown in Figure 3.2. After the down-conversion, the signals will be filtered by a low-pass filter which can remove other frequency tones. The equivalent received signal at the BS from user k at the i-th potential AoA is given by

$$r_{k,i}^{\mathrm{BS}} = \sqrt{\frac{\upsilon_k}{\upsilon_k + 1}} \boldsymbol{\gamma}_{k,i}^T \mathbf{h}_{\mathrm{L},k}^{\mathrm{BS}} + \sqrt{\frac{1}{\upsilon_k + 1}} \boldsymbol{\gamma}_{k,i}^T \mathbf{h}_{\mathrm{S},k} + \boldsymbol{\gamma}_{k,i}^T \mathbf{z}_{\mathrm{BS}}. \tag{3.9}$$

The potential AoA, which leads to the maximum value among the J observation directions, i.e.,

$$\widetilde{\boldsymbol{\gamma}}_k = \underset{\forall \gamma_{k,i},\ i \in \{1,\cdots,J\}}{\arg\max} \left| r_{k,i}^{\mathrm{BS}} \right|, \tag{3.10}$$

is considered as the strongest AoA of user k. In addition, the strongest AoA estimation shown in Equations (3.9) and (3.10) can be performed by using either a series of analog comparators in analog domain or digital buffer in digital domain. Besides, vector $\widetilde{\boldsymbol{\gamma}}_k$ corresponding to the AoA with the maximum value in (3.10) will be exploited for the design of the analog beamforming vector of user k at the BS. As a result, we can also estimate all other users' uplink AoAs at the BS from their corresponding transmitted signals simultaneously. For notational simplicity, we denote $\mathbf{F}_{\mathrm{RF}} = \left[\begin{array}{ccc} \widetilde{\boldsymbol{\gamma}}_1, \ldots, \widetilde{\boldsymbol{\gamma}}_N \end{array} \right] \in \mathbb{C}^{M \times N}$ as the BS analog beamforming matrix.

Step 2, <u>Line 4 in Algorithm 3.1</u>: The BS sends unique frequency tones to all the users exploiting analog beamforming matrix[3] $\mathbf{F}_{\mathrm{RF}}$ obtained in Step 1. This facilitates the downlink AoAs estimation at the users and this AoA information will be used to design analog beamforming vectors to be adopted at the users. The received signal $\mathbf{y}_k^{\mathrm{UE}}$ at user k can be expressed as

$$\mathbf{y}_k^{\mathrm{UE}} = \left[\mathbf{G}_{\mathrm{L},k} \mathbf{h}_{\mathrm{L},k}^* \left(\mathbf{h}_{\mathrm{L},k}^{\mathrm{BS}} \right)^T + \mathbf{G}_{\mathrm{S},k} \mathbf{H}_{\mathrm{S},k}^T \right] \widetilde{\boldsymbol{\gamma}}_k x_k + \mathbf{z}_{\mathrm{MS}}, \tag{3.11}$$

where $\mathbf{z}_{\mathrm{MS}}$ denotes the thermal noise at the antenna array of the users, $\mathbf{z}_{\mathrm{MS}} \sim \mathcal{CN}\left(\mathbf{0}, \sigma_{\mathrm{MS}}^2 \mathbf{I}\right)$, and σ_{MS}^2 is the noise variance for all the users.

The AoA detection matrix for user k, $\boldsymbol{\omega}_k \in \mathbb{C}^{P \times J}$, which also contains J estimation column vectors, is implemented at phase shifters of user k. The i-th column vector of matrix $\boldsymbol{\omega}_k$ for user k, $\boldsymbol{\omega}_{k,i} \in \mathbb{C}^{P \times 1}, i \in \{1, \cdots, J\}$, is given by

$$\boldsymbol{\omega}_{k,i} = \frac{1}{\sqrt{P}} \left[\begin{array}{ccc} 1, & \ldots, & e^{j 2\pi (P-1) \frac{d}{\lambda} \cos(\widehat{\phi}_i)} \end{array} \right]^T, \tag{3.12}$$

---

[3]This procedure can be done simultaneously in all the RF chains for all the users.

where $\widehat{\phi}_i = (i-1)\frac{180}{J}, i \in \{1, \cdots, J\}$, is the i-th potential AoA of user k and $\boldsymbol{\omega}_{k,i}^H \boldsymbol{\omega}_{k,i} = 1$. With similar procedures as shown in Step 1, the equivalent received signal from the BS at user k of the i-th potential AoA is given by

$$r_{k,i}^{\mathrm{UE}} = \boldsymbol{\omega}_{k,i}^H \sqrt{\frac{\upsilon_k}{\upsilon_k+1}} \mathbf{h}_{\mathrm{L},k}^* \left(\mathbf{h}_{\mathrm{L},k}^{\mathrm{BS}}\right)^T \widetilde{\gamma}_k + \omega_{k,i}^H \sqrt{\frac{1}{\upsilon_k+1}} \mathbf{H}_{\mathrm{S},k}^T \widetilde{\gamma}_k + \boldsymbol{\omega}_{k,i}^H \mathbf{z}_{\mathrm{MS}}. \qquad (3.13)$$

Similarly, we search for the maximum value among J observation directions and design the analog beamforming vector based on the estimated AoA of user k. The beamforming vector for user k is given by

$$\widetilde{\boldsymbol{\omega}}_k^* = \underset{\forall \boldsymbol{\omega}_{k,i},\ i \in \{1, \cdots, J\}}{\arg \max} \left| r_{k,i}^{\mathrm{UE}} \right| \qquad (3.14)$$

and we denote $\mathbf{Q}_{\mathrm{RF}} = \left[\ \widetilde{\boldsymbol{\omega}}_1^*, \ldots, \widetilde{\boldsymbol{\omega}}_N^*\ \right] \in \mathbb{C}^{P \times N}$ as the users analog beamforming matrix.

Step 3, <u>Line 7 in Algorithm 3.1</u>: The BS and users analog beamforming matrices based on estimated uplink AoAs and downlink AoAs are designed via Step 1 and Step 2, respectively. After that, all the users transmit orthogonal pilot sequences to the BS via user beamforming vectors $\widetilde{\boldsymbol{\omega}}_k^*$. In the meanwhile, the BS receives pilot sequences via the BS analog beamforming matrix $\mathbf{F}_{\mathrm{RF}}^T$. With the analog beamforming matrices, we have the equivalent channel between the BS and the users along the strongest AoA paths[4].

We denote the pilot sequences of user k in the cell as $\boldsymbol{\Phi}_k = [\vartheta_k(1), \ldots, \vartheta_k(N)]^T$, $\boldsymbol{\Phi}_k \in \mathbb{C}^{N \times 1}$, stands for N symbols transmitted across time. The pilot symbols used for the equivalent channel estimation are transmitted in sequence from symbol $\vartheta_k(1)$ to symbol $\vartheta_k(N)$. The pilot symbols for all the N users form a matrix, $\boldsymbol{\Psi} \in \mathbb{C}^{N \times N}$, where $\boldsymbol{\Phi}_k$ is a column vector of matrix $\boldsymbol{\Psi}$ given by $\boldsymbol{\Psi} =$

---

[4]The equivalent channel consists of the BS analog beamforming matrix, the mmWave channel, and the users analog beamforming matrix.

$\sqrt{E_{\mathrm{P}}} \left[\begin{array}{ccc} \boldsymbol{\Phi}_1, & \ldots, & \boldsymbol{\Phi}_N \end{array} \right]$, $\boldsymbol{\Phi}_i^H \boldsymbol{\Phi}_j = 0$, $\forall i \neq j$, $i, j \in \{1, \ldots, N\}$, where E_{P} repre-
sents the transmitted pilot symbol energy. Note that $\boldsymbol{\Psi}^H \boldsymbol{\Psi} = E_{\mathrm{P}} \mathbf{I}_N$. Meanwhile,
the BS analog beamforming matrix $\mathbf{F}_{\mathrm{RF}}$ is utilized to receive pilot sequences at
all the RF chains. As the length of the pilot sequences is equal to the number of
users, we obtain an $N \times N$ observation matrix from all the RF chains at the BS.
In particular, the received signal at the k-th RF chain at the BS is $\mathbf{s}_k^T \in \mathbb{C}^{1 \times N}$,
which is given by

$$\mathbf{s}_k^T = \widetilde{\boldsymbol{\gamma}}_k^T \sum_{i=1}^{N} \mathbf{H}_i \widetilde{\boldsymbol{\omega}}_i^* \sqrt{E_{\mathrm{P}}} \boldsymbol{\Phi}_i^T + \widetilde{\boldsymbol{\gamma}}_k^T \mathbf{Z}, \tag{3.15}$$

where $\mathbf{Z} \in \mathbb{C}^{M \times N}$ denotes the additive white Gaussian noise matrix at the BS
and the entries of $\mathbf{Z}$ are modeled by i.i.d. random variable with distribution
$\mathcal{CN}\left(0, \sigma_{\mathrm{BS}}^2\right)$.

After $\left[\begin{array}{ccc} \mathbf{s}_1, \ldots, \mathbf{s}_N \end{array} \right]$ is obtained, we then adopt the LS method for our equiv-
alent channel estimation. We note here, the LS method is widely used in practice
since it does not require any prior channel information. Subsequently, with the
help of orthogonal pilot sequences, we can construct an equivalent uplink chan-
nel matrix $\widehat{\mathbf{H}}_{\mathrm{eq}} \in \mathbb{C}^{N \times N}$ formed by the proposed scheme via the LS estimation
method. Then, by exploiting the channel reciprocity, the equivalent downlink
channel of the hybrid system $\widehat{\mathbf{H}}_{\mathrm{eq}}^T$ can be expressed as:

$$\widehat{\mathbf{H}}_{\mathrm{eq}}^T = \mathbf{\Psi}^H \begin{bmatrix} \mathbf{s}_1 & \dots & \mathbf{s}_N \end{bmatrix} = \begin{bmatrix} \widehat{\mathbf{h}}_{\mathrm{eq},1}^T \\ \vdots \\ \widehat{\mathbf{h}}_{\mathrm{eq},N}^T \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} \widetilde{\boldsymbol{\omega}}_1^H \mathbf{H}_1^T \mathbf{F}_{\mathrm{RF}} \\ \vdots \\ \widetilde{\omega}_N^H \mathbf{H}_N^T \mathbf{F}_{\mathrm{RF}} \end{bmatrix}}_{\mathbf{H}_{\mathrm{eq}}^T} + \frac{1}{\sqrt{E_{\mathrm{P}}}} \underbrace{\begin{bmatrix} \mathbf{\Phi}_1^H \mathbf{Z}^T \mathbf{F}_{\mathrm{RF}} \\ \vdots \\ \mathbf{\Phi}_N^H \mathbf{Z}^T \mathbf{F}_{\mathrm{RF}} \end{bmatrix}}_{\text{Effective noise}}, \qquad (3.16)$$

where $\widehat{\mathbf{h}}_{\mathrm{eq},k}$ is the k-th column vector of matrix $\widehat{\mathbf{H}}_{\mathrm{eq}}$. From Equation (3.16), we observe that the proposed channel estimation algorithm can obtain all users' equivalent CSI simultaneously.

3.3.2 Performance Analysis of Proposed Channel Estimation

In the high SNR regime, the effective noise component is negligible and Equation (3.16) can be simplified[5] as

$$\mathbf{H}_{\mathrm{eq}}^T = \begin{bmatrix} \widetilde{\boldsymbol{\omega}}_1^H \mathbf{H}_1^T \mathbf{F}_{\mathrm{RF}} \\ \vdots \\ \widetilde{\omega}_N^H \mathbf{H}_N^T \mathbf{F}_{\mathrm{RF}} \end{bmatrix} = \begin{bmatrix} \widetilde{\boldsymbol{\omega}}_1^H & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \widetilde{\omega}_N^H \end{bmatrix} \begin{bmatrix} \mathbf{H}_1^T \\ \vdots \\ \mathbf{H}_N^T \end{bmatrix} \underbrace{\begin{bmatrix} \widetilde{\boldsymbol{\gamma}}_1 & \dots & \widetilde{\boldsymbol{\gamma}}_N \end{bmatrix}}_{\mathbf{F}_{\mathrm{RF}}}. \quad (3.17)$$

From Equation (3.17), we can see that without the impact of noise, we can perfectly estimate the equivalent channels, which consists of the BS analog beamforming matrix $\mathbf{F}_{\mathrm{RF}}$, the users analog beamforming matrix $\mathbf{Q}_{\mathrm{RF}}^T$, and mmWave channels between the BS and all the users.

Now, we analyze the performance of the proposed channel estimation. With

---

[5]The performance degradation due to the high SNR assumption will be verified by analysis and simulation in the following sections.

the normalization factor $\frac{1}{\sqrt{MP}}$, the normalized mean squared error (NMSE) of equivalent channel estimation $\widehat{\mathbf{h}}_{\mathrm{eq},k}^{T}$ is given by

$$
\begin{aligned}
\mathrm{NMSE}&_{\mathrm{eq},k} \\
&= \frac{1}{N}\mathrm{tr}\left\{ \mathbb{E}_{\mathbf{h}_{\mathrm{S},k}}\left[\left(\frac{1}{\sqrt{MP}}\widehat{\mathbf{h}}_{\mathrm{eq},k}^{T} - \frac{1}{\sqrt{MP}}\mathbf{h}_{\mathrm{eq},k}^{T}\right)^{H}\left(\frac{1}{\sqrt{MP}}\widehat{\mathbf{h}}_{\mathrm{eq},k}^{T} - \frac{1}{\sqrt{MP}}\mathbf{h}_{\mathrm{eq},k}^{T}\right)\right]\right\} \\
&= \frac{\sigma_{\mathrm{BS}}^{2}\mathrm{tr}\left[\mathbf{F}_{\mathrm{RF}}^{H}\mathbf{F}_{\mathrm{RF}}\right]}{E_{\mathrm{P}}NMP} = \frac{\sigma_{\mathrm{BS}}^{2}}{E_{\mathrm{P}}MP}.
\end{aligned} \tag{3.18}
$$

From (3.18), we observe that the NMSE of the equivalent channel of user k decreases with an increasing transmitted pilot symbol power E_{P} as well as the numbers of antennas equipped at the BS and at each user. As the numbers of antennas M and P approach infinity, the impact of noise will vanish asymptotically. In contrast, the channel estimation errors caused by noise in conventional fully digital massive MIMO systems cannot be mitigated by increasing the number of antennas equipped at the BS and the users [28, 53]. Therefore, the proposed channel estimation for the hybrid system outperforms the conventional pilot-aided channel estimation for the fully digital system in terms of noise mitigation. The NMSE analysis result will be verified via simulation. Here, we would point out that the proposed channel estimation scheme can exploit large array gains offered by the antenna arrays via using the BS analog beamforming matrix $\mathbf{F}_{\mathrm{RF}}$ and the users analog beamforming matrix $\mathbf{Q}_{\mathrm{RF}}$ to enhance the receive SNR of pilot symbols [170]. Therefore, it is expected that the performance of the proposed channel estimation scheme improves with increasing the numbers of antennas equipped at the BS and each user. For a multi-cell scenario, the received pilot sequences at the RF chains of the desired BS will be affected by the reused pilot sequences from neighboring cells [28, 53, 86]. The threat of pilot contamination attack will be detailed and discussed in the next chapter.
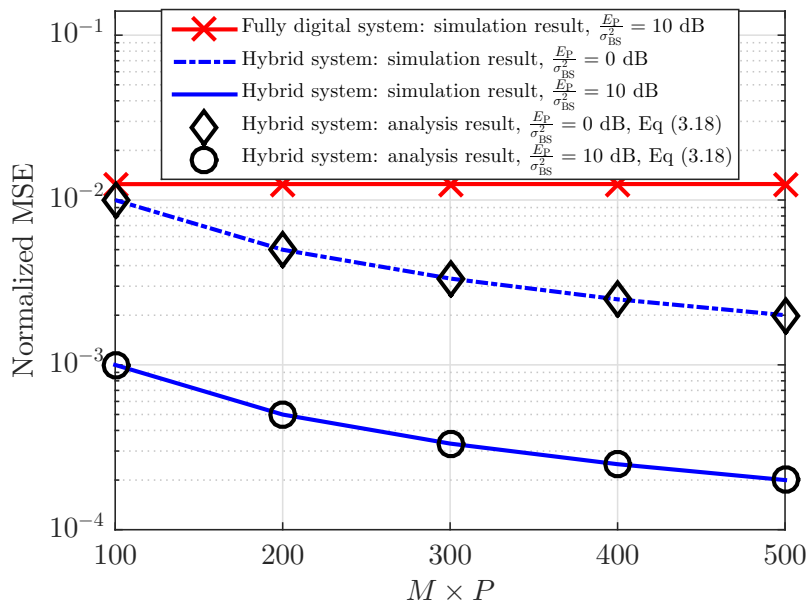
Figure 3.4: The NMSE performance comparison between the proposed pilot-aided channel estimation algorithm for the hybrid system and the conventional pilot-aided LS for the fully digital system.

Figure 3.4 shows the NMSE of the equivalent channel estimation of user k versus the total numbers of antennas equipped at the BS and the user. In Figure 3.4, we see that the simulation results match with analytical results derived in Equation (3.18). For the SNR of the transmitted pilot at $\frac{E_\text{P}}{\sigma^2_\text{BS}} = 10$ dB, it can be seen that the NMSE of the proposed algorithm decreases with an increasing $M \times P$. On the contrary, we have the observation that the numbers of antennas equipped at the BS and users have no significant impact on the conventional pilot-aided LS channel estimation for the fully digital system. Furthermore, Figure 3.4 also shows that a high transmitted pilot energy is helpful to lower the NMSE of the proposed algorithm, since the impact of noise is reduced proportional to $\frac{1}{E_\text{P} M P}$ as shown in Equation (3.18). It is interesting to note that, to meet a certain required NMSE of channel estimation, we can either increase the number of antennas equipped at the BS or the number of antennas at each user.

This indicates that increasing the number of antennas equipped at the BS can always improve the system performance, despite the possibly limited numbers of antennas equipped at the users.

3.4   ZF Precoding and Performance Analysis

In this section, we illustrate and analyze the achievable rate performance per user of the considered hybrid mmWave system under digital ZF downlink transmission. The digital ZF downlink precoding is based on the estimated equivalent channel $\widehat{\mathbf{H}}_{\text{eq}}$, which is a concatenation of the BS analog beamforming matrix $\mathbf{F}_{\text{RF}}$ and the users analog beamforming matrix $\mathbf{Q}_{\text{RF}}$. We derive a closed-form upper bound of achievable rate per user of the ZF precoding in the considered hybrid system. Also, we compare the system achievable rate upper bound obtained by the fully digital system exploiting the ZF precoding for a large number of antennas. The achievable rate performance gap between the considered hybrid mmWave system and the fully digital system is characterized, which is verified via analysis and simulation results.

3.4.1   ZF Precoding

Now, we utilize the estimated equivalent channel for downlink digital ZF precoding. To study the best achievable rate performance, we first assume that the equivalent channel is estimated in the high SNR regime. In this case, the equivalent channel is considered as perfectly estimated as the proposed channel estimation is only affected by noise as shown in Equation (3.18). Therefore, the baseband digital ZF precoder $\overline{\mathbf{W}}_{\text{eq}} \in \mathbb{C}^{N \times N}$ based on $\mathbf{H}_{\text{eq}}$ is given by

$$\overline{\mathbf{W}}_{\text{eq}} = \mathbf{H}_{\text{eq}}^* (\mathbf{H}_{\text{eq}}^T \mathbf{H}_{\text{eq}}^*)^{-1} = \left[\; \overline{\mathbf{w}}_{\text{eq},1}, \ldots, \overline{\mathbf{w}}_{\text{eq},N} \; \right], \tag{3.19}$$

where $\overline{\mathbf{w}}_{\mathrm{eq},k} \in \mathbb{C}^{N \times 1}$ is the k-th column of ZF precoder for user k. As each user is equipped with only one RF chain, one superimposed signal is received at each user at each time instant with hybrid transceivers. The received signal at user k after beamforming can be expressed as:

$$y_{\mathrm{ZF}}^k = \underbrace{\widetilde{\boldsymbol{\omega}}_k^H \mathbf{H}_k^T \mathbf{F}_{\mathrm{RF}} \overline{\beta} \overline{\mathbf{w}}_{\mathrm{eq},k} x_k}_{\text{desired signal}} + \underbrace{\widetilde{\boldsymbol{\omega}}_k^H \mathbf{H}_k^T \sum_{j=1,j \neq k}^N \mathbf{F}_{\mathrm{RF}} \overline{\beta} \overline{\mathbf{w}}_{\mathrm{eq},j} x_j}_{\text{interference}} + \underbrace{\widetilde{\boldsymbol{\omega}}_k^H \mathbf{z}_{\mathrm{MS},k}}_{\text{noise}}, \qquad (3.20)$$

where $x_k \in \mathbb{C}^{1 \times 1}$ is the transmitted symbol energy from the BS to user k, $\mathbb{E}\left[|x_k^2|\right] = E_s$, E_s is the average transmitted power for each user, $\overline{\beta} = \sqrt{\frac{1}{\mathrm{tr}(\overline{\mathbf{W}}_{\mathrm{eq}} \overline{\mathbf{W}}_{\mathrm{eq}}^H)}}$ is the transmission power normalization factor, and the effective noise part $\mathbf{z}_{\mathrm{MS},k} \sim \mathcal{CN}\left(\mathbf{0}, \sigma_{\mathrm{MS}}^2 \mathbf{I}\right)$. Due to the fact that the MU interference within the AoA directions can be suppressed by the digital ZF precoder, thus

$$\mathbf{h}_{\mathrm{eq},i}^T \overline{\mathbf{w}}_{\mathrm{eq},j} = 0, \ \forall i \neq j, \ \text{and} \ \widetilde{\boldsymbol{\omega}}_k^H \mathbf{H}_k^T \sum_{j=1,j \neq k}^N \mathbf{F}_{\mathrm{RF}} \left(\overline{\mathbf{w}}_{\mathrm{eq},j}\right) x_j = 0. \qquad (3.21)$$

Then we express the SINR of user k as

$$\mathrm{SINR}_{\mathrm{ZF}}^k = \frac{\overline{\beta}^2 E_s}{\sigma_{\mathrm{MS}}^2}. \qquad (3.22)$$

In the sequal, we study the performance of the considered hybrid mmWave systems. For simplicity, we assume that channels of all the users have the same Rician K-factor, i.e., $\upsilon_k = \upsilon, \forall k$.

3.4.2 Performance Upper Bound of ZF Precoding

Now, exploiting the SINR expression in Equation (3.22), we summarize the upper bound of achievable rate per user of the digital ZF precoding with the proposed channel estimation algorithm in the following theorem.

Theorem 3.1. *The achievable rate per user of the proposed ZF precoding is bounded by*

$$R_{\text{HB}} \leqslant R_{\text{HB}}^{\text{upper}}$$

$$= \log_2 \left\{ 1 + \left[\left(\frac{\upsilon}{\upsilon+1} \right) MP \| \mathbf{F}_{\text{RF}}^H \mathbf{F}_{\text{RF}} \|_{\text{F}}^2 + \left(\frac{1}{\upsilon+1} \right) N^2 \right] \frac{1}{N^2} \frac{E_s}{\sigma_{\text{MS}}^2} \right\}. \quad (3.23)$$

Proof: Please refer to Appendix C.   ∎

From Equation (3.23), we see that the upper bound of achievable rate per user of the proposed ZF precoding depends on the Rician K-factor, υ. Also, we can further observe that the upper bound of the achievable rate per user also depends on the BS analog beamforming matrix $\mathbf{F}_{\text{RF}}$ designed in Step 2 of the proposed channel estimation algorithm. We note that since the analog beamforming only allows the BS to transmit each user's signal via its strongest AoA direction, the proposed scheme can utilize the transmission power more effectively. In addition, the interference outside the strongest AoA directions is reduced as less transmission power will leak to undesired users. On other hand, with an increasing number of antennas at the BS, the communication channels are more likely to be orthogonal to each other. Therefore, it is interesting to evaluate the asymptotic upper bound $R_{\text{HB}}^{\text{upper}}$ for the case of a large number of antennas. We note that, even if the number of antennas equipped at the BS is sufficiently large, the required number of RF chains is still only required to equal to the number of users in the hybrid mmWave system structures.

Corollary 3.1. *In the large numbers of antennas regime, i.e., $M \to \infty$, such that $\mathbf{F}_{\text{RF}}^H \mathbf{F}_{\text{RF}} \overset{a.s.}{\to} \mathbf{I}_N$, the asymptotic achievable rate per user of the hybrid system is bounded above by*

$$R_{\text{HB}}^{\text{upper}} \overset{a.s.}{\underset{M \to \infty}{\to}} \log_2 \left\{ 1 + \left[\frac{MP}{N} \frac{\upsilon}{\upsilon+1} + \frac{1}{\upsilon+1} \right] \frac{E_s}{\sigma_{\text{MS}}^2} \right\}. \quad (3.24)$$

Proof: The result follows by substituting $\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}} \underset{M\to\infty}{\overset{a.s.}{\to}} \mathbf{I}_N$ into (3.23). ∎

From Equation (3.24), we have an intuitive observation that the asymptotic performance of the proposed precoding is mainly determined by the numbers of equipped antennas and RF chains.

3.4.3 Comparison with Fully Digital Systems

In this section, we derive the achievable rate performance of a fully digital mmWave system in the large numbers of antennas regime. The obtained analytical results in this section will be used as a reference for comparing to the proposed hybrid system. To this end, for the fully digital mmWave system, we assume that each user is equipped with one RF chain and P antennas. The P antenna array equipped at each user can provide $10\log(P)$ dB array gain. We note that, the number of antennas equipped at the BS is M and the number of RF chains equipped at the BS is equal to the number of antennas. The channel matrix from the BS to user k is given by

$$\mathbf{H}_k^T = \mathbf{h}_k^*\mathbf{h}_{\mathrm{BS},k}^T. \tag{3.25}$$

We assume that the CSI is perfectly known to the users and the BS. The BS with the fully digital system is adopted to illustrate the maximal performance gap in terms of achievable rate between the fully digital system and the considered hybrid system. The CSI at the BS for the downlink information transmission is given by

$$\mathbf{H}_{\mathrm{FD}}^T = \begin{bmatrix} \mathbf{h}_{\mathrm{BS},1}^T & ... & \mathbf{h}_{\mathrm{BS},k}^T & ... & \mathbf{h}_{\mathrm{BS},N}^T \end{bmatrix}. \tag{3.26}$$

The ZF precoder for the equivalent channel $\mathbf{H}_{\mathrm{FD}}^T$ is denoted as

$$\mathbf{W}_{\mathrm{FD}} = \mathbf{H}_{\mathrm{FD}}^* \left(\mathbf{H}_{\mathrm{FD}}^T\mathbf{H}_{\mathrm{FD}}^*\right)^{-1}. \tag{3.27}$$

Therefore, the achievable rate per user of the fully digital system is bounded by

$$R_{\text{FD}} = \log_2 \left[1 + \frac{P}{\text{tr} \left[\mathbf{W}_{\text{FD}} \mathbf{W}_{\text{FD}}^H \right]} \frac{E_s}{\sigma_{\text{MS}}^2} \right]$$

$$\overset{(c)}{\leqslant} \log_2 \left[1 + \frac{P}{N^2} \text{tr} \left[\mathbf{H}_{\text{FD}}^H \mathbf{H}_{\text{FD}} \right] \frac{E_s}{\sigma_{\text{MS}}^2} \right] = R_{\text{FD}}^{\text{upper}}, \qquad (3.28)$$

where (c) follows (C.4) in Appendix C.

Corollary 3.2. *In the large numbers of antennas regime, the asymptotic achievable rate per user of the fully digital system is bounded by*

$$R_{\text{FD}} \leqslant R_{\text{FD}}^{\text{upper}} \overset{a.s.}{\underset{M \to \infty}{\to}} \log_2 \left[1 + \frac{MP}{N} \frac{E_s}{\sigma_{\text{MS}}^2} \right]. \qquad (3.29)$$

Proof: The result follows by substituting $\frac{1}{M} \mathbf{H}_{\text{FD}}^H \mathbf{H}_{\text{FD}} \overset{a.s.}{\underset{M \to \infty}{\to}} \mathbf{I}_N$ into (3.28).

∎

Based on Equations (3.24) and (3.29), we further quantify the achievable rate performance gap between the considered hybrid system and the fully digital system in the large numbers of antennas regime.

Corollary 3.3. *In the large numbers of antennas regime, the gap between the achievable rate upper bounds for the hybrid system and the fully digital system can be expressed as:*

$$\Delta_{\text{GAP}} = R_{\text{HB}}^{\text{upper}} - R_{\text{FD}}^{\text{upper}} \overset{(S)}{\underset{M \to \infty}{\approx}} \log_2 \left\{ \frac{v}{v+1} \right\} \leqslant 0, \qquad (3.30)$$

where (S) *stands for* $\text{SNR} \to \infty$.

Proof: The result follows by substituting (3.24) and (3.29) into (3.30). ∎

In the large numbers of antennas regime, based on Equations (3.24) and (3.29), it is interesting to observe that with an increasing Rician K-factor v, the performance upper bounds of the two considered systems will coincide. Intuitively, as the Rician K-factor increases, the LOS component becomes the dominant element

of the communication channel, as shown in Equation (3.1). Therefore, the BS analog beamforming matrix based on the estimated strongest AoA will allocate a smaller portion of the transmitted signal energy to the scattering component. At the same time, the interference caused by other users is suppressed by the digital baseband ZF precoding, $\overline{\mathbf{W}}_{\text{eq}}$.

In Figure 3.5, we present a comparison between the achievable rate per user of the hybrid system and the fully digital system for $M = 100$, $N = 10$, and a Rician K-factor of $v_k = 2, \forall k$. Firstly, our simulation results verify the tightness of derived upper bounds in Equations (3.24) and (3.29). It can be observed from Figure 3.5 that even for a small value of Rician K-factor, e.g. $v = 2$, our proposed channel estimation scheme with ZF precoding can achieve a considerable high sum rate performance due to its interference suppression capability.
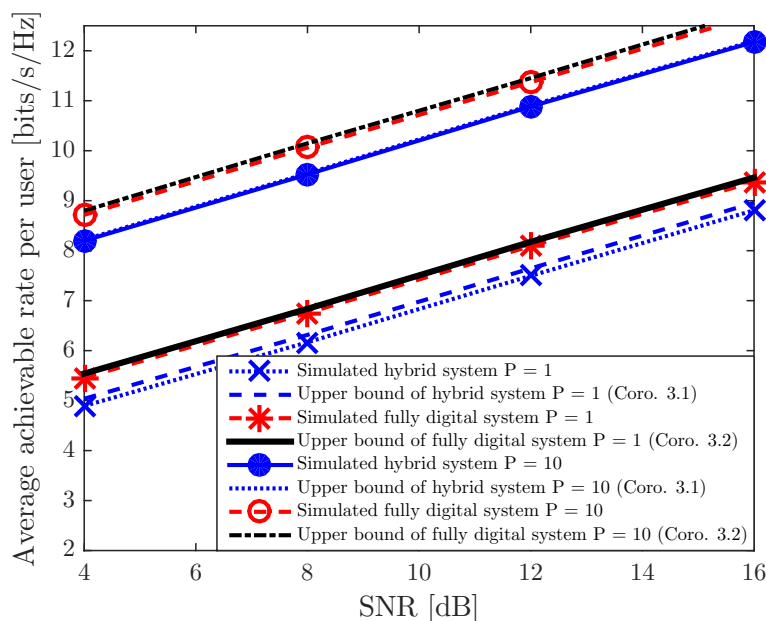
Figure 3.5: Average achievable rate per user [bits/s/Hz] versus SNR for the hybrid and the fully digital systems with $M = 100$, $N = 10$ and $v = 2$.
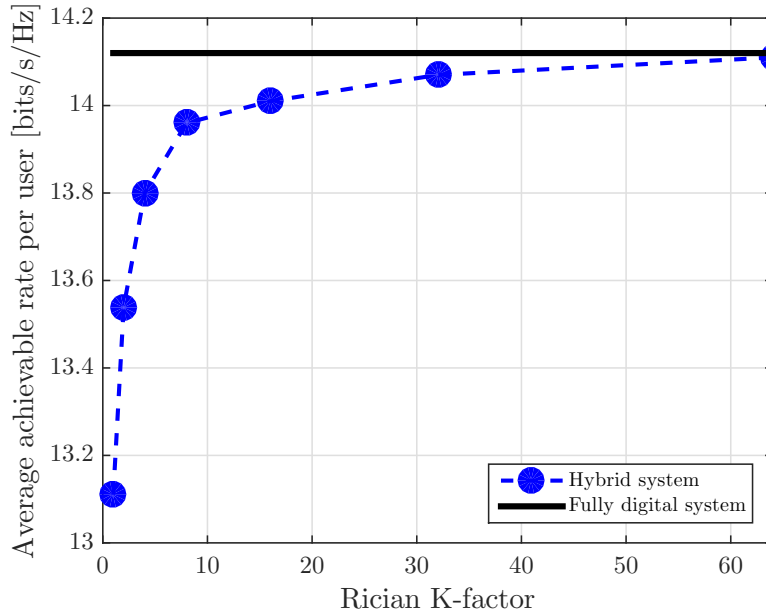
Figure 3.6: Average achievable rate [bits/s/Hz] versus Rician K-factor for different systems with SNR = 20 dB, $M = 100$, $P = 16$, and $N = 4$.

In Figure 3.6, the achievable rate performance gap between the fully digital system and the hybrid system decreases with the increasing Rician K-factor, which is predicted by Equation (3.30). In particular, with a sufficiently large Rician K-factor, the achievable rates of these two systems will coincide.

3.5 Performance Analysis with Hardware Impairments

In the last section, we study the proposed mmWave hybrid system in ideal hardware and ideal estimation conditions. In practice, hardware components may have various types of impairments that may degrade the achievable rate performance[6],

---

[6]The authors of [152] proved that for a fully digital massive MIMO system, the additive distortion caused by hardware impairments create finite ceilings on the channel estimation accuracy and on the uplink/downlink capacity, which are irrespective of the SNR and the number of base station antennas. In addition, work [153] concluded that the impact of phase

e.g. phase errors in phase shifters induced by thermal noise, transceiver RF beamforming errors caused by AoA estimation errors, and channel estimation errors affected by independent additive distortion noises [83, 152, 153]. In this section, we analyze the rate performance degradation under hardware impairments.

3.5.1 Transceiver Beamforming Errors and Random Phase Errors

Here, we first discuss the scenario that the equivalent channels are estimated in the high pilot transmit power regime, i.e., $\frac{E_{\mathrm{P}}}{\sigma_{\mathrm{BS}}^2} \to \infty$, where errors caused by thermal noise are negligible. As a result, we focus on the effect of the phase errors in phase shifters and transceiver analog beamforming errors on the achievable rate performance.

Firstly, we start from quantifying the impact of transceiver beamforming errors, which is caused by AoA estimation errors. As the number of antennas equipped at the BS is sufficiently large, the beamwidth of the antenna array is narrow. Therefore, even a small AoA estimation error may cause significant impacts on the system performance. Here, the BS analog beamforming error matrix of user k, $\boldsymbol{\Delta}_k \in \mathbb{C}^{M \times M}$, which is caused by the AoA estimation error, can be expressed as

$$\boldsymbol{\Delta}_k = \frac{1}{\sqrt{M}} \begin{bmatrix} 1 & & \\ & \ddots & \\ & & e^{j2\pi(M-1)\frac{d}{\lambda}\cos\left(\Delta\theta_{\mathrm{BS},k}\right)} \end{bmatrix}, \qquad (3.31)$$

where the AoA estimation error $\Delta\theta_{\mathrm{BS},k}$ at the BS for user k is modeled by i.i.d. random variable with the distribution $\mathcal{CN}\left(0, \varrho_{\mathrm{BS}}^2\right)$. In addition, the user beam-

---

error on hybrid beamforming is a further reduction on the number of effective antenna per user. In fact, the achievable rate degradation caused by phase errors can be compensated by simply employing more transmit antennas.

forming error matrix $\boldsymbol{\Theta}_k \in \mathbb{C}^{P \times P}$, which is caused by AoA estimation error $\Delta\theta_{\text{MS},k}$, can also be similarly formulated. Now, we assume that the hybrid system is only affected by AoA estimation errors. Therefore, the received pilot symbols at the BS during the third step of channel estimation is expressed as:

$$\left(\widehat{\mathbf{s}}_k^{\text{BE}}\right)^T = \widetilde{\boldsymbol{\gamma}}_k^T \boldsymbol{\Delta}_k \sum_{i=1}^{N} \mathbf{H}_i \boldsymbol{\Theta}_i \widetilde{\boldsymbol{\omega}}_i^* \sqrt{E_{\text{P}}} \boldsymbol{\Phi}_i^T + \underbrace{\widehat{\mathbf{z}}_{\text{eq}}}_{\text{effective noise}} , \tag{3.32}$$

where the entries of effective noise part $\widehat{\mathbf{z}}_{\text{eq}} = \mathbf{f}_k^T \boldsymbol{\Delta}_k \boldsymbol{\Xi}_k \mathbf{Z}$ can still be modeled by i.i.d. random variable with distribution $\mathcal{CN}\left(0, \sigma_{\text{BS}}^2\right)$. As we assumed that $\dfrac{E_{\text{P}}}{\sigma_{\text{BS}}^2} \to \infty$, the estimated equivalent channel under transceiver analog beamforming errors, $\mathbf{H}_{\text{eq}}^{\text{BE}}$, is given by

$$\left(\mathbf{H}_{\text{eq}}^{\text{BE}}\right)^T = \begin{bmatrix} \widetilde{\boldsymbol{\omega}}_1^H \boldsymbol{\Theta}_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \widetilde{\boldsymbol{\omega}}_N^H \boldsymbol{\Theta}_N \end{bmatrix} \begin{bmatrix} \mathbf{H}_1^T \\ \vdots \\ \mathbf{H}_N^T \end{bmatrix} \begin{bmatrix} \boldsymbol{\Delta}_1 \widetilde{\boldsymbol{\gamma}}_1 \ldots \boldsymbol{\Delta}_N \widetilde{\boldsymbol{\gamma}}_N \end{bmatrix}. \tag{3.33}$$

Following the similar signal processing procedures as in Equations (3.19)–(3.21), the average achievable rate per user under the BS and users analog beamforming errors can be expressed as

$$\widehat{R}_{\text{HB}}^{\text{BE}} = \mathbb{E}_{\mathbf{H}_{\text{S}}, \Delta\theta_{\text{MS}}, \Delta\theta_{\text{BS}}} \left\{ \log_2 \left[1 + \frac{(\widehat{\beta}^{\text{BE}})^2 E_s}{\sigma_{\text{MS}}^2} \right] \right\}, \tag{3.34}$$

where $\widehat{\beta}^{\text{BE}} = \sqrt{\frac{1}{\text{tr}[(\widehat{\mathbf{W}}_{\text{eq}}^{\text{BE}})(\widehat{\mathbf{W}}_{\text{eq}}^{\text{BE}})^H]}}$ is the transmission power normalization factor and $\widehat{\mathbf{W}}_{\text{eq}}^{\text{BE}}$ is the downlink ZF precoder based on $\mathbf{H}_{\text{eq}}^{\text{BE}}$. From Equation (3.34), it is interesting to observe that the achievable rate performance is not bounded by above. This is due to the fact that the impact of transceiver analog beamforming errors on different RF chains can be estimated by the pilot matrix $\boldsymbol{\Psi}$, treated as parts of the channel, and compensated by the digital ZF transmission. There-

fore, it is expected that our proposed scheme is robust against transceiver analog beamforming errors.

Observing Equations (3.34) and (3.22), we see that the performance degradation caused by AoA estimation errors can be quantified by comparing power normalization factors $\widehat{\beta}^{\mathrm{BE}}$ to $\overline{\beta}$. In particular, the hybrid system with hardware impairments incurs a power loss in the received SINR at the users. However, it is difficult to provide an explicit mathematical expression to quantify the loss. Therefore, we borrow the "power loss" concept from the literature of array signal processing. Specifically, the power loss due to AoA estimation errors is related to the HPBW [111] and the HPBW can be approximated (p. 48, [111])

$$\mathrm{HPBW} \approx \frac{1.782}{M}, \tag{3.35}$$

where M is the number of antennas equipped in the array. Hence, we can approximate the impact of AoA estimation errors by introducing a power loss coefficient $\xi \in (0, 1]$, which is determined by the beam pattern of the array and the variance of the distribution of AoA estimation errors. For example, if the variance of AoA estimation errors is assumed no larger than half of the HPBW of the antenna array[7], i.e., $\varrho_{\mathrm{BS}}^2 \leqslant \frac{1.782}{2M}$, the power loss coefficient is given by $\xi \approx 0.5$ according to the half power loss principle [111]. With an increasing AoA estimation error variance ϱ_{BS}^2, ξ decreases significantly[8]. With the help of ξ, we now can express the approximation of $\widehat{R}_{\mathrm{HB}}^{\mathrm{BE}}$, the average achievable rate per user under transceiver

---

[7]In general, this assumption is valid for practical communication systems with a sufficiently large number of antennas.

[8]The calculation of ξ for different numbers of antennas and AoA errors is well studied for linear arrays and interested readers may refer to [111] for a detailed discussion.

beamforming errors, in the large numbers of antennas regime as

$$\widehat{R}_{\text{HB}}^{\text{BE}} \overset{M \to \infty}{\approx} \log_2 \left\{ 1 + \left[\frac{\upsilon}{\upsilon + 1} \frac{MP}{N} \xi + \frac{1}{\upsilon + 1} \right] \frac{E_s}{\sigma_{\text{MS}}^2} \right\}. \tag{3.36}$$

From (3.36), it is interesting to note that, the beamforming errors matrices $\boldsymbol{\Theta}_k$ and $\boldsymbol{\Delta}_k$ only lead to a certain power loss, which depends on ξ. Besides, this loss can be compensated at the expense of a higher transmit power, i.e., increase E_{S}.

Now, we discuss the impact of phase errors, which are caused by the AWGN and the limited quantization resolution of the phase shifters. At the users sides, phase errors of user k are modeled by [153]

$$\boldsymbol{\Lambda}_k = \text{diag} \left\{ e^{j\Delta\phi_{k,p}} \right\} \in \mathbb{C}^{P \times P}, \ p \in \{1, \ldots, P\}, \tag{3.37}$$

where phase errors $\Delta\phi_{k,p}, \forall p$, are uniformly distributed over $[-a, \ a]$ and $a > 0$ is the maximal phase error of user k. Similarly, for the phase shifters connected with the k-th RF chain in the BS, the associated phase errors are given by

$$\boldsymbol{\Xi}_k = \text{diag} \left\{ e^{j\Delta\psi_{k,m}} \right\} \in \mathbb{C}^{M \times M}, \ m \in \{1, \ldots, M\}, \tag{3.38}$$

where the errors $\Delta\psi_{k,m}, \forall m$, are uniformly distributed over $[-b, \ b]$ and $b > 0$ is the maximal phase error of the BS. The property of $\boldsymbol{\Lambda}_k$ and $\boldsymbol{\Xi}_k$ can be expressed as

$$\mathbb{E}\left[\boldsymbol{\Lambda}_k\right] = \int_{-a}^{a} \frac{1}{2a} e^{j\Delta\phi_{k,p}} \, d\Delta\phi_{k,p} = \frac{\sin(a)}{a} \mathbf{I}_{\text{P}} \ \text{and}$$

$$\mathbb{E}\left[\boldsymbol{\Xi}_k\right] = \int_{-b}^{b} \frac{1}{2b} e^{-j\Delta\psi_{k,m}} \, d\Delta\psi_{k,m} = \frac{\sin(b)}{b} \mathbf{I}_{\text{M}}, \tag{3.39}$$

respectively. Then, the received pilot symbols $\widehat{\mathbf{s}}_k^T$ used for the equivalent channel

estimation at the k-th RF chain, which is under the impact of phase errors and transceiver analog beamforming errors, can be expressed as

$$\widehat{\mathbf{s}}_k^T = \widetilde{\boldsymbol{\gamma}}_k^T \boldsymbol{\Delta}_k \boldsymbol{\Xi}_k \sum_{i=1}^{N} \mathbf{H}_i \boldsymbol{\Lambda}_i \boldsymbol{\Theta}_i \widetilde{\boldsymbol{\omega}}_i^* \sqrt{E_{\mathrm{P}}} \boldsymbol{\Phi}_i^T + \underbrace{\widehat{\mathbf{z}}_{\mathrm{eq}}}_{\text{effective noise}} . \qquad (3.40)$$

Similarly, we can express the average achievable rate per user as

$$\widehat{R}_{\mathrm{HB}} = \mathbb{E}_{\mathbf{H}_{\mathrm{S}},\Delta\phi,\Delta\psi,\Delta\theta_{\mathrm{MS}},\Delta\theta_{\mathrm{BS}}} \left\{ \log_2 \left[1 + \frac{\widehat{\beta}^2 E_s}{\sigma_{\mathrm{MS}}^2} \right] \right\}, \qquad (3.41)$$

where $\widehat{\beta}$ is the transmission power normalization factor under phase errors and transceiver beamforming errors. Based on Equations (3.24), (3.36), and (3.39), the approximation of the average achievable rate per user $\widehat{R}_{\mathrm{HB}}$ in the large numbers of antennas regime is given by

$$\widehat{R}_{\mathrm{HB}} \overset{M\to\infty}{\approx} \log_2 \left\{ 1 + \left[\frac{\upsilon}{\upsilon+1} \frac{MP}{N} \widehat{\xi} + \frac{1}{\upsilon+1} \right] \frac{E_s}{\sigma_{\mathrm{MS}}^2} \right\}, \qquad (3.42)$$

where $\widehat{\xi} = \left(\frac{\sin(a)}{a} \right)^2 \left(\frac{\sin(b)}{b} \right)^2 \xi$ is the equivalent power loss coefficient. From Equation (3.42), we note that the joint impact of random phase errors and transceiver beamforming errors on the average achievable rate cause performance degradation compared to the case of perfect hardware. Besides, using extra transmission power can compensate the performance degradation.

3.5.2 Hardware Impairment and Imperfect Channel Estimation

In this section, we further study the system performance by jointly taking account the impact of hardware impairments and equivalent channel estimation errors

caused by noise. The estimated equivalent channel $\widetilde{\mathbf{H}}_{\text{eq}}^T$ can be expressed as

$$\widetilde{\mathbf{H}}_{\text{eq}}^T = \widehat{\mathbf{H}}_{\text{eq}}^T + \Delta \widehat{\mathbf{H}}_{\text{eq}}^T, \tag{3.43}$$

where $\widehat{\mathbf{H}}_{\text{eq}}$ is the equivalent channel under random phase errors and transceiver beamforming errors and the entries of the normalized channel estimation error $\frac{1}{\sqrt{MP}}\Delta \widehat{\mathbf{H}}_{\text{eq}}^T$ are modeled by i.i.d. random variables with distribution $\mathcal{CN}\left(0, \delta^2\right)$, $\delta^2 = \text{MSE}_{\text{eq}}$, which is given by Equation (3.18). The ZF precoding matrix for the considered hybrid system based on the imperfect CSI is given by

$$\widetilde{\mathbf{W}}_{\text{eq}} = \widetilde{\mathbf{H}}_{\text{eq}}^* \left[\widetilde{\mathbf{H}}_{\text{eq}}^T \widetilde{\mathbf{H}}_{\text{eq}}^* \right]^{-1} = \widehat{\mathbf{W}}_{\text{eq}} + \Delta \widehat{\mathbf{W}}_{\text{eq}} = \left[\begin{array}{ccc} \widetilde{\mathbf{w}}_{\text{eq},1} & \cdots & \widetilde{\mathbf{w}}_{\text{eq},N} \end{array} \right], \tag{3.44}$$

where $\widetilde{\mathbf{w}}_{\text{eq},k} \in \mathbb{C}^{N \times 1}$ is the k-th column vector of $\widetilde{\mathbf{W}}_{\text{eq}}$ and $\widehat{\mathbf{W}}_{\text{eq}}$ is the precoder based on $\widehat{\mathbf{H}}_{\text{eq}}$. The received signal at user k under imperfect CSI is given by

$$\widetilde{y}_{\text{ZF}}^k = \underbrace{\widetilde{\beta}x_k}_{\text{desired signal}} + \underbrace{\widetilde{\beta}\widehat{\boldsymbol{\omega}}_k^H \mathbf{H}_k^T \widehat{\mathbf{F}}_{\text{RF}} \Delta \widehat{\mathbf{W}}_{\text{eq}}\mathbf{x}}_{\text{intra-cell interference}} + z_k, \tag{3.45}$$

where $\widetilde{\beta} = \sqrt{\frac{1}{\text{tr}\left(\widetilde{\mathbf{W}}_{\text{eq}}\widetilde{\mathbf{W}}_{\text{eq}}^H\right)}}$ is the power normalization factor. $\Delta \widehat{\mathbf{w}}_{\text{eq},j} \in \mathbb{C}^{M \times 1}$ denotes the j-th column vector of the ZF precoder error matrix $\Delta \widehat{\mathbf{W}}_{\text{eq}} = \widetilde{\mathbf{W}}_{\text{eq}} - \widehat{\mathbf{W}}_{\text{eq}}$ and $\mathbf{x} = [x_1, x_2, \ldots, x_N]^T$ denotes the transmitted signal for all users.

We then express the SINR of user k as

$$\widetilde{\text{SINR}}_{\text{ZF}}^k = \frac{E_s\widetilde{\beta}^2}{\widetilde{\beta}^2 E_s \widehat{\mathbf{h}}_{\text{eq},k}^T \text{E}_{\Delta \widehat{\mathbf{H}}_{\text{eq}}}\left[\Delta \widehat{\mathbf{W}}_{\text{eq}}\Delta \widehat{\mathbf{W}}_{\text{eq}}^H\right]\widehat{\mathbf{h}}_{\text{eq},k}^* + \sigma_{\text{MS}}^2}. \tag{3.46}$$

Now we summarize the achievable rate per user in the high SNR regime in the following theorem.

Theorem 3.2. *As the receive* $\text{SNR} = \dfrac{E_s}{\sigma_{\text{MS}}^2}$ *approaches infinity, the approximated*

achievable rate of user k of ZF precoding under imperfect hybrid CSI is given by

$$\widetilde{R}_{\text{ZF},k} \approx \log_2 \left\{ 1 + \left[\left(\sqrt{1+\delta^2} - 1 \right)^2 - 2\sqrt{1+\delta^2} \left(\sqrt{1+\delta^2} - 1 \right) \delta^2 N \eta_{kk} \right. \right.$$
$$\left. \left. + \left(\sqrt{1+\delta^2} \right) \left(2 - \sqrt{1+\delta^2} \right) \delta^2 \text{tr} \left(\mathbf{K}^{-1} \right) \right]^{-1} \right\}, \tag{3.47}$$

where $\mathbf{K} = \widehat{\mathbf{H}}_{\text{eq}}^T \widehat{\mathbf{H}}_{\text{eq}}^*$, η_{kk} *represents the k-th diagonal element of* $\mathbf{K}^{-1}$, *and* δ^2 *is the NMSE of channel estimation given by Equation (3.18).*

Proof: Please refer to Appendix D. ∎

Corollary 3.4. *In the large numbers of antennas regime, i.e.,* $M \to \infty$, *the asymptotic average achievable rate per user of the hybrid system is approximated by*

$$\widetilde{R}_{\text{ZF}} \approx \log_2 \left\{ 1 + \left[\left(\sqrt{1+\delta^2} - 1 \right)^2 \right. \right.$$
$$- 2\sqrt{1+\delta^2} \left(\sqrt{1+\delta^2} - 1 \right) \frac{\delta^2 N}{\widehat{\xi} M P} \frac{\upsilon + 1}{\upsilon}$$
$$\left. \left. + \left(\sqrt{1+\delta^2} \right) \left(2 - \sqrt{1+\delta^2} \right) \frac{\delta^2 N}{\widehat{\xi} M P} \frac{\upsilon + 1}{\upsilon} \right]^{-1} \right\}. \tag{3.48}$$

Proof: The result follows by substituting $\mathbf{K} \xrightarrow[M \to \infty]{a.s.} \widehat{\xi} M P \frac{\upsilon}{\upsilon+1} \mathbf{I}_N$ into (3.47). ∎

Based on (3.42) and (3.48), the additional achievable rate performance degradation, which is caused jointly by random phase errors and transceiver beamforming errors, is further summarized in the following corollary.

Corollary 3.5. *In the large numbers of antennas regime, i.e.,* $M \to \infty$, *the approximated achievable rate per user performance gap between the system with ideal hardware and the system under phase errors and transceiver beamforming*
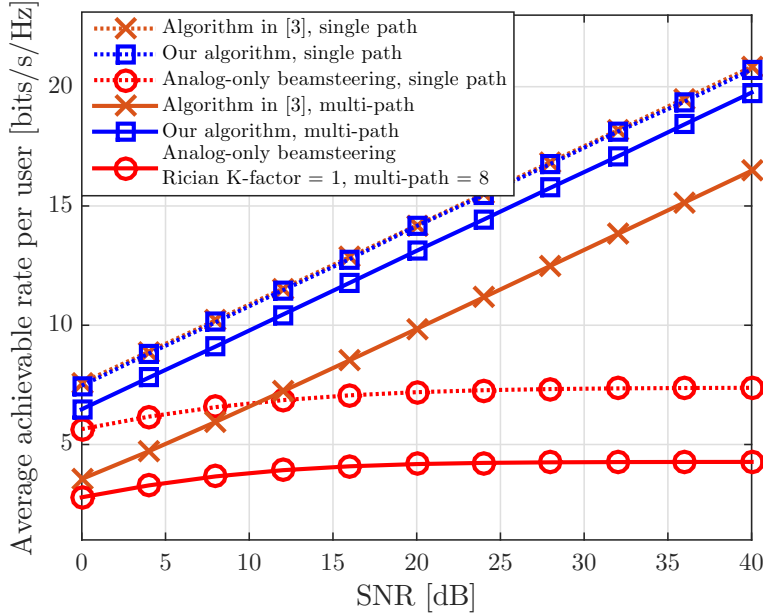
Figure 3.7: The average achievable rate per user [bits/s/Hz] versus SNR for algorithm proposed in [3] and our proposed algorithm.

errors is given by

$$\Delta\text{Gap} \approx \log_2 \left[\frac{1}{\bar{\xi}} \right]. \tag{3.49}$$

Proof: The result comes after some mathematical manipulation on (3.42) and (3.48). ■

3.6 Simulation and Discussion

In this section, we present further numerical results to validate our analysis. We consider a small single-cell hybrid mmWave system. In Figure 3.7, we compare the achievable rates using the proposed algorithm and the algorithm proposed by [3] for sparse and non-sparse mmWave channels. We assume perfect channel estimation with $M = 100$, $N = 4$, and $P = 16$. For non-sparse mmWave channels, $\upsilon_k = 1, \forall k$. Firstly, we illustrate the effectiveness of the proposed channel

estimation algorithm over non-sparse mmWave channels. For sparse single-path channels, the achievable rate of the proposed algorithm matches with the algorithm proposed in [3]. For non-sparse mmWave channels, e.g. with the number of multi-paths $N_l = 8$, we observe that the proposed algorithm achieves a better system performance than that of the algorithm proposed in [3]. The reasons are two-fold. Firstly, the proposed algorithm can effectively reduce the MU interference via the proposed analog beamformers and digital ZF precoder. In particular, the analog beamformers adopted at the BS and the users allow transmission and receiving along the strongest AoA directions, respectively. Therefore, the interference outside the strongest AoA directions is reduced. Secondly, the digital ZF precoder designed based on the equivalent channels can remove the MU interference within the strongest AoAs. We note that since the analog beamforming only allows the transmission along the strongest AoA directions, the proposed scheme can utilize the transmission power more efficiently. In contrast, the algorithm proposed in [3], which aims to maximize the desired signal energy, does not suppress the MU interference as effective as our proposed algorithm. Furthermore, Figure 3.7 also illustrates that a significant achievable rate gain can be brought by the proposed channel estimation and the digital ZF precoding over a simple analog-only beamforming steering scheme.

In Figure 3.8, we illustrate two sets of comparisons to validate our derived results in Equations (3.42) and (3.48), and to show the impact of hardware impairments on system performance. In the simulations, we set the Rician K-factor as $\upsilon_k = 2, \forall k$, the number of antennas equipped at the BS is $M = 100$, the number of antennas equipped at each user is $P = 8$, and the number of users is $N = 8$. We assume that the maximum phase error value of phase shifters is $\sigma_\Delta = 3$ degrees and the variance of AoA estimation errors at the BS side is $\varrho_{\mathrm{BS}}^2 = \dfrac{1.782}{2M}$.
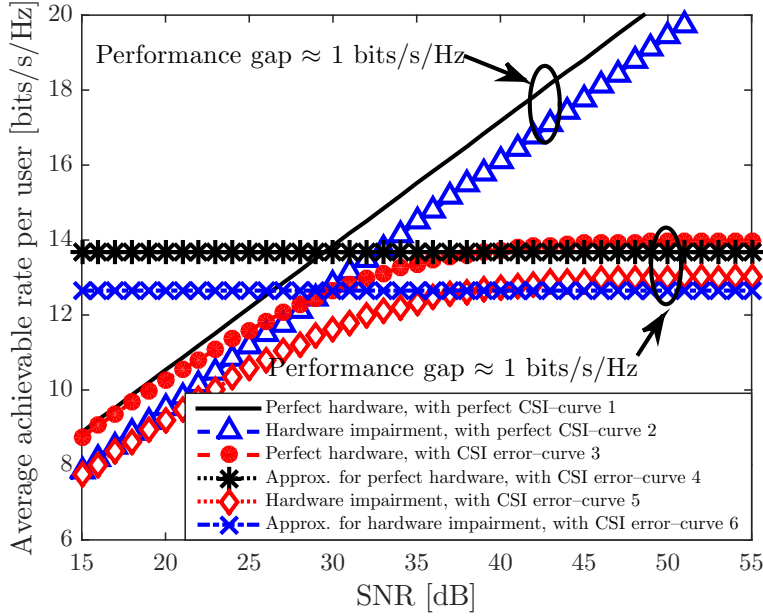
Figure 3.8: The average achievable rate per user [bits/s/Hz] versus SNR for the hybrid system.

In this case, the equivalent coefficient is $\widehat{\xi} \approx 0.5$, which predicts that there is a 3-dB loss in SNR caused by AoA estimation errors and phase errors. First, we observe that even with transceiver beamforming errors and random phase errors, the achievable rate of the proposed scheme (curve 2) scales with increasing SNR and is unbounded by above. The small gap between the achievable rate per user with perfect hardware (curve 1) and the achievable rate per user with hardware impairment (curve 2) confirms the robustness of the proposed scheme against random phase errors as well as transceiver beamforming errors, which is predicted by Equation (3.42). In addition, the 3 dB of extra power consumption caused by hardware impairments is also verified via the comparison between curves 1 and 2. Figure 3.8 also illustrates a comparison between the proposed scheme with imperfect equivalent CSI under perfect hardware (curve 3) and the proposed scheme with imperfect CSI under random phase errors and transceiver beamforming errors (curve 5), which validates the correctness of Theorem 3.2.
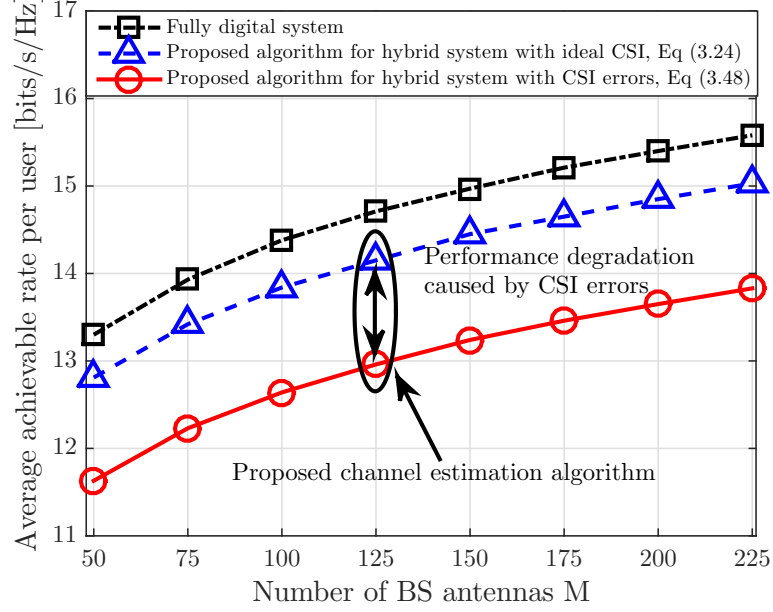
Figure 3.9: Average achievable rate [bits/s/Hz] versus different number of BS antennas for different structures with $P = 8$, the number of users $N = 8$, the normalized CSI NMSE = 0.005, SNR = 30 dB, and the Rician K-factor of 2.

These two simulation curves assume an identical normalized CSI error variance δ^2. In particular, we set the NMSE of CSI as $\delta^2 = 0.005$, which is achieved by the proposed channel estimation scheme as shown in Figure 3.4. In the high receive SNR and large numbers of antennas regimes, curve 3 and curve 5 can be asymptotically approximated by curve 4 and curve 6, respectively. Interestingly, the performance gaps between curve 3 and curve 5 as well as curve 1 and curve 2 are approximately $\Delta \text{Gap} \approx \log_2 \left(\dfrac{1}{\tilde{\xi}} \right) = 1$ bits/s/Hz, which are accurately characterized by Equation (3.49).

Now, we evaluate the achievable rate versus the number of antennas in Figures 3.9 and 3.10. The setup in Figure 3.9 is considered at receive SNR = 30 dB and the number of antennas equipped at each user is $P = 8$, with different numbers of BS antennas. Figure 3.9 shows that the achievable rate performance of the proposed digital ZF precoding increases with increasing numbers of antennas at
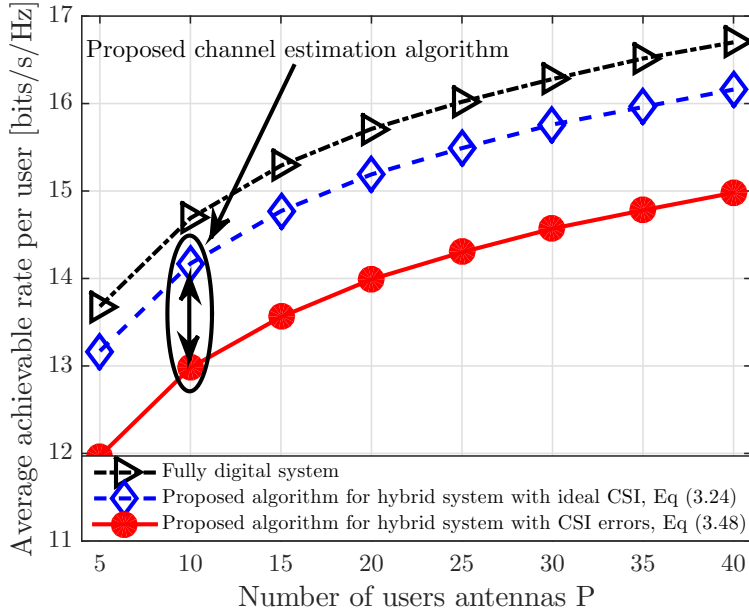
Figure 3.10: Average achievable rate [bits/s/Hz] versus different number of users antennas for different structures with $M = 100$, the number of users $N = 8$, the normalized CSI NMSE $= 0.005$, SNR $= 30$ dB, and the Rician K-factor of 2.

the BS, despite the existence of CSI estimation errors. This is mainly due to the fact that equipping more antennas at the BS and the users can lead to higher array gains. It can be observed that the achievable rate of the proposed system scales with the number of BS antennas with a similar slope as the fully digital system, which shows the effectiveness of the proposed scheme for the hybrid system in exploiting the spatial degrees of freedom. In Figure 3.10, the SNR setup and $M = 100$ are the same as Figure 3.9 but with different numbers of users antennas. Similar phenomena are observed as in Figure 3.9. With the same value of MP and identical NMSE δ^2, it is shown in Figure 3.9 and Figure 3.10 that the achievable rate performance degradations (the gap between the dash lines and the solid lines) caused by CSI errors are identical.

3.7   Summary

In this chapter, we proposed a low-complexity mmWave channel estimation algorithm exploiting the strongest AoA for the MU hybrid mmWave systems, which is applicable for both sparse and non-sparse mmWave channel environments. The NMSE performance of the proposed channel estimation was analyzed and verified via numerical simulation. The achievable rate performance of designed analog beamforming and digital ZF precoding based on the proposed channel estimation scheme was derived and compared to that of the fully digital system. The analytical and simulation results indicated that the proposed scheme can approach the rate performance achieved by the fully digital system with sufficiently large Rician K-factors. By taking into consideration of the effects of random phase errors, transceiver beamforming errors, and CSI errors in the rate performance analysis, closed-form approximation of achievable rate in the high SNR regime was derived and verified via simulation. Our results showed that the proposed scheme is robust against random phase errors and transceiver beamforming errors.

Chapter 4

Multi-cell Hybrid mmWave MIMO Systems: Pilot Contamination and Interference Mitigation

4.1 Introduction

In this chapter, we consider a multi-cell MU hybrid mmWave system. In the previous chapter, we propose and detail a novel non-feedback non-iterative channel estimation algorithm which is applicable to both non-sparse and sparse mmWave channels. In particular, we apply the non-feedback TDD-based mmWave channel estimation algorithm proposed in Chapter 3 and [1] to the considered multi-cell scenario and study the corresponding performance of channel estimation under the impact of pilot contamination. In addition, we analyze the downlink achievable rate performance of the hybrid mmWave system by taking into account the CSI

errors caused by pilot contamination and inter-cell interferences originated from the BSs in neighboring cells. Note that, for pilot contamination mitigation, this chapter does not require any information of covariance matrices of pilot-sharing users in neighbouring cells as required by the multi-cell MMSE-based precoding algorithm proposed in [144]. In addition, the algorithm adopted in this chapter can work well while mean values of cross-cell channels from pilot-sharing users in neighbouring cells to the desired BS are non-zero. Our main contributions are summarized as follows:

- We study the impact of pilot contamination on the uplink mmWave channel estimation due to the reuse of orthogonal pilot symbols among different cells. Our results reveal that in the phase of channel estimation, the receive analog beamforming matrix adopted at the desired BS forms a spatial filter which blocks the signal reception of the undesired pilot symbols from neighboring cells. In particular, with an increasing number of antennas equipped at each RF chain, the main lobe beamwidth of the spatial filter, which aligns to the strongest AoA direction, becomes narrower and the amplitude of sidelobes becomes lower. Thus, the impact of pilot contamination caused by the users outside strongest AoA directions can be mitigated. We mathematically prove that the normalized MSE performance of the channel estimation algorithm improves proportionally with the increasing number of antennas equipped at each RF chain, which is different from previous results in [28, 53].

- We adopt ZF precoding for the downlink transmission based on the estimated CSI. Taking into account the impact of pilot contamination and the inter-cell interference from the neighboring BSs, we analyze and derive the closed-form approximation of the average achievable rate performance

in the large number of antennas regimes. In addition, we show that the achievable rate scales with M in the multi-cell scenario, where M is the number of antennas equipped at each RF chain. We derive a scaling law of the achievable rate per users, $\log_2 M$. In contrast, the average achievable rate of fully digital systems adopting the conventional LS channel estimation algorithm suffers from the severe impact of pilot contamination. All the derived analytical results are verified via simulations.

The rest of the chapter is organized as follows. Section 4.2 describes the system model considered throughout the chapter. In Section 4.3, we detail the uplink channel estimation under the pilot contamination. In Section 4.4, we investigate the downlink transmission rate performance under pilot contamination. Numerical results and related discussions are presented in Section 4.5. Finally, Section 4.6 summarizes the chapter.

4.2 System Model

In this chapter, a multi-cell MU hybrid subarray mmWave system is considered. The system consists of L neighboring cells and there are one BS and N users in each cell, cf. Figure 4.1. The BS in each cell is equipped with N_{RF} RF chains serving the N users simultaneously. We assume that each RF chain equipped at the BS can access to a ULA with M antennas by using M phase shifters. Besides, each user is equipped with one RF chain and a P-antenna array. In addition, we focus on $M \geqslant N_{\text{RF}}$, which exploits a large antenna array gain with limited number of RF chains.

Work [20] proved that when the number of RF chains equipped at the BS is more than twice of the number of users, the hybrid systems can realize the rate performance of the fully digital systems. In general, for the case of $N_{\text{RF}} = N$, the

hybrid system can provide enough degrees of freedom to serve N users simultaneously. If the BS can equip with more RF chains, more degrees of freedom can be exploited for the design of analog beamformers and digital precoders. In addition, the extra RF chains adopted by the BS can be used for user scheduling design and multipath diversity exploitation. Thus, for the case that $N_{\mathrm{RF}} > N$, hybrid systems can achieve higher rate performance than that of $N_{\mathrm{RF}} = N$. In addition, authors of [48,171,172] employed RF chains with sets of digitally controlled phase paired phase-shifters and switches to achieve the same rate performance as that of the fully digital system. Furthermore, with the support of an advanced hybrid architecture [48, 171, 172], hybrid systems can achieve higher rate performance than that of adopting a simple ULA antenna array case.

However, how to exploit extra RF chains as well as advanced hybrid architectures to further improve the rate performance of hybrid systems is beyond the scope of this chapter. We may consider the case of $N_{\mathrm{RF}} > N$ and the case of adopting advanced hybrid architectures for channel estimation and downlink transmission precoding in our future works. To simplify the analysis in the following sections, without loss of generality, we set $N_{\mathrm{RF}} = N$ and each cell has the same number of RF chains equipped at the BS.

According to the widely adopted setting for multi-cell TDD in uplink channel estimation and downlink data transmission, we assume that the users and the BSs in all cells are fully synchronized in time [28, 53]. We denote $\mathbf{H}_k \in \mathbb{C}^{M \times P}$ as the uplink channel matrix between the desired BS and user k in the desired cell. Besides, $\mathbf{H}_k$ is a narrowband slow time-varying block fading channel. Recent field tests show that both strongest AoA components and non-negligible scattering components may exist in mmWave propagation channels [29, 102], especially in the urban areas. In this chapter, without loss of generality, $\mathbf{H}_k$, consists of a
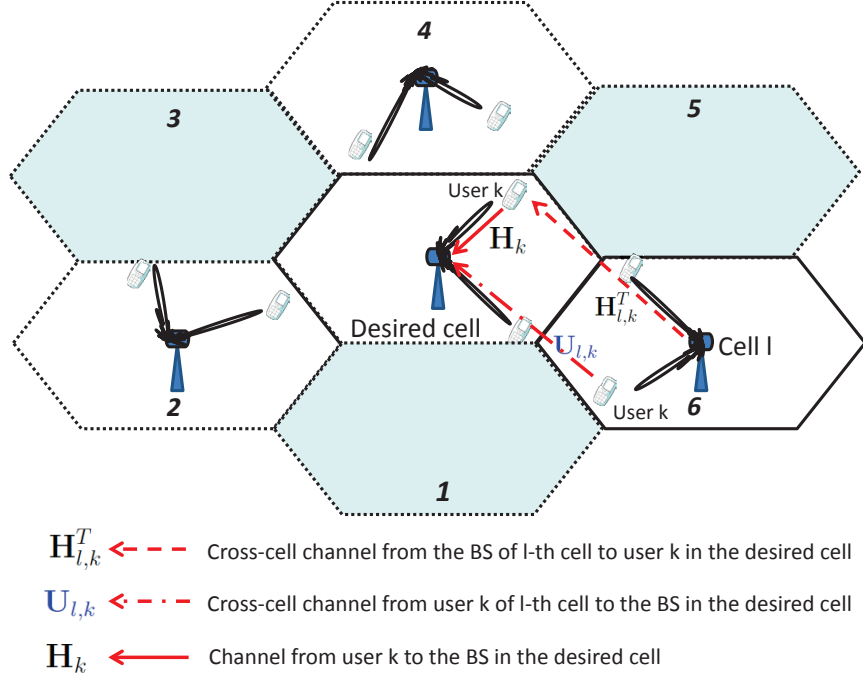
Figure 4.1: A multi-cell MU mmWave cellular system with $L = 6$ neighboring cells.

strongest AoA component $\mathbf{H}_{\mathrm{SAoA},k} \in \mathbb{C}^{M \times P}$ and N_{cl} scattering clusters/paths $\mathbf{H}_{\mathrm{S},k,i} \in \mathbb{C}^{M \times P}$, $i \in \{1, \cdots, N_{\mathrm{cl}}\}$ [146, 168, 173–175], which can be expressed as

$$\mathbf{H}_k = \sqrt{\frac{\varpi_k}{|\alpha_{k,0}|^2 + \frac{1}{N_{\mathrm{cl}}}\sum_{i=1}^{N_{\mathrm{cl}}}|\alpha_{k,i}|^2}}\left[\underbrace{\alpha_{k,0}\mathbf{H}_{\mathrm{SAoA},k}}_{\text{Strongest AoA}} + \underbrace{\sqrt{\frac{1}{N_{\mathrm{cl}}}}\sum_{i=1}^{N_{\mathrm{cl}}}\alpha_{k,i}\mathbf{H}_{\mathrm{S},k,i}}_{\text{Scattering components}}\right], \qquad (4.1)$$

where ϖ_k accounts for the corresponding large-scale path loss, $\alpha_{k,0}$ is the complex gain corresponding to the strongest AoA component and $\alpha_{k,i} \sim \mathcal{CN}(0,1)$, $i \in \{1, \cdots, N_{\mathrm{cl}}\}$ represents the complex gain corresponding to the i-th cluster. We can also assume that $\{\alpha_{k,i}\}$ are in non-increasing order, i.e., $|\alpha_{k,0}| \geq |\alpha_{k,1}| \geq \cdots \geq |\alpha_{k,N_{\mathrm{cl}}}|$. We note here, the strongest AoAs can be either from the LOS components or the NLOS components. Then, the strongest AoA component of

user k in the desired cell, $\mathbf{H}_{\text{SAoA},k}$, can be expressed as [170]

$$\mathbf{H}_{\text{SAoA},k} = \mathbf{h}_{\text{SAoA},k}^{\text{BS}} \mathbf{h}_{\text{SAoA},k}^{H}, \tag{4.2}$$

where $\mathbf{h}_{\text{SAoA},k}^{\text{BS}} \in \mathbb{C}^{M \times 1}$ and $\mathbf{h}_{\text{SAoA},k} \in \mathbb{C}^{P \times 1}$ are the antenna array response vectors of the BS and user k, respectively. In particular, $\mathbf{h}_{\text{SAoA},k}^{\text{BS}}$ and $\mathbf{h}_{\text{SAoA},k}$ can be expressed as [170]

$$\mathbf{h}_{\text{SAoA},k}^{\text{BS}} = \left[1, e^{-j2\pi \frac{d}{\lambda} \cos(\theta_k)}, \dots, e^{-j2\pi(M-1)\frac{d}{\lambda} \cos(\theta_k)} \right]^{T} \text{ and} \tag{4.3}$$

$$\mathbf{h}_{\text{SAoA},k} = \left[1, e^{-j2\pi \frac{d}{\lambda} \cos(\phi_k)}, \dots, e^{-j2\pi(P-1)\frac{d}{\lambda} \cos(\phi_k)} \right]^{T}, \tag{4.4}$$

respectively, where d is the distance between the neighboring antennas at the BS and users and λ is the wavelength of the carrier frequency. Variables $\theta_k \in [0, \pi]$ and $\phi_k \in [0, \pi]$ are the angle of incidence of the strongest path at antenna arrays of the desired BS and user k, respectively. As commonly adopted in the literature [170], we set $d = \frac{\lambda}{2}$ for convenience. Similarly, the scattering components of user k in the desired cell, $\mathbf{H}_{\text{S},k}$, can be expressed as

$$\mathbf{H}_{\text{S},k} = \sqrt{\frac{1}{N_{\text{cl}}}} \sum_{i=1}^{N_{\text{cl}}} \alpha_{k,i} \mathbf{H}_{\text{S},k,i} = \sqrt{\frac{1}{N_{\text{cl}}}} \sum_{i=1}^{N_{\text{cl}}} \alpha_{k,i} \mathbf{h}_{k,i}^{\text{BS}} \mathbf{h}_{k,i}^{H}, \tag{4.5}$$

where $\mathbf{h}_{k,i}^{\text{BS}} \in \mathbb{C}^{M \times 1}$ and $\mathbf{h}_{k,i} \in \mathbb{C}^{P \times 1}$ are the antenna array response vectors of the BS and user k associated to the i-th propagation path, respectively. By introducing the power ratio of the strongest cluster power over the other scattered paths of user k[1], $\varsigma_k = \frac{|\alpha_{k,0}|^2}{\frac{1}{N_{\text{cl}}} \sum_{i=1}^{N_{\text{cl}}} |\alpha_{k,i}|^2}$, we can rewrite Equation (4.1) as

---

[1]According to field test results summarized in Table I of [146] in the case of LOS and NLOS environments, the power ratio of the strongest cluster power over the other scattered paths' power, ς_k, is larger than 1.

$$\mathbf{H}_k = \sqrt{\varpi_k} \left[\underbrace{\sqrt{\frac{\varsigma_k}{\varsigma_k + 1}} \mathbf{h}_{\text{SAoA},k}^{\text{BS}} \mathbf{h}_{\text{SAoA},k}^H}_{\text{Strongest AoA}} + \underbrace{\sqrt{\frac{1}{\varsigma_k + 1}} \sqrt{\frac{1}{N_{\text{cl}}}} \sum_{i=1}^{N_{\text{cl}}} \alpha_{k,i} \mathbf{h}_{k,i}^{\text{BS}} \mathbf{h}_{k,i}^H}_{\text{Scattering components}} \right]. \quad (4.6)$$

With the increasing number of clusters, the path attenuation coefficients and the AoAs between the users and the BS behave randomly [102]. Let $\mathbf{U}_{l,k} \in \mathbb{C}^{M \times P}$ be the inter-cell mmWave uplink channel between user k in the l-th neighboring cell and the desired BS, $l \in \{1, \ldots, L\}$, cf. Figure 4.1. Let $\mathbf{H}_{l,k}^T \in \mathbb{C}^{P \times M}$ be the inter-cell downlink mmWave channel between the BS of the l-th neighboring cell and user k in the desired cell, cf. Figure 4.1. Since the inter-cell distance is shortened in small-cell systems, the inter-cell channels usually contain the strongest AoA components. Thus, the inter-cell uplink and downlink mmWave channels can be expressed as

$$\mathbf{U}_{l,k} = \sqrt{\overline{\varpi}_{l,k}} \underbrace{\sqrt{\frac{\varsigma_{l,k}}{\varsigma_{l,k} + 1}} \mathbf{U}_{\text{SAoA},l,k}}_{\text{Strongest AoA}} + \sqrt{\overline{\varpi}_{l,k}} \underbrace{\sqrt{\frac{1}{\varsigma_{l,k} + 1}} \mathbf{U}_{\text{S},l,k}}_{\text{Scattering component}} \text{ and} \quad (4.7)$$

$$\mathbf{H}_{l,k}^T = \sqrt{\widetilde{\varpi}_{l,k}} \underbrace{\sqrt{\frac{\widetilde{\varsigma}_{l,k}}{\widetilde{\varsigma}_{l,k} + 1}} \mathbf{H}_{\text{SAoA},l,k}^T}_{\text{Strongest AoA}} + \sqrt{\widetilde{\varpi}_{l,k}} \underbrace{\sqrt{\frac{1}{\widetilde{\varsigma}_{l,k} + 1}} \mathbf{H}_{\text{S},l,k}^T}_{\text{Scattering component}} \quad , \quad (4.8)$$

respectively, where $\overline{\varpi}_{l,k}$ and $\widetilde{\varpi}_{l,k}$ are the corresponding large-scale path loss coefficients. We note that $\varsigma_{l,k}$ and $\widetilde{\varsigma}_{l,k}$ are the power ratios of the strongest cluster power over the other scattered paths' power for the uplink and downlink inter-cell channels, respectively. The strongest AoA components $\mathbf{U}_{\text{SAoA},l,k} = \mathbf{u}_{\text{SAoA},l,k}^{\text{BS}} \mathbf{u}_{\text{SAoA},l,k}^H$, and $\mathbf{H}_{\text{SAoA},l,k}^T = \mathbf{h}_{\text{SAoA},l,k}^* \left(\mathbf{h}_{\text{SAoA},l,k}^{\text{BS}} \right)^T$ follow similar assumptions as in Equations $(4.2) - (4.4)$. Besides, the scattering components $\mathbf{U}_{\text{S},l,k}$ and $\mathbf{H}_{\text{S},l,k}^T$ follow similar assumption as in Equation (4.5). All these mentioned inter-cell propagation path loss coefficients are related to the propagation distance and modeled as

in [29,101,102]. According to recent field measurements, e.g. [29,101,102,155,156], the typical values of the power ratios of the strongest cluster power over the other scattered paths' power $\varsigma_{l,k}$ and $\widetilde{\varsigma}_{l,k}$ for $\mathbf{U}_{l,k}$ and $\mathbf{H}_{l,k}^{T}$ are in (0, 5], respectively.

4.3   Multi-cell Uplink Channel Estimation Performance Analysis

In this section, we adopt the algorithm proposed in [1] for the estimation of an equivalent mmWave channel which comprise the physical mmWave channels and analog beamforming matrices adopted at the desired transceivers. The proposed algorithm is suitable for both the conventional fully digital systems and the emerging hybrid systems with fully access and subarray implementation structures. For the sake of presentation, we provide a summary of the algorithm proposed in [1] in the following. For illustration, we adopt the subarray structure as an example, as shown in Figure 4.2. Specifically, each RF chain can access to the M antennas via a phase shifter network.

There are three steps in the algorithm proposed in [1]. In first and second step, strongest AoAs between the desired BS and the users are estimated. Then, the desired BS and the users design their analog beamforming matrices by aligning the beamforming direction to the estimated strongest AoAs. In the third step, orthogonal pilot symbols are transmitted from the users to the BS by using the pre-designed analog beamforming matrices. Then, by exploiting the channel reciprocity, the equivalent downlink channel can be estimated and adopted at the BS as the input for the digital baseband precoder.

Note that, orthogonal pilot symbols are used for the estimation of the equivalent channels in the third step. For the multi-cell scenario, pilot symbols are
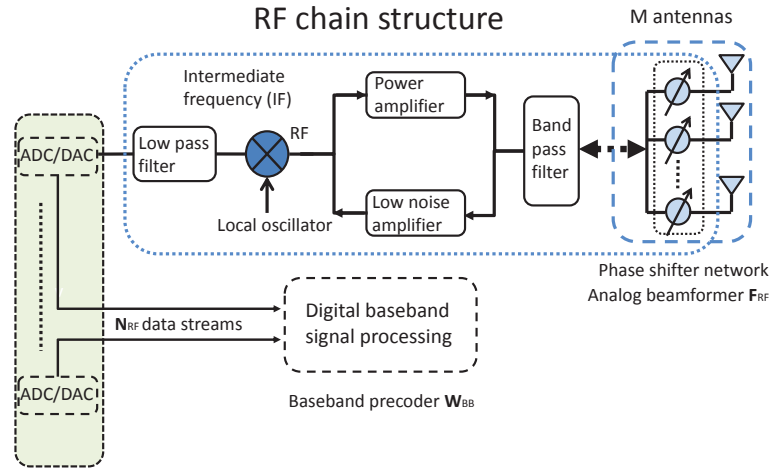
Figure 4.2: A block diagram of an RF chain of a subarray antenna structure.

reused among different cells which results in pilot contamination and cause a severe impact on the equivalent channel estimation performance. We note that only single-cell scenario was considered in [1] and it is unclear if the channel estimation algorithm provides robustness against pilot contamination. Furthermore, the performance analysis studied in [1] does not take into account the impact of potential out-of-cell interference on the performance of channel estimation. In the following sections, we investigate the impact of pilot contamination on the mmWave channel estimation performance for small-cell scenarios.

4.3.1 Channel Estimation and Pilot Contamination Analysis

Basically, step one and step two of the proposed algorithm in [1] provide the analog beamforming matrices at the desired BS and the desired users to facilitate the estimation of equivalent channel. In particular, the analog beamforming matrices pair the desired BS and the users and align the directions of data stream

transmission. Due to the inter-cell large-scale propagation path loss, the impact of multi-cell interference on the design of the analog beamforming matrices is usually negligible[2]. Besides, simulation results show that the pilot contamination does not affect strongest AoAs estimation result[3]. Therefore, to facilitate the performance analysis of the multi-cell equivalent channel estimation, we assume that strongest AoAs among the users and the BS are perfectly estimated and the desired signals always fall in the main lobe.

Based on the strongest AoAs, which are perfectly estimated at the users and the BS, the analog receive beamforming vector of user k adopted at the desired BS is given by

$$\widehat{\boldsymbol{\nu}}_k^T \in \mathbb{C}^{1 \times M} = \frac{1}{\sqrt{M}} \left[\ 1, \dots, e^{j2\pi(M-1)\frac{d}{\lambda}\cos(\theta_k)} \ \right] \qquad (4.9)$$

and the analog transmit beamformer of user i in the desired cell is given by

$$\widehat{\boldsymbol{\omega}}_i^* \in \mathbb{C}^{P \times 1} = \frac{1}{\sqrt{P}} \left[\ 1, \dots, e^{j2\pi(P-1)\frac{d}{\lambda}\cos(\phi_i)} \ \right]^H . \qquad (4.10)$$

In addition, we denote the analog beamforming matrix at the desired BS as

$$\mathbf{F}_{\mathrm{RF}} \in \mathbb{C}^{M \times N} = \left[\ \widehat{\boldsymbol{\nu}}_1, \dots, \widehat{\boldsymbol{\nu}}_N \ \right] . \qquad (4.11)$$

Let $\boldsymbol{\Phi}_k \in \mathbb{C}^{N \times 1}$ denote the pilot symbols of user k in the desired cell. The pilot symbols for all the N users in the desired cell form a matrix, $\boldsymbol{\Psi} \in \mathbb{C}^{N \times N}$, where $\boldsymbol{\Phi}_k$ is a column vector of matrix $\boldsymbol{\Psi}$ given by $\boldsymbol{\Psi} = \sqrt{E_{\mathrm{P}}} \left[\ \boldsymbol{\Phi}_1, \dots, \boldsymbol{\Phi}_N \ \right]$, $\boldsymbol{\Phi}_i^H \boldsymbol{\Phi}_j = 0$, $i \neq j$, $i, j \in \{1, \dots, N\}$, where E_{P} represents the transmitted pilot

---

[2]For strongest AoAs estimation at the desired BS in a multi-cell scenario, the received power of the reused pilot symbols transmitted from the users in neighboring cells is smaller than that of the desired pilot symbols transmitted from the users in the desired cell. In addition, strongest AoAs estimation may not rely on pilot symbols [2].

[3]The simulation result of the strongest AoAs estimation is omitted here due to the space limitation. However, the impact of pilot contamination on the strongest AoAs estimation will be captured in the final simulation.

symbol energy.

In the equivalent channel estimation, user k transmits the pilot symbols $\mathbf{\Phi}_k$ via transmit beamformer $\widehat{\boldsymbol{\omega}}_i^*$ and the desired BS receives the pilot symbols utilizing the analog beamforming matrix $\mathbf{F}_{\mathrm{RF}}$. However, the reuse of pilot symbols in neighboring cells affects the performance of equivalent channel estimation. The received signal of the k-th RF chain at the desired BS in the uplink is given by

$$\widehat{\mathbf{s}}_k^T = \boldsymbol{\nu}_k^T \sum_{i=1}^{N} \mathbf{H}_i \widehat{\boldsymbol{\omega}}_i^* \sqrt{E_{\mathrm{P}}} \mathbf{\Phi}_i^T + \boldsymbol{\nu}_k^T \underbrace{\sum_{l=1}^{L} \sum_{i=1}^{N} \left(\mathbf{U}_{l,i} \widehat{\boldsymbol{\omega}}_{l,i}^* \sqrt{E_{\mathrm{P}}} \mathbf{\Phi}_i^T \right)}_{\text{Pilot contamination}} + \boldsymbol{\nu}_k^T \mathbf{Z}, \qquad (4.12)$$

where $\widehat{\boldsymbol{\omega}}_{l,i} \in \mathbb{C}^{P \times 1} = \frac{1}{\sqrt{P}} \left[\; 1, \ldots, e^{j2\pi(P-1)\frac{d}{\lambda}\cos(\phi_{l,k})} \; \right]^T$ is the analog beamforming vector of user i in the l-th cell, the entries of noise matrix, $\mathbf{Z}$, are modeled by i.i.d. random variables with distribution $\mathcal{CN}\left(0, \sigma_{\mathrm{BS}}^2\right)$. To facilitate the investigation of channel estimation and downlink transmission, we assume that long-term power control is performed to compensate the different strongest AoAs' path loss among different desired users in the desired cell. As a result, it can be considered that large-scale propagation path losses of different users in the desired cell are identical. Thus, we can express the estimated equivalent downlink channel $\widehat{\mathbf{H}}_{\mathrm{eq}}^T \in \mathbb{C}^{N \times N}$ at the desired BS under the impact of pilot contamination as

$$\widehat{\mathbf{H}}_{\mathrm{eq}}^T = \mathbf{B} \left(\mathbf{H}_{\mathrm{eq}}^T + \Delta\widehat{\mathbf{H}}_{\mathrm{eq}}^T \right) = \mathbf{B} \left(\frac{\mathbf{\Psi}^H}{E_{\mathrm{P}}} \left[\; \widehat{\mathbf{s}}_1 \quad \ldots \quad \widehat{\mathbf{s}}_N \; \right] \right)$$

$$=\mathbf{B} \left(\underbrace{\begin{bmatrix} \widehat{\boldsymbol{\omega}}_1^H \mathbf{H}_1^T \mathbf{F}_{\mathrm{RF}} \\ \vdots \\ \widehat{\boldsymbol{\omega}}_N^H \mathbf{H}_N^T \mathbf{F}_{\mathrm{RF}} \end{bmatrix}}_{\mathbf{H}_{\mathrm{eq}}^T} + \frac{1}{\sqrt{E_{\mathrm{P}}}} \underbrace{\begin{bmatrix} \mathbf{\Phi}_1^H \mathbf{Z}^T \mathbf{F}_{\mathrm{RF}} \\ \vdots \\ \mathbf{\Phi}_N^H \mathbf{Z}^T \mathbf{F}_{\mathrm{RF}} \end{bmatrix}}_{\text{Effective noise}} + \underbrace{\begin{bmatrix} \sum_{l=1}^{L} \left(\widehat{\boldsymbol{\omega}}_{l,1}^H \mathbf{U}_{l,1}^T \right) \mathbf{F}_{\mathrm{RF}} \\ \vdots \\ \sum_{l=1}^{L} \left(\widehat{\boldsymbol{\omega}}_{l,N}^H \mathbf{U}_{l,N}^T \right) \mathbf{F}_{\mathrm{RF}} \end{bmatrix}}_{\text{Pilot contamination}} \right),$$

$$(4.13)$$

where $\Delta\widehat{\mathbf{H}}_{\mathrm{eq}}^T$ is the equivalent channel estimation error caused by pilot contamination and noise, and the path loss compensation matrix $\mathbf{B} \in \mathbb{C}^{N \times N}$ is given by

$$
\mathbf{B} = \begin{bmatrix} \frac{1}{\sqrt{\varpi_1}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\sqrt{\varpi_N}} \end{bmatrix}.
\tag{4.14}
$$

In the following, for notational simplicity, we denote $\widehat{\rho}_{l,k} = \sqrt{\frac{\widehat{\varpi}_{l,k}}{\varpi_k}}$ as the inter-cell propagation path loss coefficients. Now, to evaluate the impact of pilot contamination, we introduce a theorem which reveals the normalized MSE performance of equivalent channel estimation.

Theorem 4.1. *The normalized MSE of the equivalent channel estimation with respect to the k-th RF chain under the impacts of pilot contamination and noise can be approximated as*

$$
\begin{aligned}
\mathrm{NMSE}_{\mathrm{eq},k} &= \frac{1}{N}\mathbb{E}_{\mathbf{U}_{\mathrm{l,i}}}\left[\left(\frac{1}{\sqrt{MP}}\Delta\widehat{\mathbf{h}}_{\mathrm{eq},k}^T\right)\left(\frac{1}{\sqrt{MP}}\Delta\widehat{\mathbf{h}}_{\mathrm{eq},k}^*\right)\right] \\
&\approx \frac{1}{NMP}\left[\sum_{l=1}^{L}\frac{\widehat{\rho}_{l,k}^2\varsigma_{l,k}}{\varsigma_{l,k}+1}N\right] + \frac{1}{MP}\sum_{l=1}^{L}\frac{\widehat{\rho}_{l,k}^2}{\varsigma_{l,k}+1} + \frac{\sigma_{\mathrm{BS}}^2\mathrm{tr}\left[\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}}\right]}{\varpi_k E_{\mathrm{P}}NMP} \\
&= \underbrace{\frac{1}{MP}\sum_{l=1}^{L}\left(\widehat{\rho}_{l,k}^2\right)}_{\mathrm{Multi-cell\ pilot\ contamination}} + \underbrace{\frac{\sigma_{\mathrm{BS}}^2}{\varpi_k E_{\mathrm{P}}MP}}_{\mathrm{Noise}}.
\end{aligned}
\tag{4.15}
$$

In particular, when the number of antennas equipped at the desired BS and the users are sufficiently large, we have

$$
\lim_{M,P\to\infty}\mathrm{NMSE}_{\mathrm{eq},k} \approx 0 \quad .
\tag{4.16}
$$

Proof: Please refer to Appendix A.   ∎

In Equation (4.15), the impact of the multi-cell pilot contamination term

on the normalized MSE performance is inversely proportional to the number of antennas M and P. In addition, the noise term decreases with the increasing transmit pilot symbol energy and the number of antennas, M and P. It is important to note that, the impact of noise on channel estimation will vanish in the high SNR regime, e.g. $E_p \gg 1$. However, the impact of pilot contamination on the MSE performance cannot be mitigated by simply increasing the transmit pilot symbol energy E_p.

It is known that the conventional massive MIMO pilot-aided LS channel estimation performance under the impact of pilot contamination cannot be improved by increasing the number of antennas equipped at the BS [28, 53]. Interestingly, the result of Theorem 4.1 unveils that the impacts of pilot contamination and noise on the equivalent channel estimation will vanish asymptotically with the increasing number of antennas equipped at each RF chain, M and P. Actually, the numbers of antennas M and P have an identical effect on the normalized MSE performance.

This is because the direction of analog beamforming matrices adopted at the desired BS and the desired users align with the strongest AoA path. Hence, the analog beamforming matrices adopted at the desired BS and the users form a pair of spatial filters which block the pilot signals from undesired users to the desired BS via non-strongest paths. In addition, transmitting the pilot signals from the desired users via the analog beamforming matrix can reduce the potential energy leakage to other undesired cells, which further reduces the impact of pilot contamination.

In order to help the readers to understand Theorem 4.1 from the physical point of view, we provide a detail discussion in this subsection on the result of Equation (4.15) via an illustrative example of a single-antenna user in the high

SNR regime.

In the single-antenna user scenario, we focus on the normalized estimated equivalent channel of user k, $\widehat{\mathbf{h}}_{\text{eq},k}^T \in \mathbb{C}^{1 \times N}$, in the high SNR regime which is given by

$$\frac{1}{\sqrt{M}} \widehat{\mathbf{h}}_{\text{eq},k}^T = \frac{1}{\sqrt{M}} \left[\mathbf{h}_k^T \underbrace{\begin{bmatrix} \widehat{\boldsymbol{\nu}}_1 & \dots & \widehat{\boldsymbol{\nu}}_N \end{bmatrix}}_{\mathbf{F}_{\text{RF}}} + \sum_{l=1}^{L} \mathbf{u}_{l,k}^T \underbrace{\begin{bmatrix} \widehat{\boldsymbol{\nu}}_1 & \dots & \widehat{\boldsymbol{\nu}}_N \end{bmatrix}}_{\text{Pilot contamination}} \right]. \quad (4.17)$$

From Equation (4.17), it is shown that the equivalent channel $\widehat{\mathbf{h}}_{\text{eq},k}^T$ is formed by projecting the actual mmWave channel $\mathbf{h}_k$ on the analog beamforming matrix $\mathbf{F}_{\text{RF}}$. Then, $\eta_{k,i} \in \mathbb{C}^{1 \times 1}$, the i-th entry of the estimated equivalent channel vector $\widehat{\mathbf{h}}_{\text{eq},k}^T$, representing the projection of channel of user k in the desired cell on the i-th strongest AoA direction, is given by

$$\begin{aligned}
\eta_{k,i} =& \frac{1}{\sqrt{M}} \mathbf{h}_k^T \widehat{\boldsymbol{\nu}}_i + \frac{1}{\sqrt{M}} \sum_{l=1}^{L} \mathbf{u}_{l,k}^T \widehat{\boldsymbol{\nu}}_i \\
=& \underbrace{\sqrt{\frac{\varsigma_k}{\varsigma_k + 1}} \frac{\sin \left[M\pi \frac{d}{\lambda} \left(\cos(\theta_k) - \cos(\theta_i) \right) \right]}{M \sin \left[\pi \frac{d}{\lambda} \left(\cos(\theta_k) - \cos(\theta_i) \right) \right]}}_{\text{Strongest AoA}} + \underbrace{\sqrt{\frac{1}{\varsigma_k + 1}} \frac{1}{\sqrt{M}} \mathbf{h}_{\text{S},k}^T \widehat{\boldsymbol{\nu}}_i}_{\text{Scattering component}} \\
&+ \underbrace{\sum_{l=1}^{L} \sqrt{\frac{\varsigma_{l,k}}{\varsigma_{l,k} + 1}} \frac{\sin \left[M\pi \frac{d}{\lambda} \left(\cos(\theta_{l,k}) - \cos(\theta_i) \right) \right]}{M \sin \left[\pi \frac{d}{\lambda} \left(\cos(\theta_{l,k}) - \cos(\theta_i) \right) \right]} + \sum_{l=1}^{L} \sqrt{\frac{1}{\varsigma_{l,k} + 1}} \frac{1}{\sqrt{M}} \mathbf{u}_{\text{S},l,k}^T \widehat{\boldsymbol{\nu}}_i}_{\text{Pilot contamination}},
\end{aligned}$$

$$(4.18)$$

where $\theta_{l,k}$ is the AoA from user k in the l-th cell to the desired BS, $\mathbf{h}_{\text{S},k}^T$ is the scattering component of the channel of user k in the desired cell, and $\mathbf{u}_{\text{S},l,k}^T$ is the scattering component of the channel from user k in the l-th cell to the desired BS. For the projection of channel of user k on the k-th RF chain's analog beamforming vector $\widehat{\boldsymbol{\nu}}_k$, we have $\widehat{\boldsymbol{\nu}}_k^T = \left(\mathbf{h}_{\text{SAoA},k}^{\text{BS}} \right)^H$. Thus, the k-th entry of the vector $\widehat{\mathbf{h}}_{\text{eq},k}^T$ can be expressed as

$$\eta_{k,k} = \left[\underbrace{\sqrt{\frac{\varsigma_k}{\varsigma_k + 1}}}_{\text{Strongest AoA}} + \underbrace{\varepsilon_k \frac{1}{\sqrt{M}} \sqrt{\frac{1}{\varsigma_k + 1}}}_{\text{Scattering component}} + \underbrace{\frac{1}{\sqrt{M}} \mu_k}_{\text{Pilot contamination}} \right], \tag{4.19}$$

where $\varepsilon_k = \mathbf{h}_{\mathrm{S},k}^T \widehat{\boldsymbol{\nu}}_k$ denotes the projection of the scattering component $\sqrt{\frac{1}{\varsigma_k + 1}} \mathbf{h}_{\mathrm{S},k}$ on $\widehat{\boldsymbol{\nu}}_k$ and $\mu_k = \sum_{l=1}^{L} \mathbf{u}_{l,k}^T \widehat{\boldsymbol{\nu}}_k$ denotes the projection of inter-cell channels $\sum_{l=1}^{L} \mathbf{u}_{l,k}^T$ on $\widehat{\boldsymbol{\nu}}_k$.

The computation of Equation (4.19) can be illustrated graphically via the concept of vector space with projection as shown in Figure 4.3(a). The first term of Equation (4.19), $\sqrt{\frac{\varsigma_k}{\varsigma_k + 1}}$, which is the projection of the strongest AoA component on the analog beamforming direction, is a constant and is independent of M. For the second term and the third term of Equation (4.19), ε_k and μ_k, which consist of multiple paths with various AoAs, cannot enjoy the array gain M. Actually, $\frac{\varepsilon_k}{\sqrt{M}} \sqrt{\frac{1}{\varsigma_k + 1}}$ and $\frac{\mu_k}{\sqrt{M}}$, these two terms will vanish asymptotically with the increasing number of antennas M.

We note that for the conventional pilot-aided channel estimation algorithms, e.g. LS-based algorithms, they estimate the channels from all the directions. Thus, BSs adopting these algorithms receiving reused pilot symbols from the undesired users and cannot be distinguished from the desired pilot symbols. It is known as pilot contamination. However, adopting analog beamforming matrices for receiving pilot symbols at the desired BS via the strongest AoA directions forms a spatial filter, which blocks the undesired pilot symbols from neighboring cells via different AoA paths. Furthermore, the "blocking capability" improves with the increasing number of antennas equipped at the BS. Specifically, the beamwidth of main lobe becomes narrower and the magnitude of sidelobes is lower, which is illustrated in Figure 4.3(b). In fact, this is an important feature for mitigating the impact of pilot contamination.

(a) Channel estimation algorithm explanation via a vector space model.

(b) Antenna array beam pattern response.

Figure 4.3: (a) The illustration of equivalent channel estimation under the impact of pilot contamination. (b) The illustration of sidelobe suppression for different numbers of antennas M.

Therefore, with the equivalent channel estimation proposed in [1], the impact of pilot contamination vanishes asymptotically with an increasing number of antennas M equipped at the desired BS.

4.4 Multi-cell Downlink Achievable Rate

In the last section, we mathematically prove that the impact of pilot contamination on the equivalent channel estimation can be effectively mitigated. Analytical and simulation results show that CSI errors caused by pilot contamination vanish asymptotically with the increasing numbers of antennas equipped at each RF chain of the BS and the users, M and P.

Now, we aim to study the downlink performance of multi-cell hybrid mmWave networks, which takes into account the impact of inter-cell interference caused by the neighboring BSs as well as the intra-cell interference caused by channel estimation errors. In this section, we illustrate and derive the closed-form approximations of achievable rate per user under the impact of intra-cell and the inter-cell interference of the considered hybrid system. Our results reveal that the intra-cell interference as well as the inter-cell interference can be mitigated effectively by simply increasing the number of antennas equipped at the desired BS.

4.4.1 Downlink Achievable Rate Performance

We now detail the received information signal at user k in the desired cell. Due to the channel reciprocity, we utilize the estimated equivalent channel $\widehat{\mathbf{H}}_{\mathrm{eq}}^{T}$ in Equation (4.13) for the digital baseband precoder design during the downlink transmission. To suppress the inter-user interference among different users, we adopt a ZF precoder for the downlink transmission. In addition, equal power

allocation is used among different data streams of the users [3, 46, 115]. The
desired baseband digital ZF precoder $\overline{\mathbf{W}}_{\text{eq}} \in \mathbb{C}^{N \times N}$ is given by

$$\overline{\mathbf{W}}_{\text{eq}} = \mathbf{H}_{\text{eq}}^*(\mathbf{H}_{\text{eq}}^T \mathbf{H}_{\text{eq}}^*)^{-1} = \left[\begin{array}{c} \overline{\mathbf{w}}_{\text{eq},1}, \dots, \overline{\mathbf{w}}_{\text{eq},N} \end{array} \right]. \qquad (4.20)$$

The received signal at user k in the desired cell under the intra-cell and inter-cell
interference is given by

$$y_{\text{ZF}}^k = \underbrace{\widehat{\boldsymbol{\omega}}_k^H \mathbf{H}_k^T \mathbf{F}_{\text{RF}} \overline{\beta} \overline{\mathbf{w}}_{\text{eq},k} x_k}_{\text{Desired signal}} + \underbrace{\widehat{\boldsymbol{\omega}}_k^H \mathbf{H}_k^T \mathbf{F}_{\text{RF}} \sum_{j=1}^{N} \overline{\beta} \Delta \overline{\mathbf{w}}_{\text{eq},j} x_j}_{\text{Intra}-\text{cell interference}}$$

$$+ \underbrace{\widehat{\boldsymbol{\omega}}_k^H \sum_{l=1}^{L} \mathbf{H}_{l,k}^T \mathbf{F}_{\text{RF},l} \widehat{\beta}_l \widehat{\mathbf{W}}_{\text{eq},l} \mathbf{x}_l}_{\text{Inter}-\text{cell interference}} + \underbrace{\widehat{\boldsymbol{\omega}}_k^H \mathbf{z}_{\text{MS},k}}_{\text{Noise}}, \qquad (4.21)$$

where $x_k \in \mathbb{C}^{1 \times 1}$ is the transmitted symbol from the desired BS to the desired
user k, $\mathbb{E}[|x_k^2|] = E_s$ is the average transmitted symbol energy for each user,
$\overline{\beta} = \sqrt{\frac{1}{\text{tr}(\overline{\mathbf{W}}_{\text{eq}} \overline{\mathbf{W}}_{\text{eq}}^H)}}$ is the transmission power normalization factor of the desired
BS, $\Delta \overline{\mathbf{w}}_{\text{eq},j} \in \mathbb{C}^{N \times 1}$ denotes the j-th column vector of the ZF precoder error
matrix $\Delta \overline{\mathbf{W}}_{\text{eq}} = \widehat{\mathbf{W}}_{\text{eq}} - \overline{\mathbf{W}}_{\text{eq}} = \left[\begin{array}{c} \Delta \overline{\mathbf{w}}_{\text{eq},1}, \dots, \Delta \overline{\mathbf{w}}_{\text{eq},N} \end{array} \right]$, $\widehat{\mathbf{W}}_{\text{eq}} = \widehat{\mathbf{H}}_{\text{eq}}^*(\widehat{\mathbf{H}}_{\text{eq}}^T \widehat{\mathbf{H}}_{\text{eq}}^*)^{-1}$
is the designed ZF precoder based on the estimated equivalent channel, and the
effective noise $\mathbf{z}_{\text{MS},k} \sim \mathcal{CN}(\mathbf{0}, \sigma_{\text{MS}}^2 \mathbf{I})$. In addition, $\widehat{\mathbf{W}}_{\text{eq},l} = \widehat{\mathbf{H}}_{\text{eq},l}^*(\widehat{\mathbf{H}}_{\text{eq},l}^T \widehat{\mathbf{H}}_{\text{eq},l}^*)^{-1}$
is the digital precoder of the BS in cell l based on the estimated equivalent
channel $\widehat{\mathbf{H}}_{\text{eq},l}$, $\mathbf{F}_{\text{RF},l}$ is the analog beamforming matrix of the BS in cell l, $\mathbf{x}_l = [x_{l,1}, x_{l,2}, \dots, x_{l,N}]^T$ denotes the transmitted signal for all the users in cell l,
and $\widehat{\beta}_l = \sqrt{\frac{1}{\text{tr}(\widehat{\mathbf{W}}_{\text{eq},l} \widehat{\mathbf{W}}_{\text{eq},l}^H)}}$ is the power normalization factor of the BS in cell l.
Compared to the system model in [1], the intra-cell interference, which is caused
by CSI errors due to pilot contamination effects, is captured and investigated.
In addition, the model considered here also captures the inter-cell interference

caused by the simultaneous transmission from the neighboring BSs. We then express the average receive SINR expression of user k in the desired cell as

$$\mathbb{E}_{\Delta\widehat{\mathbf{H}}_{\text{eq}}^T, \mathbf{H}_{l,k}}\left[\widehat{\text{SINR}}_{\text{ZF}}^k\right] \approx$$

$$\frac{\overline{\beta}^2\varpi_k E_s}{\underbrace{\overline{\beta}^2\varpi_k E_s\widehat{\mathbf{h}}_{\text{eq},k}^T\mathbb{E}_{\Delta\widehat{\mathbf{H}}_{\text{eq}}^T}\left[\Delta\overline{\mathbf{W}}_{\text{eq}}\Delta\overline{\mathbf{W}}_{\text{eq}}^H\right]\widehat{\mathbf{h}}_{\text{eq},k}^*}_{\text{Intra}-\text{cell interference } \Upsilon_k} + \underbrace{\mathbb{E}_{\mathbf{H}_{l,k}}\left[\left|\widehat{\boldsymbol{\omega}}_k^H\sum_{l=1}^L\mathbf{H}_{l,k}^T\mathbf{F}_{\text{RF},l}\widehat{\beta_l}\widehat{\mathbf{W}}_{\text{eq},l}\mathbf{x}_l\right|^2\right]}_{\text{Inter}-\text{cell interference } \Omega_k} + \sigma_{\text{MS}}^2}.$$

$$(4.22)$$

4.4.2 Intra-cell and Inter-cell Interference Performance Analysis

In this section, we aim at deriving a closed-form expression of the average achievable rate per user. To this end, we first focus on the intra-cell interference term Υ_k shown in Equation (4.22). Note that the intra-cell interference Υ_k is caused by CSI errors, which is due to the impact of pilot contamination on channel estimation. Then we detail and analyze the inter-cell interference Ω_k. The expression of intra-cell interference, Υ_k, can be summarized in the following lemma.

Lemma 4.1. *In the large number of antennas regimes, the asymptotic intra-cell interference under the impact of pilot contamination can be approximated by*

$$\Upsilon_k = \overline{\beta}^2\varpi_k E_s\widehat{\mathbf{h}}_{\text{eq},k}^T\mathbb{E}_{\Delta\widehat{\mathbf{H}}_{\text{eq}}^T}\left[\Delta\overline{\mathbf{W}}_{\text{eq}}\Delta\overline{\mathbf{W}}_{\text{eq}}^H\right]\widehat{\mathbf{h}}_{\text{eq},k}^*$$

$$\overset{M\to\infty}{\approx}\overline{\beta}^2\varpi_k E_s\left[\left(\sqrt{1+\sum_{l=1}^L\frac{\widehat{\rho}_{l,k}^2}{MP}}-1\right)^2 + \left(1+\sum_{l=1}^L\frac{\widehat{\rho}_{l,k}^2}{MP}\right)\sum_{l=1}^L\frac{\widehat{\rho}_{l,k}^2 N}{MP}\frac{\varsigma_k+1}{\varsigma_k}\right].$$

$$(4.23)$$

Proof: The proof follows a similar approach as in [1]. Due to the page

limitation, we omit the proof here. ∎

Now, we introduce the following lemma which simplifies the expression of the intra-cell interference.

Lemma 4.2. *In the large numbers of antennas regime, the asymptotic inter-cell interference, which is caused by the BSs in neighboring cells, can be approximated as*

$$\Omega_k = \mathbb{E}_{\mathbf{H}_{l,k}} \left[\left| \widehat{\boldsymbol{\omega}}_k^H \sum_{l=1}^{L} \mathbf{H}_{l,k}^T \mathbf{F}_{\mathrm{RF},l} \widehat{\beta_l} \widehat{\mathbf{W}}_{\mathrm{eq},l} \mathbf{x}_l \right|^2 \right] \overset{M \to \infty}{\approx} E_s \left[\sum_{l=1}^{L} (\widetilde{\varpi}_{l,k}) \right]. \quad (4.24)$$

Proof: Please refer to Appendix B. ∎

From Equation (4.24), the inter-cell interference is mainly contributed by the downlink transmission of neighboring BSs. Besides, the power normalization factor $\overline{\beta}$ can be approximated as [1]

$$\overline{\beta} \approx \sqrt{\frac{\varsigma_k}{\varsigma_k + 1} \frac{MP}{N}}. \quad (4.25)$$

Then, we summarize the average achievable rate of user k in the large number of antennas and high SNR regime in the following theorem.

Theorem 4.2. *For a large number of antennas regime, the average achievable rate of user k adopting the ZF precoding under the intra-cell and the inter-cell interference is approximated by*

$$\widetilde{R}_{\mathrm{ZF}}^k \approx \log_2 \left\{ 1 + \left[\left(\sqrt{1 + \frac{1}{MP} \sum_{l=1}^{L} (\widehat{\rho}_{l,k}^2)} - 1 \right)^2 + \frac{N(\varsigma_k + 1)}{\varsigma_k MP} \sum_{l=1}^{L} (\zeta_{l,k}^2) \right. \right.$$

$$\left. \left. + \left[1 + \frac{1}{MP} \sum_{l=1}^{L} (\widehat{\rho}_{l,k}^2) \right] \frac{N(\varsigma_k + 1)}{\varsigma_k MP} \sum_{l=1}^{L} (\widehat{\rho}_{l,k}^2) + \frac{\sigma_{\mathrm{MS}}^2}{\overline{\beta}^2 \varpi_k E_s} \right]^{-1} \right\}, \quad (4.26)$$

where $\zeta_{l,k} = \sqrt{\dfrac{\widetilde{\varpi}_{l,k}}{\varpi_k}}$.

Proof: The result follows by substituting Equations (4.23), (4.24), and (4.25) into (4.22). ∎

In the high SNR and large antenna regimes, the average achievable rate under the impact of intra-cell interference caused by CSI errors and the inter-cell interference caused by neighboring BSs can be mitigated by simply increasing the number of antennas M. Thus, the average achievable rate per user can increase unboundedly with an increasing antenna $\log_2 M$.

For comparison, we also derive the achievable rate of user k in a single-cell hybrid system. The performance of this system, $R_{\mathrm{HB}}^{\mathrm{upper}_k}$, serves as a performance upper bound of the considered multi-cell system as there is no out-of-cell interference. The expression of $R_{\mathrm{HB}}^{\mathrm{upper}_k}$ is given by (Equation (24) of Corollary 1 in [1])

$$R_{\mathrm{HB}}^{\mathrm{upper}_k} \underset{M\to\infty}{\overset{a.s.}{\longrightarrow}} \log_2 \left\{ 1 + \left[\frac{MP}{N} \frac{\varsigma_k}{\varsigma_k + 1} + \frac{1}{\varsigma_k + 1} \right] \frac{E_s}{\sigma_{\mathrm{MS}}^2} \right\}. \tag{4.27}$$

Corollary 4.1. *Now, we derive the scaling law of the average achievable rate of user k in the high SNR regime which is given by*

$$\lim_{M\to\infty} \frac{\widetilde{R}_{\mathrm{ZF}}^k}{\log_2 M} = \lim_{\mathrm{M}\to\infty} \frac{R_{\mathrm{HB}}^{\mathrm{upper}_k}}{\log_2 M} = 1. \tag{4.28}$$

Proof: The result comes after some straightforward mathematical manipulation. Due to the page limitation, we omit the proof here. ∎

It is found that the system data rate scales with $\log_2 M$. In other words, pilot contamination is not a fundamental problem for massive MIMO with hybrid mmWave systems, when the CSI estimation algorithm proposed in [1] is adopted. In contrast, the average achievable rate performance of multi-cell MU massive MIMO networks adopting a conventional LS-based channel estimation

algorithm is limited by the effects of pilot contamination, e.g. [21]. Note that the algorithm proposed in [1] does not require any information of covariance matrices of pilot-sharing users in neighbouring cells as required by the multi-cell MMSE-based precoding algorithm proposed in [144]. In addition, the algorithm proposed in [1] can work well while mean values of cross-cell channels from pilot-sharing users in neighbouring cells to the desired BS are non-zero.

Remark 4.1. *Hybrid mmWave systems are the special case of fully digital systems, the algorithm proposed in [1] can be extended to the case of fully digital systems. Under such circumstance, the derived analysis can be applied to the latter systems.*

4.5   Discussions and Simulations

In this section, we will discuss and verify the derived results via simulations. In the following simulations unless specified otherwise, simulation settings are listed as follows. The transmit antenna gain of the BS antennas is assumed as 14 dBi. The maximum BS transmit power is set as 46 dBm. We set the number of neighboring cells as $L = 6$ and the number of users per cell $N = 10$. For the adopted simulation parameters, the large-scale path loss coefficients[4], α and ϱ, may have different values for different scenarios. We have $\alpha = 1.9$ and $\varrho = 20$ for corresponding large-scale path loss calculation [101]. In addition, the receiving thermal noise power is $\kappa = -90$ dBm[5]. The channel estimation errors of a user

---

[4]The commonly adopted empirical propagation path loss model is considered as a function of distance and carrier frequency, which is given by $\varpi_k = 10^{-\left(\frac{10\alpha \log_{10} d_k + \varrho \log_{10}\left(4\pi \frac{1}{\lambda}\right)}{10}\right)}$, where d_k is the distance between the desired BS and user k in the desired cell, α and ϱ are the least square fits of floating intercept and slope over the measured distance [101].

[5]Thermal noise power is determined by the signal bandwidth B_W and the noise spectral density level $N_0 = B_W e_B T_k$, where $B_W = 250$ MHz, $T_k = 300$ K, and $e_B = 1.38 \times 10^{-23}$ J/K is the Boltzmann constant [101].

Figure 4.4: The illustration of multi-cell normalized MSE performance under the impact of pilot contamination versus the number of antennas M in the high SNR regime, i.e., maximum transmit power 46 [dBm], for $N_{\mathrm{RF}} = N = 10$, $\varsigma_{l,k} = 5$, $k \in \{1, \cdots, N\}$ and channel estimation error $\xi^2 = 0.01$.

caused by pilot contamination is given by $\xi^2 = \sum_{l=1}^{L} \left(\widehat{\rho}_{l,k}^2 \right)$, $k \{1, \ldots, N\}$. Besides, we define the average achievable rate per user as $\widetilde{R}_{\mathrm{AVE}} = \frac{1}{N} \sum_{k=1}^{N} \widetilde{R}_{\mathrm{ZF}}^k$.

To verify the correctness of the analytical results derived in Equation (4.15), here, we provide some simulation results in Figure 4.4, which illustrates the normalized MSE performance of the equivalent channel estimation versus the number of antennas equipped at the desired BS, M, under the impact of pilot contamination. In the simulation, we take into account the inter-cell propagation path loss and AoA estimation errors in estimating the strongest AoA paths. For illustration, channel estimation errors caused by pilot contamination is set as $\xi^2 = 0.01$. Due to the existence of the inter-cell strongest AoA components, this setting can be considered as the worst case scenario of pilot contamination as the impact of inter-cell interference is magnified. In Figure 4.4, we can observe that with an
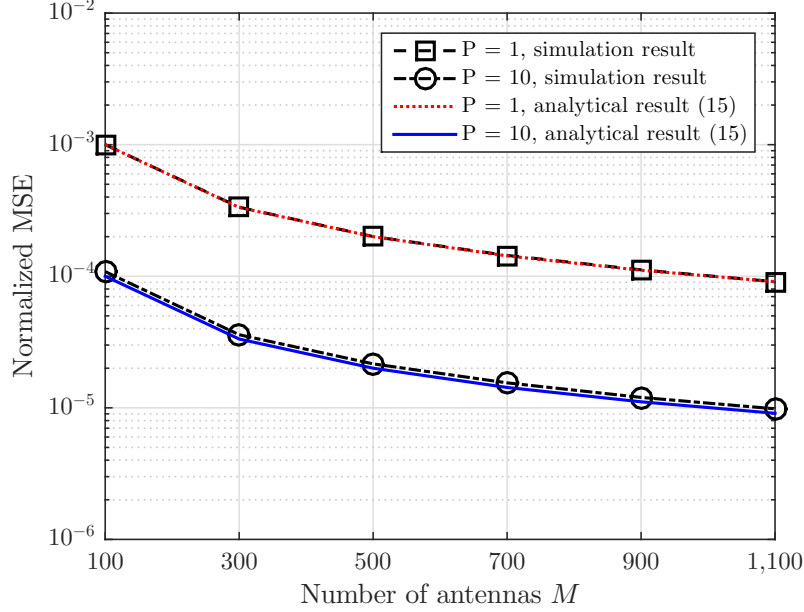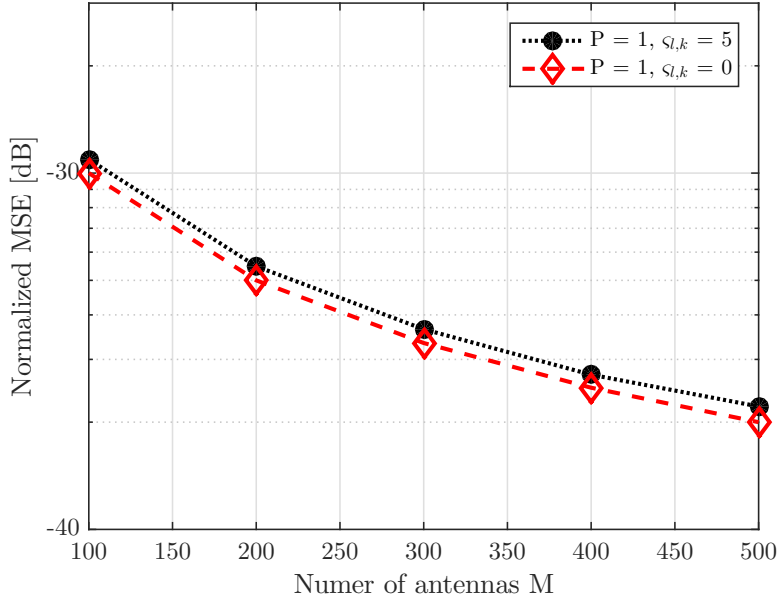
Figure 4.5: The illustration of multi-cell normalized MSE performance under the impact of pilot contamination versus the number of antennas M in the high SNR regime with different $\varsigma_{l,k}$ settings.

increasing number of antennas M, the normalized MSE decreases monotonically. In addition, an increasing number of antennas equipped at the users P can also improve the normalized MSE. Besides, the simulation results match the analytical results derived in Equation (4.15). Thus, the impact of pilot contamination on the channel estimation of multi-cell hybrid mmWave systems vanishes asymptotically, for a sufficiently large number of antennas equipped at the desired BS M.

In Chapter 4, to simplify the performance analysis, we assume the same average large scale fading coefficient for LOS components and scattering components. Thus, the change of cross-cell LOS to NLOS power ratio $\varsigma_{l,k}$ does not change the MSE performance of the channel estimation, which is demonstrated in Equation (4.15) of Theorem 4.1. However, in practice, large scale propagation path loss coefficients for cross-cell LOS and NLOS components are supposed to be different. In Figure 4.5, we set a simulation with different large scale propagation path loss
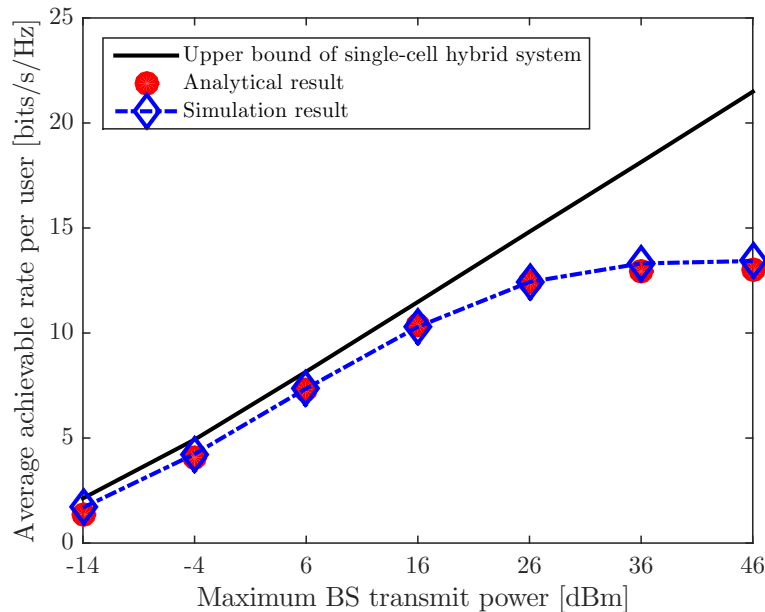
Figure 4.6: The average achievable rate per user [bits/s/Hz] under the impact of pilot contamination and inter-cell downlink transmission interference versus maximum transmit power at the desired BS [dBm] with $M = 200$.

coefficients for cross-cell LOS and NLOS components (the large scale path loss coefficient of NLOS components is larger than that of LOS components). The simulation results reveal that with increasing the cross-cell LOS to NLOS power ratio $\varsigma_{l,k}$, the normalized MSE performance decreases.

In Figure 4.6, we verify the tightness of the derived approximation in Equation (4.26) of Theorem 4.2. Figure 4.6 illustrates the average achievable rate per user under the intra-cell and inter-cell interference versus SNR. Besides, we illustrate the corresponding upper performance which is based on a single-cell hybrid mmWave system. In Figure 4.6, the number of antennas equipped at the users is set to $P = 10$, $\varsigma_k = 4$, $\varsigma_{l,k} = 2$, $k \in \{1, \ldots, N\}$, the channel estimation error caused by pilot contamination is set as $\xi^2 = 0.01$. We can observe that the derived approximation is tight for a wide range of SNR. It also can be observed that, with a finite number of antennas, there is still a ceiling of the achievable

rate performance in the high SNR regime.

The setup in Figure 4.7 is considered at $\varsigma_k = 4$, $\varsigma_{l,k} = 2$, $k \in \{1, \ldots, N\}$, and the number of antennas equipped at each user is $P = 10$. In Figure 4.7, we first illustrate the upper bounds of the average achievable rate per user of the single-cell hybrid mmWave system and the single-cell fully digital system as the system benchmarks. Without any intra-cell interference and inter-cell interference, the corresponding upper performance of the fully digital system as well as the hybrid mmWave system increases linearly with the increasing number of BS antennas M. Then, we compare the achievable rate per user of the fully digital system using a conventional LS-based CSI estimation algorithm to that of the hybrid system adopting CSI estimation algorithm proposed in [1] in the high SNR regime. For the fully digital system, we assume that the receive CSI at the users' side is perfectly known. Thus, the P antenna array equipped at each user can provide $10 \log_{10} P$ dB array gain.

In addition, we assume that channel estimation errors of a user at the desired BS due to the impact of pilot contamination effects for both the fully digital system and the hybrid system are set as $\xi^2 = 0.2$. As expected, in the large number of antennas and high SNR regimes, the average achievable rate of the hybrid system increases unboundedly and logarithmically with the increasing number of antennas M, despite the existence of pilot contamination as well as inter-cell interference. The average achievable rate of the hybrid system scales with the number of BS antennas in the order of $\log_2 M$, which is the same as that of the upper performance of the single-cell hybrid system, and verifies the correctness of the derived scaling law (Equation (4.28) in Corollary 4.1).
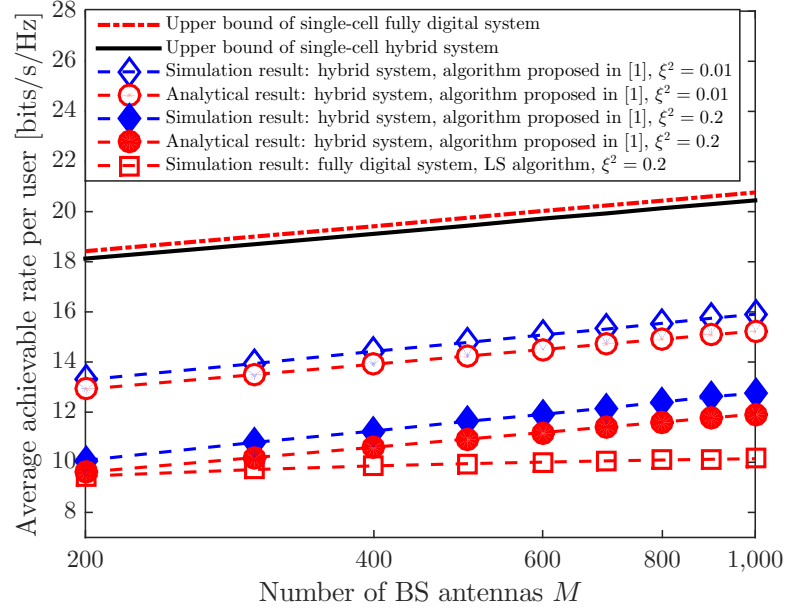
(a)
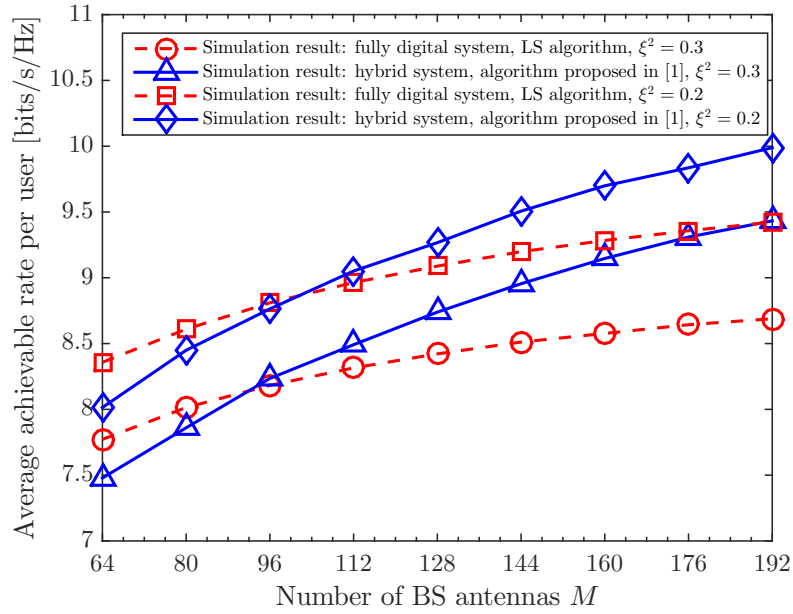
(b)

Figure 4.7: The achievable rate per user [bits/s/Hz] under the impact of pilot contamination effects and inter-cell interference versus the number of antennas M in the high SNR regime (maximum BS transmit power is 46 [dBm]).

However, the average achievable rate of the fully digital system adopting the

conventional LS-based CSI estimation algorithm under the impact of pilot con-
tamination is saturated with the increasing number of antennas at the desired
BS. Thus, pilot contamination is the performance bottleneck of massive MIMO
systems when the conventional LS-based channel estimation algorithm is adopt-
ed [21]. In contrast, when the channel estimation algorithm proposed in [1] is
applied to the multi-cell MU mmWave hybrid system, the impacts of CSI errors
caused by pilot contamination and the inter-cell interference caused by neigh-
boring BSs on the achievable rate performance will vanish asymptotically for a
sufficiently large number of antennas equipped at the desired BS. For a practical
number of antennas setup, $M \in [64, \ 192]$, the average rate performance achieved
by the algorithm proposed in [1] can also increase with an increasing number of
antennas M and outperforms that of a fully digital system, which is illustrated in
Figure 4.7(b). Meanwhile, the number of RF chains required in the hybrid system
is significantly lower than that of fully digital system. In addition, our simulation
results in Figure 4.7(a) also verify the tightness of the analytical approximation
derived in Theorem 4.2.

4.6   Summary

In this chapter, we investigated the normalized MSE performance of the channel
estimation proposed in [1] for multi-cell hybrid mmWave systems. The derived
closed-form approximation of the normalized channel estimation MSE perfor-
mance revealed that the channel estimation error caused by the impact of pilot
contamination and noise would vanish asymptotically with the increasing number
of antennas. Furthermore, based on the estimated CSI, we adopted ZF precoding
for the downlink transmission and derived the closed-form approximation of the
average achievable rate per user in the multi-cell scenario. The analytical and

simulation results showed that the intra-cell interference caused by pilot contamination as well as the inter-cell interference incurred by neighboring BSs can be mitigated effectively with an increasing antenna array gain. It is an excellent feature for multi-cell hybrid mmWave systems with small-cell radius for improving the network spectral efficiency.

Chapter 5

Conclusions and Future Works

We conclude this thesis by summarizing our contributions and discussing some potential future works.

5.1 Conclusions

This thesis mainly focused on channel estimation, analog beamformer and digital precoder design, and performance analysis that address key challenges in mmWave systems, by considering, the hardware constraints, the hardware impairments, non-sparse mmWave channels, the channel acquisition overhead, the analog beamformer/digital precoder design complexity, and the cross-cell interference caused by users and BSs in neighbouring cells.

First, we proposed a tone-based AoA estimation channel estimation algorithm for mmWave channels to reduce the channel acquisition overhead significantly. Then, we proposed efficient analog beamformer/digital precoder technologies for MU hybrid mmWave MIMO systems. Finally, we applied this solution to multi-small-cell scenarios. We proved that this channel estimation and analog beamformer/digital precoder design scheme could effectively mitigate cross-cell

133

interference during uplink and downlink phases.

In Chapter 2, for the channel estimation problem of the massive MIMO system in mmWave frequencies over Rician fading channels, we developed a tone-based AoA estimation algorithm leveraging the angular resolution. Based on estimated CSIs, we proposed an MU SLOS-based MRT strategy to achieve higher rate performance than that of the conventional PAC-based MRT strategy. The rate performance gap between the SLOS-based MRT and the PAC-based MRT was analytically evaluated and verified via simulation. Besides, the impact of phase quantization errors on the system rate performance had also been analyzed.

In Chapter 3, we proposed a low complexity channel estimation algorithm for MU TDD hybrid mmWave MIMO systems during the uplink phase, which was applicable to both non-sparse and sparse mmWave channel environments. The proposed channel estimation solution leverages the characteristics of the mmWave channels (the strongest AoAs), the orthogonal pilot symbols, and the hybrid structures. Exploiting the estimated strongest AoAs, analog beamforming matrices could be designed. Then, leveraging the designed analog beamforming matrices and orthogonal pilot symbols, the BS could estimate the equivalent mmWave channels for the design of digital precoding matrices. The performance of the proposed channel estimation algorithm was analyzed. Analytical and simulation results indicated that the downlink achievable rate based on the proposed estimated CSI could obtain a considerable achievable rate of fully digital systems. Furthermore, the robustness of the proposed channel estimation and MU downlink precoding scheme against hardware impairments was illustrated via analytical and simulation results.

In Chapter 4, we investigated the system rate performance of a multi-small-cell MU hybrid mmWave MIMO network. We applied the channel estimation algo-

rithm and the analog beamforming/digital precoding design scheme proposed in Chapter 3 to multi-small-cell scenarios. Different than the conventional massive MIMO system with the conventional LS channel estimation algorithm, our analytical and simulation results showed that the channel estimation error incurred by the impact of pilot contamination and noise vanished asymptotically with an increasing number of antennas equipped at the desired BS. The results unveiled that even facing severe intra-cell interference and inter-cell interference, the average achievable rate performance does not saturate with increasing the number of antennas equipped at the desired BS. It was an excellent feature for multi-cell hybrid mmWave systems with small-cell radius for improving the network spectral efficiency. The obtained analytical and simulation results could be generalized and extended for sub-6 GHz massive MIMO systems in Rician fading channels. In other words, pilot contamination was no longer the fundamental problem for massive MIMO systems in Rician fading channels.

5.2 Future Works

There are several possible directions for future research.

Hybrid architectures with extra RF chains:

The hybrid architecture adopted in Chapter 3 considered that the number of RF chains is equal to the number of users. This assumption is adopted to simplify the notation and the derivation. In general, for the case of $N_{\mathrm{RF}} = N$, the hybrid system can provide enough degrees of freedom to serve N users simultaneously. However, in the work [20], the authors proved that when the number of RF chains equipped at the BS is more than twice of the number of users, the hybrid systems can realize the rate performance of the fully digital systems. If the BS

can equip with more RF chains, more degrees of freedom can be exploited for the design of analog and digital precoders. In addition, the extra RF chains adopted by the BS can be used for user scheduling design, anti-blocking design, and multipath diversity exploitation. Thus, for the case that $N_{\mathrm{RF}} > N$, hybrid systems can achieve a higher rate than that of $N_{\mathrm{RF}} = N$. However, how to exploit the extra RF chains to further improve the rate performance and the channel estimation performance of hybrid systems is not considered in this thesis. The case of $N_{\mathrm{RF}} > N$ should be considered in the future works. In addition, authors of [48,171,172] employed RF chains with sets of digitally controlled phase paired phase-shifters and switches to achieve the same rate performance as that of the fully digital system. Furthermore, with the support of an advanced hybrid architecture [48, 171, 172], hybrid systems can achieve higher rate performance than that of adopting a simple ULA antenna array case. However, how to exploit extra RF chains as well as advanced hybrid architectures to further improve the rate performance of hybrid systems is still an unexplored/interesting research topic and is beyond the scope of this thesis. We may consider the case of $N_{\mathrm{RF}} > N$ and the case of adopting advanced hybrid architectures, e.g. phase shifts with digital controlled phase and amplitude, for channel estimation and downlink transmission precoding in our future works.

Distributed hybrid mmWave MIMO network to recapture the geographic degree of freedom:

The multi-small-cell hybrid mmWave MIMO network adopted in Chapter 4 pointed out that the inter-cell and intra-cell interference during the downlink transmission can be suppressed by simply increasing the number of antennas equipped at each BS. These features are useful for small cell scenarios and the dense deploy-

ment of BSs. Thus, cooperating multiple BSs across a large geographic area with small cell radius can facilitate the reuse of the same piece of spectrum and further improve the network spectral efficiency [122]. Besides, with the dense deployment of BSs, beamforming transmission technology offers robustness against blockages via inter-BS beam switching. In addition, the desired user may have multiple paths to multiple BSs. If users are equipped with multiple RF chains, the SNR maximization principle [3] for the design of analog beamforming matrices does not exploit the full channel capacity. Thus, it would be interesting to develop analog beamforming matrices and channel estimation solutions jointly to strike a balance among the channel capacity, the robustness of mmWave systems, the inter-BS handover, and the energy efficiency.

Extension to the wideband frequency selective :

In general, wideband mmWave channels are assumed as frequency selective. Since the scalable OFDM structure has been accepted as the standard of 5G NR and mmWave systems are attractive due to their wide bandwidth (in practice, mmWave are mostly operated at wideband frequencies to support higher data streams), developing efficient mmWave channel estimation for frequency selective channels is of great importance.

Current works in the literature assume that the analog beamforming matrices are frequency flat and are generated using a fully connected architecture with digitally controlled quantized phase shifters. However, analog beamforming matrices are frequency selective and have been discussed in the literature about optimal array signal processing. To extend the proposed tone-based algorithm, which estimates mmWave channels by sending narrow bandwidth pulses, to frequency selective wideband mmWave channels, there are several possible directions for

future research: 1) It is necessary to discuss the spatial beam pattern distortion of the analog beamforming in frequency selective mmWave channels. 2) If the digitally controlled quantized amplifier can be introduced to the hybrid architecture (combined with digitally controlled quantized phase shifters), can advanced array signal processing algorithms be introduced for the frequency selective compensation/equalization?

Appendix A

Proof of Theorem 2.1

The SINR expression of SLPS-based ZF is given by

$$
\mathrm{SINR}_{\mathrm{ZF}}^{k} = \frac{\widehat{\beta}^2 E_s}{\widehat{\beta}^2 E_s \mathbf{h}_k^T \mathbb{E}_{\Delta \mathbf{H}} \left[\Delta \mathbf{W} \Delta \mathbf{W}^H \right] \mathbf{h}_k^* + \sigma^2}. \tag{A.1}
$$

To express $\mathbb{E}_{\Delta \mathbf{H}}[\Delta \mathbf{W} \Delta \mathbf{W}^H]$, first we present the expression of $\Delta \mathbf{W}$ as

$$
\begin{aligned}
\Delta \mathbf{W} &= \widehat{\mathbf{W}} - \mathbf{W} \\
&= \sqrt{1 + \delta^2} (\mathbf{H} + \Delta \mathbf{H})^* \left[(\mathbf{H} + \Delta \mathbf{H})^T (\mathbf{H} + \Delta \mathbf{H})^* \right]^{-1} - \mathbf{H}^* \left(\mathbf{H}^T \mathbf{H}^* \right)^{-1},
\end{aligned} \tag{A.2}
$$

$$
(\mathbf{H}^T \mathbf{H}^*) = \mathbf{K} = (\mathbf{H}^H \mathbf{H}), \tag{A.3}
$$

$$
\mathbf{D} = (\mathbf{H}^H \Delta \mathbf{H} + \Delta \mathbf{H}^H \mathbf{H} + \Delta \mathbf{H}^H \Delta \mathbf{H}), \tag{A.4}
$$

and $\mathbf{H}^* \left(\mathbf{H}^H \mathbf{H} \right)^{-1} = \mathbf{W} = \mathbf{H}^* \mathbf{K}^{-1}$. Then, exploiting the matrix inversion approximation [176]

$$
(\mathbf{K} + \mathbf{D})^{-1} \approx \left[\mathbf{K}^{-1} - \mathbf{K}^{-1} \left(\mathbf{I} + \mathbf{D} \mathbf{K}^{-1} \right)^{-1} \mathbf{D} \mathbf{K}^{-1} \right], \tag{A.5}
$$

we re-express (A.2) as

$$
\Delta \mathbf{W} = \sqrt{1 + \delta^2} \left(\mathbf{H}^* + \Delta \mathbf{H}^* \right) \left[\mathbf{K}^{-1} - \mathbf{K}^{-1} \left(\mathbf{I} + \mathbf{D} \mathbf{K}^{-1} \right)^{-1} \mathbf{D} \mathbf{K}^{-1} \right] - \mathbf{H}^* \left(\mathbf{H}^H \mathbf{H} \right)^{-1}. \tag{A.6}
$$

Finally, we could express the closed-form approximation of

$$
\begin{aligned}
\mathbb{E}_{\Delta\mathbf{H}} = [\Delta\mathbf{W}\Delta\mathbf{W}^H] \overset{(a)}{=\approx} & (\sqrt{1+\delta^2}-1)^2\mathbf{W}\mathbf{W}^H \\
& + 2(\sqrt{1+\delta^2}-1-\delta^2)\delta^2\left[M\mathbf{W}\mathbf{K}^{-1}\left(\mathbf{I}+\mathbf{D}\mathbf{K}^{-1}\right)^{-1}\mathbf{W}^H\right] \\
& + 2(\sqrt{1+\delta^2}-1-\delta^2)\delta^2\left[\mathrm{tr}(\mathbf{K}^{-1})\mathbf{W}\left(\mathbf{I}+\mathbf{D}\mathbf{K}^{-1}\right)^{-1}\mathbf{W}^H\right] \\
& + (\sqrt{1+\delta^2})^2\delta^2\mathrm{tr}\left[\mathbf{K}^{-1}\left(\mathbf{I}+\mathbf{D}\mathbf{K}^{-1}\right)^{-1}\right]\mathbf{W}\mathbf{W}^H \\
\overset{(b)}{\approx} & (\sqrt{1+\delta^2}-1)^2\mathbf{W}\mathbf{W}^H \\
& + \frac{2\sqrt{1+\delta^2}(1-\sqrt{1+\delta^2})\delta^2\mathrm{tr}\left(\mathbf{K}^{-1}\right)}{1+\delta^2\mathrm{tr}\left(\mathbf{K}^{-1}\right)}\mathbf{W}\mathbf{W}^H \\
& + \frac{2\sqrt{1+\delta^2}(1-\sqrt{1+\delta^2})\delta^2 M}{1+\delta^2\mathrm{tr}(\mathbf{K}^{-1})}\mathbf{W}\mathbf{K}^{-1}\mathbf{W}^H \\
& + (\sqrt{1+\delta^2})^2\delta^2\frac{\mathrm{tr}(\mathbf{K}^{-1})}{1+\delta^2\mathrm{tr}(\mathbf{K}^{-1})}\mathbf{W}\mathbf{W}^H. \qquad (A.7)
\end{aligned}
$$

In (a), we omit some negligibly small random scattering correlation parts. Also, due to the reason that $(\mathbf{I}+\mathbf{D}\mathbf{K})^{-1}$ is symmetric and small, we can obtain the approximation $(\mathbf{I}+\mathbf{D}\mathbf{K})^{-1} \approx \frac{1}{1+\delta^2\mathrm{tr}(\mathbf{K}^{-1})}\mathbf{I}$ in (b). We substitute (A.7) into (2.26), then the result comes immediately after some straight forward mathematical manipulation.

Appendix B

Proof of Corollary 2.1

Now we prove Corollary 2.1. The desired signal part Φ_k of the received signal at user k is given by

$$\Phi_k = \beta^2 E_s \mathbf{h}_k^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{h}_k = M^2 \beta^2 E_s. \tag{B.1}$$

Then we express the interference part Ω_k as

$$\Omega_k = \beta^2 \mathbb{E}_{\mathbf{h}_S, \Delta\mathbf{h}} \left[\left(\sum_{j=1, j\neq k}^{N} \mathbf{h}_j^H x_j + \sum_{j=1}^{N} \Delta\mathbf{h}_j^H x_j \right) \mathbf{h}_k \mathbf{h}_k^H \left(\sum_{j=1, j\neq k}^{N} x_j^H \mathbf{h}_j + \sum_{j=1}^{N} x_j^H \Delta\mathbf{h}_j \right) \right]$$

$$= \beta^2 E_s \mathbb{E}_{\mathbf{h}_S} \underbrace{\left[\sum_{j=1, j\neq k}^{N} \left(\mathbf{h}_j^H \mathbf{h}_k \mathbf{h}_k^H \mathbf{h}_j \right) \right]}_{Q} + \beta^2 E_s \mathbb{E}_{\mathbf{h}_S} \underbrace{\left[\sum_{j=1}^{N} E_{\Delta\mathbf{h}} \left(\Delta\mathbf{h}_j^H \mathbf{h}_k \mathbf{h}_k^H \Delta\mathbf{h}_j \right) \right]}_{U}, \tag{B.2}$$

where Q is given by

$$Q = \left(\frac{\upsilon_k}{\upsilon_k + 1} \right)^2 \sum_{j=1, j\neq k}^{N} \mathbf{h}_{L,j}^H \mathbf{h}_{L,k} \mathbf{h}_{L,k}^H \mathbf{h}_{L,j} + \frac{1}{\upsilon_k + 1} \sum_{j=1, j\neq k}^{N} \mathbb{E}_{\mathbf{h}_{ds}} \left[\mathbf{h}_j^H \mathbf{h}_{S,k} \mathbf{h}_{S,k}^H \mathbf{h}_j \right]$$

$$+ \frac{\upsilon_k}{(\upsilon_k + 1)^2} \sum_{j=1, j\neq k}^{N} \mathbb{E}_{\mathbf{h}_S} \left[\mathbf{h}_{S,j}^H \mathbf{h}_{L,k} \mathbf{h}_{L,k}^H \mathbf{h}_{S,j} \right] + \frac{1}{(\upsilon_k + 1)^2} \sum_{j=1, j\neq k}^{N} \mathbb{E}_{\mathbf{h}_S} \left[\mathbf{h}_{S,j}^H \mathbf{h}_{S,k} \mathbf{h}_{S,k}^H \mathbf{h}_{S,j} \right]$$

$$\stackrel{(c)}{\approx} \left(\frac{\upsilon_k}{\upsilon_k + 1} \right)^2 \sum_{j=1, j\neq k}^{N} \mathbf{h}_{L,j}^H \mathbf{h}_{L,k} \mathbf{h}_{L,k}^H \mathbf{h}_{L,j} + \frac{(N-1)M(2\upsilon_k + 1)}{(\upsilon_k + 1)^2} \tag{B.3}$$

and U is given by

$$U = \sum_{j=1}^{N} \mathbb{E}_{\Delta\mathbf{h}} \left[\left(\Delta\mathbf{h}_j^H \mathbf{h}_k \mathbf{h}_k^H \Delta\mathbf{h}_j \right) \right]$$
$$= \text{tr} \left(\mathbf{h}_k^H \mathbf{h}_k \right) N\delta^2 = MN\delta^2. \tag{B.4}$$

In (c), we omit the negligibly small random scattering correlation part. Then, the expression in Equation (2.34) comes immediately after some straight forward mathematical manipulation.

Appendix C

Proof of Theorem 3.1

Now we prove Theorem 3.1.

$$R_{\mathrm{HB}} = \mathbb{E}_{\mathbf{H}_{\mathrm{S}}} \left\{ \log_2 \left[1 + \left[\mathrm{tr} \left[(\mathbf{H}_{\mathrm{eq}}^T \mathbf{H}_{\mathrm{eq}}^*)^{-1} \right] \right]^{-1} \frac{E_s}{\sigma_{\mathrm{MS}}^2} \right] \right\}. \tag{C.1}$$

First, we introduce some preliminaries.

Since $\mathbf{H}_{\mathrm{eq}}^T \mathbf{H}_{\mathrm{eq}}^*$ is a positive definite Hermitian matrix, by eigenvalue decomposition, it can be decomposed as $\mathbf{H}_{\mathrm{eq}}^T \mathbf{H}_{\mathrm{eq}}^* = \mathbf{U} \mathbf{\Lambda} \mathbf{V}^H$, $\mathbf{\Lambda} \in \mathbb{C}^{N \times N}$ is the positive diagonal eigenvalue matrix, while $\mathbf{V} \in \mathbb{C}^{N \times N}$ and $\mathbf{U} \in \mathbb{C}^{N \times N}$ are unitary matrixes, $\mathbf{U} = \mathbf{V}^H$.

The trace of the eigenvalues of $\mathbf{H}_{\mathrm{eq}}^T \mathbf{H}_{\mathrm{eq}}^*$ equals to the trace of matrix $\mathbf{\Lambda}$.

Then we can rewrite the power normalization factor in (C.1) as

$$\frac{N}{\mathrm{tr} \left[(\mathbf{H}_{\mathrm{eq}}^T \mathbf{H}_{\mathrm{eq}}^*)^{-1} \right]} = N \left[\mathrm{tr} \left[\mathbf{U} \mathbf{\Lambda} \mathbf{U}^H \right]^{-1} \right]^{-1} \tag{C.2}$$

$$= N \left[\mathrm{tr} \left(\mathbf{\Lambda}^{-1} \right) \right]^{-1} = \left[\sum_{i=1}^{N} \frac{1}{N} \lambda_i^{-1} \right]^{-1},$$

In addition, $f(x) = x^{-1}$, $x > 0$, is a strictly decreasing convex function and exploiting the convexity, we have the following results [177]

$$\left[\sum_{i=1}^{N} \frac{1}{N} \lambda_i^{-1} \right]^{-1} \leqslant \sum_{i=1}^{N} \frac{1}{N} \left[(\lambda_i^{-1})^{-1} \right] = \sum_{i=1}^{N} \frac{1}{N} \lambda_i. \tag{C.3}$$

Therefore, based on (C.2) and (C.3), we have the following inequality:

$$\frac{1}{\operatorname{tr}\left[\left(\mathbf{H}_{\mathrm{eq}}^{T}\mathbf{H}_{\mathrm{eq}}^{*}\right)^{-1}\right]} \leqslant \sum_{i=1}^{N}\frac{1}{N^2}\lambda_i = \frac{1}{N^2}\sum_{i=1}^{N}\lambda_i = \frac{1}{N^2}\operatorname{tr}\left[\mathbf{H}_{\mathrm{eq}}^{T}\mathbf{H}_{\mathrm{eq}}^{*}\right]. \qquad (C.4)$$

Based on (C.4), Equation (C.1) can be rewritten as

$$
\begin{aligned}
R_{\mathrm{HB}} &\overset{(a)}{\leqslant} \mathbb{E}_{\mathbf{H}_{\mathrm{S}}}\left\{\log_2\left[1 + \frac{1}{N^2}\operatorname{tr}\left[\mathbf{H}_{\mathrm{eq}}^{T}\mathbf{H}_{\mathrm{eq}}^{*}\right]\frac{E_s}{\sigma_{\mathrm{MS}}^2}\right]\right\} \\
&\overset{(b)}{\leqslant} \log_2\left\{1 + \frac{1}{N^2}\mathbb{E}_{\mathbf{H}_{\mathrm{S}}}\left[\operatorname{tr}\left(\mathbf{H}_{\mathrm{eq}}^{T}\mathbf{H}_{\mathrm{eq}}^{*}\right)\right]\frac{E_s}{\sigma_{\mathrm{MS}}^2}\right\} \\
&= \log_2\left\{1 + \frac{1}{N^2}\frac{E_s}{\sigma_{\mathrm{MS}}^2}\left[\left(\frac{\upsilon}{\upsilon+1}\right)MP\|\mathbf{F}_{\mathrm{RF}}^{H}\mathbf{F}_{\mathrm{RF}}\|_{\mathrm{F}}^2\right] + \left(\frac{1}{\upsilon+1}\right)\frac{E_s}{\sigma_{\mathrm{MS}}^2}\right\}.
\end{aligned}
$$
$$(C.5)$$

In (a), we follow (C.4) and in (b), we adopt the Jensen's inequality. This completes the proof.

Appendix D

Proof of Theorem 3.2

Now we prove Theorem 3.2.

The receive SINR of user k is given by

$$\widetilde{\text{SINR}}_{\text{ZF}}^k = \frac{\widetilde{\beta}^2 E_s}{\underbrace{\widetilde{\beta}^2 E_s \widehat{\mathbf{h}}_{\text{eq},k}^T \mathbb{E}_{\Delta\widehat{\mathbf{H}}_{\text{eq}}} \left[\Delta\widehat{\mathbf{W}}_{\text{eq}} \Delta\widehat{\mathbf{W}}_{\text{eq}}^H \right] \widehat{\mathbf{h}}_{\text{eq},k}^* + \sigma_{\text{MS}}^2}_{\text{Interference term due to errors}}}. \tag{D.1}$$

To express $\mathbb{E}_{\Delta\widehat{\mathbf{H}}_{\text{eq}}} \left[\Delta\widehat{\mathbf{W}}_{\text{eq}} \Delta\widehat{\mathbf{W}}_{\text{eq}}^H \right]$, first we present the normalized expression of $\Delta\widehat{\mathbf{W}}_{\text{eq}} =$

$$\sqrt{1+\delta^2}(\widehat{\mathbf{H}}_{\text{eq}}^* + \Delta\widehat{\mathbf{H}}_{\text{eq}}^*) \left[(\widehat{\mathbf{H}}_{\text{eq}} + \Delta\widehat{\mathbf{H}}_{\text{eq}})^T (\widehat{\mathbf{H}}_{\text{eq}}^* + \Delta\widehat{\mathbf{H}}_{\text{eq}}^*) \right]^{-1} - \widehat{\mathbf{H}}_{\text{eq}}^* \left(\widehat{\mathbf{H}}_{\text{eq}}^T \widehat{\mathbf{H}}_{\text{eq}} \right)^{-1}, \tag{D.2}$$

where $\mathbf{K} = (\widehat{\mathbf{H}}_{\text{eq}}^T \widehat{\mathbf{H}}_{\text{eq}}^*)$, $\mathbf{D} = (\Delta\widehat{\mathbf{H}}_{\text{eq}}^T \widehat{\mathbf{H}}_{\text{eq}}^* + \widehat{\mathbf{H}}_{\text{eq}}^T \Delta\widehat{\mathbf{H}}_{\text{eq}}^* + \Delta\widehat{\mathbf{H}}_{\text{eq}}^T \Delta\widehat{\mathbf{H}}_{\text{eq}}^*)$, and $\widehat{\mathbf{W}}_{\text{eq}} = \widehat{\mathbf{H}}_{\text{eq}}^* \mathbf{K}^{-1}$.

The matrix inversion approximation is given by

$$(\mathbf{K} + \mathbf{D})^{-1} \approx \left[\mathbf{K}^{-1} - \mathbf{K}^{-1}\mathbf{D}\mathbf{K}^{-1} \right]. \tag{D.3}$$

In this case, we re-express (D.2) as

$$\Delta\widehat{\mathbf{W}}_{\text{eq}} \approx \sqrt{1+\delta^2} \left(\widehat{\mathbf{H}}_{\text{eq}}^* + \Delta\widehat{\mathbf{H}}_{\text{eq}}^* \right) \left(\mathbf{K}^{-1} - \mathbf{K}^{-1}\mathbf{D}\mathbf{K}^{-1} \right) - \widehat{\mathbf{H}}_{\text{eq}}^* \mathbf{K}^{-1}. \tag{D.4}$$

Finally, we have

$$
E_{\Delta \widehat{\mathbf{H}}_{\text{eq}}} \left[\Delta \widehat{\mathbf{W}}_{\text{eq}} \Delta \widehat{\mathbf{W}}_{\text{eq}}^H \right]
$$

$$
\overset{(b)}{\approx} (\sqrt{1+\delta^2} - 1)^2 \left[\widehat{\mathbf{W}}_{\text{eq}} \widehat{\mathbf{W}}_{\text{eq}}^H \right] + 2\sqrt{1+\delta^2}(1 - \sqrt{1+\delta^2})\delta^2 N \left[\widehat{\mathbf{W}}_{\text{eq}} \mathbf{K}^{-1} \widehat{\mathbf{W}}_{\text{eq}}^H \right]
$$

$$
+ (2\sqrt{1+\delta^2} - \sqrt{1+\delta^2})\delta^2 \left[\text{tr}(\mathbf{K}^{-1}) \widehat{\mathbf{W}}_{\text{eq}} \widehat{\mathbf{W}}_{\text{eq}}^H \right] \tag{D.5}
$$

$$
\overset{(d)}{\approx} (\sqrt{1+\delta^2} - 1)^2 \left[\widehat{\mathbf{W}}_{\text{eq}} \widehat{\mathbf{W}}_{\text{eq}}^H \right] + \frac{2N}{\xi M P} \sqrt{1+\delta^2}(1 - \sqrt{1+\delta^2})\delta^2 \widehat{\mathbf{W}}_{\text{eq}} \mathbf{G}_{\text{L}}^{-1} \widehat{\mathbf{W}}_{\text{eq}}^H
$$

$$
+ (2\sqrt{1+\delta^2} - \sqrt{1+\delta^2})\delta^2 \left[\frac{N}{\xi M P} \widehat{\mathbf{W}}_{\text{eq}} \mathbf{G}_{\text{L}}^{-1} \widehat{\mathbf{W}}_{\text{eq}}^H \right]. \tag{D.6}
$$

In (b), we omit some negligibly small parts which neither dominate the performance nor scale with M. In (d), while the number of antennas $M \to \infty$, $\mathbf{K} = \widehat{\mathbf{H}}_{\text{eq}}^T \widehat{\mathbf{H}}_{\text{eq}}^* \overset{a.s.}{\underset{M\to\infty}{\approx}} \xi M P \mathbf{G}_{\text{L}}$ holds, where $\xi \in (0, 1]$. We substitute (D.5) into (3.46) and substitute (D.6) into (3.46), the Equations (3.47) and (3.48) come immediately after some straight forward mathematical manipulation.

Appendix E

Proof of Theorem 4.1

The normalized MSE performance of equivalent channel estimation under the impact of pilot contamination is given by

$$
\begin{aligned}
\mathrm{NMSE}_{\mathrm{eq},k} &= \frac{1}{N}\mathbb{E}_{\mathbf{U}_{l,i}}\left[\left(\frac{1}{\sqrt{MP}}\Delta\widehat{\mathbf{h}}_{\mathrm{eq},k}^{T}\right)\left(\frac{1}{\sqrt{MP}}\Delta\widehat{\mathbf{h}}_{\mathrm{eq},k}^{*}\right)\right] \\
&\overset{(a)}{\approx} \frac{1}{\varpi_{k}NMP}\mathbb{E}_{\mathbf{U}_{l,i}}\underbrace{\left[\sum_{l=1}^{L}\left(\widehat{\boldsymbol{\omega}}_{l,k}^{H}\mathbf{U}_{l,k}^{T}\mathbf{F}_{\mathrm{RF}}\right)\right]\left[\sum_{l=1}^{L}\left(\mathbf{F}_{\mathrm{RF}}^{H}\mathbf{U}_{l,k}^{*}\widehat{\boldsymbol{\omega}}_{l,k}\right)\right]}_{\text{Pilot contamination}} \\
&\quad + \underbrace{\frac{\sigma_{\mathrm{BS}}^{2}\mathrm{tr}\left[\mathbf{F}_{\mathrm{RF}}^{H}\mathbf{F}_{\mathrm{RF}}\right]}{\varpi_{k}E_{\mathrm{P}}NMP}}_{\text{Effective noise}}.
\end{aligned}
\tag{E.1}
$$

In (a), we omit some small part. Due to the small-cell radius, the inter-cell uplink propagation channels contain the strongest AoA components from the users in the neighboring cells to the desired BS. Thus, the part associated with pilot

contamination can be expressed as

$$\frac{1}{\varpi_k}\mathbb{E}_{\mathbf{U}_{l,k}}\left[\sum_{l=1}^{L}\left(\widehat{\boldsymbol{\omega}}_{l,k}^{H}\mathbf{U}_{l,k}^{T}\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{RF}}^{H}\mathbf{U}_{l,k}^{*}\widehat{\boldsymbol{\omega}}_{l,k}\right)\right]$$

$$=\mathbb{E}_{\mathbf{U}_{l,k}}\left[\sum_{l=1}^{L}\widehat{\boldsymbol{\omega}}_{l,k}^{H}\left(\sqrt{\frac{\varsigma_{l,k}}{\varsigma_{l,k}+1}}\widehat{\rho}_{l,k}\mathbf{U}_{\mathrm{SAoA},l,k}^{T}+\sqrt{\frac{1}{\varsigma_{l,k}+1}}\widehat{\rho}_{l,k}\mathbf{U}_{\mathrm{S},l,k}^{T}\right)\mathbf{F}_{\mathrm{RF}}\right.$$

$$\times\left.\mathbf{F}_{\mathrm{RF}}^{H}\left(\sqrt{\frac{\varsigma_{l,k}}{\varsigma_{l,k}+1}}\widehat{\rho}_{l,k}\mathbf{U}_{\mathrm{SAoA},l,k}^{*}+\sqrt{\frac{1}{\varsigma_{l,k}+1}}\widehat{\rho}_{l,k}\mathbf{U}_{\mathrm{S},l,k}^{*}\right)\widehat{\boldsymbol{\omega}}_{l,k}\right]$$

$$\approx\underbrace{\mathbb{E}_{\mathbf{U}_{\mathrm{SAoA},l,k}}\left[\sum_{l=1}^{L}\left(\frac{\widehat{\rho}_{l,k}^{2}\varsigma_{l,k}}{\varsigma_{l,k}+1}\widehat{\boldsymbol{\omega}}_{l,k}^{H}\mathbf{U}_{\mathrm{SAoA},l,k}^{T}\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{RF}}^{H}\mathbf{U}_{\mathrm{SAoA},l,k}^{*}\widehat{\boldsymbol{\omega}}_{l,k}\right)\right]}_{\mathrm{Inter-cell\ interference\ caused\ by\ strongest\ AoA\ components}}$$

$$+\underbrace{\mathbb{E}_{\mathbf{U}_{\mathrm{S},l,k}}\left[\sum_{l=1}^{L}\left(\frac{\widehat{\rho}_{l,k}^{2}}{\varsigma_{l,k}+1}\widehat{\boldsymbol{\omega}}_{l,k}^{H}\mathbf{U}_{\mathrm{S},l,k}^{T}\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{RF}}^{H}\mathbf{U}_{\mathrm{S},l,k}^{*}\widehat{\boldsymbol{\omega}}_{l,k}\right)\right]}_{\mathrm{Inter-cell\ interference\ caused\ by\ scattering\ components}},\qquad(\mathrm{E.2})$$

where $\widehat{\rho}_{l,k}=\sqrt{\frac{\widehat{\varpi}_{l,k}}{\varpi_k}}$. Then, the inter-cell interference caused by scattering component can be further approximated as

$$\left[\sum_{l=1}^{L}\frac{\widehat{\rho}_{l,k}^{2}}{\varsigma_{l,k}+1}\widehat{\boldsymbol{\omega}}_{l,k}^{H}\mathbb{E}_{\mathbf{U}_{\mathrm{S},l,k}}\left(\mathbf{U}_{\mathrm{S},l,k}^{T}\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{RF}}^{H}\mathbf{U}_{\mathrm{S},l,k}^{*}\right)\widehat{\boldsymbol{\omega}}_{l,k}\right]$$

$$\approx\left[\sum_{l=1}^{L}\frac{\widehat{\rho}_{l,k}^{2}}{\varsigma_{l,k}+1}\widehat{\boldsymbol{\omega}}_{l,k}^{H}\mathrm{tr}\left(\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{RF}}^{H}\right)\mathbf{I}_{\mathrm{P}}\widehat{\boldsymbol{\omega}}_{l,k}\right]=\sum_{l=1}^{L}\frac{\widehat{\rho}_{l,k}^{2}}{\varsigma_{l,k}+1}N.\qquad(\mathrm{E.3})$$

Now, we would like to approximate the inter-cell interference caused by the multi-cell strongest AoA components

$$\mathbb{E}_{\mathbf{U}_{\mathrm{SAoA},l,k}}\left[\sum_{l=1}^{L}\left(\frac{\widehat{\rho}_{l,k}^{2}\varsigma_{l,k}}{\varsigma_{l,k}+1}\widehat{\boldsymbol{\omega}}_{l,k}^{H}\mathbf{U}_{\mathrm{SAoA},l,k}^{T}\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{RF}}^{H}\mathbf{U}_{\mathrm{SAoA},l,k}^{*}\widehat{\boldsymbol{\omega}}_{l,k}\right)\right]$$

$$=\mathbb{E}_{\mathbf{U}_{\mathrm{SAoA},l,k}}\left[\sum_{l=1}^{L}\left[\frac{\widehat{\rho}_{l,k}^{2}\varsigma_{l,k}}{\varsigma_{l,k}+1}\left(\mathbf{u}_{\mathrm{SAoA},l,k}^{\mathrm{BS}}\right)^{T}\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{RF}}^{H}\left(\mathbf{u}_{\mathrm{SAoA},l,k}^{\mathrm{BS}}\right)^{*}\mathbf{u}_{\mathrm{SAoA},l,k}^{T}\widehat{\boldsymbol{\omega}}_{l,k}\widehat{\boldsymbol{\omega}}_{l,k}^{H}\mathbf{u}_{\mathrm{SAoA},l,k}^{*}\right]\right].$$

$$(\mathrm{E.4})$$

In Equation (E.4), we have

$$\widehat{\boldsymbol{\omega}}_{l,k}^{H} = \frac{1}{\sqrt{P}} \left[\begin{array}{cccc} 1, & e^{j2\pi\frac{d}{\lambda}\cos(\phi_{l,k})}, & \ldots, & e^{j2\pi(P-1)\frac{d}{\lambda}\cos(\phi_{l,k})} \end{array} \right]^{*} \text{ and} \quad \text{(E.5)}$$

$$\mathbf{u}_{\text{SAoA},l,k}^{*} = \left[\begin{array}{cccc} 1, & e^{-j2\pi\frac{d}{\lambda}\cos(\Delta\phi_{l,k})}, & \ldots, & e^{-j2\pi(P-1)\frac{d}{\lambda}\cos(\Delta\phi_{l,k})} \end{array} \right]^{H}, \quad \text{(E.6)}$$

where variables $\phi_{l,k} \in [0, \pi]$ is the angle of incidence of the strongest AoA path at antenna arrays of user k in cell l, and $\Delta\phi_{l,k} \in [0, \pi]$ is the angle of incidence of the inter-cell strongest AoA path at antenna arrays from user k of cell l to the desired BS. By defining the array gain function $G_{\text{act},P}[x]$, cf. [170], as

$$G_{\text{act},P}[x] = \frac{\left\{ \sin\left[P\pi\frac{d}{\lambda}(x) \right] \right\}^{2}}{P\left\{ \sin\left[\pi\frac{d}{\lambda}(x) \right] \right\}^{2}}, \quad \text{(E.7)}$$

where $\frac{d}{\lambda} = \frac{1}{2}$. Then, we have:

$$\mathbf{u}_{\text{SAoA},l,k}^{T}\widehat{\boldsymbol{\omega}}_{l,k}\widehat{\boldsymbol{\omega}}_{l,k}^{H}\mathbf{u}_{\text{SAoA},l,k}^{*} = G_{\text{act},P}[\cos(\phi_{l,k}) - \cos(\Delta\phi_{l,k})]$$

$$= \frac{P\left\{ \text{sinc}\left[P\frac{\pi}{2}(\cos(\phi_{l,k}) - \cos(\Delta\phi_{l,k})) \right] \right\}^{2}}{\left\{ \text{sinc}\left[\frac{\pi}{2}(\cos(\phi_{l,k}) - \cos(\Delta\phi_{l,k})) \right] \right\}^{2}}. \quad \text{(E.8)}$$

Similarly, we can have following preliminaries, i.e.,

$$\left(\mathbf{u}_{\text{SAoA},l,k}^{\text{BS}}\right)^{T} = \left[\begin{array}{cccc} 1, & e^{-j2\pi\frac{d}{\lambda}\cos(\theta_{l,k})}, & \ldots, & e^{-j2\pi(M-1)\frac{d}{\lambda}\cos(\theta_{l,k})} \end{array} \right] \text{ and}$$

$$\widehat{\boldsymbol{\nu}}_{i} = \frac{1}{\sqrt{M}} \left[\begin{array}{cccc} 1, & e^{j2\pi\frac{d}{\lambda}\cos(\theta_{i})}, & \ldots, & e^{j2\pi(M-1)\frac{d}{\lambda}\cos(\theta_{i})} \end{array} \right]^{T}. \quad \text{(E.9)}$$

Based on expressions mentioned above, we rewrite $\left(\mathbf{u}_{\text{SAoA},l,k}^{\text{BS}}\right)^{T}\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{RF}}^{H}\left(\mathbf{u}_{\text{SAoA},l,k}^{\text{BS}}\right)^{*}$ as

$$\left[\begin{array}{ccc} \left(\mathbf{u}_{\text{SAoA},l,k}^{\text{BS}}\right)^{T}\widehat{\boldsymbol{\nu}}_{1}, & \ldots, & \left(\mathbf{u}_{\text{SAoA},l,k}^{\text{BS}}\right)^{T}\widehat{\boldsymbol{\nu}}_{N} \end{array} \right]$$

$$\times \left[\begin{array}{ccc} \widehat{\boldsymbol{\nu}}_{1}^{H}\left(\mathbf{u}_{\text{SAoA},l,k}^{\text{BS}}\right)^{*}, & \ldots, & \widehat{\boldsymbol{\nu}}_{N}^{H}\left(\mathbf{u}_{\text{SAoA},l,k}^{\text{BS}}\right)^{*} \end{array} \right]^{T}$$

$$= \sum_{i=1}^{N} \frac{\left\{ \sin\left[\frac{\pi}{2}M(\cos(\theta_{l,k}) - \cos(\theta_{i})) \right] \right\}^{2}}{M\left\{ \sin\left[\frac{\pi}{2}\cos(\theta_{l,k}) - \cos(\theta_{i}) \right] \right\}^{2}} = \sum_{i=1}^{N} G_{\text{act},M}[\cos(\theta_{l,k}) - \cos(\theta_{i})]. \quad \text{(E.10)}$$

Then, Equation (E.4) can be re-expressed as

$$
\mathbb{E}_{\mathbf{U}_{\mathrm{SAoA},l,k}}\left[\sum_{l=1}^{L}\left[\frac{\widehat{\rho}_{l,k}^{2}\varsigma_{l,k}}{\varsigma_{l,k}+1}\left(\mathbf{u}_{\mathrm{SAoA},l,k}^{\mathrm{BS}}\right)^{T}\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{RF}}^{H}\left(\mathbf{u}_{\mathrm{SAoA},l,k}^{\mathrm{BS}}\right)^{*}\mathbf{u}_{\mathrm{SAoA},l,k}^{T}\widehat{\boldsymbol{\omega}}_{l,k}\widehat{\boldsymbol{\omega}}_{l,k}^{H}\mathbf{u}_{\mathrm{SAoA},l,k}^{*}\right]\right]
$$

$$
=\mathbb{E}_{\phi_{l,k},\phi_{l,k},\theta_{l,k},\theta_{i}}\left\{\sum_{l=1}^{L}\left[\frac{\widehat{\rho}_{l,k}^{2}\varsigma_{l,k}}{\varsigma_{l,k}+1}G_{\mathrm{act},P}[\cos\left(\phi_{l,k}\right)-\cos\left(\Delta\phi_{l,k}\right)]\right.\right.
$$

$$
\left.\left.\times\sum_{i=1}^{N}G_{\mathrm{act},M}[\cos\left(\theta_{l,k}\right)-\cos\left(\theta_{i}\right)]\right]\right\},\qquad(\mathrm{E}.11)
$$

where $\cos\left(\phi_{l,k}\right)$, $\cos\left(\Delta\phi_{l,k}\right)$, $\cos\left(\theta_{l,k}\right)$, and $\cos\left(\theta_{i}\right)$ are uniformly distributed over $[-1,\ 1]$. Besides, they are independent with each other. Exploiting the periodic property of function $e^{j2\pi x}$, the linear antenna array gain $G_{\mathrm{act},P}[\cos\left(\phi_{l,k}\right)-\cos\left(\Delta\phi_{l,k}\right)]$ is equal in distribution to $G_{\mathrm{act},P}[\mu_{l,k}]$ and $G_{\mathrm{act},M}[\cos\left(\theta_{l,k}\right)-\cos\left(\theta_{i}\right)]$ is equal in distribution to $G_{\mathrm{act},M}[\epsilon_{l,k,i}]$, where $\mu_{l,k}$, $k\in\{1,\ldots,N\}$, and $\epsilon_{l,k,i}$, $i\in\{1,\ldots,N\}$ are uniformly distributed over $[-1,\ 1]$ (Lemma 1 of [178]). In addition, $\mu_{l,k}$ and $\epsilon_{l,k,i}$ is independent with each other. Thus, we re-express Equation (E.11) as

$$
\sum_{l=1}^{L}\frac{\widehat{\rho}_{l,k}^{2}\varsigma_{l,k}}{\varsigma_{l,k}+1}\mathbb{E}_{\phi_{l,k},\phi_{l,k}}\left[G_{\mathrm{act},P}[\cos\left(\phi_{l,k}\right)-\cos\left(\Delta\phi_{l,k}\right)]\right]
$$

$$
\times\sum_{i=1}^{N}\mathbb{E}_{\theta_{l,k},\theta_{i}}\left[G_{\mathrm{act},M}[\cos\left(\theta_{l,k}\right)-\cos\left(\theta_{i}\right)]\right]
$$

$$
=\sum_{l=1}^{L}\frac{\widehat{\rho}_{l,k}^{2}\varsigma_{l,k}}{\varsigma_{l,k}+1}\mathbb{E}_{\mu_{l,k}}\left[G_{\mathrm{act},P}[\mu_{l,k}]\right]\mathbb{E}_{\epsilon_{l,k,i}}\left(G_{\mathrm{act},M}[\epsilon_{l,k,i}]\right)
$$

$$
=\sum_{l=1}^{L}\frac{\widehat{\rho}_{l,k}^{2}\varsigma_{l,k}}{\varsigma_{l,k}+1}\mathbb{E}_{\mu_{l,k}}\left(\frac{\left[\mathrm{sinc}\left(\frac{\pi}{2}P\mu_{l,k}\right)\right]^{2}P}{\left[\mathrm{sinc}\left(\frac{\pi}{2}\mu_{l,k}\right)\right]^{2}}\right)\mathbb{E}_{\epsilon_{l,k,i}}\left(\sum_{i=1}^{N}\frac{\left[\mathrm{sinc}\left(\frac{\pi}{2}M\epsilon_{l,k,i}\right)\right]^{2}M}{\left[\mathrm{sinc}\left(\frac{\pi}{2}\epsilon_{l,k,i}\right)\right]^{2}}\right)
$$

$$
\overset{(b)}{\geqslant}\sum_{l=1}^{L}\frac{\widehat{\rho}_{l,k}^{2}\varsigma_{l,k}}{\varsigma_{l,k}+1}\mathbb{E}_{\mu_{l,k}}\left(\left[\mathrm{sinc}\left(\frac{\pi}{2}P\mu_{l,k}\right)\right]^{2}P\right)\mathbb{E}_{\epsilon_{l,k,i}}\left(\sum_{i=1}^{N}\left[\mathrm{sinc}\left(\frac{\pi}{2}M\epsilon_{l,k,i}\right)\right]^{2}M\right)
$$

$$
\overset{(c)}{\approx}\sum_{l=1}^{L}\frac{\widehat{\rho}_{l,k}^{2}\varsigma_{l,k}}{\varsigma_{l,k}+1}N.\qquad(\mathrm{E}.12)
$$

In (b), we exploit the fact that

$$
[\mathrm{sinc}(x)]^{2}=\left(\frac{\sin x}{x}\right)^{2}\leqslant 1.\qquad(\mathrm{E}.13)
$$

In (c), we explore the law of integration of sinc function for the number of antennas M is sufficiently large, i.e.,

$$\mathbb{E}_{\epsilon_{l,k,i}}\left[\left[\text{sinc}\left(\frac{\pi}{2}M\epsilon_{l,k,i}\right)\right]^2 M\right] = M\int_{-1}^{1}\frac{1}{2}\left[\text{sinc}\left(\frac{\pi}{2}M\epsilon_{l,k,i}\right)\right]^2 d\epsilon_{l,k,i}$$

$$=\frac{M}{M\pi}\int_{-1}^{1}\left[\text{sinc}\left(\frac{\pi}{2}M\epsilon_{l,k,i}\right)\right]^2 d\frac{\pi}{2}M\epsilon_{l,k,i} \overset{M\to\infty}{\approx} \frac{1}{\pi}\int_{-\infty}^{\infty}[\text{sinc}(\chi)]^2 d\chi = 1, \quad \text{(E.14)}$$

where $\chi = \frac{\pi}{2}M\epsilon_{l,k,i}$. We substitute Equation (E.3) and (E.12) into (E.1), the expression in (4.15) comes immediately after some straightforward mathematical manipulation.

Appendix F

Proof of Lemma 4.2

The inter-cell downlink transmission interference is given by

$$
\Omega_k \geqslant
$$

$$
E_s \mathbb{E}_{\mathbf{H}_{\mathrm{SAoA},l,k}} \underbrace{\left[\sum_{l=1}^{L} \widetilde{\varpi}_{l,k} \left(\widehat{\beta}_l \right)^2 \left(\frac{\varsigma_{l,k}}{\varsigma_{l,k}+1} \right) \widehat{\boldsymbol{\omega}}_k^H \mathbf{H}_{\mathrm{SAoA},l,k}^T \mathbf{F}_{\mathrm{RF},l} \widehat{\mathbf{W}}_{\mathrm{eq},l} \widehat{\mathbf{W}}_{\mathrm{eq},l}^H \mathbf{F}_{\mathrm{RF},l}^H \mathbf{H}_{\mathrm{SAoA},l,k}^* \widehat{\boldsymbol{\omega}}_k \right]}_{\mathrm{Inter-cell\ interfernce\ caused\ by\ strongest\ AoAs}}
$$

$$
+ E_s \mathbb{E}_{\mathbf{H}_{\mathrm{S},l,k}} \underbrace{\left[\sum_{l=1}^{L} \widetilde{\varpi}_{l,k} \left(\widehat{\beta}_l \right)^2 \left(\frac{1}{\varsigma_{l,k}+1} \right) \widehat{\boldsymbol{\omega}}_k^H \mathbf{H}_{\mathrm{S},l,k}^T \mathbf{F}_{\mathrm{RF},l} \widehat{\mathbf{W}}_{\mathrm{eq},l} \widehat{\mathbf{W}}_{\mathrm{eq},l}^H \mathbf{F}_{\mathrm{RF},l}^H \mathbf{H}_{\mathrm{S},l,k}^* \widehat{\boldsymbol{\omega}}_k \right]}_{\mathrm{Inter-cell\ interference\ caused\ by\ scattering\ component}}.
$$

$$(\mathrm{F.1})$$

Following similar approaches shown in Equations $(\mathrm{E.4}) - (\mathrm{E.10})$, we can re-express the inter-cell interference caused by the strongest AoA components as

$$
\mathbb{E}_{\mathbf{H}_{\mathrm{SAoA},l,k}} \left\{ \sum_{l=1}^{L} \widetilde{\varpi}_{l,k} \left(\widehat{\beta}_l \right)^2 \left(\frac{\varsigma_{l,k}}{\varsigma_{l,k}+1} \right) \right.
$$

$$
\times \mathrm{tr} \left[\widetilde{\boldsymbol{\omega}}_k^H \mathbf{H}_{\mathrm{SAoA},l,k}^T \mathbf{F}_{\mathrm{RF},l} \widehat{\mathbf{W}}_{\mathrm{eq},l} \widehat{\mathbf{W}}_{\mathrm{eq},l}^H \mathbf{F}_{\mathrm{RF},l}^H \mathbf{H}_{\mathrm{SAoA},l,k}^* \widetilde{\boldsymbol{\omega}}_k \right] \right\}
$$

$$
= \sum_{l=1}^{L} \left\{ \widetilde{\varpi}_{l,k} \left(\widehat{\beta}_l \right)^2 \left(\frac{\varsigma_{l,k}}{\varsigma_{l,k}+1} \right) \mathbb{E}_{\widehat{\eta}_{l,k}} \left[\frac{\mathrm{sinc}^2 \left[P \frac{\pi}{2} \left(\widehat{\eta}_{l,k} \right) \right] P}{\mathrm{sinc}^2 \left[\frac{\pi}{2} \left(\widehat{\eta}_{l,k} \right) \right]} \right] \right.
$$

$$
\left. \times \mathrm{tr} \left[\widehat{\mathbf{W}}_{\mathrm{eq},l} \widehat{\mathbf{W}}_{\mathrm{eq},l}^H \cdot \mathbb{E}_{\tau_{l,k,i}} \left[\mathbf{F}_{\mathrm{RF},l}^H \left(\mathbf{h}_{\mathrm{SAoA},l,k}^{\mathrm{BS}} \right)^* \left(\mathbf{h}_{\mathrm{SAoA},l,k}^{\mathrm{BS}} \right)^T \mathbf{F}_{\mathrm{RF},l} \right] \right]^T \right\}, \quad (\mathrm{F.2})
$$

where $i \in \{1, \ldots, N\}$, $\widehat{\eta}_{l,k}$ and $\tau_{l,k,i}$ are also independent uniformly distributed in $[-1, 1]$. In addition, variables $\Delta \kappa_{l,k} \in [0, \pi]$ is the angle of incidence of

the inter-cell strongest AoA path at antenna arrays of user k in the desired cell from the BS in cell l, and $\psi_{l,k} \in [0,\ \pi]$ is the angle of incidence of the inter-cell strongest AoA path at antenna arrays of the BS of cell l for user k. Further, in the large number of antennas regime, $\mathbb{E}_{\tau_{l,k,i}} \left[\mathbf{F}_{\mathrm{RF},l}^H \left(\mathbf{h}_{\mathrm{SAoA},l,k}^{\mathrm{BS}}\right)^* \left(\mathbf{h}_{\mathrm{SAoA},l,k}^{\mathrm{BS}}\right)^T \mathbf{F}_{\mathrm{RF},l} \right],\ \ i \in \{1,\dots,N\}$, can be approximated as

$$
\begin{bmatrix}
\mathbb{E}_{\tau_{l,k,1}} \left[\dfrac{\operatorname{sinc}^2\left[M\frac{\pi}{2}\left(\tau_{l,k,1}\right)\right]M}{\operatorname{sinc}^2\left[\frac{\pi}{2}\left(\tau_{l,k,1}\right)\right]} \right] & \cdots & \mathbb{E}_{\tau_{l,k,1},\tau_{l,k,N}} \left[\dfrac{\sin\left[M\frac{\pi}{2}\left(\tau_{l,k,1}\right)\right]}{\sqrt{M}\sin\left[\frac{\pi}{2}\left(\tau_{l,k,1}\right)\right]} \dfrac{\sin\left[M\frac{\pi}{2}\left(\tau_{l,k,N}\right)\right]}{\sqrt{M}\sin\left[\frac{\pi}{2}\left(\tau_{l,k,N}\right)\right]} \right] \\
\vdots & \ddots & \vdots \\
\mathbb{E}_{\tau_{l,k,1},\tau_{l,k,N}} \left[\dfrac{\sin\left[M\frac{\pi}{2}\left(\tau_{l,k,1}\right)\right]}{\sqrt{M}\sin\left[\frac{\pi}{2}\left(\tau_{l,k,1}\right)\right]} \dfrac{\sin\left[M\frac{\pi}{2}\left(\tau_{l,k,N}\right)\right]}{\sqrt{M}\sin\left[\frac{\pi}{2}\left(\tau_{l,k,N}\right)\right]} \right] & \cdots & \mathbb{E}_{\tau_{l,k,N}} \left[\dfrac{\operatorname{sinc}^2\left[M\frac{\pi}{2}\left(\tau_{l,k,N}\right)\right]M}{\operatorname{sinc}^2\left[\frac{\pi}{2}\left(\tau_{l,k,N}\right)\right]} \right]
\end{bmatrix} .
$$

$$\text{(F.3)}$$

In the large number of antennas regime, $\mathbb{E}_{\tau_{l,k,i}} \left[\dfrac{\operatorname{sinc}^2\left[M\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]M}{\operatorname{sinc}^2\left[\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]} \right], i \in \{1,\dots,N\}$, can be asymptotically approximated as

$$
\mathbb{E}_{\tau_{l,k,i}} \left[\frac{\operatorname{sinc}^2 \left[M\frac{\pi}{2}\left(\tau_{l,k,i}\right) \right] M}{\operatorname{sinc}^2 \left[\frac{\pi}{2}\left(\tau_{l,k,i}\right) \right]} \right] \approx 1. \tag{F.4}
$$

As variables $\tau_{l,k,i}$ and $\tau_{l,k,j}$ are independent with each other, here we first derive $\mathbb{E}_{\tau_{l,k,i}} \left[\dfrac{\sin\left[M\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]}{\sqrt{M}\sin\left[\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]} \right]$. We rewrite $\dfrac{\sin\left[M\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]}{\sqrt{M}\sin\left[\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]}$ as

$$
\frac{\sin\left[M\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]}{\sqrt{M}\sin\left[\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]} = \sqrt{M}\frac{\dfrac{\sin\left[M\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]}{M\frac{\pi}{2}\left(\tau_{l,k,i}\right)}}{\dfrac{\sin\left[\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]}{\frac{\pi}{2}\left(\tau_{l,k,i}\right)}} = \sqrt{M}\frac{\operatorname{sinc}\left[M\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]}{\operatorname{sinc}\left[\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]}. \tag{F.5}
$$

Since the value of $\tau_{l,k,i}$ is uniformly distributed over $[-1,\ 1]$, we can obtain the following result:

$$
\operatorname{sinc}\left(\frac{\pi}{2}\right) \leq \operatorname{sinc}\left[\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right] \leq 1. \tag{F.6}
$$

Based on Equation (F.6) and Equation (F.5), we have the following inequality

$$
\sqrt{M}\operatorname{sinc}\left[M\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right] \leq \sqrt{M}\frac{\operatorname{sinc}\left[M\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]}{\operatorname{sinc}\left[\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]} < 2\sqrt{M}\operatorname{sinc}\left[M\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]. \tag{F.7}
$$

Now, we can obtain the following result by exploring the law of integration of sinc function for a sufficiently large number of antennas M:

$$
\begin{aligned}
& \mathrm{E}_{\tau_{l,k,i}}\left[\sqrt{M}\operatorname{sinc}\left[M\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]\right] \\
=&\sqrt{M}\int_{-1}^{1}\frac{1}{2}\operatorname{sinc}\left[\frac{\pi}{2}M\left(\tau_{l,k,i}\right)\right]d\tau_{l,k,i}=\frac{\sqrt{M}}{M\pi}\int_{-1}^{1}\operatorname{sinc}\left[\frac{\pi}{2}M\left(\tau_{l,k,i}\right)\right]d\frac{\pi}{2}M\tau_{l,k,i} \\
\approx&\frac{1}{\sqrt{M}\pi}\int_{-\infty}^{\infty}\operatorname{sinc}\left(\chi_{2}\right)d\chi_{2}\stackrel{M\to\infty}{\approx}0 \\
=&\mathrm{E}_{\tau_{l,k,i}}\left[2\sqrt{M}\operatorname{sinc}\left[M\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]\right],
\end{aligned}
\tag{F.8}
$$

where $\chi_{2}=\frac{\pi}{2}M\tau_{l,k,i}$.

According to the squeeze theorem [179], we have

$$
\mathrm{E}_{\tau_{l,k,i}}\left[\frac{\sin\left[M\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]}{\sqrt{M}\sin\left[\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]}\right]\stackrel{M\to\infty}{\approx}0.
\tag{F.9}
$$

Thus, for $i\neq j$, $i,\ j\in\{1,\ldots,N\}$ and the large number of antennas regime, $\mathbb{E}_{\tau_{l,k,i},\tau_{l,k,j}}\left[\frac{\sin\left[M\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]}{\sqrt{M}\sin\left[\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]}\frac{\sin\left[M\frac{\pi}{2}\left(\tau_{l,k,j}\right)\right]}{\sqrt{M}\sin\left[\frac{\pi}{2}\left(\tau_{l,k,j}\right)\right]}\right]$ can be approximated as

$$
\mathbb{E}_{\tau_{l,k,i}}\left[\frac{\sin\left[M\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]}{\sqrt{M}\sin\left[\frac{\pi}{2}\left(\tau_{l,k,i}\right)\right]}\right]\mathbb{E}_{\tau_{l,k,j}}\left[\frac{\sin\left[M\frac{\pi}{2}\left(\tau_{l,k,j}\right)\right]}{\sqrt{M}\sin\left[\frac{\pi}{2}\left(\tau_{l,k,j}\right)\right]}\right]\approx0.
\tag{F.10}
$$

In the large number of antennas regime, Equation (F.3) can be asymptotically approximated as

$$
\mathbb{E}_{\tau_{l,k,i}}\left[\mathbf{F}_{\mathrm{RF},l}^{H}\left(\mathbf{h}_{\mathrm{SAoA},l,k}^{\mathrm{BS}}\right)^{*}\left(\mathbf{h}_{\mathrm{SAoA},l,k}^{\mathrm{BS}}\right)^{T}\mathbf{F}_{\mathrm{RF},l}\right]\approx\mathbf{I}_{\mathrm{N}},\ i\in\{1,\ldots,N\}.
\tag{F.11}
$$

Thus, Equation (F.2) can be asymptotically approximated as

$$
\begin{aligned}
& \mathbb{E}_{\mathbf{H}_{\mathrm{SAoA},l,k}}\left[\sum_{l=1}^{L}\left(\frac{\widetilde{\varpi}_{l,k}\varsigma_{l,k}}{\varsigma_{l,k}+1}\right)\left[\frac{\operatorname{sinc}^{2}\left[P\frac{\pi}{2}\left(\eta_{l,k}\right)\right]P}{\operatorname{sinc}^{2}\left[\frac{\pi}{2}\left(\eta_{l,k}\right)\right]}\right]\widehat{\beta}^{2}\operatorname{tr}\left(\widehat{\mathbf{W}}_{\mathrm{eq},l}\widehat{\mathbf{W}}_{\mathrm{eq},l}^{H}\right)\right] \\
& \stackrel{M\to\infty}{\approx}\sum_{l=1}^{L}\left(\frac{\widetilde{\varpi}_{l,k}\varsigma_{l,k}}{\varsigma_{l,k}+1}\right)\mathbb{E}_{\eta_{l,k}}\left[\frac{\operatorname{sinc}^{2}\left[P\frac{\pi}{2}\left(\eta_{l,k}\right)\right]P}{\operatorname{sinc}^{2}\left[\frac{\pi}{2}\left(\eta_{l,k}\right)\right]}\right] \\
& \approx\sum_{l=1}^{L}\left(\frac{\widetilde{\varpi}_{l,k}\varsigma_{l,k}}{\varsigma_{l,k}+1}\right).
\end{aligned}
\tag{F.12}
$$

On the other hand, we re-express the inter-cell interference caused by scattering components in the large number of antennas regime as

$$\sum_{l=1}^{L} \widetilde{\boldsymbol{\omega}}_k^H \left(\frac{\widetilde{\varpi}_{l,k}}{\varsigma_{l,k}+1} \right) \frac{1}{\text{tr}\left(\widehat{\mathbf{W}}_{\text{eq},l} \widehat{\mathbf{W}}_{\text{eq},l}^H \right)} \text{tr}\left(\mathbf{F}_{\text{RF},l} \widehat{\mathbf{W}}_{\text{eq},l} \widehat{\mathbf{W}}_{\text{eq},l}^H \mathbf{F}_{\text{RF},l}^H \right) \widetilde{\boldsymbol{\omega}}_k$$

$$\overset{M \to \infty}{\approx} \sum_{l=1}^{L} \left(\frac{\widetilde{\varpi}_{l,k}}{\varsigma_{l,k}+1} \right). \tag{F.13}$$

By substituting Equations (F.12) and (F.13) into Equation (F.1), we have

$$\Omega_k \approx E_s \left[\sum_{l=1}^{L} \left(\frac{\widetilde{\varpi}_{l,k}}{\varsigma_{l,k}+1} \right) + \sum_{l=1}^{L} \left(\frac{\widetilde{\varpi}_{l,k}\varsigma_{l,k}}{\varsigma_{l,k}+1} \right) \right] = E_s \left[\sum_{l=1}^{L} \left(\widetilde{\varpi}_{l,k} \right) \right]. \tag{F.14}$$

This completes the proof.

Bibliography

[1] L. Zhao, D. W. K. Ng, and J. Yuan, "Multi-user precoding and channel estimation for hybrid millimeter wave systems," *IEEE J. Select. Areas Commun.*, vol. 35, no. 7, pp. 1576–1590, Jul. 2017.

[2] L. Zhao, G. Geraci, T. Yang, D. W. K. Ng, and J. Yuan, "A tone-based AoA estimation and multiuser precoding for millimeter wave massive MIMO," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5209–5225, Dec. 2017.

[3] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6481–6494, Nov. 2015.

[4] L. Zhao, Z. Wei, D. W. K. Ng, J. Yuan, and M. C. Reed, "Multi-cell hybrid millimeter wave systems: Pilot contamination and interference mitigation," 2018. [Online]. Available: https://arxiv.org/abs/1502.07228

[5] M. Kokshoorn, H. Chen, P. Wang, Y. Li, and B. Vucetic, "Millimeter wave MIMO channel estimation using overlapped beam patterns and rate adaptation," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 601–616, Feb. 2017.

[6] L. Zhao, Z. Wei, D. W. K. Ng, J. Yuan, and M. C. Reed, "Mitigating pilot contamination in multi-cell hybrid millimeter wave systems," 2018. [Online]. Available: https://arxiv.org/abs/1801.09176

[7] L. Zhao, D. W. K. Ng, and J. Yuan, "Multiuser precoding and channel estimation for hybrid millimeter wave MIMO systems," in *Proc. IEEE Intern. Commun. Conf. (ICC)*, May 2017, pp. 1–7.

[8] L. Zhao, T. Yang, G. Geraci, and J. Yuan, "Downlink multiuser massive MIMO in Rician channels under pilot contamination," in *Proc. IEEE Intern. Commun. Conf. (ICC)*, May 2016, pp. 1–6.

[9] Z. Wei, L. Zhao, J. Guo, D. W. K. Ng, and J. Yuan, "A multi-beam NOMA framework for hybrid mmwave systems," 2018, accepted, ICC 2018. [Online]. Available: https://arxiv.org/abs/1804.08303

[10] Z. Wei, D. W. K. Ng, J. Yuan, and H. M. Wang, "Optimal resource allocation for power-efficient MC-NOMA with imperfect channel state information," *IEEE Trans. Commun.*, vol. 65, no. 9, pp. 3944–3961, Sep. 2017.

[11] Z. Wei, D. W. K. Ng, and J. Yuan, "Joint pilot and payload power control for uplink MIMO-NOMA with MRC-SIC receivers," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 692–695, Apr. 2018.

[12] L. Xiang, D. W. K. Ng, R. Schober, and V. W. S. Wong, "Secure video streaming in heterogeneous small cell networks with untrusted cache helpers," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2645–2661, Apr. 2018.

[13] S. He, Y. Wu, D. W. K. Ng, and Y. Huang, "Joint optimization of analog beam and user scheduling for millimeter wave communications," *IEEE Commun. Lett.*, vol. 21, no. 12, pp. 2638–2641, Dec. 2017.

[14] N. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. D. Renzo, "Safeguarding 5G wireless communication networks using physical layer security," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 20–27, Apr. 2015.

[15] L. Dai, Z. Wang, and Z. Yang, "Spectrally efficient time-frequency training OFDM for mobile large-scale MIMO systems," *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 251–263, Feb. 2013.

[16] J. Zhu, R. Schober, and V. K. Bhargava, "Secure transmission in multicell massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4766–4781, Sep. 2014.

[17] J. A. Zhang, X. Huang, V. Dyadyuk, and Y. J. Guo, "Massive hybrid antenna array for millimeter-wave cellular communications," *IEEE Wireless Commun.*, vol. 22, no. 1, pp. 79–87, Feb. 2015.

[18] V. W. S. Wong, R. Schober, D. W. K. Ng, and L.-C. Wang, *Key Technologies for 5G Wireless Systems.* Cambridge Univ. Press, 2017.

[19] S. Akbar, Y. Deng, A. Nallanathan, M. Elkashlan, and G. K. Karagiannidi, "Massive multiuser MIMO in heterogeneous cellular networks with full duplex small cells," *IEEE Trans. Commun.*, vol. PP, no. 99, pp. 1–1, 2017.

[20] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Select. Topics in Signal Process.*, vol. 10, no. 3, pp. 501–513, Apr. 2016.

[21] F. Rusek, D. Persson, B. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO : Opportunities and challenges with very large arrays," *IEEE Trans. Signal Process.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[22] J. Andrews, S. Buzzi, C. Wan, S. Hanly, A. L. A. Soong, and J. Zhang, "What will 5G be?" *IEEE J. Select. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[23] M. Zhao, Z. Shi, and M. C. Reed, "Iterative turbo channel estimation for OFDM system over rapid dispersive fading channel," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3174–3184, Aug. 2008.

[24] H. Lin, F. Gao, S. Jin, and G. Y. Li, "A new view of multi-user hybrid massive MIMO: Non-orthogonal angle division multiple access," *IEEE J. Select. Areas Commun.*, vol. 35, no. 10, pp. 2268–2280, Oct. 2017.

[25] C. Lin, G. Y. Li, and L. Wang, "Subarray-based coordinated beamforming training for mmWave and sub-THz communications," *IEEE J. Select. Areas Commun.*, vol. 35, no. 9, pp. 2115–2126, Sep. 2017.

[26] G. Wang, F. Gao, W. Chen, and C. Tellambura, "Channel estimation and training design for two-way relay networks in time-selective fading environments," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2681–2691, Aug. 2011.

[27] 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.

[28] 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.

[29] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3029–3056, Sep. 2015.

[30] W. Roh, J. Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 106–113, Feb. 2014.

[31] G. C. Ferrante, G. Geraci, T. Q. S. Quek, and M. Z. Win, "Group-blind detection with very large antenna arrays in the presence of pilot contamination," in *Proc. IEEE Intern. Conf. on Acoustics, Speech and Signal Process. (ICASSP)*, 2016.

[32] H. Xie, F. Gao, S. Zhang, and S. Jin, "A unified transmission strategy for TDD/FDD massive MIMO systems with spatial basis expansion model," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3170–3184, Apr. 2017.

[33] T. E. Bogale and L. B. Le, "Massive MIMO and millimeter wave for 5G wireless HetNet: Potentials and challenges," *IEEE Veh. Technol. Mag.*, vol. 11, no. 1, pp. 64–75, Mar. 2016.

[34] Y. Deng, L. Wang, K. K. Wong, A. Nallanathan, M. Elkashlan, and S. Lambotharan, "Safeguarding massive MIMO aided hetnets using physical layer

security," in *Intern. Conf. on Wireless Commun. Signal Process. (WCSP)*, Oct. 2015, pp. 1–5.

[35] T. Yang, Q. T. Sun, J. A. Zhang, and J. Yuan, "A linear network coding approach for uplink distributed MIMO systems: Protocol and outage behavior," *IEEE J. Select. Areas Commun.*, vol. 33, no. 2, pp. 250–263, Feb. 2015.

[36] L. Dai, X. Gao, X. Su, S. Han, C. L. I, and Z. Wang, "Low-complexity soft-output signal detection based on Gauss-Seidel method for uplink multiuser large-scale MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 64, no. 10, pp. 4839–4845, Oct. 2015.

[37] 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.

[38] J. Zhang, L. Dai, X. Zhang, E. Björnson, and Z. Wang, "Achievable rate of Rician large-scale MIMO channels with transceiver hardware impairments," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8800–8806, Oct. 2016.

[39] M. Elkashlan, T. Q. Duong, and H. H. Chen, "Millimeter-wave communications for 5G: Fundamentals: Part I [Guest Editorial]," *IEEE Communications Magazine*, vol. 52, no. 9, pp. 52–54, Sep. 2014.

[40] A. Rahimian and F. Mehran, "RF link budget analysis in urban propagation microcell environment for mobile radio communication systems link planning," in *Proc. Intern. Conf. on Wireless Commun. and Signal Process. (WCSP)*, 2011, pp. 1–5.

[41] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, "Millimeter-wave massive MIMO: the next wireless revolution?" *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 56–62, Sep. 2014.

[42] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with large numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3292–3304, Sep. 2012.

[43] J. Hoydis, K. Hosseini, S. T. Brink, and M. Debbah, "Making smart use of excess antennas: Massive MIMO, small cells, and TDD," *Bell Labs Technical Journal*, vol. 18, no. 2, pp. 5–21, Sep. 2013.

[44] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: how many antennas do we need?" *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

[45] R. R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *IEEE J. Select. Topics in Signal Process.*, vol. 8, no. 5, pp. 773–786, Oct. 2014.

[46] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 172–179, Mar. 2013.

[47] S. Mumtaz, J. Rodriguez, and L. Dai, *mmWave Massive MIMO: A Paradigm for 5G.* Elsevier Inc., 2017.

[48] T. E. Bogale, L. B. Le, A. Haghighat, and L. Vandendorpe, "On the number of RF chains and phase shifters, and scheduling design with hybrid analog-digital beamforming," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3311–3326, May 2016.

[49] 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.

[50] Z. Xiang, M. Tao, and X. Wang, "Massive MIMO multicasting in noncooperative cellular networks," *IEEE J. Select. Areas Commun.*, vol. 32, no. 6, pp. 1180–1193, Jun. 2014.

[51] S. Wu, C. X. Wang, H. Haas, e. H. M. Aggoune, M. M. Alwakeel, and B. Ai, "A non-stationary wideband channel model for massive MIMO communication systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1434–1446, Mar. 2015.

[52] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[53] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.

[54] A. Ozgur, O. Leveque, and D. Tse, "Spatial degrees of freedom of large distributed MIMO systems and wireless Ad Hoc networks," *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 202–214, Feb. 2013.

[55] 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.

[56] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems," *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 192–201, Feb. 2013.

[57] J. Choi, D. J. Love, and P. Bidigare, "Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory," *IEEE J. Select. Areas Commun.*, vol. 8, no. 5, pp. 802–814, Oct. 2014.

[58] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[59] S. Noh, M. D. Zoltowski, Y. Sung, and D. J. Love, "Pilot beam pattern design for channel estimation in massive MIMO systems," *IEEE J. Select. Topics in Signal Process.*, vol. 8, no. 5, pp. 787–801, Oct. 2014.

[60] J. Ma and L. Ping, "Data-aided channel estimation in large antenna systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3111–3124, Jun. 2014.

[61] D. Fan, F. Gao, G. Wang, Z. Zhong, and A. Nallanathan, "Angle domain signal processing-aided channel estimation for indoor 60-ghz TDD/FDD massive MIMO systems," *IEEE J. Select. Areas Commun.*, vol. 35, no. 9, pp. 1948–1961, Sep. 2017.

[62] K. Ghavami and M. Naraghi-Pour, "Blind channel estimation and symbol detection for multi-cell massive MIMO systems by expectation propagation," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 943–954, Feb. 2018.

[63] H. Yin, L. Cottatellucci, D. Gesbert, R. R. Mller, and G. He, "Robust pilot decontamination based on joint angle and power domain discrimination," *IEEE Trans. Signal Process.*, vol. 64, no. 11, pp. 2990–3003, Jun. 2016.

[64] P. Švač, F. Meyer, E. Riegler, and F. Hlawatsch, "Soft-heuristic detectors for large MIMO systems," *IEEE Trans. Signal Process.*, vol. 61, no. 18, pp. 4573–4586, Sep. 2013.

[65] K. V. Vardhan, S. K. Mohammed, A. Chockalingam, and B. S. Rajan, "A low-complexity detector for large MIMO systems and multicarrier CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 26, no. 3, pp. 473–485, Apr. 2008.

[66] S. Wu, L. Kuang, Z. Ni, J. Lu, D. Huang, and Q. Guo, "Low-complexity iterative detection for large-scale multiuser MIMO-OFDM systems using approximate message passing," *IEEE J. Select. Topics in Signal Process.*, vol. 8, no. 5, pp. 902–915, Oct. 2014.

[67] M. Wu, B. Yin, G. Wang, C. Dick, J. R. Cavallaro, and C. Studer, "Large-scale MIMO detection for 3GPP LTE: Algorithms and FPGA implementations," *IEEE J. Select. Topics in Signal Process.*, vol. 8, no. 5, pp. 916–929, Oct. 2014.

[68] O. N. Alrabadi, E. Tsakalaki, H. Huang, and G. F. Pedersen, "Beamforming via large and dense antenna arrays above a clutter," *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 314–325, Feb. 2013.

[69] M. Hong, R. Sun, H. Baligh, and Z. Q. Luo, "Joint base station clustering and beamformer design for partial coordinated transmission in heterogeneous networks," *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 226–240, Feb. 2013.

[70] A. Alkhateeb, G. Leus, and R. W. Heath, "Multi-layer precoding: A potential solution for full-dimensional massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5810–5824, Sept 2017.

[71] T. E. Bogale and L. B. Le, "Beamforming for multiuser massive MIMO systems: Digital versus hybrid analog-digital," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2014, pp. 4066–4071.

[72] A. Liu and V. Lau, "Phase only RF precoding for massive MIMO systems with limited RF chains," *IEEE Trans. Signal Process.*, vol. 62, no. 17, pp. 4505–4515, Sep. 2014.

[73] J. Choi, "Multiuser precoding with limited cooperation for large-scale MIMO multicell downlink," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1295–1308, Mar. 2015.

[74] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Massive MIMO performance evaluation based on measured propagation data," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3899–3911, Jul. 2015.

[75] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, "Massive MIMO in real propagation environments: Do all antennas contribute equally?" *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 3917–3928, Nov. 2015.

[76] D. W. K. Ng, E. S. Lo, and R. Schober, "Dynamic resource allocation in MIMO-OFDMA systems with full-duplex and hybrid relaying," *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1291–1304, May 2012.

[77] Y. Huang, C. W. Tan, and B. D. Rao, "Joint beamforming and power control in coordinated multicell: Max-min duality, effective network and large system transition," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2730–2742, Jun. 2013.

[78] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, Jun. 2015.

[79] E. Björnson, L. Sanguinetti, and M. Kountouris, "Deploying dense networks for maximal energy efficiency: Small cells meet massive MIMO," *IEEE J. Select. Areas Commun.*, vol. 34, no. 4, pp. 832–847, Apr. 2016.

[80] 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.

[81] A. Pitarokoilis, S. K. Mohammed, and E. G. Larsson, "Effect of oscillator phase noise on uplink performance of large MU-MIMO systems," in *2012 50th Annual Allerton Conf. on Commun., Control, and Computing (Allerton)*, Oct. 2012, pp. 1190–1197.

[82] J. Choi, "Downlink multiuser beamforming with compensation of channel reciprocity from rf impairments," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2158–2169, Jun. 2015.

[83] J. Zhang, L. Dai, X. Zhang, E. Björnson, and Z. Wang, "Achievable rate of Rician large-scale MIMO channels with transceiver hardware impairments," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8800–8806, Oct. 2016.

[84] Y. Wu, D. W. K. Ng, C. K. Wen, R. Schober, and A. Lozano, "Low-complexity MIMO precoding for finite-alphabet signals," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4571–4584, Jul. 2017.

[85] J. Zhu, D. W. K. Ng, N. Wang, R. Schober, and V. K. Bhargava, "Analysis and design of secure massive MIMO systems in the presence of hardware impairments," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 2001–2016, Mar. 2017.

[86] Y. Wu, R. Schober, D. W. K. Ng, C. Xiao, and G. Caire, "Secure massive MIMO transmission with an active eavesdropper," *IEEE Trans. Inf. Theory*, vol. 62, no. 7, pp. 3880–3900, Jul. 2016.

[87] J. Zhao, F. Gao, W. Jia, S. Zhang, S. Jin, and H. Lin, "Angle domain hybrid precoding and channel tracking for millimeter wave massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 10, pp. 6868–6880, Oct. 2017.

[88] S. Zarei, W. H. Gerstacker, J. Aulin, and R. Schober, "Multi-cell massive MIMO systems with hardware impairments: Uplink-downlink duality and

downlink precoding," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5115–5130, Aug. 2017.

[89] A. K. Saxena, I. Fijalkow, and A. L. Swindlehurst, "Analysis of one-bit quantized precoding for the multiuser massive MIMO downlink," *IEEE Trans. Signal Process.*, vol. 65, no. 17, pp. 4624–4634, Sep. 2017.

[90] J. Zhang, L. Dai, Z. He, S. Jin, and X. Li, "Performance analysis of mixed-ADC massive MIMO systems over Rician fading channels," *IEEE J. Select. Areas Commun.*, vol. 35, no. 6, pp. 1327–1338, Jun. 2017.

[91] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "The multicell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2350–2361, Jun. 2013.

[92] H. Huh, S. H. Moon, Y. T. Kim, I. Lee, and G. Caire, "Multi-cell MIMO downlink with cell cooperation and fair scheduling: A large-system limit analysis," *IEEE Trans. Inf. Theory*, vol. 57, no. 12, pp. 7771–7786, Dec. 2011.

[93] H. Yin, D. Gesbert, and L. Cottatellucci, "Dealing with interference in distributed large-scale MIMO systems: A statistical approach," *IEEE J. Select. Topics in Signal Process.*, vol. 8, no. 5, pp. 942–953, Oct. 2014.

[94] C.-N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 637–650, Mar. 2002.

[95] D. Aktas, M. N. Bacha, J. S. Evans, and S. V. Hanly, "Scaling results on the sum capacity of cellular networks with MIMO links," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3264–3274, Jul. 2006.

[96] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large system analysis of linear precoding in correlated MISO broadcast channels under limited feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4509–4537, Jul. 2012.

[97] W. Yang, G. Durisi, and E. Riegler, "On the capacity of large-MIMO block-fading channels," *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 117–132, Feb. 2013.

[98] T. Yoneyama and S. Nishida, "Nonradiative dielectric waveguide for millimeter-wave integrated circuits," *IEEE Trans. on Microwave Theory and Techniques*, vol. 29, no. 11, pp. 1188–1192, Nov. 1981.

[99] D. Pozar, "Considerations for millimeter wave printed antennas," *IEEE Trans. on Antennas and Propagation*, vol. 31, no. 5, pp. 740–747, Sep. 1983.

[100] T. Rappaport, S. Shu, R. Mayzus, Z. Hang, Y. Azar, K. Wang, G. Wong, J. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: it will work," *IEEE Access Magazine*, vol. 1, pp. 335–349, May 2013.

[101] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Select. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.

[102] S. Hur, S. Baek, B. Kim, Y. Chang, A. F. Molisch, T. S. Rappaport, K. Haneda, and J. Park, "Proposal on millimeter-wave channel modeling for 5G cellular system," *IEEE J. Select. Topics in Signal Process.*, vol. 10, no. 3, pp. 454–469, Apr. 2016.

[103] M. Xiao, S. Mumtaz, Y. Huang, L. Dai, Y. Li, M. Matthaiou, G. K. Karagiannidis, E. Bjrnson, K. Yang, I. Chih-Lin, and A. Ghosh, "Millimeter wave communications for future mobile networks (guest editorial), part i," *IEEE J. Select. Areas Commun.*, vol. 35, no. 7, pp. 1425–1431, Jul. 2017.

[104] K. Haneda, L. Tian, and et. al, "5G 3GPP-like channel models for outdoor urban microcellular and macrocellular environments," 2016. [Online]. Available: https://arxiv.org/abs/1602.07533

[105] E. Björnson, M. Kountouris, and M. Debbah, "Massive MIMO and small cells: Improving energy efficiency by optimal soft-cell coordination," in *ICT 2013*, May 2013, pp. 1–5.

[106] G. R. MacCartney and T. S. Rappaport, "A flexible millimeter-wave channel sounder with absolute timing," *IEEE J. Select. Areas Commun.*, vol. 35, no. 6, pp. 1402–1418, Jun. 2017.

[107] V. Raghavan, A.Partyka, A. Sampath, S. Subramanian, and et al, "Millimeter wave MIMO prototype: Measurements and experimental results," 2017. [Online]. Available: https://arxiv.org/abs/1710.09449

[108] M. K. Samimi and T. S. Rappaport, "3-D millimeter-wave statistical channel model for 5G wireless system design," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 7, pp. 2207–2225, Jul. 2016.

[109] J. Ko, Y. J. Cho, S. Hur, T. Kim, J. Park, A. F. Molisch, K. Haneda, M. Peter, D. J. Park, and D. H. Cho, "Millimeter-wave channel measurements and analysis for statistical spatial channel model in in-building and urban environments at 28 GHz," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5853–5868, Sep. 2017.

[110] H. C. Nguyen, G. R. MacCartney, T. Thomas, T. S. Rappaport, B. Vejlgaard, and P. Mogensen, "Evaluation of empirical ray-tracing model for an urban outdoor scenario at 73 GHz E-band," in *2014 IEEE 80th Veh. Techn. Conf. (VTC2014-Fall)*, Sep. 2014, pp. 1–6.

[111] H. L. V. Trees, *Optimum array processing: Part IV of detection, estimation, and modulation theory.* John Wiley & Sons, Inc., 2002.

[112] R. W. Heath, N. G. Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Select. Topics in Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[113] J. G. Andrews, T. Bai, M. N. Kulkarni, A. Alkhateeb, A. K. Gupta, and R. W. Heath, "Modeling and analyzing millimeter wave cellular systems," *IEEE Trans. Commun.*, vol. 65, no. 1, pp. 403–430, Jan. 2017.

[114] R. Mendez-Rial, C. Rusu, A. Alkhateeb, N. Gonzalez-Prelcic, and R. W. Heath, "Channel estimation and hybrid combining for mmWave: phase shifters or switches?" in *Proc. IEEE Inf. Theory and Applications Workshop (ITA)*, 2015, pp. 90–97.

[115] W. Ni and X. Dong, "Hybrid block diagonalization for massive multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 201–211, Jan. 2016.

[116] S. Han, C. l. I, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 186–194, Jan. 2015.

[117] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[118] X. Gao, L. Dai, S. Han, C. L. I, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays," *IEEE J. Select. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, Apr. 2016.

[119] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave (mmWave) communications for 5G: Opportunities and challenges," 2015, to appear.

[120] S. Park and R. W. Heath, "Frequency selective hybrid precoding in millimeter wave OFDMA systems," in *2015 IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–6.

[121] C. U. Bas, R. Wang, S. Sangodoyin, and et. al, "28 GHz microcell measurement campaign for residential environment," 2017. [Online]. Available: https://arxiv.org/abs/1711.00170

[122] G. Miao and G. Song, *Energy and spectrum efficient wireless network design.* Cambridge Univ. Press, 2014.

[123] V. Petrov, M. Komarov, D. Moltchanov, J. M. Jornet, and Y. Koucheryavy, "Interference and SINR in millimeter wave and terahertz communication systems with blocking and directional antennas," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1791–1808, Mar. 2017.

[124] C. Liu, M. Li, S. V. Hanly, I. B. Collings, and P. Whiting, "Millimeter wave beam alignment: Large deviations analysis and design insights," *IEEE J. Select. Areas Commun.*, vol. 35, no. 7, pp. 1619–1631, Jul. 2017.

[125] Z. Xiao, T. He, P. Xia, and X. G. Xia, "Hierarchical codebook design for beamforming training in millimeter-wave communication," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3380–3392, May. 2016.

[126] W. Shen, L. Dai, B. Shim, Z. Wang, and R. W. H. Jr, "Channel feedback based on aod-adaptive subspace codebook in fdd massive MIMO systems," 2017. [Online]. Available: https://arxiv.org/abs/1704.00658

[127] S. Noh, M. D. Zoltowski, and D. J. Love, "Multi-resolution codebook and adaptive beamforming sequence design for millimeter wave beam alignment," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5689–5701, Sep. 2017.

[128] P. Wang, Y. Li, Y. Peng, S. C. Liew, and B. Vucetic, "Non-uniform linear antenna array design and optimization for millimeter-wave communications," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7343–7356, Nov. 2016.

[129] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar. 2006.

[130] S. He, J. Wang, Y. Huang, B. Ottersten, and W. Hong, "Codebook-based hybrid precoding for millimeter wave multiuser systems," *IEEE Trans. Signal Process.*, vol. 65, no. 20, pp. 5289–5304, Oct. 2017.

[131] X. Gao, L. Dai, S. Han, C. L. I, and X. Wang, "Reliable beamspace channel estimation for millimeter-wave massive MIMO systems with lens antenna array," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6010–6021, Sep. 2017.

[132] X. Gao, L. Dai, C. Yuen, and Z. Wang, "Turbo-like beamforming based on Tabu search algorithm for millimeter-wave massive MIMO systems," vol. 65, no. 7, pp. 5731–5737, Jul. 2016.

[133] J. Wang, "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," *IEEE J. Select. Areas Commun.*, vol. 27, no. 8, pp. 1390–1399, Oct. 2009.

[134] A. Adhikary, J. Nam, J. Ahn, and G. Caire, "Joint spatial division and multiplexing – the large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, Oct. 2013.

[135] A. Alkhateeb, O. E. Ayach, G. Leuz, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics in Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[136] P. Raviteja, Y. Hong, and E. Viterbo, "Millimeter wave analog beamforming with low resolution phase shifters for multiuser uplink," vol. PP, no. 99, pp. 1–1, 2017.

[137] S. Yang and L. Hanzo, "Fifty years of MIMO detection: The road to large-scale MIMOs," *IEEE Communications Surveys Tutorials*, vol. 17, no. 4, pp. 1941–1988, Fourthquarter 2015.

[138] L. Sanguinetti, A. Kammoun, and M. Debbah, "Asymptotic analysis of multicell massive MIMO over Rician fading channels," in *Proc. IEEE Intern. Conf. on Acoustics, Speech and Signal Process. (ICASSP)*, Mar. 2017, pp. 3539–3543.

[139] L. Zhao, T. Yang, G. Geraci, and J. Yuan, "Downlink multiuser massive MIMO in Rician channels under pilot contamination," in *Proc. IEEE Intern. Commun. Conf. (ICC)*, May 2016, pp. 1–6.

[140] N. Akbar, N. Yang, P. Sadeghi, and R. A. Kennedy, "Multi-cell multiuser massive MIMO networks: User capacity analysis and pilot design," *IEEE Trans. Commun.*, vol. 64, no. 12, pp. 5064–5077, Dec. 2016.

[141] W. A. W. M. Mahyiddin, P. A. Martin, and P. J. Smith, "Performance of synchronized and unsynchronized pilots in finite massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 6763–6776, Dec. 2015.

[142] T. E. Bogale, L. B. Le, X. Wang, and L. Vandendorpe, "Pilot contamination mitigation for wideband massive MMO: Number of cells vs multipath," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–6.

[143] A. Farhang, A. Aminjavaheri, N. Marchetti, L. E. Doyle, and B. Farhang-Boroujeny, "Pilot decontamination in cmt-based massive MI-MO networks," in *2014 11th Intern. Symposium on Wireless Commun. Systems (ISWCS)*, Aug. 2014, pp. 589–593.

[144] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO has unlimited capacity," *IEEE Trans. Wireless Commun.*, vol. PP, no. 99, pp. 1–1, 2017.

[145] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, "Femtocells: Past, present, and future," *IEEE J. Select. Areas Commun.*, vol. 30, no. 3, pp. 497–508, Apr. 2012.

[146] S. Hur, Y. J. Cho, J. Lee, N.-G. Kang, J. Park, and H. Benn, "Synchronous channel sounder using horn antenna and indoor measurements on 28 GHz," in *2014 IEEE Intern. Black Sea Conf. on Commun. and Networking (Black-SeaCom)*, May 2014, pp. 83–87.

[147] M. A. Richards, *Fundamentals of radar signal processing, second edition.* McGraw-Hill Education, 2014.

[148] J. He, T. Kim, H. Ghauch, K. Liu, and G. Wang, "Millimeter wave MI-MO channel tracking systems," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2014, pp. 416–421.

[149] D. Araujo, A. de Almeida, J. Axnas, and J. Mota, "Channel estimation for millimeter-wave very-large MIMO systems," in *Proc. Europ. Signal Process. Conf. (EUSIPCO)*, Sep. 2014, pp. 81–85.

[150] W. Shen, L. Dai, B. Shim, S. Mumtaz, and Z. Wang, "Joint CSIT acquisition based on low-rank matrix completion for FDD massive MIMO systems," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2178–2181, Dec. 2015.

[151] J. Flordelis, F. Rusek, F. Tufvesson, E. G. Larsson, and O. Edfors, "Massive MIMO performance - TDD versus FDD: What do measurements say?" *IEEE Trans. Wireless Commun.*, vol. PP, no. 99, pp. 1–1, 2018.

[152] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4353–4368, Aug. 2015.

[153] D. Ying, F. W. Vook, T. A. Thomas, and D. J. Love, "Hybrid structure in massive MIMO: Achieving large sum rate with fewer RF chains," in *Proc. IEEE Intern. Commun. Conf. (ICC)*, Jun. 2015, pp. 2344–2349.

[154] J. Mo and R. W. Heath, "Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5498–5512, Oct. 2015.

[155] Z. Al-Daher, L. P. Ivrissimtzis, and A. Hammoudeh, "Electromagnetic modeling of high-frequency links with high-resolution terrain data," *IEEE Antennas and Wireless Propagation Lett.*, vol. 11, pp. 1269–1272, Oct. 2012.

[156] Z. Muhi-Eldee, L. P. Ivrissimtzis, and M. Al-Nuaimi, "Modeling and measurements of millimeter wave propagation in urban environments," *IET Microw. Antennas Propagation*, vol. 4, no. 9, pp. 1300–1309, Sep. 2010.

[157] A. Alkhateeb, G. Leusz, and R. W. Heath, "Compressed sensing based multi-user millimeter wave systems: How many measurements are needed?" in *Proc. IEEE Intern. Conf. on Acoustics, Speech and Signal Process. (ICASSP)*, Apr. 2015, pp. 2909–2913.

[158] J. Vieira, S. Malkowsky, K. Nieman, Z. Miers, N. Kundargi, L. Liang, I. Wong, V. Owall, O. Edfors, and F. Tufvesson, "A flexible 100-antenna testbed for massive MIMO," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, 2014, pp. 287–293.

[159] D. W. Yue, Y. Zhang, and Y. Jia, "Beamforming based on specular component for massive MIMO systems in Ricean fading," *IEEE Commun. Lett.*, vol. 4, no. 2, pp. 197–200, Apr. 2015.

[160] S. L. H. Nguyen, K. Haneda, J. Jarvelainen, A. Karttunen, and J. Putkonen, "On the mutual orthogonality of millimeter-wave massive MIMO channels," in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, 2015, pp. 1–5.

[161] N. Hung and H. L. V. Trees, "Comparison of performance bounds for DoA estimation," in *Proc. IEEE Seventh SP Workshop on Statistical Signal and Array Process.*, 1994, pp. 313–316.

[162] "Keysight technologies E8257D PSG microwave analog signal generator," Tech. Rep., 2016. [Online]. Available: http://literature.cdn.keysight.com/litweb/pdf/5989-0698EN.pdf?id=484534

[163] Nokia Siemens Networks, "Long term HSPA evolution: mobile broadband evolution beyond 3GPP release 10," 2010. [Online]. Available: http://lteworld.org/whitepaper/...-3gpp-release-10

[164] M. Giordani, M. Polese, A. Roy, D. Castor, and M. Zorzi, "A tutorial on beam management for 3GPP NR at mmWave frequencies," 2018. [Online]. Available: https://arxiv.org/abs/1804.01908

[165] S. Buzzi and C. D'Andrea, "Doubly massive mmWave MIMO systems: Using very large antenna arrays at both transmitter and receiver," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2016.

[166] A. Khansefid and H. Minn, "On channel estimation for massive MIMO with pilot contamination," *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1660–1663, Sep. 2015.

[167] H. Falconet, L. Sanguinetti, A. Kammoun, and M. Debbah, "Asymptotic analysis of downlink MISO systems over Rician fading channels," in *Proc. IEEE Intern. Conf. on Acoustics, Speech and Signal Process. (ICASSP)*, Mar. 2016, pp. 3926–3930.

[168] L. Dai, X. Gao, S. Han, I. Chih-Lin, and X. Wang, "Beamspace channel estimation for millimeter-wave massive MIMO systems with lens antenna array," in *2016 IEEE/CIC Intern. Conf. on Commun. in China (ICCC)*, Jul. 2016, pp. 1–6.

[169] T. Bai and R. W. Heath, "Coverage and rate analysis for millimeter-wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 1100–1114, Feb. 2015.

[170] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge University Press, 2005.

[171] X. Zhang, A. F. Molisch, and S.-Y. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," vol. 53, no. 11, pp. 4091–4103, Nov. 2005.

[172] E. Zhang and C. Huang, "On achieving optimal rate of digital precoder by RF-baseband codesign for MIMO systems," in *Proc. IEEE Veh. Techn. Conf. (VTC)*, Sep. 2014, pp. 1–5.

[173] A. A. M. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE J. Select. Areas Commun.*, vol. 5, no. 2, pp. 128–137, Feb. 1987.

[174] V. Raghavan, J. Cezanne, S. Subramanian, A. Sampath, and O. Koymen, "Beamforming tradeoffs for initial UE discovery in millimeter-wave MIMO systems," *IEEE J. Select. Topics in Signal Process.*, vol. 10, no. 3, pp. 543–559, Apr. 2016.

[175] V. Raghavan, S. Subramanian, J. Cezanne, A. Sampath, O. H. Koymen, and J. Li, "Single-user versus multi-user precoding for millimeter wave MIMO systems," *IEEE J. Select. Areas Commun.*, vol. 35, no. 6, pp. 1387–1401, Jun. 2017.

[176] K. B. Petersen and M. S. Pedersen, *The Matrix Cookbook*, 2012. [Online]. Available: http://matrixcookbook.com

[177] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: Wiley, 1991.

[178] X. Yu, J. Zhang, M. Haenggi, and K. B. Letaief, "Coverage analysis for millimeter wave networks: The impact of directional antenna arrays," *IEEE J. Select. Areas Commun.*, vol. 35, no. 7, pp. 1498–1512, Jul. 2017.

[179] H. H. Sohrab, *Basic Real Analysis.* Springer, 2014. [Online]. Available: https://books.google.com.au/books?id=QnpqBQAAQBAJ