Analysis and Design of Millimeter Wave Cellular Networks

Simin Xu

February 2021

A thesis submitted for the degree of
Master of Philosophy of
The Australian National University

School of Engineering
College of Engineering and Computer Science
The Australian National University

© Copyright by Simin Xu 2021

Declaration

The contents of this thesis are the results of original research and have not been submitted for a higher degree to any other university or institution.

The work in this thesis has been published or has been accepted for publication as conference papers.

The research work presented in this thesis has been performed jointly with A/Prof. Nan Yang (The Australian National University), Dr. Shihao Yan (Macquarie University), Dr. Biao He (MediaTek USA Inc.), and Prof. Hamid Jafarkhani (University of California, Irvine). The substantial majority of this work was my own.

Simin Xu
School of Engineering,
College of Engineering and Computer Science,
The Australian National University,
Canberra, ACT 2601,
AUSTRALIA

Acknowledgments

This work would not have been possible without the help of a number of people.

- First and foremost, I would like to express my deepest gratitude to my primary supervisor, A/Prof. Nan Yang, for his continuous support and encouragement throughout my entire MPhil study. He has always made himself available for discussion, provided me a lot of sound advices on research problems and paper writing. His rigor and conciseness shaped my research principles. It is my pleasure and honor to be his student.

- Second, I would like to thank Dr. Shihao Yan (Macquarie University), Dr. Biao He (MediaTek USA Inc.), and Prof. Hamid Jafarkhani (University of California, Irvine) for sharing their knowledge, which inspired much of the work in this thesis.

- Third, I would like to thank Prof. Thushara Abhayapala, Assoc. Prof. Salman Durrani, and Assoc. Prof. Xiangyun Zhou for being on my supervisory panel and providing me invaluable suggestions to strengthen my research skills.

- Fourth, I would like to thank The Australian National University for providing such a friendly and supportive environment. Special thanks to Prof. Parastoo Sadeghi, Prof. Rodney A. Kennedy, Dr. Abbas Koohian, Dr. Alice Bates, Chunhui Li, Fei Ma, Dr. Hanchi Chen, Hang Yuan, Huiyuan Sun, Jianwei Hu, Dr. Jihui Zhang, Dr. Jing Guo, Khurram Shahzad, Dr. Mingchao Yu, Muhammed Akram Mohamed Shafie, Nicole Sawyer, Nilupuli Senadhira, Dr. Noman Akbar, Rabbia Saleem, Sahar Idrees, Dr. Shama Islam, Sheeraz Alvi, Usama Elahi, Wageesha Mananperi, Dr. Wanchun Liu, Xiaofang Sun, Xiang Wu, Dr. Xiaohui Zhou, Xiaolun Jia, Xinyu Huang, Yanyan Wang, Dr. Yifei Huang, Yiran Wang, Dr. Yirui Cong, Yonggang Hu, Youhong Feng, Yuting Fang, Zhifeng Tang, Dr. Zhuo Sun, Dr. Zohair Abu Shaban. I would also like to acknowledge scholarship and grant supports from The Australian National University.

- Last but not least, I would like to thank my mother Song Ding, my father Xinling Xu, my partner Simon Bergens, and my older brother Yan Xu for their love and support. Without them, I would not be able to achieve what I have achieved now.

Abstract

Millimeter wave (mmWave) communications has been widely acknowledged as an attractive strategy for the rapidly growing data rate requirements of cellular user equipments (UEs), due to the vast amounts of available frequencies at the mmWave band. However, the unique propagation characteristics of mmWave, including 1) high path loss, 2) extreme sensitivity to blockage, and 3) rapid channel fluctuations, bring serious challenges to the deployment of mmWave cellular networks. Against this background, this thesis focuses on the analysis and design of mmWave cellular networks.

In Chapter 1, the motivation of the studies presented in this thesis is described. Moreover, a literature review of several key research topics is presented, including mmWave channel models, mmWave-enabled heterogeneous networks (HetNets), mmWave precoding, mmWave-based non-orthogonal multiple access (NOMA), and mmWave prototypes. Furthermore, an overview of this thesis is provided.

In Chapter 2, a two-tier mmWave cellular HetNet is considered. As pointed out by the 3rd Generation Partnership Project (3GPP), a major issue in the HetNet is that high-power BSs are often heavily loaded, while low-power BSs are always lightly loaded and therefore not fully exploited. This load disparity inevitably leads to suboptimal resource allocation across the network, where a large number of UEs may be associated with one high-power BS but experience poor date rates. To increase the load of low-power BSs and strike a load balance between high-power BSs and low-power BSs, an association bias factor needs to be added to increase the possibility that UEs are associated with low-power BSs. In this chapter, we conduct novel analysis to assess the impact of the bias factor on the rate coverage performance of the considered network. In order to obtain tractable analytical results on the rate coverage probability, we model the considered network using a stochastic geometry based approach. We first analyze the loads of high-power BSs and low-power BSs, based on which we derive a new expression for the rate coverage probability of the network. Through numerical results, we demonstrate the correctness of our analysis. In addition, we thoroughly examine the impact of load balancing and various network parameters on the rate coverage probability in various scenarios, offering valuable guidelines on the design of practical mmWave HetNets.

In Chapter 3, a relay assisted mmWave cellular network is considered. In this network, the BS adopts either the direct mode to transmit to the destination UE, or the relay mode if the direct mode fails, where the BS transmits to the relay and then the relay transmits to the destination UE. To address the drastic rotational movements of destination UEs in practice, we propose to adopt selection combining at destination UEs. Similar to Chapter 2, in order

to obtain tractable analytical results on the system-level coverage probability, we model the system using a stochastic geometry based approach. New expression is derived for the signal-to-interference-plus-noise ratio (SINR) coverage probability of the network. Using numerical results, we first demonstrate the accuracy of our new expression. Then we show that ignoring spatial correlation, which has been commonly adopted in the literature, leads to severe overestimation of the SINR coverage probability. Furthermore, we show that introducing relays into a mmWave cellular network vastly improves the coverage performance. In addition, we show that the optimal BS density maximizing the SINR coverage probability can be determined by using our analysis.

In Chapter 4, a summary of the conclusions drawn from this thesis is presented. Moreover, a number of future research directions are identified, including integrated mmWave/sub-6 GHz cellular networks, the mobility support in mmWave cellular networks, ultra-low latency mmWave cellular networks, and the transport layer design of mmWave cellular networks.

List of Publications

The work in this thesis has been published or has been accepted for publication as conference papers. These papers are:

C1. **S. Xu**, N. Yang, and S. Yan, ”Impact of load balancing on rate coverage performance in millimeter wave cellular heterogeneous networks,” in *Proc. IEEE ICC 2018 Workshop on Evolutional Technologies & Ecosystems for 5G Phase II*, Kansas City, MO, May 2018, pp. 1–6.

C2. **S. Xu**, N. Yang, B. He, and H. Jafarkhaniand, ”Coverage analysis of relay assisted millimeter wave cellular networks with spatial correlation,” accepted for publication in *Proc. IEEE WCNC 2020*, Seoul, South Korea, Apr. 2020, pp. 1–8.

The papers C1 and C2 are *used* in the thesis.

Acronyms

AF	amplified-and-forward
AoA	azimuth of arrival
AoD	azimuth of departure
AWGN	additive white Gaussian noise
BS	base station
CCDF	complementary cumulative distribution function
CDF	cumulative distribution function
CDMA	code division multiple access
CRE	cell range expansion
D2D	device-to-device
DF	decode-and-forward
EoA	elevation of arrival
EoD	elevation of departure
FCC	Federal Communications Commission
FDMA	frequency division multiple access
GPS	Global Position System
HCPP	hard core point process
HetNets	heterogeneous networks
i.i.d.	independent and identically distributed
IoT	Internet of Things
ITU	International Telecommunication Union
LoS	line-of-sight

METIS	Mobile and wireless communications Enablers for the Twenty-twenty Information Society
MGF	moment generation function
MIMO	multiple input multiple output
MiWEBA	Millimeter-Wave Evolution for Backhaul and Access
mmWave	millimeter wave
NLoS	non-line-of-sight
NOMA	non-orthogonal multiple access
OFDM	orthogonal frequency division multiplexing
OFDMA	orthogonal frequency division multiple access
PAN	personal area network
PDF	probability density function
PPP	Poisson point process
QoS	quality of service
RBUR	resource block utility ratio
RF	radio frequency
RMS	root mean square
RRUR	radio resource utilization ratio
SAR	specific absorption rate
SIC	successive interference cancellation
SINR	signal-to-interference-plus-noise ratio
SIRCIM	Simualtion of Indoor Radio Channel Impulse response Model
SMRCIM	Simualtion of Mobile Radio Channel Impulse response Model
SV	Saleh-Valenzuela
TDMA	time division multiple access
UE	user equipment
w. p.	with probability

3GPP	3rd Generation Partnership Project

Notations

$\Pr(\cdot)$	probability
$\Pr(\cdot\|\cdot)$	conditional probability
$x \in X$	x is an element of X
$\min\{\cdot\}$	minimum value of a set
Φ	point process
$\mathbb{E}\left[\cdot\right]$	expectation
$\mathbb{E}_x\left[\cdot\right]$	expectation with respect to x
e^x	exponential function
$M_x\left(\cdot\right)$	MGF of x
$F_x\left(\cdot\right)$	CDF of x
$f_x\left(\cdot\right)$	PDF of x
$\triangleq$	defined as
$\prod$	product
$\sum$	summation

Contents

Declaration **iii**

Acknowledgments **v**

Abstract **vii**

List of Publications **ix**

Acronyms **xi**

Notations **xv**

1 Introduction **1**

- 1.1 Motivation 1
- 1.2 Literature Review 4
  - 1.2.1 mmWave Channel Models 4
    - 1.2.1.1 Sub-6 GHz Channel Models 5
    - 1.2.1.2 mmWave Channel Models 6
  - 1.2.2 mmWave-enabled Heterogeneous Networks 7
    - 1.2.2.1 Mobility 8
    - 1.2.2.2 Decoupled Uplink Downlink Cell Association 8
    - 1.2.2.3 Downlink Cell Association with Load Balancing 9
    - 1.2.2.4 Backhaul Bottleneck 9
    - 1.2.2.5 Next Generation mmWave-enabled HetNets 10
  - 1.2.3 mmWave Precoding 10
    - 1.2.3.1 Fully Digital Precoding 10
    - 1.2.3.2 Analog Precoding 12
    - 1.2.3.3 Hybrid Precoding 15
  - 1.2.4 mmWave-based Non-orthogonal Multiple Access 16
  - 1.2.5 mmWave Prototypes and Commercial mmWave Cellular Networks . . 17
    - 1.2.5.1 mmWave Prototypes 18
    - 1.2.5.2 Commercial mmWave Cellular Networks 18
- 1.3 Thesis Overview and Contributions 20

1.3.1 Coverage Analysis of mmWave HetNets with Load Balancing 20
1.3.2 Coverage Analysis of Relay Assisted mmWave Cellular Networks with Spatial Correlation 22

2 Coverage Analysis of mmWave Cellular HetNets with Load Balancing 25
2.1 mmWave Cellular Heterogeneous Network 25
2.1.1 Blockage Effect 26
2.1.2 Directional Gains 27
2.1.3 Small Scale Fading 28
2.2 Load and Rate Coverage Probability Analysis 29
2.2.1 Load Analysis 29
2.2.2 Rate Coverage Probability Analysis 31
2.3 Numerical Results and Discussions 40
2.4 Summary 46

3 Coverage Analysis of Relay Assisted mmWave Networks with Spatial Correlation 47
3.1 Relay Assisted mmWave Cellular Network 47
3.1.1 Locations of BSs and UEs 48
3.1.2 Directional Gains 49
3.1.3 Association Strategy 51
3.1.4 SINR Formulation 51
3.2 SINR Coverage Probability Analysis 54
3.3 Numerical Results and Discussions 58
3.4 Summary 64

4 Conclusions and Future Research Directions 65
4.1 Conclusions 65
4.2 Future Research Directions 66

A Appendix A 69
A.1 Proof of Theorem 3.1 69
A.2 Proof of Theorem 3.2 71
A.3 Proof of Theorem 3.3 72

Bibliography 75

List of Figures

1.1 Illustration of the fully digital precoding architecture, where N_s is the number of data streams and N_t is the number of antennas at the transmitter antenna array. 11

1.2 Illustration of the analog precoding architecture, where N_t is the number of antennas at the transmitter antenna array. 11

1.3 Illustration of the beam directions of the transmitter and the receiver, where the beamwidths of the transmitter and the receiver are 30° and 60°, respectively. 12

1.4 Illustration of the fully connected hybrid precoding architecture, where N_s is the number of data streams, N_{RF} is the number of RF chains, and N_t is the number of antennas at the transmitter antenna array. Each RF chain is connected with all transmitter antennas. There are $N_{RF} \times N_t$ phase shifters in the architecture. 15

1.5 Illustration of the partially connected hybrid precoding architecture, where N_s is the number of data streams, N_{RF} is the number of RF chains, and N_t is the number of antennas at the transmitter antenna array. Each RF chain is connected with $\frac{N_t}{N_{RF}}$ transmitter antennas. There are N_t phase shifters in the architecture. 15

1.6 Illustration of a typical two-UE power domain NOMA system. 17

2.1 Illustration of our considered mmWave cellular HetNet. The typical UE, denoted by UE A, is located at the origin, $\mathbf{o}$. All LoS high-power BSs are located within $\mathcal{B}(\mathbf{o}, \mu_m)$ and all LoS low-power BSs are located within $\mathcal{B}(\mathbf{o}, \mu_s)$. 26

2.2 The association probabilities, P_{t_m} and P_{t_s}, versus the bias factor, A_s, for different values of λ_s. 41

2.3 The loads, L_s and L_m, versus the bias factor, A_s, for different values of λ_s. . . . 41

2.4 The loads, L_s and L_m, versus the bias factor, A_s, for different values of λ_m. . . . 42

2.5 The rate coverage probability, P_c, versus the rate threshold, δ, for different values of θ_s, when $\lambda_s = 8 \times 10^{-2}/\mathrm{m}^2$ and $\lambda_m = 10^{-3}/\mathrm{m}^2$. 42

2.6 The rate coverage probability, P_c, versus the rate threshold, δ, for different values of θ_m, when $\lambda_s = 10^{-2}/\mathrm{m}^2$, $\lambda_m = 2 \times 10^{-3}/\mathrm{m}^2$, and $\mu_s = 100$ m. 43

2.7 The rate coverage probability based on the sectored gain model and the rate coverage probability based on the actual gain, versus the rate threshold, δ, for different values of P_s, N_m, and N_s, where N_m and N_s are the numbers of elements in the uniform linear array at each high-power BS and low-power BS, respectively. 43

2.8 The rate coverage probability, P_c, versus the rate threshold, δ, for different values of λ_s, when $\lambda_m = 10^{-4}/\mathrm{m}^2$, $\lambda_u = 10^{-2}/\mathrm{m}^2$, $\mu_m = 1000$ m, $\mu_s = 100$ m, and $\theta_s = 0.1$ rad. 44

2.9 The rate coverage probability, P_c, versus the rate threshold, δ, for different values of λ_m, when $\lambda_s = 10^{-2}/\mathrm{m}^2$, $\lambda_u = 10^{-2}/\mathrm{m}^2$, $\mu_m = 1000$ m, $\mu_s = 100$ m, and $\theta_s = 0.1$ rad. 44

2.10 The rate coverage probability, P_c, versus the bias factor, A_s, for different values of μ_m and μ_s, when $\delta = 10^{6.5}$ bits per second, $P_m = 40$ dBm, $P_s = 20$ dBm, $\lambda_s = 10^{-4}/\mathrm{m}^2$, $S_m = 10^9$ Hz, and $S_s = 10^9$ Hz. 45

2.11 The rate coverage probability, P_c, versus the bias factor, A_s, for different values of λ_s, when $\delta = 10^{6.5}$ bits per second, $P_m = 40$ dBm, $P_s = 20$ dBm, $\mu_m = 1000$ m, $\mu_s = 100$ m, $S_m = 10^9$ Hz, and $S_s = 10^9$ Hz. 45

3.1 The two transmission modes of the considered relay assisted mmWave cellular network. 48

3.2 The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the SINR threshold, τ, for different values of N_u, when $\eta = 2$, $P_b = 30$ dBm, $P_u = 20$ dBm, and $N_b = 16$. . . 58

3.3 The SINR coverage probability based on the sectored gain model and the SINR coverage probability based on the actual gain, versus the SINR threshold, τ, for different values of P_u, N_b, and N_u. 58

3.4 The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the number of antennas at each BS, N_b, for different values of τ, when $N_u = 2$. 59

3.5 The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the SINR threshold, τ, for different values of N_b, when $N_u = 2$. 60

3.6 The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the number of antennas at each UE, N_u, for different values of τ. 60

3.7 The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the SINR threshold, τ, for different values of N_u. 61

3.8 The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the UE transmit power, P_u, for different values of τ, when $N_u = 2$. 61

3.9 The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the BS transmit power, P_b, for different values of τ, when $N_u = 2$. 62

3.10 The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the LoS BS density, λ_b, for different values of N_b, when $\tau = 10$ dB. 63
3.11 The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the BS transmit power, P_b, for different values of N_u, when $\tau = 10$ dB. 63
3.12 The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the UE transmit power, P_u, for different values of N_u, when $\tau = 10$ dB. 64

Chapter 1

Introduction

1.1 Motivation

The rapidly growing demand for mobile data brings unprecedented challenges for cellular operators [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. According to reports from Cisco, monthly global mobile data traffic will reach 77 exabytes by 2022 [11]. As widely acknowledged within the telecommunications community, the most effective and promising solution to satisfying the growing demand for mobile data is to improve spectral efficiency. Against this background, advanced technologies have been adopted in cellular networks for spectral efficiency improvement, such as orthogonal frequency division multiplexing (OFDM) and multiple input multiple output (MIMO) [4]. Apart from improving spectral efficiency, another solution to addressing the growing demand for mobile data is to exploit the rich spectrum resources at higher frequencies. Thus, the millimeter wave (mmWave) band has recently attracted numerous attention, since there is an abundance of bandwidth available in mmWave frequencies. According to the definition from the International Telecommunication Union (ITU), the mmWave band refers to frequencies from 30 GHz to 300 GHz. Moreover, in both academia and industry, neighboring frequencies such as 28 GHz are also referred to as the mmWave band.

It is noted that in earlier generations of cellular networks, cellular operators rejected to use the mmWave band, due to two major reasons. First, some necessary components in mmWave cellular networks were either extremely expensive or unavailable since the semiconductor industry did not have the technical capability to produce the components. Second, earlier generations of cellular networks adopt high-power base stations (BSs) that cover wide areas on the order of a few kilometers. MmWave is practically not feasible to cover such wide areas [8]. However, the past decade has witnessed great advancements in the cellular industry. First, the semiconductor industry has made great progress to produce devices that can be used at the mmWave band. This implies that the necessary components of mmWave cellular networks are now available at reasonable prices. Second, the concept of small cells has been widely adopted by cellular operators and a large number of low-power BSs have been installed worldwide. As a result, the two problems that prevented cellular operators from using the mmWave band in earlier generations have been resolved and mmWave BSs are being widely adopted by cellular operators in the 5G era. Specifically, cellular operators in the U.S. have deployed their 5G commercial mmWave cellular networks and cellular operators in a number of other countries are preparing to deploy their commercial mmWave cellular networks in the near future. The deployment progress of commercial mmWave cellular networks is thoroughly described in Chapter 1.2.5.2.

As aforementioned, mmWave has been acknowledged as an attractive solution for the rapidly growing data rate requirements of cellular user equipments (UEs), due to the vast amounts of available frequencies at the mmWave band [1, 2, 3, 9, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24]. However, the unique propagation characteristics of mmWave bring serious challenges to the analysis and design of mmWave cellular networks. These unique propagation characteristics of mmWave are summarized as follows:

- High path loss: As described in the Friis' law, the far-field received power in wireless communications is given by

  $$P_r = P_t G_t G_r \left(\frac{\lambda}{4\pi d}\right)^2, \tag{1.1}$$

  where P_r and P_t are the received power and the transmit power, respectively, G_t and G_r are the transmitter and receiver antenna gains, respectively, λ is the wavelength, and d is the distance between the transmitter and the receiver. Due to that mmWave communications has small wavelengths, it suffers severe path loss. To compensate for the severe path loss, large antenna arrays with high gains are adopted in mmWave cellular networks [1, 2, 3]. The adoption of large antenna arrays has significantly changed the precoding architecture of cellular networks, which is thoroughly analyzed in Chapter 1.2.3.

- Extreme sensitivity to blockage: MmWave communications has extremely high penetration loss, as shown in Table 1.1. The high penetration loss makes mmWave communications extremely sensitive to blockage, bringing serious challenges to the signal-to-interference-plus-noise ratio (SINR) coverage performance of mmWave cellular networks. Various approaches have been proposed to resolve the blockage problem. For example, the multi-access-point scheme was proposed in [13], where the multi-access-point diversity was explored.

- Rapid channel fluctuations: A wide gulf of properties between line-of-sight (LoS) and non-line-of-sight (NLoS) channels has been observed in mmWave measurement results [9, 14, 15, 16, 17, 18, 19, 20, 21, 23, 24]. Thus, mmWave is extremely sensitive to small variations in the communication environment. For example, altering the position

Table 1.1: Outdoor and indoor penetration loss at 28 GHz [9]

Environment	Material	Thickness (cm)	Penetration loss (dB)
Outdoor	Tinted glass	3.8	40.1
Indoor	Clear glass	<1.3	3.9
	Tinted glass	<1.3	24.5
	Wall	38.1	6.8

of the transceiver by 1 cm can reduce the received signal strength by as much as 46.7 dB [22]. Moreover, the dynamic degree of fading reaches 35 dB when a person moves between fixed transceivers [23]. Due to that the received signal strength can be highly fluctuating in mmWave cellular networks, mobility support is more difficult in mmWave cellular networks than sub-6 GHz cellular networks, which is thoroughly discussed in Chapter 1.2.2.

It is noted that another challenge in the analysis and design of mmWave cellular networks is to model the location of BSs. The conventional approach to modeling the location of BSs is to place BSs on a grid deterministically, such as the traditional hexagonal grid model. This approach has two disadvantages. First, it is highly idealized since in real cellular networks, the location of BSs does not exactly follow a grid. Second, it does not lead to tractable analytical results on system-level performance metrics such as the SINR coverage probability [25, 26]. Thus, complex Monte Carlo simulations are required to obtain system-level performance metrics, which are typically extremely time-consuming.

Against this background, a novel approach based on stochastic geometry was proposed to model the location of BSs in [25, 27]. This approach not only captures the topological characteristics of real cellular networks but also leads to tractable analytical results on system-level performance metrics. In this approach, the location of BSs follows a Poisson point process (PPP). Specifically, in a PPP, the number of points in an enclosed area follows Poisson distribution and the location of each point follows independent and identically distributed (i.i.d.) uniform distribution within the enclosed area. Over the past years, PPP has been widely adopted in existing studies, due to that the analytical results obtained based on the PPP model can closely approximate the measurement results obtained from real cellular networks [25, 28, 29, 30]. Despite this, given that the locations of points are independent from each other, it is observed that two points in a PPP can be extremely close with each other, which is against the deployment principle of real cellular networks. This problem is resolved in the hard core point process (HCPP) where the distance between any two points is higher than a threshold distance value. Thus, HCPP captures the location of BSs in real cellular networks more accurately than PPP. However, HCPP does not have the probability generating functional which is essential in deriving certain system-level performance metrics such as the SINR coverage probability. Thus, the exploitation of HCPP is still an open problem in the research community [26].

Among various system-level performance metrics, the SINR/rate coverage probability is the most fundamental one and it is closely related to a number of other performance metrics, such as energy efficiency and spectral efficiency. Specifically, the SINR/rate coverage probability is the probability that a typical UE is able to achieve a certain threshold SINR/rate, i.e., it is the complementary cumulative distribution function (CCDF) of SINR/rate. As aforementioned, the adoption of the PPP enables the analysis of the SINR/rate coverage probability, although such analysis is complex. This is due to the additional source of randomness introduced by the PPP, i.e., the position of BSs. As a result, the derived SINR/rate coverage probability is typically not a closed-form expression but involves computable integrals. Valuable insights can be obtained based on the SINR/rate coverage probability, providing useful guidelines into the design of cellular networks in practical environments.

1.2 Literature Review

In this chapter, an overview of the state-of-the-art studies on mmWave communications is presented, including mmWave channel models, mmWave-enabled heterogeneous networks (HetNets), mmWave precoding, mmWave-based NOMA, and mmWave prototypes.

1.2.1 mmWave Channel Models

In this subchapter, mmWave channel models are summarized. Existing mmWave channel models are divided into two categories: 1) geometry-based stochastic models and 2) map-based models [31]. A major difference between geometry-based stochastic models and map-based models is that the former characterizes channels using probability theories, while the latter characterizes channels in a deterministic manner based on the predefined layout of a network. A necessary precondition for map-based models is the known layout of a network. However, such layout is not always accessible. Thus, this subchapter focuses on geometry-based stochastic models. It is noted that the geometry-based stochastic models for conventional sub-6 GHz cannot be directly applied to the mmWave band, due to the fundamental differences between sub-6 GHz communications and mmWave communications. Two major fundamental differences are: 1) highly directional beams applied in mmWave communications have significant impact on channel characteristics such as the root mean square (RMS) delay spread and 2) blocking effect has more impact on the received signal strength in mmWave channels than in sub-6 GHz channels, due to the extremely high penetration loss of the mmWave band. Given that mmWave geometry-based stochastic models are generally developed based on conventional sub-6 GHz geometry-based stochastic models, in this subchapter, the widely-adopted conventional sub-6 GHz geometry-based stochastic models are first summarized. Then, existing mmWave geometry-based stochastic models are summarized.

1.2.1.1 Sub-6 GHz Channel Models

Existing widely-adopted sub-6 GHz geometry-based stochastic models include:

- 3GPP model: The introduction of multiple antennas into 3G motivated the studies on modeling MIMO channels. Particularly, the 3rd Generation Partnership Project (3GPP) MIMO channel model is one of the most popular MIMO channel models. The target of this model is to describe three general scenarios: 1) suburban macrocells where the BS to BS distance is approximately 3 km, 2) urban macrocells where the BS to BS distance is approximately 3 km but there is significant scattering, and 3) urban microcells where the BS to BS distance is approximately 1 km [32, 33]. In a macrocell which typically covers a wide area on the order of a few kilometers, only a very small area around the macro BS has the chance to have LoS channels. Thus, for the simplicity of the model, all channels between macro BSs and macro UEs are assumed to be NLoS. However, in the third urban microcells scenario, the channels between micro BSs and micro UEs can be either LoS or NLoS. Particularly, the microcell LoS model is introduced to describe the channel's probability of being LoS. This is due to that a micro BS covers a small area and a large fraction of this area has a chance to own LoS channels. Thus, it is essential to take into account LoS channels in microcells.

- COST 2100 model: Comparing with the aforementioned 3GPP MIMO channel model, the COST 2100 channel model is not at the system-level where multiple BSs and multiple UEs are considered [34]. The COST 2100 channel model was originally proposed for simulating the channel between a multi-antenna BS and a multi-antenna UE. A key concept defined in the COST 2100 channel model is *visibility region* [34, 35]. In order to explain the concept of visibility region, another key concept of *cluster* is explained as follows: A cluster is defined as a group of scatterers that generate multipath components with similar delay, azimuth of departure (AoD), elevation of departure (EoD), azimuth of arrival (AoA), and elevation of arrival (EoA). Given that the channel impulse response can be obtained by adding all multipath components received at the UE, the channel impulse response is actually determined by cluster activities. Motivated by this, the concept of visibility region was proposed, which describes the cluster activity within a limited geographical area. Specifically, given the fixed location of the multi-antenna BS, each cluster has its own visibility region which is a circular region. When an UE moves into the visibility region of a cluster, the corresponding cluster has impact on the channel between the UE and the BS. When an UE moves out of the visibility region, the corresponding cluster has no impact on the channel between the UE and the BS. The visibility regions of seperate clusters may overlap with each other.

- SIRCIM/SMRCIM model: The Simualtion of Indoor Radio Channel Impulse response

Model (SIRCIM) and the Simualtion of Mobile Radio Channel Impulse response Model (SMRCIM) are based on the measurement results of channel impulse responses. SIRCIM and SMRCIM were successfully implemented from thousands of collected channel impulse responses at the carrier frequency of 1.3 GHz. These models were popular within the telecommunications industry when digital cellular networks and WiFi were first proposed.

1.2.1.2 mmWave Channel Models

As previously mentioned, conventional sub-6 GHz geometry-based stochastic models cannot be directly applied to the mmWave band. Let us take the 3GPP MIMO channel model for example. The 3GPP MIMO model is defined for a 5 MHz bandwidth CDMA system in the 2 GHz band. In the 3GPP MIMO model, there are fixed 6 paths in each of the aforementioned three scenarios, where each path represents a Dirac function in the delay domain. Moreover, each path is made up of 20 spatially separated subpaths according to the sum-of-sinusoids method [36]. However, as pointed out in [37], the number of subpaths is smaller in the mmWave band. Specifically, [37] proposed a mmWave model where the number of subpaths is 10.

Existing mmWave geometry-based stochastic models include:

- MiWEBA model: The Millimeter-Wave Evolution for Backhaul and Access (MiWEBA) model characterizes mmWave channels in a quasi-deterministic manner [35]. Specifically, a few strong paths are generated deterministically based on the heights of the transmitter and the receiver and the distance between the transmitter and the receiver, while a few relatively weak paths are generated randomly based on statistical distributions. Then the channel is obtained by adding all generated paths. Similar to sub-6 GHz channel models, the MiWEBA model assumes that the arrival of random paths follows a Poisson process, with exponentially distributed inter-arrival times. The amplitude of the random path follows i.i.d. Rayleigh distribution and the phase of the random path follows i.i.d. uniform distribution between 0 and 2π. Furthermore, the MiWEBA model incorporates the fact that mmWave paths are extremely sensitive to blockage and the signal strength of each path highly depends on the LoS probability of the path. Thus, the MiWEBA model considers the LoS probability of every path of the mmWave channel.
- METIS model: As a project mainly targeted at laying the foundation of 5G, the Mobile and wireless communications Enablers for the Twenty-twenty Information Society (METIS) model supports frequency bands up to 70 GHz [38]. The METIS model employs a combination of map-based and stochastic channel models to estimate path amplitudes and thus is suitable for evaluating massive MIMO and advanced beamforming algorithms through Monte-Carlo simulations. Specifically, the METIS model uses

ray tracing techniques to obtain large scale fading characteristics, such as the omnidirectional RMS delay spread, the azimuth spread, the shadowing fading, and the Rician K-factor for LoS channels, and uses measurement-based results to model small scale fading. The statistics of small scale fading are determined by following the approaches described in the sub-6 GHz 3GPP MIMO channel model.

- Saleh-Valenzuela (SV) model: Apart from that mmWave transmission is carried out at higher frequencies, another fundamental difference between sub-6 GHz communications and mmWave communications is that the latter has wider bandwidth than the former. Particularly, the bandwidth of mmWave communications is on the order of GHz [4, 9, 39]. Thus, mmWave geometry-based stochastic models need to consider high delay resolution, i.e., subpaths within a path may be resolvable and the number of subpaths within a path may vary. In the SV channel model, which was originally proposed for 1.5 GHz indoor multipath propagation, the number of subpaths within each path is assumed to follow a Poisson distribution. The SV channel model has been used to evaluate the system performance of the IEEE wireless personal area network (PAN) which supports a bandwidth over 500 MHz [40]. Since mmWave channels are expected to have bandwidths that are wider than 500 MHz, the SV channel model has been adopted to model mmWave channels in existing studies [31, 41, 42, 43].

1.2.2 mmWave-enabled Heterogeneous Networks

In this subchapter, the rationale behind adopting mmWave-enabled HetNets is explored, followed by a summary of the challenges in mmWave-enabled HetNets. In the end, a brief discussion on next generation mmWave-enabled HetNets is provided.

In traditional cellular networks, there are only one type of BSs and all BSs have the same transmit power which is very high. Differing from such traditional networks, HetNets consist of different types of BSs with different transmit power [44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60]. Over the past decade, cellular networks have shifted from voice-centric networks which have low data rate requirements to data-centric networks which have high data rate requirements. As a result, HetNets have been adopted worldwide, due to the fact that HetNets are a key enabler for high data rate transmission. Specifically, traditional cellular networks have mainly targeted at providing universal coverage, connecting people worldwide virtually everywhere on the planet [52, 53]. However, with the rapidly growing demand for mobile data, the data rate is elevated to a central problem and the data rate provided by traditional cellular networks cannot meet the demand of multimedia services and Internet applications. Thus, over the past years, a large number of different types of BSs have been installed to form HetNets. The rationale behind installing different types of BSs is distilled into two major factors:

- Easy installation: High-power BSs can cover large areas. However, high-power BSs are difficult to install, since high-power BSs are typically installed on tall towers. In many urban areas, installing more high-power BSs is not viable, due to the lack of available towers. Also, many cities or neighborhood associations are not very cooperative about opening up new tower locations [52]. This triggers the installation of low-power BSs, since low-power BSs can be easily installed on lampposts or at building corners.

- Low financial cost: The financial cost of a low-power BS is much less than a high-power BS [54]. Thus, low-power BSs are widely adopted to cover small areas with extreme data rate needs.

Due to the aforementioned rationale, HetNets are attractive solutions for high data rates. It is noted that another attractive solution for high data rates is mmWave, due to the vast amounts of available frequencies at the mmWave band. It is widely acknowledged that mmWave can be perfectly integrated with HetNets [47, 49]. However, there are indeed a number of challenges to resolve in mmWave-enabled HetNets, which are described as follows.

1.2.2.1 Mobility

Mobility support is more difficult in mmWave-enabled HetNets than in traditional cellular networks, due to the following reasons:

- Highly fluctuating signal strength: MmWave communications is extremely sensitive to blockage. Blockers such as buildings and human bodies can cause severe reduction in signal strength [61]. Given that there are commonly a huge amount of blockers in the communication environment, especially in urban areas, the received signal strength at a moving UE in mmWave-enabled HetNets can be highly fluctuating.

- Frequent handoffs: Low-power BSs have small coverage areas. Thus, a moving UE can frequently move into and out of the coverage areas of low-power BSs, which triggers frequent handoffs. 3GPP has proposed the concept of minimum time of stay to prevent frequent handoffs [52]. The minimum time of stay defines a threshold of time that a UE being associated to a BS, below which a handoff should not occur. However, this proposed concept has some disadvantages. For example, if a UE moves into the coverage area of a low-power BS without associating itself to the low-power BS, the UE causes very strong uplink interference to the low-power BS and receives very strong downlink interference from the low-power BS.

1.2.2.2 Decoupled Uplink Downlink Cell Association

In mmWave-enabled HetNets, a UE can be associated with two different BSs in the uplink and downlink, which is referred to as decoupled uplink downlink cell association [48, 62]. The rationale behind decoupled uplink downlink cell association is two-fold. First, in downlink cell association, a UE is more likely to be associated with a high-power BS even though the high-power BS does not have the closest distance. This is because the transmit power disparities between high-power BSs and low-power BSs lead to that high-power BSs have high downlink signal strength and low-power BSs have low downlink signal strength. Second, in uplink cell association, a UE is more likely to be associated with the closest BS, regardless of the transmit power of the BS. This is due to that UEs are transmitters and BSs are receivers in uplink and the transmit power of BSs has no impact on uplink signal strength. Decoupled uplink downlink cell association changes many entrenched ideas about cellular networks. Specifically, the challenges brought by decoupled uplink downlink cell association include: 1) uplink traffic and downlink traffic are routed from and to different BSs in the core network, 2) the uplink and downlink of the same UE have different quality of service (QoS), and 3) channel models that assume symmetry in uplink and downlink are not applicable.

1.2.2.3 Downlink Cell Association with Load Balancing

In traditional cellular networks, there are only one type of BSs and all BSs have roughly the same amount of traffic load [52]. Thus, it is acceptable to ignore the load information of BSs in downlink cell association and a UE is typically associated with the BS that provides the maximum signal strength. However, in mmWave-enabled HetNets, due to the transmit power disparities between high-power BSs and low-power BSs, the traditional maximum signal strength cell association strategy always leads to load imbalance where low-power BSs are lightly loaded and therefore not fully exploited, as pointed out by 3GPP [57]. In order to offload more UEs to low-power BSs, the biased maximum signal strength cell association strategy has been proposed [49, 57, 58]. In this association strategy, different bias factors are allocated to different types of BSs. It is highlighted that this association strategy helps cell range expansion since the coverage range of low-power BSs is effectively expanded when large bias factors are allocated to low-power BSs and small bias factors are allocated to high-power BSs.

1.2.2.4 Backhaul Bottleneck

In traditional cellular networks, BSs are equipped with wired high-speed backhaul connections to the core network. Such wired backhaul connections can easily support the data traffic flowing to and from BSs and are hardly ever a data rate bottleneck. However, in mmWave-enabled HetNets, many low-power BSs installed on lampposts do not have wired backhaul connections. In such cases, a desirable alternative is wireless backhaul connections using unlicensed frequencies. However, the quality of such wireless backhaul connections is typically very low, which makes the wireless backhaul a bottleneck to the peak data rate in mmWave-enabled HetNets [46].

1.2.2.5 Next Generation mmWave-enabled HetNets

It has been envisaged that the BSs in the next generation mmWave-enabled HetNets are equipped with mobile edge caching and mobile edge computing capabilities. The rationale behind this envisagement is that in cellular networks, some popular contents are repeatedly requested by a large number of UEs [63, 64, 65]. Also, popular contents are usually requested by different types of UEs with different resolution requirements [63, 66]. To fit different types of UEs, popular contents, especially video contents, are transformed into up to 40 different versions [66, 67, 68]. With mobile edge caching, BSs can pre-fetch popular contents and store them. Thus, when a UE requests a popular content, one of the neighboring BSs that have stored the requested content can send the content directly to the UE [63, 64, 65]. Moreover, with mobile edge computing, BSs can transform contents into different versions to fit different types of UEs. The advantages of equipping BSs with mobile edge caching and mobile edge computing capabilities include: 1) reducing the content retrieval delay, 2) improving the network capacity, and 3) reducing the backhaul traffic load.

1.2.3 mmWave Precoding

In this subchapter, mmWave precoding architectures are summarized. Existing mmWave precoding architectures can be divided into three categories: 1) fully digital precoding, 2) analog precoding, and 3) hybrid precoding. All three precoding architectures are further explored as follows.

1.2.3.1 Fully Digital Precoding

Large antenna arrays with high gains are adopted in mmWave networks to compensate for the high path loss of mmWave communications. Specifically, the number of antennas in a mmWave antenna array commonly ranges from 8 to 256, while the number of antennas in a sub-6 GHz antenna array is much smaller, commonly ranging from 2 to 8 [2, 69]. One major implication of this difference in the number of antennas is that fully digital precoding, which is widely adopted in sub-6 GHz networks, is not practically feasible in mmWave networks, due to the following reasons:

- High power consumption: As shown in Fig. 1.1, each antenna is equipped with a dedicated radio frequency (RF) chain in the fully digital precoding architecture. Specifically, an RF chain is a cascade of electronic components such as the digital-to-analog converter. The power consumption of an RF chain is high, since the power consumption of the digital-to-analog converter alone reaches up to 795 mW at mmWave frequencies [2, 70, 71, 72, 73, 74]. Given that mmWave antenna arrays commonly have a large

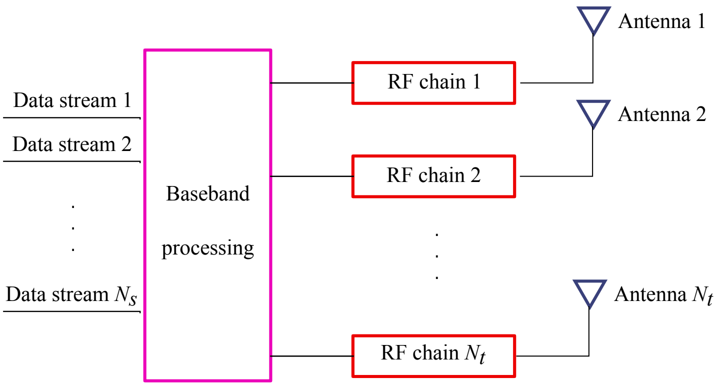

Data stream 1
Data stream 2
Data stream N_s
Baseband
processing
RF chain 1
RF chain 2
RF chain N_t
Antenna 1
Antenna 2
Antenna N_t

Figure 1.1: Illustration of the fully digital precoding architecture, where N_s is the number of data streams and N_t is the number of antennas at the transmitter antenna array.

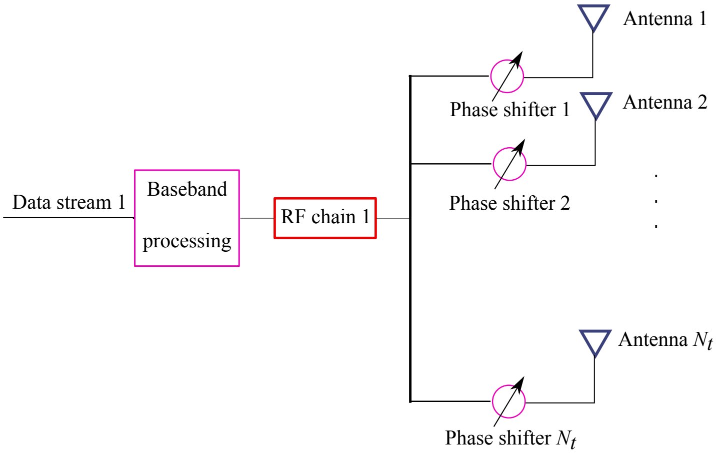

Data stream 1
Baseband
processing
RF chain 1
Phase shifter 1
Phase shifter 2
Phase shifter N_t
Antenna 1
Antenna 2
Antenna N_t

Figure 1.2: Illustration of the analog precoding architecture, where N_t is the number of antennas at the transmitter antenna array.

number of antennas, the power consumption of the fully digital precoding architecture is extremely high [75, 76, 77, 78, 79, 80].

- High financial cost: The financial cost of an RF chain is high. Deploying an RF chain for each antenna brings massive financial burden to mmWave networks [75, 76, 77, 78, 79, 80].

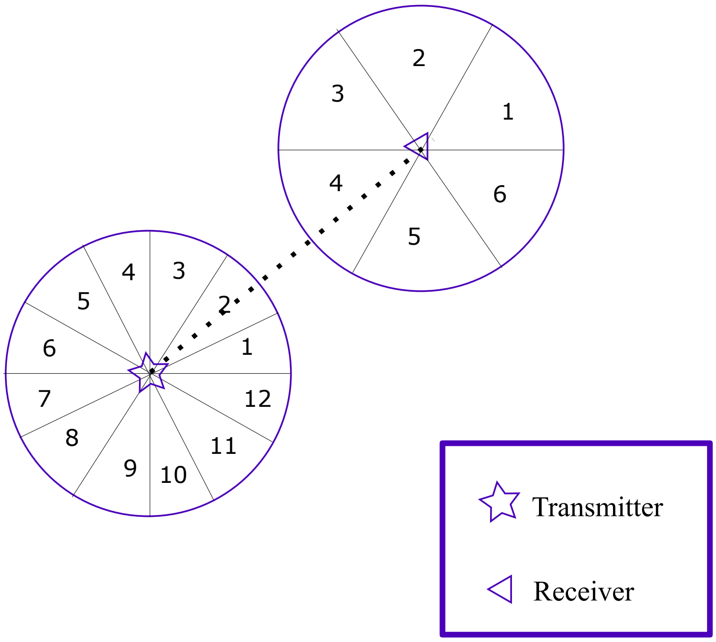


Figure 1.3: Illustration of the beam directions of the transmitter and the receiver, where the beamwidths of the transmitter and the receiver are 30° and 60°, respectively.

1.2.3.2 Analog Precoding

Unlike fully digital precoding, analog precoding (also known as analog beamforming) is practically feasible in mmWave networks. Thus, analog precoding is widely adopted in mmWave standards such as IEEE 802.11ad [2]. As shown in Fig. 1.2, only one RF chain is required in the analog precoding architecture, regardless of the number of antennas in the transmitter antenna array. Thus, the analog precoding architecture does not have the high power consumption and high financial cost problems that the fully digital precoding architecture has. However, the analog precoding architecture is not flawless. The disadvantages of the analog precoding architecture are listed as follows:

- The analog precoding architecture serves one UE and one data stream. Neither multiple UEs nor multiple data streams can be supported in the analog precoding architecture [2].
- As shown in Fig. 1.2, a phase shifter is allocated to each antenna in the analog precoding architecture to adjust the phase of the signal transmitted at the corresponding antenna. Amplitude adjustment is not available in the analog precoding architecture. Thus, unlike the fully digital precoding architecture, the analog precoding architecture does not guarantee the optimal precoding performance.

In mmWave networks, the analog precoding architecture commonly leads to very narrow beams. Specifically, if there are 16 antennas in the transmitter antenna array, the 3-dB-beamwidth of the transmitter is approximately 22.5° [81]. Thus, it is crucial to correctly steer the transmitter beam towards the desired receiver. Assuming that the receiver antenna array deploys the analog precoding architecture as well, as shown in Fig. 1.3, and the beamwidths of the transmitter and the receiver are 30° and 60°, respectively, the numbers of beam directions at the transmitter and the receiver are 12 and 6, respectively, to cover all directions. Thus, there are $12 \times 6 = 72$ possible beam pairs. Two different beam steering strategies are described next to find the optimal beam pair which is transmitter beam 2 and receiver beam 4, as shown in Fig. 1.3. In the first beam steering strategy, all 72 beam pairs are examined exhaustively and the optimal beam pair is chosen based on the examination results. In the second beam steering strategy, which is commonly referred to as the multi-step beam steering strategy, beams which are much wider than the original transmitter beams are generated and examined first. In the next step, narrower beams covered by the selected wider beams in the previous step are examined. This step is repeated until the beamwidth of the examined beams is the same as the original beamwidth of the transmitter. Compared with the first beam steering strategy, the second beam steering strategy takes less time, especially when the original beamwidth of the transmitter is extremely narrow and there are hundreds of beam pairs.

It is noted that the optimal beam pair between a BS and a UE stays valid if the communication environment is static and the UE is not moving. However, in practical mmWave cellular networks, the communication environment is commonly very dynamic and UEs are frequently moving. In this environment, there are three major factors that make the optimal beam pair invalid, i.e., the optimal beam pair for the previous transmission period is not the optimal beam pair for the current transmission period. These factors are described as follows:

- UE displacement: UEs may move when people holding UEs are walking or in a vehicle.
- UE rotational movements: As pointed out in [81], UEs experience frequent drastic rotational movements in practical usage scenarios. For example, the rotational movement when people play video games can be up to 80° per 100 ms [81].
- Moving blockers: MmWave communications is extremely sensitive to blockage. Hence, moving blockers such as vehicles and human bodies can vastly change the communication environment between a BS and a UE, causing about 20 dB reduction in signal strength [61].

The signal strength may vastly decrease in the presence of the aforementioned factors. A brute-force solution is to restart the beam steering process whenever the signal strength is below a threshold. This solution consumes a huge amount of network resources, especially in networks with extremely narrow beams where there are hundreds of beam pairs. To reduce the consumption of resources, alternative solutions have been proposed by [81] to consume less resources by identifying the factor that causes signal strength reduction first and then resolving

Table 1.2: Characteristics adopted in identifying the factor that causes signal strength reduction

Factors	Rate of signal strength reduction	Validity of beam selection at the BS
UE displacement	Slow	Invalid
UE rotational movements	Fast	Valid
Moving blockers	Fast	Invalid

the beam steering error based on the identified factor. Such solutions typically consist of the following two steps:

- Step 1: Identify the factor that causes signal strength reduction. It is easy to identify UE displacement and UE rotational movements with external aids of the Global Position System (GPS) and the accelerometer installed at the UE. However, such external aids are not always available. A novel method to identify the factor without external aids has been proposed in [81]. Specifically, two characteristics, including the rate of signal strength reduction and the validity of beam selection at the BS, are analyzed to identify the factor that causes signal strength reduction, as shown in Table 1.2. If the signal strength reduction is caused by UE displacement, the rate of signal strength reduction depends on three elements: 1) the velocity of the UE displacement, 2) the direction of the UE displacement, and 3) the distance between the BS and the UE. For example, the rate of signal strength reduction is slower at UEs far from the BS than at UEs close to the BS. Moreover, the rate of signal strength reduction caused by UE displacement is much slower than UE rotational movements and moving blockers. Furthermore, in terms of the validity of beam selection at the BS, the beam selection becomes invalid if the UE moves out of the coverage area of the selected beam. Therefore, if signal strength reduction is caused by UE rotational movements, the beam selection at the BS stays valid, due to that the UE stays within the coverage area of the selected beam. Alternatively, if signal strength reduction is caused by moving blockers, the beam selection at the BS becomes invalid, due to that the selected beam is blocked by moving blockers.

- Step 2: Remedies are applied to find the new optimal beam pair based on the identified factor. If signal strength reduction is caused by UE displacement, the new optimal BS/UE beam is within the set of neighboring beams and it is unnecessary to deploy an exhaustive examination of all BS/UE beams, due to that the UE displacement is a continuous process. If signal strength reduction is caused by UE rotational movements, the beam selection at the BS stays valid and only the UE beam needs to be adjusted. If signal strength reduction is caused by moving blockers, the BS and the UE switch to the backup beam pair which has often been determined before blockage happens.

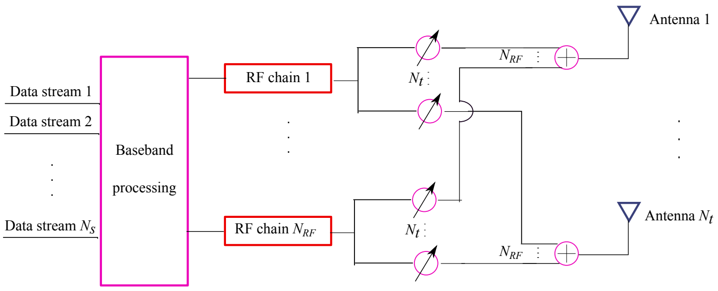

Antenna 1
Data stream 1
Data stream 2
Data stream N_s
Baseband
processing
RF chain 1
RF chain N_{RF}
N_t
N_{RF}
Antenna N_t

Figure 1.4: Illustration of the fully connected hybrid precoding architecture, where N_s is the number of data streams, N_{RF} is the number of RF chains, and N_t is the number of antennas at the transmitter antenna array. Each RF chain is connected with all transmitter antennas. There are $N_{RF} \times N_t$ phase shifters in the architecture.

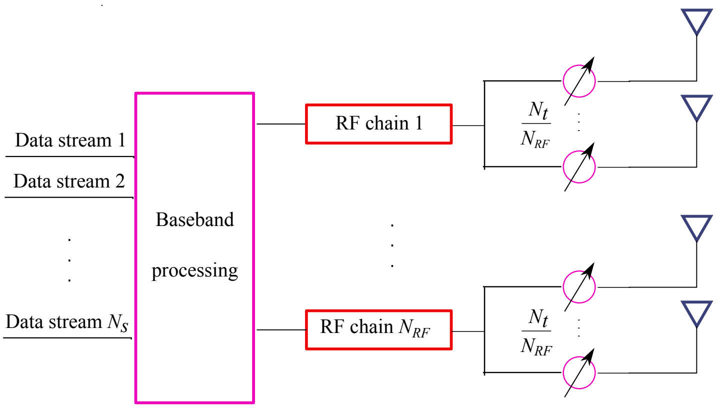

Data stream 1
Data stream 2
Data stream N_s
Baseband
processing
RF chain 1
RF chain N_{RF}
$\frac{N_t}{N_{RF}}$

Figure 1.5: Illustration of the partially connected hybrid precoding architecture, where N_s is the number of data streams, N_{RF} is the number of RF chains, and N_t is the number of antennas at the transmitter antenna array. Each RF chain is connected with $\frac{N_t}{N_{RF}}$ transmitter antennas. There are N_t phase shifters in the architecture.

1.2.3.3 Hybrid Precoding

Hybrid precoding architectures have recently attracted numerous research attention, which is because that the fully digital precoding architecture is not practically feasible and the analog precoding architecture has limited performance due to its lack of amplitude adjustments [75, 76, 77, 78, 79, 80]. Depending on whether each RF chain is connected with all antennas, hybrid precoding architectures are divided into two categories: 1) the fully connected hybrid precoding architecture where each RF chain is connected with all transmitter antennas and 2) the partially connected hybrid precoding architecture where each RF chain is connected with a proportion of transmitter antennas, as shown in Fig. 1.4 and Fig. 1.5, respectively.

In hybrid precoding architectures, the BS first processes data streams digitally at baseband using an $N_{RF} \times N_s$ digital precoder, $\mathbf{V}_{\mathrm{D}}$, and then up-converts the processed data to the carrier frequency by passing through N_{RF} RF chains. After that, the BS uses an $N_t \times N_{RF}$ analog precoder, $\mathbf{V}_{\mathrm{A}}$, to construct the transmitted signal. Mathematically, the transmitted signal can be written as

$$\mathbf{x} = \mathbf{V}_{\mathrm{A}} \mathbf{V}_{\mathrm{D}}\, \mathbf{s}, \tag{1.2}$$

where $\mathbf{x}$ is the $N_t \times 1$ transmitted signal vector and $\mathbf{s}$ is the $N_s \times 1$ data stream vector. It is widely acknowledged that hybrid precoding achieves great balance between performance and cost, and is very promising for future mmWave communications [75, 76, 77, 78, 79, 80].

1.2.4 mmWave-based Non-orthogonal Multiple Access

In this subchapter, the rationale behind adopting mmWave-based non-orthogonal multiple access (NOMA) is explored first. Then an overview of the state-of-the-art studies on mmWave-based NOMA is presented.

In the history of cellular networks, the multiple access scheme has been the key technology to allow multiple UEs to share the same communication resource. Specifically, the multiple access schemes that have been adopted in 1G, 2G, 3G, and 4G include frequency division multiple access (FDMA), time division multiple access (TDMA), code division multiple access (CDMA), and orthogonal frequency division multiple access (OFDMA) [82]. Over the past few years, NOMA has been widely acknowledged as a promising multiple access scheme for future cellular networks and thus attracted increased attention from standardization bodies such as 3GPP [82]. The benefits of NOMA include: 1) NOMA supports high spectral efficiency and 2) NOMA serves more UEs than conventional orthogonal multiple access schemes, which is important in future wireless applications for the Internet of Things (IoT). The reason why NOMA serves more UEs than conventional orthogonal multiple access schemes is that one resource block, which can be in frequency, time, or code domain, is allocated to one UE in conventional orthogonal multiple access schemes in order to avoid inter-UE interference. However, one resource block can be allocated to multiple UEs in NOMA, as shown in Fig. 1.6. There is indeed inter-UE interference in NOMA. However, with the adoption of successive interference cancellation (SIC), inter-UE interference is well controlled. In addition, when NOMA is integrated with mmWave, the high path loss of mmWave communications effectively reduces the inter-UE interference [83].

Considering a typical two-UE power domain NOMA system, where UE 1 is closer to the BS than UE 2, as shown in Fig. 1.6, the transmitted signal at the BS is given by $x = \sqrt{a_1 P}\, x_1 + \sqrt{a_2 P}\, x_2$, where x_i, $i \in \{1,2\}$, is the signal required by UE i, P is the transmit power of the BS, and a_i is the power allocation coefficient of UE i, denoting the proportion of the transmit power allocated to UE i [84]. Following the principle of power domain NOMA, given that UE 1 is closer to the BS than UE 2, it can be obtained that $a_1 < a_2$ and $a_1 + a_2 \leq 1$.

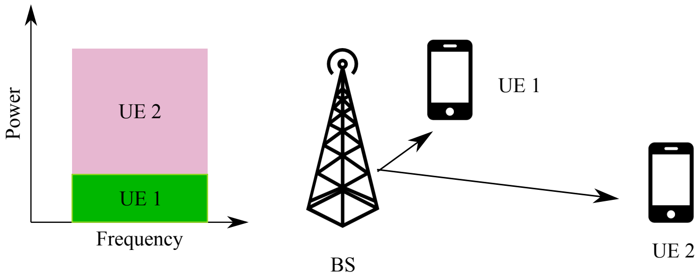


Figure 1.6: Illustration of a typical two-UE power domain NOMA system.

The received signal at each UE consists of both the desired signal and the signal of the other UE. In order to resolve the inter-UE interference, NOMA adopts SIC. Specifically, at UE 1, the signal of UE 2 is detected and removed from the received signal first, and then the desired signal is detected. At UE 2, the signal of UE 1 is weak so UE 2 can directly detect the desired signal, treating the signal of UE 1 as noise.

Several studies have been conducted on the power allocation algorithm in mmWave-based NOMA systems [85, 86, 87, 88]. [85] proposed a power allocation algorithm to maximize the energy efficiency of the system, where the total power consumption of the system includes: 1) the transmit power which is adjustable with the power allocation algorithm and 2) the power consumed by other components of the system, which is not adjustable with the power allocation algorithm. [86, 87, 88] proposed power allocation algorithms to maximize the sum rate of the system, where [87] and [87] studied the downlink and uplink of a two-UE mmWave-based NOMA system, respectively, and [88] studied the downlink of a multi-UE mmWave-based NOMA system. Apart from power allocation, another extensively studied topic in mmWave-based NOMA systems is to extend single-beam mmWave-based NOMA systems to multi-beam mmWave-based NOMA systems [86, 89]. This is because single-beam mmWave-based NOMA systems can only serve a very limited number of UEs due to the narrow beamwidth of mmWave, which reduces the potential performance gain brought by NOMA. With such extension, multi-beam mmWave-based NOMA systems can serve a large number of UEs distributed at separate directions [86, 89]. In addition, simultaneous wireless information and power transfer in mmWave-based NOMA systems was studied in [90]. Finite resolution analog precoding in mmWave-based NOMA systems was studied in [91]. In the end, unmanned-aerial-vehicle-based mmWave-based NOMA systems were studied in [92].

1.2.5 mmWave Prototypes and Commercial mmWave Cellular Networks

In this subchapter, mmWave prototypes are described followed by a survey of the deployment progress of commercial mmWave cellular networks worldwide.

1.2.5.1 mmWave Prototypes

Over the past few years, mmWave prototypes have been developed in both academia and industry [9, 14, 15, 16, 17, 69, 93, 94]. Using these mmWave prototypes, measurement campaigns have been conducted to study the propagation characteristics of mmWave, the data rate and error rate of mmWave communications, and the biological implications of mmWave cellular phones [9, 14, 15, 16, 17, 69, 93, 94].

Among the mmWave propagation characteristic measurement campaigns, a number of the best-known ones were conducted by the NYU WIRELESS research center. Using the mmWave channel sounders developed by the research center, extensive mmWave propagation measurements in urban environments in New York City and suburban environments in Austin were conducted [9, 93, 94]. Moreover, among the measurement campaigns to study the data rate and error rate of mmWave communications, a good example was carried out by Samsung Electronics, South Korea [69]. In this measurement campaign, a mmWave BS transmitted data to a mmWave cellular phone with the data rate of 528 Mb/s. The mmWave BS was equipped with a 16×16 antenna array and the mmWave cellular phone was equipped with a 16×1 antenna array. The distance between the BS and the cellular phone was fixed at 6 m. The channel between the BS and the cellular phone was LoS. The carrier frequency was 27.925 GHz and the bandwidth was 520 MHz. The measurement results showed that the block error rate reached 10^{-6}. Furthermore, the biological implications of mmWave cellular phones on human bodies such as the level of electromagnetic absorption by the human body were studied in [69]. It was reported that the maximum level of electromagnetic absorption by the human body, namely the specific absorption rate (SAR), is strictly and universally regulated by governmental bodies. The measurement results showed that the average SAR at 28 GHz is less than at 1.9 GHz. The measurement results also showed that the penetrated skin depth at 28 GHz is 3 mm, which implies that the majority of the absorbed electromagnetic energy is confined to the epidermis at 28 GHz. In contrast, the penetrated skin depth at 1.9 GHz exceeds 45 mm.

1.2.5.2 Commercial mmWave Cellular Networks

Commercial mmWave cellular networks are one of keystones in 5G. Cellular operators in the U.S. have deployed their 5G commercial mmWave cellular networks and cellular operators in a number of other countries such as Australia are preparing to deploy their commercial mmWave cellular networks.

In the U.S., the Federal Communications Commission (FCC) has allocated a large number of mmWave frequencies to 5G. Specifically, in 2019 the FCC has auctioned vast amounts of mmWave frequencies in the 24 GHz, 28 GHz, upper 37 GHz, 39 GHz, and 47 GHz bands to cellular operators for 5G deployment [95]. Using the mmWave frequencies allocated by the FCC, three major cellular operators in the U.S., namely Verizon, T-Mobile, and AT&T, have deployed their mmWave cellular networks. Specifically, in April 2019 Verizon launched its 5G Ultra Wideband cellular network in Chicago and Minneapolis, making it the world's first commercial mmWave cellular network. T-Mobile and AT&T deployed their mmWave cellular networks soon after. By the end of December 2019, Verizon's 5G Ultra Wideband cellular network has expanded to cover 20 cities in the U.S..

Along with the launch of commercial mmWave cellular networks, mmWave cellular phones have become available in the U.S. such as Samsung Galaxy S10 5G. However, the production of mmWave cellular phones is considerably more difficult than traditional cellular phones and there are only a limited number of companies that are capable of producing the chipsets that are used in mmWave cellular phones. Existing chipsets that support mmWave include Intel XMM 8160, Qualcomm Snapdragon X55, and Samsung Exynos Modem 5100.

Using mmWave cellular phones, measurement campaigns have been conducted to examine the performance of Verizon's 5G Ultra Wideband cellular network [96]. Firstly, the performance of the network when the LoS channel between the BS and the cellular phone is blocked was examined. The measurement results showed that when the LoS channel was blocked by a human body, the blockage triggered the 5G-to-4G handoff, leading to significant performance degradation because 4G has much lower data rates than 5G. In contrast, in 4G cellular networks, when the LoS channel was blocked by a human body, the cellular phone did not experience noticeable performance degradation. The measurement results also showed that when the LoS channel was blocked by a cardboard box, the blockage did not trigger the 5G-to-4G handoff. The measurement results further showed that the blockage caused by clear glasses did not trigger the 5G-to-4G handoff either and the 5G network was working in vehicles because the 5G signal can penetrate the front windshield and the windows which are typically clear glasses. Secondly, the handoff frequency of the 5G network was examined. In this measurement campaign, one person was holding a cellular phone while walking at a normal speed for about 8 minutes. The measurement results showed that during this 8-minute walk, the cellular phone experienced 31 handoffs and bounced between 4G and 5G for 13 times. Such frequent handoffs make the data rate highly fluctuating, ranging from 0 to 954 Mb/s, bringing highly inconsistent user experiences.

Apart from the U.S., a number of other countries have used commercial mmWave cellular networks in 5G as well. For example, the Australian government has announced its intentions to auction 2.4 GHz of frequencies in the 26 GHz band to cellular operators for 5G deployment in early 2021, which will be the first time that Australian cellular operators have access to mmWave frequencies [97]. Moreover, the European Union has reserved the 26 GHz band for 5G throughout its territories as well, paving the way for commercial mmWave cellular networks in Europe by the end of 2020 [98]. Furthermore, according to a work plan released by the Ministry of Industry and Information Technology, China has been working on a usage plan for mmWave frequencies in 5G [99]. In the end, the largest cellular operator in South Korea, SK Telecom, will be testing and preparing mmWave cellular networks for commercialization in 2020 to cover hotspots with extreme data rate needs [100].

1.3 Thesis Overview and Contributions

The main focus of the thesis is on the analysis and design of mmWave cellular networks. Specifically, stochastic geometry is used to model the location of BSs and analytical results on system-level performance metrics such as the SINR coverage probability are obtained. Moreover, Monte Carlo simulations are conducted to validate the analytical results. In the end, the impact of various network parameters on performance metrics in different scenarios is examined, offering valuable guidelines on the design of mmWave cellular networks.

The specific motivation and contribution of each work are described as follows.

1.3.1 Coverage Analysis of mmWave HetNets with Load Balancing

Load balancing has been extensively studied, such as [101, 102, 103, 104, 105, 106, 107, 108, 109]. A commonly adopted load metric to determine an overloaded cell is the resource block utility ratio (RBUR) [101, 102, 103, 104]. Specifically, a cell with the RBUR higher than a given threshold is determined to be an overloaded cell and the load of an overloaded cell will be partially handed over to underloaded neighboring cells. In terms of the RBUR threshold, [101, 102, 103] adopted a fixed threshold to determine overloaded cells. However, due to the fact that the load situation of a network changes over time and space, a fixed threshold, which does not adapt to these changes, can be too high or low for the load situation of the network. Thus, [104] proposed an adaptive threshold which changes with the load situation of the network. [105] adopted the adaptive threshold and proposed a novel load metric named radio resource utilization ratio (RRUR). The difference between RBUR and RRUR is that RBUR cannot be used in networks where radio resources are not allocated in terms of physical resource blocks, such as non-terrestrial networks, while RRUR can. Moreover, [106] proposed a load balancing algorithm which aims at finding the optimal handover offset value between the overloaded cell and a target cell to prevent ping-pong handovers. Furthermore, [107] emphasized that the handover procedure consumes substantial network resources and aimed to minimize the number of handovers needed to achieve load balancing.

In addition to the aforementioned reactive load balancing algorithms, multiple proactive load balancing algorithms which preempt cells from being overloaded have been proposed over the past few years [108, 109]. For example, [108] proposed a proactive load balancing algorithm which uses Twitter data to prevent overloaded cells. Specifically, [108] analyzed Twitter data to detect future events and highlighted the event locations as future hotspots. Then, the prediction of future hotspots was fed to the load balancing strategy to optimize network configurations and prevent overloaded cells. [109] proposed a proactive load balancing algorithm which uses mobility prediction to prevent overloaded cells. Specifically, [109] leveraged the semi-Markov stochastic process to predict UE mobility, and hence future cell loads. Then, the prediction of future cell loads was fed to the load balancing strategy to optimize network configurations and prevent overloaded cells.

Apart from load balancing, another extensively studied topic is the system-level coverage performance. As aforementioned in Chapter 1.1, the adoption of stochastic geometry to model the network leads to tractable analytical results for system-level coverage probabilities. Thus, several studies have adopted stochastic geometry to investigate the system-level coverage performance in various networks [110, 111, 112, 113, 114]. For example, [110] investigated the downlink coverage performance of a vehicle-to-everything network where the spatial layout of roads is modeled as a Poisson line process, the location of vehicular nodes on each road is modeled as a PPP, and the location of BSs is modeled as a PPP. [111] developed a general performance analytical framework for downlink NOMA transmission in mmWave networks where the location of UEs is modeled as a PPP. [112] investigated the coverage performance of two-dimensional and three-dimensional peer-to-peer mmWave networks where the location of nodes is modeled as a PPP. [113] investigated the downlink coverage performance of an unmanned aerial vehicle network where the location of unmanned aerial vehicles is modeled as a uniform binomial point process. [114] investigated the coverage performance of multi-stream MIMO HetNets with maximum ratio combining receivers, where the location of BSs is modeled as a PPP.

In Chapter 2, we examine the impact of load balancing in a two-tier mmWave HetNet. As described in Chapter 1.2, in HetNets where the transmit power differs between high-power BSs and low-power BSs, the traditional maximum signal strength cell association strategy always leads to load imbalance such that low-power BSs are lightly loaded and therefore not fully exploited [57]. In order to offload more UEs to low-power BSs, the biased maximum signal strength cell association strategy has been proposed. This strategy is commonly referred to as the cell range expansion (CRE) strategy, since it expands the coverage range of low-power BSs [49, 57, 58]. We note that the previous studies have investigated the CRE only in sub-6 GHz HetNets, e.g., [45, 50]. However, these results cannot be easily extended to mmWave HetNets, due to fundamental differences between mmWave HetNets and sub-6 GHz HetNets. First, highly directional beams are applied in mmWave HetNets, which significantly changes the interference behavior across the network. Second, the blocking effect in mmWave HetNets reduces the received signal strength by tens of dBs, therefore cannot be ignored in mmWave HetNets. Due to these reasons, new studies are needed to evaluate and optimize the impact of CRE on the performance of mmWave HetNets, which motivates our work. This impact has not been examined in the existing studies on mmWave networks, e.g., [49, 115].

In Chapter 2, a two-tier generalized mmWave HetNet is considered where the high-power

BSs in the macro-tier and the low-power BSs in the micro-tier co-exist to serve UEs. The generality of our considered network lies in the use of the PPP to model the location of BSs and UEs, the use of the LoS ball model to characterize the probability of a channel being LoS, the use of the sectored antenna model at BSs to capture key antenna array properties, and the use of versatile Nakagami-$\mathcal{M}$ fading to model the wireless channel. We assume that a bias factor is used in the network to offload UEs to low-power BSs. For this network, we first analyze the loads of high-power BSs and low-power BSs. Based on such analysis, we derive a new expression for the rate coverage probability of the network. Using numerical results, we demonstrate the accuracy of our analytical expressions and show how the rate coverage probability is affected by the bias factor in various scenarios. We further comprehensively examine the impact of other network parameters on the rate coverage probability.

The results in this chapter have been presented in [116], which is also listed below for ease of reference:

[C1] **S. Xu**, N. Yang, and S. Yan, ”Impact of load balancing on rate coverage performance in millimeter wave cellular heterogeneous networks,” in *Proc. IEEE ICC 2018 Workshop on Evolutional Technologies & Ecosystems for 5G Phase II*, Kansas City, MO, May 2018, pp. 1–6.

1.3.2 Coverage Analysis of Relay Assisted mmWave Cellular Networks with Spatial Correlation

In Chapter 3, we evaluate the coverage performance of a relay assisted mmWave cellular network. As aforementioned, one of the most pressing challenges for mmWave cellular networks is to enhance the coverage performance, due to the unique properties of mmWave propagation such as its extreme sensitivity to blockage [4]. One promising strategy to enhance the coverage performance is to introduce relays into mmWave cellular networks for circumventing the blockage [117, 118]. The introduction of relays can be realized by either deploying infrastructure relays or allowing idle UEs to function as relays (e.g., idle UEs serve as device-to-device (D2D) transmitters) [118]. In the work presented in Chapter 3, we consider idle relay UEs since such relays have practical advantages relative to infrastructure relays, e.g., avoiding extra cost in infrastructure.

Motivated by the benefits of relays, the performance of relay assisted mmWave networks has been studied in the literature, e.g., [117, 118, 119, 120, 121]. For example, [117] investigated relay selection among multiple decode-and-forward (DF) relays between one BS and one UE. Focusing on an amplified-and-forward (AF) relay, [119] evaluated the maximum achievable rates for half-duplex and full-duplex relaying. It is noted that neither [117] nor [119] considered multiple BSs or multiple UEs. To overcome such limitations, [120] investigated the coverage performance of a relay assisted mmWave cellular network with multiple BSs and multiple UEs while ignoring the interference in the network. Recently, [118] considered the interference when analyzing the coverage performance of a relay assisted mmWave cellular network. However, small scale fading was not considered in this analysis and directional combining was assumed at UEs.

The consideration of directional combining at multi-antenna UEs, such as [118], was commonly adopted in mmWave studies, e.g., [29, 116, 122, 123]. The main rationale behind this consideration is that directional combining not only enhances the desired signal quality but also reduces the interference, under the assumption that the direction of the UE is perfectly aligned with its associated transmitter. However, as pointed out in [81], it is extremely difficult to keep this perfect alignment in practical usage scenarios, due to the frequent drastic rotational movements of UEs, especially when UEs are not idle. For example, when people play video games, the UE is frequently rotated since more and more video games need to exploit the gyroscopic sensor at the UE. Notably, the rotational movement when people play video games can be up to 80° per 100 ms [81]. When the UE adopts directional combining, its rotational movement causes a severe misalignment problem, which significantly damages the reliability of the mmWave cellular network [81]. Motivated by this, we propose to adopt *selection combining* at UEs when they receive downlink data. We emphasize that, unlike directional combining, selection combining does not incur misalignment problems since it does not require the UE to align its direction with its associated transmitter.

In Chapter 3, the coverage performance is analyzed for a relay assisted mmWave cellular network which consists of multiple BSs, multiple relay UEs, and multiple destination UEs. In this network, the BS transmits data either directly to the destination UE, or indirectly via the relay UE when the direct transmission fails. We propose a new analytical framework to derive the SINR coverage probability of the network. In this framework, we consider: 1) generalized Nakagami-$\mathcal{M}$ fading, 2) interference, 3) selection combining at destination UEs, and 4) *spatial correlation*. Here, spatial correlation addresses the phenomenon that the received signals at the destination UE antennas are not independent, which has commonly been ignored in existing studies for simplifying analysis. Using numerical results, we first demonstrate the correctness of the derived SINR coverage probability. We also show that ignoring spatial correlation leads to severe overestimation of the coverage probability, as well as severe underestimation of the minimum number of antennas needed at destination UEs to achieve an SINR coverage probability target. We further show the benefit of introducing relays for coverage performance enhancement and that the optimal BS density maximizing the SINR coverage probability can be found through our analysis.

The results in this chapter have been presented in [124], which is also listed below for ease of reference:

[C2] **S. Xu**, N. Yang, B. He, and H. Jafarkhaniand, "Coverage analysis of relay assisted millimeter wave cellular networks with spatial correlation," accepted for publication in *Proc.*

IEEE WCNC 2020, Seoul, South Korea, Apr. 2020, pp. 1–8.

Finally, Chapter 4 presents a summary of results and future research directions.

Chapter 2

Coverage Analysis of mmWave Cellular HetNets with Load Balancing

In this chapter, we consider a two-tier mmWave cellular HetNet where the high-power BSs in the macro-tier and the low-power BSs in the micro-tier co-exist to serve UEs. In order to obtain tractable analytical results on the system-level rate coverage probability, we model the considered network using the stochastic geometry based approach described in Chapter 1.1. As mentioned in Chapter 1.2.2.3, due to the transmit power disparities between high-power BSs and low-power BSs, the traditional maximum signal strength cell association strategy always leads to load imbalance where low-power BSs are lightly loaded and therefore not fully exploited. In order to offload more UEs to low-power BSs, we adopt the biased maximum signal strength cell association strategy. In this chapter, we aim to characterize the impact of the bias factor on the system-level rate coverage probability. This characterization is completed in two steps. First, we analyze the loads of high-power BSs and low-power BSs. Second, we analyze the rate coverage probability of the network. The accuracy of our analytical results is validated using numerical results. Particularly, the optimal bias factor which maximizes the rate coverage probability can be numerically found using our analysis.

This chapter is organized as follows: Subchapter 2.1 describes the system model of the considered mmWave cellular HetNet. Subchapter 2.2 presents the analyses of the load of a typical high-power BS, the load of a typical low-power BS, and the rate coverage probability of the network. Subchapter 2.3 presents numerical results to examine the impact of load balancing and network parameters on the rate coverage probability. Finally, Subchapter 2.4 presents a summary of this chapter.

2.1 mmWave Cellular Heterogeneous Network

In this chapter, we consider a mmWave cellular HetNet, as illustrated in Fig. 2.1, which consists of a macro-tier and a micro-tier. The macro-tier consists of high-power BSs and the micro-tier consists of low-power BSs. We model the high-power BS location and the low-power BS

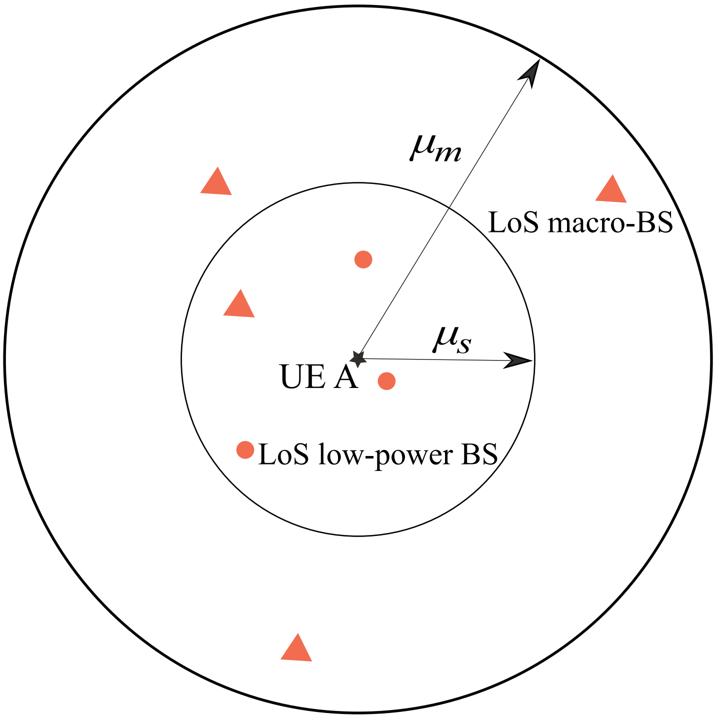


Figure 2.1: Illustration of our considered mmWave cellular HetNet. The typical UE, denoted by UE A, is located at the origin, $\mathbf{o}$. All LoS high-power BSs are located within $\mathcal{B}(\mathbf{o}, \mu_m)$ and all LoS low-power BSs are located within $\mathcal{B}(\mathbf{o}, \mu_s)$.

location as two independent homogeneous PPPs, denoted by $\Phi_{all,m}$ with the density $\lambda_{all,m}$ on $\mathbb{R}^2$ and $\Phi_{all,s}$ with the density $\lambda_{all,s}$ on $\mathbb{R}^2$, respectively. We also model the UE location as another independent PPP, denoted by Φ_u with the density λ_u on $\mathbb{R}^2$. In this chapter, we randomly select one UE and refer to it as the typical UE (i.e., UE A in Fig. 2.1). Then we establish a polar coordinate system with the typical UE at the pole. Based on the Slivnyak theorem [58], the conclusions drawn for the typical UE can be extended to other UEs.

To build a systematic performance study of the considered mmWave cellular HetNet, we need to incorporate the unique characteristics of mmWave communications into the network, as follows:

2.1.1 Blockage Effect

We adopt an accurate yet simple LoS ball model in the network, due to its high suitability for the system-level analysis of mmWave networks [29, 48, 49, 123, 125, 126]. In the LoS ball model, the probability of a channel being LoS is a function of the distance between the BS and the UE, r. In this chapter, we use $P_{LoS,\xi}(r)$ to represent this probability, where $\xi = m$ for the macro-tier and $\xi = s$ for the micro-tier. Mathematically, we express $P_{LoS,m}(r)$ as

$$P_{LoS,m}(r) = \begin{cases} \omega_m, & \text{if } 0 < r < \mu_m, \\ 0, & \text{otherwise,} \end{cases} \tag{2.1}$$

where $0 \leq \omega_m \leq 1$ is the value of the probability and $\mu_m > 0$ is depicted in Fig. 2.1. Similarly, we express $P_{LoS,s}(r)$ as

$$P_{LoS,s}(r) = \begin{cases} \omega_s, & \text{if } 0 < r < \mu_s, \\ 0, & \text{otherwise,} \end{cases} \tag{2.2}$$

where $0 \leq \omega_s \leq 1$ is the value of the probability and $\mu_s > 0$ is depicted in Fig. 2.1. In this chapter, we only consider LoS BSs, due to the extremely high outdoor penetration loss of mmWave propagation [4, 14]. According to [127, Prop. (1.3.5)], the LoS high-power BS location observed at the typical UE follows the PPP Φ_m with the density $\lambda_{all,m}\omega_m$ within the circular area $\mathcal{B}(\mathbf{o}, \mu_m)$. Moreover, the LoS low-power BS location observed at the typical UE follows the PPP Φ_s with the density $\lambda_{all,s}\omega_s$ within the circular area $\mathcal{B}(\mathbf{o}, \mu_s)$. These two PPPs are illustrated in Fig. 2.1. For the convenience of presentation, we define $\lambda_m \triangleq \lambda_{all,m}\omega_m$ and $\lambda_s \triangleq \lambda_{all,s}\omega_s$.

2.1.2 Directional Gains

We assume that each BS is equipped with an antenna array to perform directional beamforming. We also assume that each UE is equipped with a single omnidirectional antenna. As adopted in [48, 49, 126], the actual antenna array gain pattern of BSs is approximated by a sectored antenna model, which captures the key antenna array characteristics including the beamwidth, the main lobe gain, and the side lobe gain. For example, the actual antenna array gain pattern of an N element uniform linear array is given by

$$G_{act}(\varepsilon) = \frac{\sin^2\left(N\frac{\zeta}{2}\right)}{\sin^2\left(\frac{\zeta}{2}\right)}, \tag{2.3}$$

where $\zeta = \frac{2\pi}{\lambda} d \cos(\varepsilon) + \beta$, λ is the wavelength, d is the distance between elements, ε is the angle, and β is the relative phase between elements [123]. Its corresponding sectored antenna model is given by

$$G_{sec}(\theta) = \begin{cases} G_{\max,sec}, & \text{if } |\theta| \leq \frac{\theta_{sec}}{2}, \\ G_{\min,sec}, & \text{otherwise,} \end{cases} \tag{2.4}$$

where the main lobe gain, $G_{\max,sec}$, the side lobe gain, $G_{\min,sec}$, and the beamwidth, θ_{sec}, are assumed to be the maximum gain, the first minor maximum gain, and the half power beamwidth of (2.3), respectively, and θ is the angle off the boresight direction [123].

Mathematically, the antenna array gain pattern of high-power BSs is expressed as

$$G_m(\theta) = \begin{cases} G_{\max,m}, & \text{if } |\theta| \leq \frac{\theta_m}{2}, \\ G_{\min,m}, & \text{otherwise}, \end{cases} \tag{2.5}$$

where $G_{\max,m}$ and $G_{\min,m}$ are the main lobe gain and the side lobe gain of high-power BSs, respectively, and θ_m is the beamwidth of high-power BSs. Similarly, the antenna array gain pattern of low-power BSs is expressed as

$$G_s(\theta) = \begin{cases} G_{\max,s}, & \text{if } |\theta| \leq \frac{\theta_s}{2}, \\ G_{\min,s}, & \text{otherwise}, \end{cases} \tag{2.6}$$

where $G_{\max,s}$ and $G_{\min,s}$ are the main lobe gain and the side lobe gain of low-power BSs, respectively, and θ_s is the beamwidth of low-power BSs.

2.1.3 Small Scale Fading

In this chapter, we adopt independent Nakagami-$\mathcal{M}$ fading for each link to model the small scale fading h, where $\mathcal{M}$ is the fading parameter. This adoption arises from the fact that Nakagami-$\mathcal{M}$ fading is a generalized model which particularly encompasses Rician fading, a typically used model for LoS scenarios, as a special case [128, 129][1]. Against this background, Nakagami-$\mathcal{M}$ fading has recently been adopted in the performance analysis of mmWave cellular networks [49, 123].

We define $\hbar \triangleq |h|^2$ as the small scale fading power gain. Therefore, $\hbar$ follows the Gamma distribution with the shape parameter $\mathcal{M}$ and the scale parameter $1/\mathcal{M}$. It follows that the cumulative distribution function (CDF) of $\hbar$ is given by

$$F_{\hbar}(x) = 1 - e^{-\mathcal{M}x} \sum_{k=0}^{\mathcal{M}-1} \frac{(\mathcal{M}x)^k}{k!}, \tag{2.7}$$

where $\mathcal{M}$ is assumed to be a positive integer. Accordingly, the moment generation function (MGF) of $\hbar$ is given by

$$M_{\hbar}(s) = \left(1 - \frac{s}{\mathcal{M}}\right)^{-\mathcal{M}}. \tag{2.8}$$

We assume that all high-power BSs have the same transmit power P_m and all low-power BSs have the same transmit power P_s. Thus, the received power at the typical UE from a high-power BS is given by $P = P_m G_m(\theta) r^{-\alpha} \hbar$, where r is the distance between the high-power BS and the typical UE and α is the path loss exponent. Similarly, the received power at the typical UE from a low-power BS is given by $P = P_s G_s(\theta) r^{-\alpha} \hbar$.

---

[1] We clarify that Nakagami-$\mathcal{M}$ fading is a generalized fading model which can fit a variety of empirical measurements, such as Rayleigh fading by setting $\mathcal{M} = 1$ and Rician-K fading by setting $\mathcal{M} = (K+1)^2 / (2K+1)$.

In this chapter, the biased maximum signal strength cell association strategy [45] is adopted. Therefore, the typical UE chooses the nearest LoS high-power BS as its serving BS if

$$P_m r_{\min,m}^{-\alpha} G_{\max,m} > A_s P_s r_{\min,s}^{-\alpha} G_{\max,s} \quad \Rightarrow \quad r_{\min,s} > \rho r_{\min,m}, \tag{2.9}$$

where A_s is the bias factor, r_m is the distance between a LoS high-power BS and the typical UE, r_s is the distance between a LoS low-power BS and the typical UE, $r_{\min,m} = \min\{r_m\}$, $r_{\min,s} = \min\{r_s\}$, and ρ is given by $\rho \triangleq (P_m G_{\max,m} / A_s P_s G_{\max,s})^{-1/\alpha}$. If $r_{\min,s} \leq \rho r_{\min,m}$, the typical UE chooses the nearest LoS low-power BS as its serving BS. As per the cell association strategy, one BS can be chosen by multiple UEs at the same time. To avoid inter-UE interference, we adopt the TDMA scheme due to its popularity in existing mmWave standards such as IEEE 802.11ad [130], IEEE 802.15.3c [131], and ECMA-387 [132].

2.2 Load and Rate Coverage Probability Analysis

In this chapter, we aim to characterize the impact of the bias factor, A_s, on the rate coverage probability, P_c. This characterization is completed in two steps. First, we derive the load of a typical high-power BS and the load of a typical low-power BS, where the load of a BS is defined as the mean number of the UEs associated with this BS. Second, we analyze the rate coverage probability at the typical UE.

2.2.1 Load Analysis

According to the Slivnyak theorem [58], the load of a high-power BS is given by $\lambda_u P_{t_m} / \lambda_{all,m}$, where P_{t_m} is the probability that a UE is associated with a high-power BS. Similarly, the load of a low-power BS is given by $\lambda_u P_{t_s} / \lambda_{all,s}$, where P_{t_s} is the probability that a UE is associated with a low-power BS. For the convenience of presentation, we define $L_m \triangleq \lambda_u P_{t_m} / \lambda_{all,m}$ and $L_s \triangleq \lambda_u P_{t_s} / \lambda_{all,s}$. To evaluate L_m and L_s, we need to derive P_{t_m} and P_{t_s}.

To commence our derivation, we define $B_m \triangleq \Pr(\Phi_m \neq \varnothing)$ and $B_s \triangleq \Pr(\Phi_s \neq \varnothing)$. Based on [133, Eq. (2.15)], we find that

$$B_m = 1 - e^{-\lambda_m \pi \mu_m^2} \text{ and } B_s = 1 - e^{-\lambda_s \pi \mu_s^2}.$$

We also find that B_m and B_s are not necessarily equal to 1, which implies that the typical UE is possible to find no LoS high-power BSs or LoS low-power BSs to associate. Thus, we define five possible scenarios of the considered mmWave cellular HetNet, as follows:

- Scenario 1: $\Phi_m = \varnothing$ and $\Phi_s = \varnothing$.

- Scenario 2: $\Phi_m \neq \varnothing$ and $\Phi_s = \varnothing$.

- Scenario 3: $\Phi_m = \varnothing$ and $\Phi_s \neq \varnothing$.

- Scenario 4: $\Phi_m \neq \varnothing$ and $\Phi_s \neq \varnothing$, while the typical UE is associated with the nearest LoS high-power BS.

- Scenario 5: $\Phi_m \neq \varnothing$ and $\Phi_s \neq \varnothing$, while the typical UE is associated with the nearest LoS low-power BS.

Given $\Phi_m \neq \varnothing$, the CDF and the probability density function (PDF) of $r_{\min,m}$ are given by

$$F_{r_{\min,m}}(x) = \begin{cases} \frac{1-e^{-\lambda_m \pi x^2}}{B_m}, & \text{if } 0 \leq x \leq \mu_m, \\ 1, & \text{if } x > \mu_m, \end{cases} \tag{2.10}$$

and

$$f_{r_{\min,m}}(x) = \begin{cases} \frac{2\lambda_m \pi x e^{-\lambda_m \pi x^2}}{B_m}, & \text{if } 0 \leq x \leq \mu_m, \\ 0, & \text{if } x > \mu_m, \end{cases} \tag{2.11}$$

respectively. Similarly, given $\Phi_s \neq \varnothing$, the CDF and PDF of $r_{\min,s}$ are given by

$$F_{r_{\min,s}}(x) = \begin{cases} \frac{1-e^{-\lambda_s \pi x^2}}{B_s}, & \text{if } 0 \leq x \leq \mu_s, \\ 1, & \text{if } x > \mu_s, \end{cases} \tag{2.12}$$

and

$$f_{r_{\min,s}}(x) = \begin{cases} \frac{2\lambda_s \pi x e^{-\lambda_s \pi x^2}}{B_s}, & \text{if } 0 \leq x \leq \mu_s, \\ 0, & \text{if } x > \mu_s, \end{cases} \tag{2.13}$$

respectively. Using (2.10), (2.11), (2.12), and (2.13), we obtain P_{t_m} as

$$\begin{aligned} P_{t_m} &= \Pr(\text{Scenario } 2) + \Pr(\text{Scenario } 4) \\ &= B_m(1-B_s) + \Pr(\Phi_m \neq \varnothing)\Pr(\Phi_s \neq \varnothing)\Pr\left(r_{\min,m} < \frac{r_{\min,s}}{\rho} | \Phi_m \neq \varnothing, \Phi_s \neq \varnothing\right) \\ &= B_m(1-B_s) + B_m B_s \mathbb{E}_{r_{\min,s}}\left[F_{r_{\min,m}}\left(\frac{r_{\min,s}}{\rho}\right)\right] \\ &= B_m(1-B_s) + B_m B_s \int_0^{\mu_s} f_{r_{\min,s}}(x) F_{r_{\min,m}}\left(\frac{x}{\rho}\right) dx, \end{aligned} \tag{2.14}$$

where $\mathbb{E}\left[\cdot\right]$ denotes expectation. Similarly, we obtain P_{t_s} as

$$\begin{aligned}
P_{t_s} &= \Pr\left(\text{Scenario 3}\right) + \Pr\left(\text{Scenario 5}\right) \\
&= B_s\left(1-B_m\right) + \Pr\left(\Phi_m \neq \varnothing\right)\Pr\left(\Phi_s \neq \varnothing\right)\Pr\left(r_{\min,s} < \rho r_{\min,m} | \Phi_m \neq \varnothing, \Phi_s \neq \varnothing\right) \\
&= B_s\left(1-B_m\right) + B_m B_s \mathbb{E}_{r_{\min,m}}\left[F_{r_{\min,s}}\left(\rho r_{\min,m}\right)\right] \\
&= B_s\left(1-B_m\right) + B_m B_s \int_0^{\mu_m} f_{r_{\min,m}}\left(x\right) F_{r_{\min,s}}\left(\rho x\right) dx. \qquad (2.15)
\end{aligned}$$

Substituting (2.14) and (2.15) into $L_m = \lambda_u P_{t_m} / \lambda_{all,m}$ and $L_s = \lambda_u P_{t_s} / \lambda_{all,s}$, respectively, we obtain the final expressions for L_m and L_s, which are given by

$$L_m = \frac{\lambda_u}{\lambda_{all,m}}\left(B_m\left(1-B_s\right) + B_m B_s \int_0^{\mu_s} f_{r_{\min,s}}\left(x\right) F_{r_{\min,m}}\left(\frac{x}{\rho}\right) dx\right), \qquad (2.16)$$

and

$$L_s = \frac{\lambda_u}{\lambda_{all,s}}\left(B_s\left(1-B_m\right) + B_m B_s \int_0^{\mu_m} f_{r_{\min,m}}\left(x\right) F_{r_{\min,s}}\left(\rho x\right) dx\right). \qquad (2.17)$$

Such expressions determine the loads of high-power BSs and low-power BSs and allow us to analyze the rate coverage probability of the network.

2.2.2 Rate Coverage Probability Analysis

In the considered mmWave cellular network with TDMA, the maximum rate at which the information can be transmitted by a high-power BS or a low-power BS is

$$R = \frac{S_\xi}{L_\xi}\log_2\left(1+\text{SINR}\right), \qquad (2.18)$$

where S_ξ is the spectrum resource allocated to a high-power BS or a low-power BS and L_ξ is obtained in (2.16) and (2.17). Relying on (2.18), the rate coverage probability of the network, P_c, is defined as the probability that the maximum rate of the network, R, is larger than the rate threshold, δ, i.e., $P_c \triangleq \Pr\left(R > \delta\right)$. Then, we express P_c as

$$\begin{aligned}
P_c &\overset{\text{(i)}}{=} \sum_{i=1}^{5} \Pr\left(R > \delta, \text{Scenario i}\right) \\
&\overset{\text{(ii)}}{=} \Pr\left(R > \delta, \text{Scenario 2}\right) + \Pr\left(R > \delta, \text{Scenario 3}\right) \\
&\quad + \Pr\left(R > \delta, \text{Scenario 4}\right) + \Pr\left(R > \delta, \text{Scenario 5}\right) \\
&\overset{\text{(iii)}}{=} \Pr\left(\text{SINR} > 2^{\frac{\delta L_m}{S_m}} - 1, \text{Scenario 2}\right) + \Pr\left(\text{SINR} > 2^{\frac{\delta L_s}{S_s}} - 1, \text{Scenario 3}\right) \\
&\quad + \Pr\left(\text{SINR} > 2^{\frac{\delta L_m}{S_m}} - 1, \text{Scenario 4}\right) + \Pr\left(\text{SINR} > 2^{\frac{\delta L_s}{S_s}} - 1, \text{Scenario 5}\right) \\
&\triangleq P_2\left(2^{\frac{\delta L_m}{S_m}} - 1\right) + P_3\left(2^{\frac{\delta L_s}{S_s}} - 1\right) + P_4\left(2^{\frac{\delta L_m}{S_m}} - 1\right) + P_5\left(2^{\frac{\delta L_s}{S_s}} - 1\right),
\end{aligned} \tag{2.19}$$

where step (i) is due to the law of total probability, step (ii) is due to $\Pr\left(R > \delta, \text{Scenario 1}\right) = 0$, step (iii) is obtained based on (2.18). In order to obtain P_c, we next derive $P_2\left(\tau\right)$, $P_3\left(\tau\right)$, $P_4\left(\tau\right)$, and $P_5\left(\tau\right)$.

Analysis of $P_2\left(\tau\right)$

In Scenario 2, the power of the received signal at the typical UE is given by $S = \kappa_m \hbar$, where

$$\kappa_m = P_m r_{\min,m}^{-\alpha} G_{\max,m}. \tag{2.20}$$

The power of the interference at the typical UE is given by

$$I_m^* \triangleq \sum_{(r_m,\theta)\in\Phi_m/(r_{\min,m},\theta)} P_m r_m^{-\alpha} G_m\left(\theta\right)\hbar. \tag{2.21}$$

Thus, the SINR at the typical UE in Scenario 2, denoted by γ_2, is given by

$$\gamma_2 = \frac{\kappa_m \hbar}{\sum_{(r_m,\theta)\in\Phi_m/(r_{\min,m},\theta)} P_m r_m^{-\alpha} G_m\left(\theta\right)\hbar + \sigma^2}, \tag{2.22}$$

where σ^2 is the noise power at the typical UE.

Based on (2.22), we derive $P_2(\tau)$ as

$$\begin{aligned}
P_2(\tau) &\overset{(a)}{=} \Pr(\Phi_m \neq \varnothing)\Pr(\Phi_s = \varnothing)\Pr\left(\frac{S}{I_m^* + \sigma^2} > \tau | \Phi_m \neq \varnothing\right) \\
&= (1 - B_s) B_m \mathbb{E}_{r_{\min,m}, I_m^*}\left[\Pr\left(\hbar > \frac{\tau\left(I_m^* + \sigma^2\right)}{\kappa_m} | \Phi_m \neq \varnothing\right)\right] \\
&\overset{(b)}{=} (1 - B_s) B_m \mathbb{E}_{r_{\min,m}, I_m^*}\left[e^{-\frac{\psi_m}{r_{\min,m}^{-\alpha}}\left(I_m^*+\sigma^2\right)} \sum_{k=0}^{\mathcal{M}-1} \frac{\left(\psi_m\left(I_m^* + \sigma^2\right)\right)^k}{k! r_{\min,m}^{-\alpha k}}\right] \\
&= (1 - B_s) B_m \int_0^{\mu_m} \sum_{k=0}^{\mathcal{M}-1} \mathbb{E}_{I_m^*}\left[e^{-\psi_m\left(I_m^*+\sigma^2\right)x^\alpha} \frac{\left(\psi_m\left(I_m^* + \sigma^2\right)x^\alpha\right)^k}{k!}\right] f_{r_{\min,m}}(x)\, dx, \quad (2.23)
\end{aligned}$$

where we define $\psi_m \triangleq \mathcal{M}\tau / P_m G_{\max,m}$ for the convenience of presentation. In (2.23), step (a) is obtained based on the independence between Φ_m and Φ_s, and step (b) is obtained based on the CDF of $\hbar$ given in (2.7).

We now introduce the Laplace transform of $I_m^* + \sigma^2$ to obtain a simpler presentation of $P_2(\tau)$. Specifically, the Laplace transform of $I_m^* + \sigma^2$ is given by

$$\mathcal{L}_{I_m^*+\sigma^2}(a) = \mathbb{E}_{I_m^*}\left[e^{-a\left(I_m^*+\sigma^2\right)}\right]. \quad (2.24)$$

Accordingly, the kth derivative of $\mathcal{L}_{I_m^*+\sigma^2}(a)$ is given by

$$\mathcal{L}_{I_m^*+\sigma^2}^{(k)}(a) = \mathbb{E}_{I_m^*}\left[e^{-a\left(I_m^*+\sigma^2\right)}(-1)^k\left(I_m^* + \sigma^2\right)^k\right]. \quad (2.25)$$

Substituting (2.25) into (2.23), we obtain

$$P_2(\tau) = (1 - B_s) B_m \sum_{k=0}^{\mathcal{M}-1} \frac{1}{k!}(-\psi_m)^k \int_0^{\mu_m} x^{\alpha k} \mathcal{L}_{I_m^*+\sigma^2}^{(k)}(\psi_m x^\alpha) f_{r_{\min,m}}(x)\, dx. \quad (2.26)$$

As indicated by (2.26), we need to find $\mathcal{L}_{I_m^*+\sigma^2}(a)$ to obtain $P_2(\tau)$. Based on (2.24), we obtain $\mathcal{L}_{I_m^*+\sigma^2}(a)$ as

$$\begin{aligned}
\mathcal{L}_{I_m^*+\sigma^2}(a) &= \mathbb{E}_{I_m^*}\left[e^{-a\left(I_m^*+\sigma^2\right)}\right] \\
&= e^{-a\sigma^2}\mathbb{E}_{I_m^*}\left[e^{-aI_m^*}\right] \\
&\overset{(c)}{=} e^{-a\sigma^2 - \lambda_m \int_{r_{\min,m}}^{\mu_m} r \int_0^{2\pi}\left(1 - \mathbb{E}_{\hbar}\left[e^{-\hbar a P_m G_m(\theta) r^{-\alpha}}\right]\right) d\theta dr} \\
&\overset{(d)}{=} e^{-a\sigma^2 - \lambda_m \int_{r_{\min,m}}^{\mu_m} r \int_0^{2\pi}\left(1 - \left(1 + \frac{a P_m G_m(\theta) r^{-\alpha}}{\mathcal{M}}\right)^{-\mathcal{M}}\right) d\theta dr} \\
&\overset{(e)}{=} e^{-a\sigma^2 - \lambda_m \int_{r_{\min,m}}^{\mu_m} r \Omega_m dr}, \quad (2.27)
\end{aligned}$$

where step (c) is obtained based on [127, Cor. (2.3.2)], step (d) is obtained based on the MGF of $\hbar$ given in (2.8), step (e) is obtained based on the antenna gain function given in (2.5), and Ω_m is given by

$$\Omega_m \triangleq \theta_m \left[1-\left(1+\frac{aP_mG_{\max,m}}{\mathcal{M}r^{\alpha}}\right)^{-\mathcal{M}}\right]+\left(2\pi-\theta_m\right)\left[1-\left(1+\frac{aP_mG_{\min,m}}{\mathcal{M}r^{\alpha}}\right)^{-\mathcal{M}}\right]. \tag{2.28}$$

Analysis of $P_3\left(\tau\right)$

In Scenario 3, the power of the received signal at the typical UE is given by $S=\kappa_s\hbar$, where

$$\kappa_s = P_s r_{\min,s}^{-\alpha} G_{\max,s}. \tag{2.29}$$

The power of the interference at the typical UE is given by

$$I_s^* \triangleq \sum_{(r_s,\theta)\in\Phi_s/(r_{\min,s},\theta)} P_s r_s^{-\alpha} G_s\left(\theta\right)\hbar. \tag{2.30}$$

Thus, the SINR at the typical UE in Scenario 3, denoted by γ_3, is given by

$$\gamma_3 = \frac{\kappa_s\hbar}{\sum_{(r_s,\theta)\in\Phi_s/(r_{\min,s},\theta)} P_s r_s^{-\alpha} G_s\left(\theta\right)\hbar+\sigma^2}. \tag{2.31}$$

Based on (2.31), we derive $P_3\left(\tau\right)$ as

$$\begin{aligned}
P_3\left(\tau\right) &\overset{(f)}{=} \Pr\left(\Phi_m=\varnothing\right)\Pr\left(\Phi_s\neq\varnothing\right)\Pr\left(\frac{S}{I_s^*+\sigma^2}>\tau|\Phi_s\neq\varnothing\right)\\
&=\left(1-B_m\right)B_s\Pr\left(\frac{S}{I_s^*+\sigma^2}>\tau|\Phi_s\neq\varnothing\right)\\
&=\left(1-B_m\right)B_s\mathbb{E}_{r_{\min,s},I_s^*}\left[\Pr\left(\hbar>\frac{\tau\left(I_s^*+\sigma^2\right)}{\kappa_s}|\Phi_s\neq\varnothing\right)\right]\\
&\overset{(g)}{=}\left(1-B_m\right)B_s\mathbb{E}_{r_{\min,s},I_s^*}\left[e^{-\frac{\psi_s}{r_{\min,s}^{-\alpha}}\left(I_s^*+\sigma^2\right)}\sum_{k=0}^{\mathcal{M}-1}\frac{\left(\psi_s\left(I_s^*+\sigma^2\right)\right)^k}{k!r_{\min,s}^{-\alpha k}}\right]\\
&=\left(1-B_m\right)B_s\int_0^{\mu_s}\sum_{k=0}^{\mathcal{M}-1}\mathbb{E}_{I_s^*}\left[e^{-\psi_s\left(I_s^*+\sigma^2\right)x^{\alpha}}\frac{\left(\psi_s\left(I_s^*+\sigma^2\right)x^{\alpha}\right)^k}{k!}\right]f_{r_{\min,s}}\left(x\right)dx,
\end{aligned} \tag{2.32}$$

where we define $\psi_s \triangleq \mathcal{M}\tau/P_sG_{\max,s}$ for the convenience of presentation. In (2.32), step (f) is obtained based on the independence between Φ_m and Φ_s, and step (g) is obtained based on the CDF of $\hbar$ given in (2.7).

We introduce the Laplace transform of $I_s^*+\sigma^2$ to obtain a simpler presentation of $P_3\left(\tau\right)$.

Specifically, the Laplace transform of $I_s^* + \sigma^2$ is given by

$$\mathcal{L}_{I_s^*+\sigma^2}(a) = \mathbb{E}_{I_s^*}\left[e^{-a\left(I_s^*+\sigma^2\right)}\right]. \tag{2.33}$$

Accordingly, the kth derivative of $\mathcal{L}_{I_s^*+\sigma^2}(a)$ is given by

$$\mathcal{L}^{(k)}_{I_s^*+\sigma^2}(a) = \mathbb{E}_{I_s^*}\left[e^{-a\left(I_s^*+\sigma^2\right)}(-1)^k\left(I_s^*+\sigma^2\right)^k\right]. \tag{2.34}$$

Substituting (2.34) into (2.32), we obtain

$$P_3(\tau) = (1-B_m)B_s\sum_{k=0}^{\mathcal{M}-1}\frac{1}{k!}(-\psi_s)^k\int_0^{\mu_s}x^{\alpha k}\mathcal{L}^{(k)}_{I_s^*+\sigma^2}(\psi_s x^{\alpha})f_{r_{\min,s}}(x)\,dx. \tag{2.35}$$

We further derive $\mathcal{L}_{I_s^*+\sigma^2}(a)$. Based on (2.33), we obtain $\mathcal{L}_{I_s^*+\sigma^2}(a)$ as

$$\begin{aligned}
&\mathcal{L}_{I_s^*+\sigma^2}(a)\\
&= e^{-a\sigma^2}\mathbb{E}_{I_s^*}\left[e^{-aI_s^*}\right]\\
&\overset{(h)}{=} e^{-a\sigma^2-\lambda_s\int_{r_{\min,s}}^{\mu_s} r\int_0^{2\pi}\left(1-\mathbb{E}_{\hbar}\left[e^{-\hbar a P_S G_S(\theta)r^{-\alpha}}\right]\right)d\theta dr}\\
&\overset{(i)}{=} e^{-a\sigma^2-\lambda_s\int_{r_{\min,s}}^{\mu_s} r\int_0^{2\pi}\left(1-\left(1+\frac{aP_S G_S(\theta)r^{-\alpha}}{\mathcal{M}}\right)^{-\mathcal{M}}\right)d\theta dr}\\
&\overset{(j)}{=} e^{-a\sigma^2-\lambda_s\int_{r_{\min,s}}^{\mu_s} r\Omega_s dr},
\end{aligned} \tag{2.36}$$

where step (h) is obtained based on [127, Cor. (2.3.2)], step (i) is obtained based on the MGF of $\hbar$ given in (2.8), step (j) is obtained based on the antenna gain function given in (2.6), and Ω_s is given by

$$\Omega_s \triangleq \theta_s\left[1-\left(1+\frac{aP_sG_{\max,s}}{\mathcal{M}r^{\alpha}}\right)^{-\mathcal{M}}\right]+(2\pi-\theta_s)\left[1-\left(1+\frac{aP_sG_{\min,s}}{\mathcal{M}r^{\alpha}}\right)^{-\mathcal{M}}\right]. \tag{2.37}$$

Analysis of $P_4(\tau)$

In Scenario 4, the power of the received signal at the typical UE is given by $S = \kappa_m\hbar$, where κ_m is defined in (2.20). Moreover, the power of the interference at the typical UE is given by $I = I_m^* + I_s$, where I_m^* is defined in (2.21) and I_s is given by

$$I_s \triangleq \sum_{(r_s,\theta)\in\Phi_s} P_s r_s^{-\alpha} G_s(\theta)\hbar. \tag{2.38}$$

Thus, the SINR at the typical UE in Scenario 4, denoted by γ_4, is given by

$$\gamma_4 = \frac{\kappa_m \hbar}{\sum_{(r_m,\theta)\in\Phi_m/(r_{\min,m},\theta)} P_m r_m^{-\alpha} G_m(\theta)\hbar + \sum_{(r_s,\theta)\in\Phi_s} P_s r_s^{-\alpha} G_s(\theta)\hbar + \sigma^2}. \tag{2.39}$$

Based on (2.39), we derive $P_4(\tau)$ as

$$\begin{aligned}
&P_4(\tau)\\
&= \Pr\left(\gamma_4 > \tau, \Phi_m \neq \varnothing, \Phi_s \neq \varnothing, r_{\min,s} > \rho r_{\min,m}\right)\\
&= \Pr\left(\frac{\kappa_m \hbar}{I_m^* + I_s + \sigma^2} > \tau \middle| \Phi_m \neq \varnothing, \Phi_s \neq \varnothing, r_{\min,s} > \rho r_{\min,m}\right)\\
&\quad \times \Pr\left(\Phi_m \neq \varnothing, \Phi_s \neq \varnothing, r_{\min,s} > \rho r_{\min,m}\right)\\
&= B_s B_m \int_0^{\mu_m} \left(1 - F_{r_{\min,s}}(\rho x)\right) \Pr\left(\hbar > \frac{\tau\left(I_m^* + I_s + \sigma^2\right)}{\kappa_m}\right) f_{r_{\min,m}}(x)\, dx\\
&= B_s B_m \int_0^{\mu_m} \left(1 - F_{r_{\min,s}}(\rho x)\right) \mathbb{E}_{I_m^*, I_s}\left[e^{-\frac{\psi_m}{x^{-\alpha}}\left(I_m^* + I_s + \sigma^2\right)} \sum_{k=0}^{\mathcal{M}-1} \frac{\left(\psi_m\left(I_m^* + I_s + \sigma^2\right)\right)^k}{k! x^{-\alpha k}}\right] f_{r_{\min,m}}(x)\, dx\\
&= B_s B_m \int_0^{\mu_m} \left(1 - F_{r_{\min,s}}(\rho x)\right) \sum_{k=0}^{\mathcal{M}-1} \mathbb{E}_{I_m^*, I_s}\left[e^{-\frac{\psi_m}{x^{-\alpha}}\left(I_m^* + I_s + \sigma^2\right)} \frac{\left(\psi_m\left(I_m^* + I_s + \sigma^2\right)\right)^k}{k! x^{-\alpha k}}\right] f_{r_{\min,m}}(x)\, dx.
\end{aligned} \tag{2.40}$$

We now introduce the Laplace transform of $I_m^* + I_s + \sigma^2$ to obtain a simpler presentation of $P_4(\tau)$. Specifically, the Laplace transform of $I_m^* + I_s + \sigma^2$ is given by

$$\mathcal{L}_{I_m^* + I_s + \sigma^2}(a) = \mathbb{E}_{I_m^*, I_s}\left[e^{-a\left(I_m^* + I_s + \sigma^2\right)}\right]. \tag{2.41}$$

Accordingly, the kth derivative of $\mathcal{L}_{I_m^* + I_s + \sigma^2}(a)$ is given by

$$\mathcal{L}^{(k)}_{I_m^* + I_s + \sigma^2}(a) = \mathbb{E}_{I_m^*, I_s}\left[e^{-a\left(I_m^* + I_s + \sigma^2\right)}(-1)^k\left(I_m^* + I_s + \sigma^2\right)^k\right]. \tag{2.42}$$

Substituting (2.42) into (2.40), we obtain

$$\begin{aligned}
&P_4\left(\tau\right)\\
&= B_s B_m \int_0^{\mu_m} \left(1 - F_{r_{\min,s}}\left(\rho x\right)\right) \sum_{k=0}^{\mathcal{M}-1} \frac{\left(-\psi_m x^{\alpha}\right)^k}{k!} \mathcal{L}^{(k)}_{I_m^*+I_s+\sigma^2}\left(\psi_m x^{\alpha}\right) f_{r_{\min,m}}\left(x\right) dx\\
&\overset{(k)}{=} B_s B_m \int_0^{\min\left(\frac{\mu_s}{\rho},\mu_m\right)} \left(1 - F_{r_{\min,s}}\left(\rho x\right)\right) \sum_{k=0}^{\mathcal{M}-1} \frac{\left(-\psi_m x^{\alpha}\right)^k}{k!} \mathcal{L}^{(k)}_{I_m^*+I_s+\sigma^2}\left(\psi_m x^{\alpha}\right) f_{r_{\min,m}}\left(x\right) dx\\
&= B_s \int_0^{\min\left(\frac{\mu_s}{\rho},\mu_m\right)} \left(1 - \frac{1-e^{-\lambda_s \pi x^2 \rho^2}}{B_s}\right) \sum_{k=0}^{\mathcal{M}-1} \frac{\left(-\psi_m x^{\alpha}\right)^k}{k!} \mathcal{L}^{(k)}_{I_m^*+I_s+\sigma^2}\left(\psi_m x^{\alpha}\right) 2\lambda_m \pi x e^{-\lambda_m \pi x^2} dx,
\end{aligned} \tag{2.43}$$

Here, step (k) is based on the fact that for $x > \frac{\mu_s}{\rho}$, $F_{r_{\min,s}}\left(\rho x\right) = 1$.

As indicated by (2.43), we need to find $\mathcal{L}_{I_m^*+I_s+\sigma^2}\left(a\right)$ to obtain $P_4\left(\tau\right)$. Based on (2.41), we obtain $\mathcal{L}_{I_m^*+I_s+\sigma^2}\left(a\right)$ as

$$\begin{aligned}
&\mathcal{L}_{I_m^*+I_s+\sigma^2}\left(a\right)\\
&= e^{-a\sigma^2} \mathbb{E}_{I_m^*}\left[e^{-aI_m^*}\right] \mathbb{E}_{I_s}\left[e^{-aI_s}\right]\\
&= e^{-a\sigma^2-\lambda_m \int_{r_{\min,m}}^{\mu_m} r\Omega_m dr} \mathbb{E}_{I_s}\left[e^{-aI_s}\right]\\
&= e^{-a\sigma^2-\lambda_m \int_{r_{\min,m}}^{\mu_m} r\Omega_m dr} e^{-\lambda_s \int_0^{\mu_s} r \int_0^{2\pi} \left(1-\mathbb{E}_{\hbar}\left[e^{-\hbar a P_s G_s(\theta) r^{-\alpha}}\right]\right) d\theta dr}\\
&= e^{-a\sigma^2-\lambda_m \int_{r_{\min,m}}^{\mu_m} r\Omega_m dr} e^{-\lambda_s \int_0^{\mu_s} r \int_0^{2\pi} \left(1-\left(1+\frac{aP_s G_s(\theta) r^{-\alpha}}{\mathcal{M}}\right)^{-\mathcal{M}}\right) d\theta dr}\\
&= e^{-a\sigma^2-\lambda_m \int_{r_{\min,m}}^{\mu_m} r\Omega_m dr} e^{-\lambda_s \int_0^{\mu_s} r\Omega_s dr},\\
&= e^{-a\sigma^2-\lambda_m \int_{r_{\min,m}}^{\mu_m} r\Omega_m dr-\lambda_s \int_0^{\mu_s} r\Omega_s dr}.
\end{aligned} \tag{2.44}$$

For the convenience of presentation, we define

$$\rho\left(\tau\right) \triangleq \int_0^{\min\left(\frac{\mu_s}{\rho},\mu_m\right)} \left(B_s - 1 + e^{-\lambda_s \pi x^2 \rho^2}\right) x^{\alpha k+1} \mathcal{L}^{(k)}_{I_m^*+I_s+\sigma^2}\left(\psi_m x^{\alpha}\right) e^{-\lambda_m \pi x^2} dx, \tag{2.45}$$

Thus, P_4 is given by

$$P_4\left(\tau\right) = 2\lambda_m \pi \sum_{k=0}^{\mathcal{M}-1} \frac{1}{k!} \left(-\psi_m\right)^k \rho\left(\tau\right). \tag{2.46}$$

Analysis of $P_5\left(\tau\right)$

In Scenario 5, the power of the received signal at the typical UE is given by $S = \kappa_s \hbar$, where κ_s is defined in (2.29). Moreover, the power of the interference at the typical UE is given by

$I = I_s^* + I_m$, where I_s^* is defined in (2.30) and I_m is given by

$$I_m \triangleq \sum_{(r_m,\theta)\in\Phi_m} P_m r_m^{-\alpha} G_m(\theta)\hbar. \tag{2.47}$$

Thus, the SINR at the typical UE in Scenario 5, denoted by γ_5, is given by

$$\gamma_5 = \frac{\kappa_s \hbar}{\sum_{(r_s,\theta)\in\Phi_s/(r_{\min,s},\theta)} P_s r_s^{-\alpha} G_s(\theta)\hbar + \sum_{(r_m,\theta)\in\Phi_m} P_m r_m^{-\alpha} G_m(\theta)\hbar + \sigma^2}. \tag{2.48}$$

Based on (2.48), we derive $P_5(\tau)$ as

$$\begin{aligned}
&P_5(\tau)\\
&= \Pr\left(\gamma_5 > \tau, \Phi_s \neq \varnothing, \Phi_m \neq \varnothing, r_{\min,m} > \frac{r_{\min,s}}{\rho}\right)\\
&= \Pr\left(\frac{\kappa_s\hbar}{I_s^* + I_m + \sigma^2} > \tau \middle| \Phi_s \neq \varnothing, \Phi_m \neq \varnothing, r_{\min,m} > \frac{r_{\min,s}}{\rho}\right)\\
&\quad \times \Pr\left(\Phi_s \neq \varnothing, \Phi_m \neq \varnothing, r_{\min,m} > \frac{r_{\min,s}}{\rho}\right)\\
&= B_m B_s \int_0^{\mu_s} \left(1 - F_{r_{\min,m}}\left(\frac{x}{\rho}\right)\right) \Pr\left(\hbar > \frac{\tau\left(I_s^* + I_m + \sigma^2\right)}{\kappa_s}\right) f_{r_{\min,s}}(x)\,dx\\
&= B_m B_s \int_0^{\mu_s} \left(1 - F_{r_{\min,m}}\left(\frac{x}{\rho}\right)\right) \mathbb{E}_{I_s^*,I_m}\left[e^{-\frac{\psi_s}{x^{-\alpha}}\left(I_s^*+I_m+\sigma^2\right)} \sum_{k=0}^{\mathcal{M}-1} \frac{\left(\psi_s\left(I_s^* + I_m + \sigma^2\right)\right)^k}{k! x^{-\alpha k}}\right] f_{r_{\min,s}}(x)\,dx\\
&= B_m B_s \int_0^{\mu_s} \left(1 - F_{r_{\min,m}}\left(\frac{x}{\rho}\right)\right) \sum_{k=0}^{\mathcal{M}-1} \mathbb{E}_{I_s^*,I_m}\left[e^{-\frac{\psi_s}{x^{-\alpha}}\left(I_s^*+I_m+\sigma^2\right)} \frac{\left(\psi_s\left(I_s^* + I_m + \sigma^2\right)\right)^k}{k! x^{-\alpha k}}\right] f_{r_{\min,s}}(x)\,dx.
\end{aligned} \tag{2.49}$$

We now introduce the Laplace transform of $I_s^* + I_m + \sigma^2$ to obtain a simpler presentation of $P_5(\tau)$. Specifically, the Laplace transform of $I_s^* + I_m + \sigma^2$ is given by

$$\mathcal{L}_{I_s^*+I_m+\sigma^2}(a) = \mathbb{E}_{I_s^*,I_m}\left[e^{-a\left(I_s^*+I_m+\sigma^2\right)}\right]. \tag{2.50}$$

Accordingly, the kth derivative of $\mathcal{L}_{I_s^*+I_m+\sigma^2}(a)$ is given by

$$\mathcal{L}^{(k)}_{I_s^*+I_m+\sigma^2}(a) = \mathbb{E}_{I_s^*,I_m}\left[e^{-a\left(I_s^*+I_m+\sigma^2\right)}(-1)^k\left(I_s^* + I_m + \sigma^2\right)^k\right]. \tag{2.51}$$

Substituting (2.51) into (2.49), we obtain

$$
\begin{aligned}
&P_5\left(\tau\right)\\
&= B_m B_s \int_0^{\mu_s} \left(1 - F_{r_{\min,m}}\left(\frac{x}{\rho}\right)\right) \sum_{k=0}^{\mathcal{M}-1} \frac{\left(-\psi_s x^{\alpha}\right)^k}{k!} \mathcal{L}^{(k)}_{I_s^*+I_m+\sigma^2}\left(\psi_s x^{\alpha}\right) f_{r_{\min,s}}\left(x\right) dx\\
&\overset{(l)}{=} B_m B_s \int_0^{\min\left(\mu_m\rho,\mu_s\right)} \left(1 - F_{r_{\min,m}}\left(\frac{x}{\rho}\right)\right) \sum_{k=0}^{\mathcal{M}-1} \frac{\left(-\psi_s x^{\alpha}\right)^k}{k!} \mathcal{L}^{(k)}_{I_s^*+I_m+\sigma^2}\left(\psi_s x^{\alpha}\right) f_{r_{\min,s}}\left(x\right) dx\\
&= B_m \int_0^{\min\left(\mu_m\rho,\mu_s\right)} \left(1 - \frac{1-e^{-\frac{\lambda_m \pi x^2}{\rho^2}}}{B_m}\right) \sum_{k=0}^{\mathcal{M}-1} \frac{\left(-\psi_s x^{\alpha}\right)^k}{k!} \mathcal{L}^{(k)}_{I_s^*+I_m+\sigma^2}\left(\psi_s x^{\alpha}\right) 2\lambda_s \pi x e^{-\lambda_s \pi x^2} dx,
\end{aligned}
\tag{2.52}
$$

Here, step (l) is based on the fact that for $x > \mu_m\rho$, $F_{r_{\min,m}}\left(\frac{x}{\rho}\right) = 1$.

As indicated by (2.52), we need to find $\mathcal{L}_{I_s^*+I_m+\sigma^2}\left(a\right)$ to obtain $P_5\left(\tau\right)$. Based on (2.50), we obtain $\mathcal{L}_{I_s^*+I_m+\sigma^2}\left(a\right)$ as

$$
\begin{aligned}
\mathcal{L}_{I_s^*+I_m+\sigma^2}\left(a\right) &= e^{-a\sigma^2} \mathbb{E}_{I_s^*}\left[e^{-aI_s^*}\right] \mathbb{E}_{I_m}\left[e^{-aI_m}\right]\\
&= e^{-a\sigma^2 - \lambda_s \int_{r_{\min,s}}^{\mu_s} r\Omega_s dr} \mathbb{E}_{I_m}\left[e^{-aI_m}\right]\\
&= e^{-a\sigma^2 - \lambda_s \int_{r_{\min,s}}^{\mu_s} r\Omega_s dr} e^{-\lambda_m \int_0^{\mu_m} r \int_0^{2\pi}\left(1-\mathbb{E}_{\hbar}\left[e^{-\hbar a P_m G_m(\theta) r^{-\alpha}}\right]\right) d\theta dr}\\
&= e^{-a\sigma^2 - \lambda_s \int_{r_{\min,s}}^{\mu_s} r\Omega_s dr} e^{-\lambda_m \int_0^{\mu_m} r \int_0^{2\pi}\left(1-\left(1+\frac{a P_m G_m(\theta) r^{-\alpha}}{\mathcal{M}}\right)^{-\mathcal{M}}\right) d\theta dr}\\
&= e^{-a\sigma^2 - \lambda_s \int_{r_{\min,s}}^{\mu_s} r\Omega_s dr} e^{-\lambda_m \int_0^{\mu_m} r\Omega_m dr},\\
&= e^{-a\sigma^2 - \lambda_s \int_{r_{\min,s}}^{\mu_s} r\Omega_s dr - \lambda_m \int_0^{\mu_m} r\Omega_m dr}.
\end{aligned}
\tag{2.53}
$$

For the convenience of presentation, we define

$$
\varpi\left(\tau\right) \triangleq \int_0^{\min\left(\mu_s,\rho\mu_m\right)} \left(B_m - 1 + e^{-\frac{\lambda_m \pi x^2}{\rho^2}}\right) x^{\alpha k+1} \mathcal{L}^{(k)}_{I_m+I_s^*+\sigma^2}\left(\psi_s x^{\alpha}\right) e^{-\lambda_s \pi x^2} dx, \tag{2.54}
$$

Thus, P_5 is given by

$$
P_5\left(\tau\right) = 2\lambda_s \pi \sum_{k=0}^{\mathcal{M}-1} \frac{1}{k!} \left(-\psi_s\right)^k \varpi\left(\tau\right). \tag{2.55}
$$

Substituting (2.26), (2.35), (2.46), and (2.55) into (2.19), we obtain the exact rate coverage

Table 2.1: Parameters Used in Subchapter 2.3

Parameters	Value
P_m & P_s	50 & 30 dBm
λ_m, λ_s, & λ_u	10^{-5}, 10^{-3}, & $10^{-1}/\text{m}^2$
μ_m & μ_s	200 & 50 m
ω_m & ω_s	0.6 & 0.5
θ_m & θ_s	0.1 & 0.2 rad
$G_{\max,m}$ & $G_{\max,s}$	4×10^3 & 10^3
$G_{\min,m}$ & $G_{\min,s}$	1 & 1
A_s	1
α	2.2
σ^2	1
S_m & S_s	10^{10} & 10^{10} Hz

probability of the network, P_c, which is given by

$$\begin{aligned}
P_c = {} & (1-B_s)B_m \sum_{k=0}^{\mathcal{M}-1} \frac{(-\varepsilon_m)^k}{k!} \int_0^{\mu_m} x^{\alpha k} \mathcal{L}^{(k)}_{I_m^*+\sigma^2}(\varepsilon_m x^\alpha) f_{r_{\min,m}}(x)\, dx \\
& + (1-B_m)B_s \sum_{k=0}^{\mathcal{M}-1} \frac{(-\varepsilon_s)^k}{k!} \int_0^{\mu_s} x^{\alpha k} \mathcal{L}^{(k)}_{I_s^*+\sigma^2}(\varepsilon_s x^\alpha) f_{r_{\min,s}}(x)\, dx \\
& + 2\lambda_m \pi \sum_{k=0}^{\mathcal{M}-1} \frac{(-\varepsilon_m)^k}{k!} \rho\left(2^{\frac{\delta L_m}{S_m}} - 1\right) + 2\lambda_s \pi \sum_{k=0}^{\mathcal{M}-1} \frac{(-\varepsilon_s)^k}{k!} \varpi\left(2^{\frac{\delta L_s}{S_s}} - 1\right),
\end{aligned} \tag{2.56}$$

where $\varepsilon_m = \mathcal{M}\left(2^{\frac{\delta L_m}{S_m}} - 1\right)/P_m G_{\max,m}$ and $\varepsilon_s = \mathcal{M}\left(2^{\frac{\delta L_s}{S_s}} - 1\right)/P_s G_{\max,s}$. This result allows us to investigate the impact of the bias factor and other network parameters on P_c. We note that a closed-form expression for P_c is extremely difficult, if not impossible, to obtain, due to the consideration of PPPs in the network, as mentioned in Chapter 1.1. Indeed, integral-form expressions are widely used to evaluate the coverage performance of HetNets, e.g., [49, 50, 58, 125].

2.3 Numerical Results and Discussions

In this subchapter, we present numerical results to examine the impact of load balancing and network parameters on the performance. The number of samples in the Monte Carlo simulation in this Chapter is 10^6. Unless otherwise specified, the parameters used in this subchapter are summarized in Table 2.1.

First, we examine the impact of the LoS low-power BS density, λ_s, and the bias factor, A_s, on the association probabilities, P_{t_m} and P_{t_s}, in Fig. 2.2. By comparing analytical curves with the Monte Carlo simulation points marked by '△', we observe that our analytical expressions for P_{t_m} and P_{t_s}, given in (2.14) and (2.15), respectively, precisely agree with simulations. This

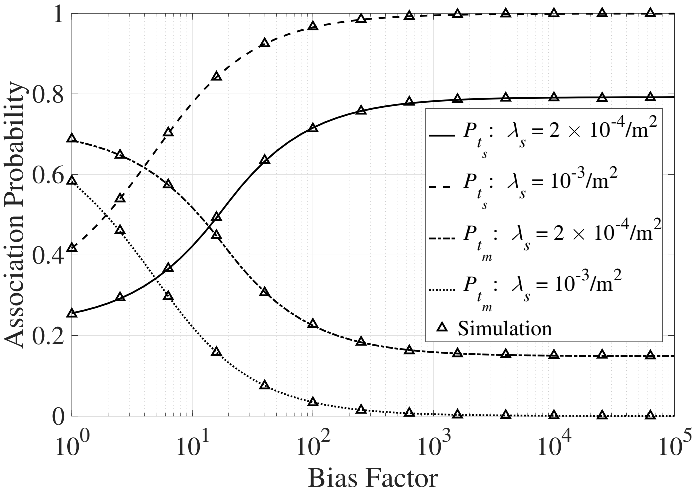
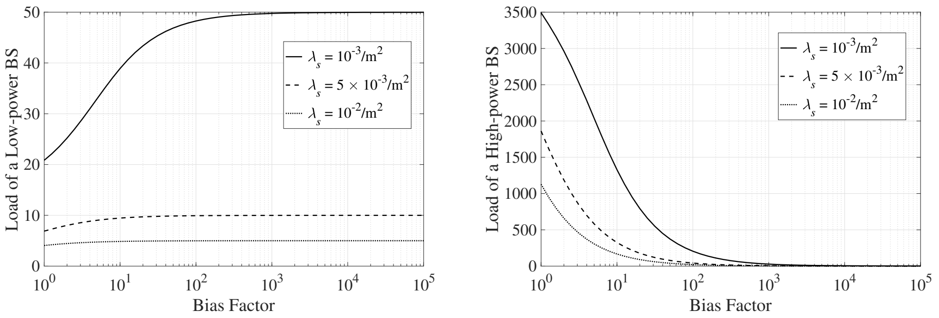


Figure 2.2: The association probabilities, P_{t_m} and P_{t_s}, versus the bias factor, A_s, for different values of λ_s.



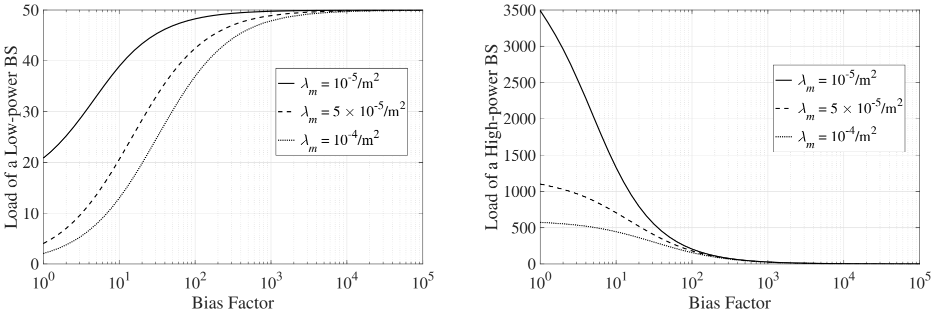
(a) Load of a Low-power BS, L_s.

(b) Load of a High-power BS, L_m.

Figure 2.3: The loads, L_s and L_m, versus the bias factor, A_s, for different values of λ_s.

corroborates the accuracy of (2.14) and (2.15). Moreover, we observe that the limit value of P_{t_m} as A_s grows large for $\lambda_s = 10^{-3}/\mathrm{m}^2$ is lower than that for $\lambda_s = 2 \times 10^{-4}/\mathrm{m}^2$. This can be analytically explained based on (2.14). Specifically, we find from (2.14) that $\lim_{A_s \to \infty} P_{t_m} = B_m \left(1 - B_s\right)$. When λ_s increases, there are more LoS low-power BSs available in the network and thus, B_s increases and $\lim_{A_s \to \infty} P_{t_m}$ decreases.

Second, we examine the impact of the LoS low-power BS density, λ_s, and the bias factor, A_s, on the loads of a low-power BS and a high-power BS, L_s and L_m, in Fig. 2.3. We observe that L_s increases when A_s increases, while L_m decreases when A_s increases. Specifically, when $\lambda_s = 10^{-3}/\mathrm{m}^2$ and A_s increases from 1 to 10 and then to 100, L_s increases from 20.84 to 38.91 and then to 48.26 and L_m decreases from 3492.29 to 1328.60 and then to 205.42. Moreover, we observe that L_s and L_m decrease when λ_s increases. This is due to the fact that, given that the number of UEs is constant, the load of each BS decreases when the number of low-power



(a) Load of a Low-power BS, L_s.

(b) Load of a High-power BS, L_m.

Figure 2.4: The loads, L_s and L_m, versus the bias factor, A_s, for different values of λ_m.

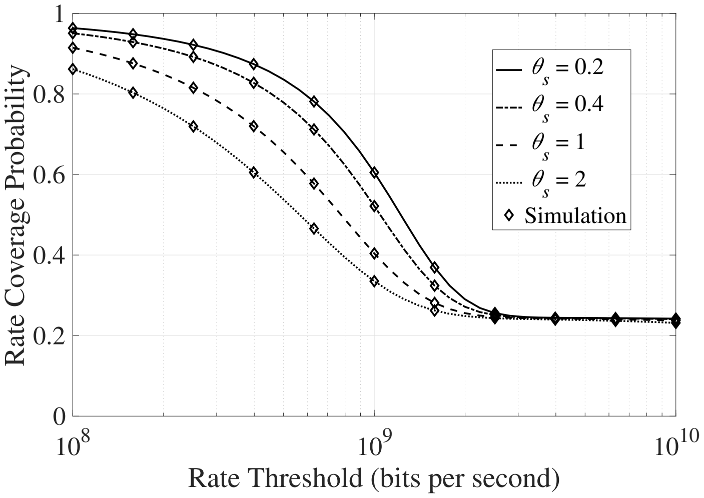


Figure 2.5: The rate coverage probability, P_c, versus the rate threshold, δ, for different values of θ_s, when $\lambda_s = 8 \times 10^{-2}/\mathrm{m}^2$ and $\lambda_m = 10^{-3}/\mathrm{m}^2$.

BSs increases.

Third, we examine the impact of the LoS high-power BS density, λ_m, and the bias factor, A_s, on the loads of a low-power BS and a high-power BS, L_s and L_m, in Fig. 2.4. Again, we observe that L_s increases when A_s increases, while L_m decreases when A_s increases. Specifically, when $\lambda_m = 5 \times 10^{-5}/\mathrm{m}^2$ and A_s increases from 1 to 10 and then to 100, L_s increases from 4.02 to 20.58 and then to 42.47 and L_m decreases from 1101.81 to 705.85 and then to 179.97. Moreover, we observe that L_s and L_m decrease when λ_m increases. This is due to the fact that, given that the number of UEs is constant, the load of each BS decreases when the number of high-power BSs increases.

Fourth, we examine the impact of θ_s and θ_m on the rate coverage probability, P_c. We first validate our analytical expression for P_c by observing the exact match between the analytical curves and the Monte Carlo simulation points marked by '$\diamond$' in Fig. 2.5 and Fig. 2.6. More-

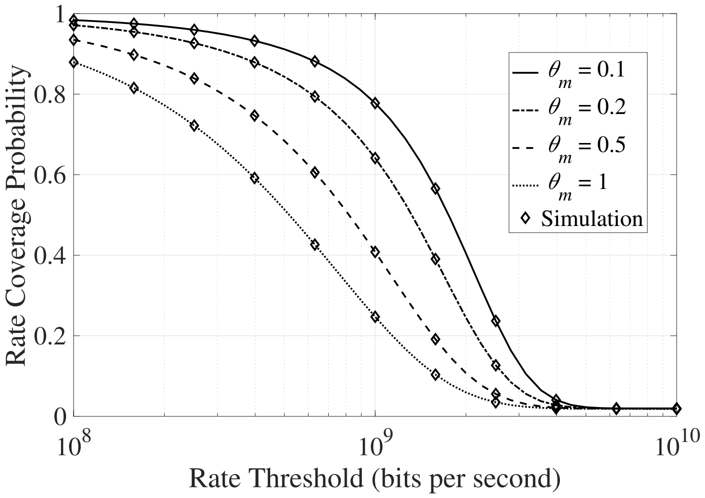


Figure 2.6: The rate coverage probability, P_c, versus the rate threshold, δ, for different values of θ_m, when $\lambda_s = 10^{-2}/\mathrm{m}^2$, $\lambda_m = 2 \times 10^{-3}/\mathrm{m}^2$, and $\mu_s = 100$ m.

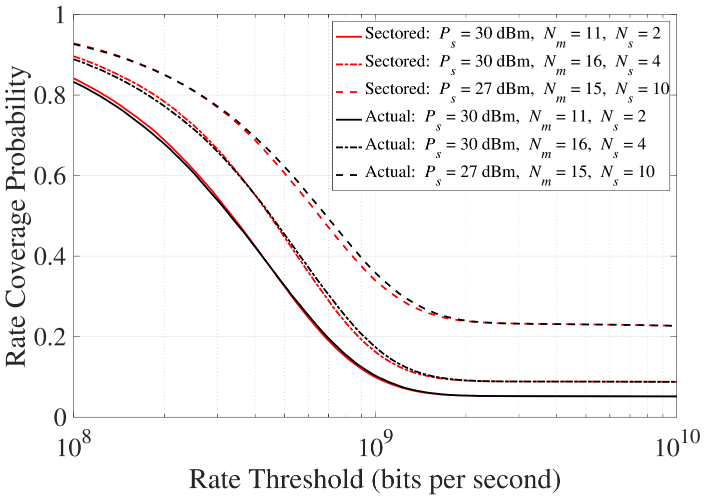


Figure 2.7: The rate coverage probability based on the sectored gain model and the rate coverage probability based on the actual gain, versus the rate threshold, δ, for different values of P_s, N_m, and N_s, where N_m and N_s are the numbers of elements in the uniform linear array at each high-power BS and low-power BS, respectively.

over, we observe that the rate coverage probability based on the sectored gain model closely approximates the rate coverage probability based on the actual gain in Fig. 2.7. Furthermore, we observe that P_c improves when the beam becomes narrower. Specifically, Fig. 2.5 illustrates that P_c increases when θ_s decreases and Fig. 2.6 illustrates that P_c increases when θ_m decreases. This is expected since the narrower the beam, the less interference caused by BSs.

Fifth, we examine the impact of λ_s and λ_m on the rate coverage probability, P_c. We observe from Fig. 2.8 that deploying more low-power BSs profoundly improves the rate coverage

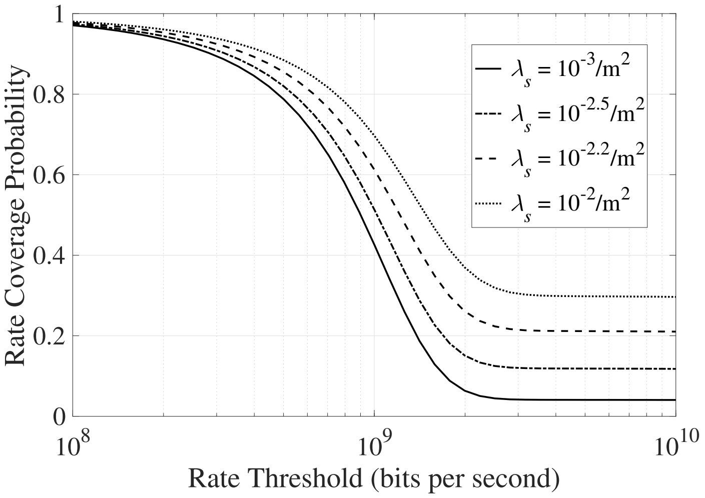


Figure 2.8: The rate coverage probability, P_c, versus the rate threshold, δ, for different values of λ_s, when $\lambda_m = 10^{-4}/\mathrm{m}^2$, $\lambda_u = 10^{-2}/\mathrm{m}^2$, $\mu_m = 1000$ m, $\mu_s = 100$ m, and $\theta_s = 0.1$ rad.

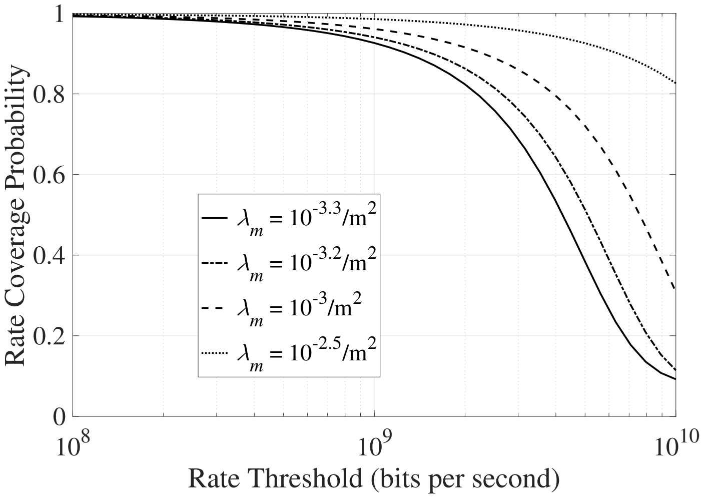


Figure 2.9: The rate coverage probability, P_c, versus the rate threshold, δ, for different values of λ_m, when $\lambda_s = 10^{-2}/\mathrm{m}^2$, $\lambda_u = 10^{-2}/\mathrm{m}^2$, $\mu_m = 1000$ m, $\mu_s = 100$ m, and $\theta_s = 0.1$ rad.

performance when the beam is narrow. Specifically, for $\delta = 10^9$ bits per second, P_c increases from 0.43 to 0.51 and then to 0.70 when λ_s increases from $10^{-3}/\mathrm{m}^2$ to $10^{-2.5}/\mathrm{m}^2$ and then to $10^{-2}/\mathrm{m}^2$. Moreover, we observe from Fig. 2.9 that deploying more high-power BSs also improves the rate coverage performance. Specifically, for $\delta = 10^9$ bits per second, P_c increases from 0.93 to 0.96 and then to 0.99 when λ_m increases from $10^{-3.3}/\mathrm{m}^2$ to $10^{-3}/\mathrm{m}^2$ and then to $10^{-2.5}/\mathrm{m}^2$.

Finally, we examine the optimal value of A_s which maximizes P_c. We first observe from Fig. 2.10 and Fig. 2.11 that P_c with $A_s > 1$ is larger than P_c with $A_s = 1$, which confirms that the rate coverage performance is improved when offloading more UEs to low-power BSs.

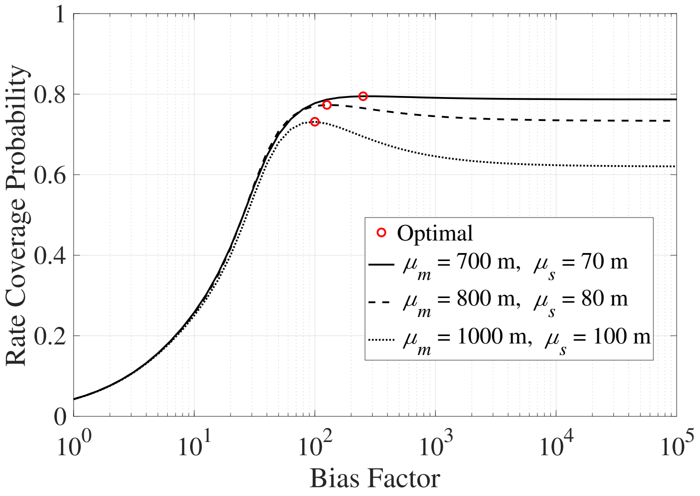


Figure 2.10: The rate coverage probability, P_c, versus the bias factor, A_s, for different values of μ_m and μ_s, when $\delta = 10^{6.5}$ bits per second, P_m = 40 dBm, P_s = 20 dBm, $\lambda_s = 10^{-4}/\text{m}^2$, $S_m = 10^9$ Hz, and $S_s = 10^9$ Hz.

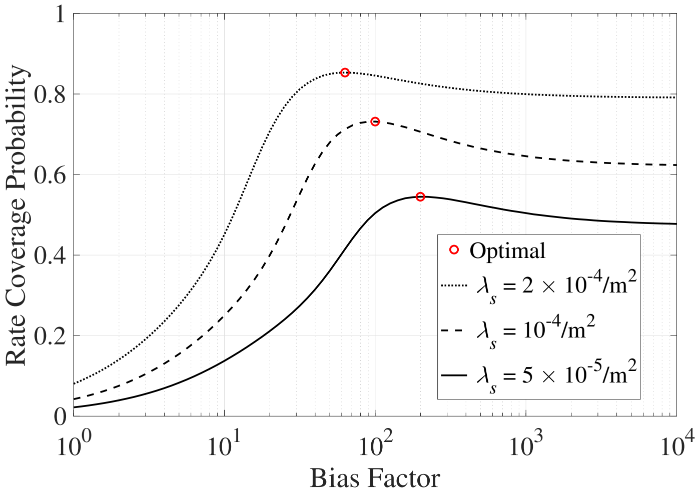


Figure 2.11: The rate coverage probability, P_c, versus the bias factor, A_s, for different values of λ_s, when $\delta = 10^{6.5}$ bits per second, P_m = 40 dBm, P_s = 20 dBm, μ_m = 1000 m, μ_s = 100 m, $S_m = 10^9$ Hz, and $S_s = 10^9$ Hz.

Moreover, we observe that P_c first increases and then decreases when A_s increases. This implies the existence of the optimal value of A_s which maximizes P_c. Particularly, the optimal A_s can be numerically found using our analysis. Furthermore, we observe from Fig. 2.11 that the optimal A_s decreases when λ_s increases. This is due to the fact that deploying more low-power BSs helps to offload more UEs to low-power BSs, as illustrated in Fig. 2.2.

2.4 Summary

In this chapter, we examined the impact of load balancing on the rate coverage performance in a two-tier mmWave cellular HetNet. Importantly, the unique features of mmWave networks, e.g., the extremely narrow beams at BSs and the vulnerability of networks to blocking, were addressed by adopting the sectored antenna gain model and the LoS ball model, respectively. For the considered mmWave cellular HetNet, we first analyzed the loads of the macro BS and the low-power BS. Then we derived a new exact expression for the rate coverage probability experienced by the typical UE in the network. This expression facilitated us to identify the impact of the bias factor on the rate coverage probability. Using numerical results, we demonstrated the correctness of our analytical expressions and evaluated the impact of various network parameters on the coverage performance of the network.

Chapter 3

Coverage Analysis of Relay Assisted mmWave Networks with Spatial Correlation

As mentioned in Chapter 1.1, one of the most pressing challenges for mmWave cellular networks is to enhance the coverage performance, due to the fact that mmWave is extremely sensitive to blockage. One promising strategy to enhance the coverage performance is to introduce relays to circumvent the blockage. Thus, in this chapter, we consider a relay assisted mmWave cellular network. Specifically, we allow idle UEs to function as relays and analyze the coverage performance of the network. Similar to Chapter 2, we model the system using the stochastic geometry based approach described in Chapter 1.1 to obtain tractable analytical results on the system-level coverage probability. Based on the system model, we propose a new analytical framework to derive the SINR coverage probability of the network. In this framework, we consider: 1) generalized Nakagami-$\mathcal{M}$ fading, 2) interference, 3) selection combining at destination UEs, and 4) *spatial correlation*. The accuracy of our analytical results is validated using numerical results. We also show the benefit of introducing relays for coverage performance enhancement. Furthermore, we find that the optimal BS density maximizing the SINR coverage probability can be determined through our analysis.

This chapter is organized as follows: Subchapter 3.1 describes the system model of the considered relay assisted mmWave cellular network. Subchapter 3.2 presents the analysis of the SINR coverage probability of the considered network. Subchapter 3.3 presents numerical results to evaluate the impact of spatial correlation and network parameters on the SINR coverage probability. Finally, Subchapter 3.4 presents a summary of this chapter.

3.1 Relay Assisted mmWave Cellular Network

In this chapter, we consider a relay assisted mmWave cellular network where BSs communicate with UEs. In this network, we assume that there are three types of UEs:

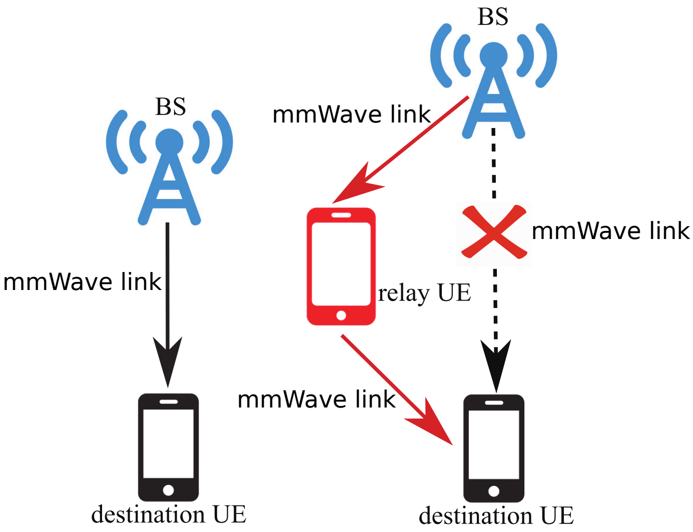


Figure 3.1: The two transmission modes of the considered relay assisted mmWave cellular network.

- Destination UEs which are the UEs that require downlink data from BSs.
- Uplink UEs which are the UEs that transmit uplink data to BSs.
- Relay UEs which are idle UEs that function as relays (or equivalently, D2D transmitters) to help the transmission between a BS and a destination UE.

As depicted in Fig. 3.1, there are two transmission modes in the considered network:

- Direct mode where the BS transmits data to the destination UE directly.
- Relay mode where the BS first transmits data to the relay UE and then the relay UE forwards data to the destination UE.

Specifically, the relay mode serves as a backup option when the direct mode fails, i.e., the SINR of the direct link from the BS to the destination UE is below a given SINR threshold.

3.1.1 Locations of BSs and UEs

We assume that the locations of BSs, relay UEs, and destination UEs follow three PPPs on $\mathbb{R}^2$, the densities of which are denoted by $\hat{\lambda}_b$, $\hat{\lambda}_r$, and $\hat{\lambda}_d$, respectively. In this chapter, we randomly select one destination UE, refer to it as the typical destination UE, and establish a polar coordinate system with the typical destination UE at the origin, $\mathbf{o}$. Based on the Slivnyak theorem [58], the conclusions drawn for the typical destination UE can be extended to other destination UEs.

We note that mmWave communications is extremely sensitive to blockage [4]. Moreover, the diffraction of mmWave propagation is weak due to the used high frequencies. Thus, in this chapter we only consider LoS links as in [123, 134]. To model the spatial distribution of LoS BSs and LoS relay UEs, we adopt an accurate and simple LoS ball model as in [48, 116, 123]. As per this model, the probability of a channel being LoS, $P_{\mathrm{LoS}}(\ell)$, is a function of the distance between the transmitter and the receiver, ℓ. Mathematically, $P_{\mathrm{LoS}}(\ell)$ is given by

$$P_{\mathrm{LoS}}(\ell) = \begin{cases} \rho_b, & \text{if } 0 < \ell < r_b, \\ 0, & \text{otherwise,} \end{cases} \tag{3.1}$$

for BSs, where r_b is the radius of the BS LoS ball and $0 \leq \rho_b \leq 1$ is the LoS probability of BSs. Similarly, $P_{\mathrm{LoS}}(\ell)$ is given by

$$P_{\mathrm{LoS}}(\ell) = \begin{cases} \rho_u, & \text{if } 0 < \ell < r_u, \\ 0, & \text{otherwise,} \end{cases} \tag{3.2}$$

for relay UEs, where r_u is the radius of the relay UE LoS ball and $0 \leq \rho_u \leq 1$ is the LoS probability of relay UEs. According to the thinning theorem [127, Prop. (1.3.5)], the LoS BSs follow a PPP Φ_b with density $\lambda_b \triangleq \rho_b \hat{\lambda}_b$ within the circular area $\mathcal{B}(\mathbf{o}, r_b)$. Similarly, the LoS relay UEs follow another PPP Φ_r with density $\lambda_r \triangleq \rho_u \hat{\lambda}_r$ within the circular area $\mathcal{B}(\mathbf{o}, r_u)$. We denote the set of LoS BSs by $\Phi_b = \{b_0, b_1, b_2, \ldots\}$, where b_0 is the associated BS, b_k is the kth interfering BS, and $k \in \{1, 2, \ldots\}$.

3.1.2 Directional Gains

We assume that an N_b element uniform linear antenna array is adopted at each BS and an N_u element uniform linear antenna array is adopted at each UE. At BSs, highly directional beams are assumed to be adopted to combat the high path loss of mmWave transmission. Then we note that destination UEs often experience rapid rotational movements. Thus, we assume that the typical destination UE does not leverage directional beams to avoid the beam misalignment caused by rapid rotational movements. Instead, it adopts selection combining to choose the signal with the maximum SINR, out of N_u received SINRs at N_u antennas. Furthermore, we assume that relay UEs leverage directional beams since relay UEs are idle such that they do not experience rapid rotational movements.

To characterize the mmWave transmission in the network, we adopt the sectored directional gain model [48, 49, 126] which incorporates key characteristics including the beamwidth, main lobe gain, and side lobe gain. Under the assumption that an N_b element uniform linear antenna array is adopted at each BS, the actual antenna array gain pattern of b_k is given by

$$G_{b_k,act}(\varepsilon) = \frac{\sin^2\left(N_b \frac{\zeta_b}{2}\right)}{\sin^2\left(\frac{\zeta_b}{2}\right)}, \tag{3.3}$$

where $\zeta_b = \frac{2\pi}{\lambda} d_b \cos(\varepsilon) + \beta_b$, λ is the wavelength, d_b is the distance between elements, ε is the angle, and β_b is the relative phase between elements [123]. Its corresponding sectored antenna model G_{b_k} is given by

$$G_{b_k} = \begin{cases} G_B, & \text{if } |\theta_o| \leq \frac{1}{2}\theta_b, \\ G_b, & \text{otherwise,} \end{cases} \tag{3.4}$$

where the main lobe gain, G_B, the side lobe gain, G_b, and the beamwidth, θ_b, are assumed to be the maximum gain, the first minor maximum gain, and the half power beamwidth of (3.3), respectively, and θ_o is the angle off the boresight direction [123].

Similarly, under the assumption that an N_u element uniform linear antenna array is adopted at each relay UE, the actual antenna array gain pattern of each relay UE is given by

$$G_{e,act}(\varepsilon) = \frac{\sin^2\left(N_u \frac{\zeta_u}{2}\right)}{\sin^2\left(\frac{\zeta_u}{2}\right)}, \tag{3.5}$$

where $\zeta_u = \frac{2\pi}{\lambda} d_u \cos(\varepsilon) + \beta_u$, d_u is the distance between elements, and β_u is the relative phase between elements [123]. Its corresponding sectored antenna model G_e is given by

$$G_e = \begin{cases} G_U, & \text{if } |\theta_o| \leq \frac{1}{2}\theta_u, \\ G_u, & \text{otherwise,} \end{cases} \tag{3.6}$$

where the main lobe gain, G_U, the side lobe gain, G_u, and the beamwidth, θ_u, are assumed to be the maximum gain, the first minor maximum gain, and the half power beamwidth of (3.5), respectively [123].

Here, we model θ_o as a uniform random variable between $-\pi$ and π. Thus, the probability mass function of G_{b_k} is given by

$$G_{b_k} = \begin{cases} G_B, & \text{with probability (w. p.) } \frac{\theta_b}{2\pi}, \\ G_b, & \text{w. p. } 1 - \frac{\theta_b}{2\pi}. \end{cases} \tag{3.7}$$

Similarly, the probability mass function of the directional gain of relay UEs, G_e, is given by

$$G_e = \begin{cases} G_U, & \text{w. p. } \frac{\theta_u}{2\pi}, \\ G_u, & \text{w. p. } 1 - \frac{\theta_u}{2\pi}. \end{cases} \tag{3.8}$$

Furthermore, we define the directional gain from b_k to the relay UE as $G_{b_k r} \triangleq G_{b_k} G_e$. Thus, the probability mass function of $G_{b_k r}$ is given by

$$G_{b_k r} = \begin{cases} G_B G_U, & \text{w. p. } \frac{\theta_b \theta_u}{4\pi^2}, \\ G_B G_u, & \text{w. p. } \frac{\theta_b}{2\pi}\left(1 - \frac{\theta_u}{2\pi}\right), \\ G_b G_U, & \text{w. p. } \left(1 - \frac{\theta_b}{2\pi}\right)\frac{\theta_u}{2\pi}, \\ G_b G_u, & \text{w. p. } \left(1 - \frac{\theta_b}{2\pi}\right)\left(1 - \frac{\theta_u}{2\pi}\right). \end{cases} \tag{3.9}$$

3.1.3 Association Strategy

We assume that all BSs have the same transmit power, P_b, and all relay UEs and uplink UEs have the same transmit power, P_u. We then assume that the maximum signal strength association strategy is applied in the network. Thus, in the direct mode, the destination UE always associates with its nearest LoS BS, while in the relay mode, the destination UE always associates with its nearest LoS relay UE and the associated relay UE always associates with the associated relay UE's nearest LoS BS [118]. Thus, throughout this chapter, the link from the associated BS to the typical destination UE is referred to as the *direct link*, the link from the associated BS to the associated relay UE is referred to as the *BR link*, and the link from the associated relay UE to the typical destination UE is referred to as the *RD link*. We denote the distances of the direct link, the BR link, and the RD link by ℓ_{bd}, ℓ_{br}, and ℓ_{rd}, respectively. If $\Phi_b \neq \emptyset$, the PDF of ℓ_{bd} is the same as the PDF of ℓ_{br}, which is given by

$$f_{\ell_{bd}}(x) = f_{\ell_{br}}(x) = \frac{2\pi\lambda_b x e^{-\pi\lambda_b x^2}}{\Xi_b}, \tag{3.10}$$

where $\Xi_b \triangleq \Pr(\Phi_b \neq \emptyset) = 1 - e^{-\pi\lambda_b r_b^2}$ and $0 < x < r_b$ [127]. Similarly, if $\Phi_r \neq \emptyset$, the PDF of ℓ_{rd} is given by

$$f_{\ell_{rd}}(x) = \frac{2\pi\lambda_r x e^{-\pi\lambda_r x^2}}{\Xi_r}, \tag{3.11}$$

where $\Xi_r \triangleq \Pr(\Phi_r \neq \emptyset) = 1 - e^{-\pi\lambda_r r_u^2}$ and $0 < x < r_u$.

3.1.4 SINR Formulation

In this subchapter, we first characterize the interference from the BS to the destination UE, the interference from the BS to the relay UE, and the interference from the relay UE to the destination UE. Then we formulate the SINRs between two nodes based on the interference characterization.

We assume that frequency division duplex is adopted, i.e., downlink frequency band and uplink frequency band are non-overlapping. Thus, the transmission using the downlink frequency band and the transmission using the uplink frequency band do not interfere with each other. In this network, we clarify that there are three types of transmitters:

- BSs that transmit to destination UEs in the direct mode and relay UEs in the relay mode.
- Relay UEs that transmit to destination UEs.
- Uplink UEs that transmit uplink data to BSs.

Similar to [118], we assume that BSs transmit using the downlink frequency band, while relay UEs and uplink UEs transmit using the uplink frequency band. Thus, the transmission from BSs and the transmissions from relay UEs and uplink UEs do not interfere with each other, but the transmission from relay UEs and the transmission from uplink UEs interfere with each other. It follows that in the direct link and the BR link, the interference solely comes from other BSs. Differently, in the RD link, the interference comes from not only other relay UEs but also uplink UEs.

We emphasize that not all relay UEs and uplink UEs interfere with the RD link. It is widely acknowledged that in a cellular network, network resources must be divided into multiple sub-channels to enable multiple access. In this chapter, we assume that the UEs using the same sub-channel interfere with each other, while the UEs using different sub-channels do not. Thus, in the RD link, only the relay UEs and uplink UEs which use the same sub-channel as the associated relay UE cause interference. In the following, we define all interfering relay UEs and uplink UEs as interfering UEs. In order to model the distribution of all interfering UEs, we define a multiplexing factor, ρ, which represents the average number of interfering UEs on each sub-channel within the cell [118]. Considering the distribution of BSs, the distribution of LoS interfering UEs can be modeled by a homogeneous PPP with density $\lambda_i \triangleq \rho_u \rho \hat{\lambda}_b$ within the circular area $\mathcal{B}(\mathbf{o}, r_u)$, $\Phi_i = \{i_1, i_2, \ldots\}$, where i_k is the kth interfering UE and $k \in \{1, 2, \ldots\}$.

Based on the aforementioned interference characterization, we obtain that the SINR at antenna n in the direct link, denoted by $\gamma_{bd,n}$ where $n \in \{1, 2, \ldots, N_u\}$, is written as

$$\gamma_{bd,n} = \frac{P_b G_B \ell_{bd}^{-\eta} |h_{nb_0}|^2}{\sum\limits_{b_k \in \tilde{\Phi}_b} P_b G_{b_k} \ell_{b_k d}^{-\eta} |h_{nb_k}|^2 + \sigma^2}, \tag{3.12}$$

where $\tilde{\Phi}_b \triangleq \Phi_b \backslash \{b_0\}$ denotes the set of LoS BSs excluding b_0, $\ell_{b_k d}$ is the distance between b_k and the typical destination UE, $\eta \geq 2$ is the path loss exponent, h_{nb_l} is the small scale fading from b_l to antenna n, $l \in \{0, k\}$, and σ^2 is the variance of the additive white Gaussian noise (AWGN) at antenna n. In this chapter, we assume that h_{nb_l} is subject to i.i.d. Nakagami-$\mathcal{M}$ fading. By defining $\hbar_{nb_l} \triangleq |h_{nb_l}|^2$, we find that $\hbar_{nb_l}$ follows the Gamma distribution with the shape parameter $\mathcal{M}_{bd}$ and the scale parameter $\frac{1}{\mathcal{M}_{bd}}$, i.e., $\hbar_{nb_l} \sim \Gamma\left(\mathcal{M}_{bd}, \frac{1}{\mathcal{M}_{bd}}\right)$. Under the assumption that $\mathcal{M}_{bd}$ is a positive integer [49, 123], the MGF of $\hbar_{nb_l}$ is given by

$$M_{\hbar_{nb_l}}(s) = \left(1 - \frac{s}{\mathcal{M}_{bd}}\right)^{-\mathcal{M}_{bd}}. \tag{3.13}$$

Moreover, as given in [135, Lemma 1], the CDF of $\hbar_{nb_l}$ can be tightly lower bounded by

$$F_{\hbar_{nb_l}}(x) = 1 - e^{-\mathcal{M}_{bd}x} \sum_{\omega=0}^{\mathcal{M}_{bd}-1} \frac{(\mathcal{M}_{bd}x)^{\omega}}{\omega!} > \left(1 - e^{-\alpha_{bd}x}\right)^{\mathcal{M}_{bd}}, \tag{3.14}$$

where $\alpha_{bd} = \mathcal{M}_{bd}\left(\mathcal{M}_{bd}!\right)^{-\frac{1}{\mathcal{M}_{bd}}}$.

In the BR link, the SINR at the relay UE, denoted by γ_{br}, is written as

$$\gamma_{br} = \frac{P_b G_B G_U \ell_{br}^{-\eta} \hbar_{b_0}}{\sum\limits_{b_k \in \tilde{\Phi}_b} P_b G_{b_k r} \ell_{b_k r}^{-\eta} \hbar_{b_k} + \sigma^2}, \tag{3.15}$$

where $\ell_{b_k r}$ is the distance between b_k and the relay UE, $\hbar_{b_l}$ is the small scale fading gain between b_l and the relay UE, $l \in \{0, k\}$. In this chapter, we assume that $\hbar_{b_l}$ is subject to i.i.d. Gamma distribution with the shape parameter $\mathcal{M}_{br}$ and the scale parameter $\frac{1}{\mathcal{M}_{br}}$, i.e., $\hbar_{b_l} \sim \Gamma\left(\mathcal{M}_{br}, \frac{1}{\mathcal{M}_{br}}\right)$. Under the assumption that $\mathcal{M}_{br}$ is a positive integer [49, 123], the MGF of $\hbar_{b_l}$ is given by

$$M_{\hbar_{b_l}}(s) = \left(1 - \frac{s}{\mathcal{M}_{br}}\right)^{-\mathcal{M}_{br}}. \tag{3.16}$$

Moreover, as given in [135, Lemma 1], the CDF of $\hbar_{b_l}$ can be tightly lower bounded by

$$F_{\hbar_{b_l}}(x) = 1 - e^{-\mathcal{M}_{br}x} \sum_{\omega=0}^{\mathcal{M}_{br}-1} \frac{(\mathcal{M}_{br}x)^{\omega}}{\omega!} > \left(1 - e^{-\alpha_{br}x}\right)^{\mathcal{M}_{br}}, \tag{3.17}$$

where $\alpha_{br} = \mathcal{M}_{br}\left(\mathcal{M}_{br}!\right)^{-\frac{1}{\mathcal{M}_{br}}}$.

In the RD link, the SINR at antenna n, $n \in \{1, 2, \ldots, N_u\}$, denoted by $\gamma_{rd,n}$, is written as

$$\gamma_{rd,n} = \frac{P_u G_U \ell_{rd}^{-\eta} \hbar_{nr}}{\sum\limits_{i_k \in \Phi_i} P_u G_e \ell_{i_k d}^{-\eta} \hbar_{ni_k} + \sigma^2}, \tag{3.18}$$

where $\ell_{i_k d}$ is the distance between the kth interfering UE and the destination UE, $\hbar_{nr}$ is the small scale fading gain from the relay UE to antenna n, $\hbar_{ni_k}$ is the small scale fading from i_k to antenna n. In this chapter, we assume that $\hbar_{nr}$ and $\hbar_{ni_k}$ are subject to i.i.d. Gamma distribution with the shape parameter $\mathcal{M}_{rd}$ and the scale parameter $\frac{1}{\mathcal{M}_{rd}}$, i.e., $\hbar_{nr} \sim \Gamma\left(\mathcal{M}_{rd}, \frac{1}{\mathcal{M}_{rd}}\right)$ and $\hbar_{ni_k} \sim \Gamma\left(\mathcal{M}_{rd}, \frac{1}{\mathcal{M}_{rd}}\right)$. Under the assumption that $\mathcal{M}_{rd}$ is a positive integer [49, 123], the MGF of $\hbar_{nr}$ and $\hbar_{ni_k}$ is given by

$$M_{\hbar_{nr}}(s) = M_{\hbar_{ni_k}}(s) = \left(1 - \frac{s}{\mathcal{M}_{rd}}\right)^{-\mathcal{M}_{rd}}. \tag{3.19}$$

Moreover, as given in [135, Lemma 1], the CDF of $\hbar_{nr}$ and $\hbar_{ni_k}$ can be tightly lower bounded by

$$F_{\hbar_{nr}}(x) = F_{\hbar_{ni_k}}(x) = 1 - e^{-\mathcal{M}_{rd}x} \sum_{\omega=0}^{\mathcal{M}_{rd}-1} \frac{(\mathcal{M}_{rd}x)^{\omega}}{\omega!} > \left(1 - e^{-\alpha_{rd}x}\right)^{\mathcal{M}_{rd}}, \tag{3.20}$$

where $\alpha_{rd} = \mathcal{M}_{rd}(\mathcal{M}_{rd}!)^{-\frac{1}{\mathcal{M}_{rd}}}$. In the analysis presented in Subchapter 3.2, we adopt the lower bounds of $F_{\hbar_{nb_l}}(x)$, $F_{\hbar_{b_l}}(x)$, $F_{\hbar_{ni_k}}(x)$, and $F_{\hbar_{nr}}(x)$, since these lower bounds enable us to derive the SINR coverage probability when the Nakagami-$\mathcal{M}$ fading parameters are higher than one.

By observing (3.12) and (3.18), we find that the interference signals at the N_u UE antennas are correlated due to the common interfering BSs' location or the common interfering UEs' location. Also, the desired signals at the N_u UE antennas are correlated due to the common associated BS' location or the common associated relay UE's location. Despite that both interference signals and desired signals are correlated in wireless networks, these spatial correlations are often ignored in the existing studies on mmWave systems to simplify the analysis. To overcome this, we address the spatial correlation when analyzing the coverage performance of the network considered in this chapter.

3.2 SINR Coverage Probability Analysis

In this subchapter, we derive the SINR coverage probability of the considered network, $\mathbb{P}_{N_u}(\tau)$. Here, $\mathbb{P}_{N_u}(\tau)$ is defined as the probability that the SINR at the N_u-antenna typical destination UE is higher than a given SINR threshold, τ. As aforementioned, we derive $\mathbb{P}_{N_u}(\tau)$ by considering spatial correlation at the destination UE. This mandates a new and challenging analytical framework for the SINR coverage probability, as pointed out in [136]. Furthermore, we aim to answer an open question in mmWave cellular networks: "*Is the analysis error caused by ignoring spatial correlation acceptable?*" To answer this question, we also derive the SINR coverage probability for the case where spatial correlation is ignored, which enables us to compare the two coverage probabilities in Subchapter 3.3.

We assume that the half-duplex DF relay protocol is adopted. Thus, in the relay mode, the SINR at the typical destination UE is higher than τ when both the SINR of the BR link and the SINR of the RD link are higher than τ. We denote the SINR coverage probability of the direct link, the BR link, and the RD link as $\mathbb{P}_{N_u,bd}(\tau)$, $\mathbb{P}_{br}(\tau)$, and $\mathbb{P}_{N_u,rd}(\tau)$, respectively.

This allows us to express $\mathbb{P}_{N_u}(\tau)$ as

$$\mathbb{P}_{N_u}(\tau) = 1 - \left[1 - \mathbb{P}_{N_u,bd}(\tau)\right]\left[1 - \mathbb{P}_{br}(\tau)\mathbb{P}_{N_u,rd}(\tau)\right]. \tag{3.21}$$

We next derive and present $\mathbb{P}_{N_u,bd}(\tau)$, $\mathbb{P}_{br}(\tau)$, and $\mathbb{P}_{N_u,rd}(\tau)$ in Theorems 3.1, 3.2, and 3.3, respectively.

Theorem 3.1. Under the assumption that $F_{\hbar_{nb_l}}(x)$ is approximated as $\left(1 - e^{-\alpha_{bd}x}\right)^{\mathcal{M}_{bd}}$, the SINR coverage probability of the direct link is derived as

$$\begin{aligned}\mathbb{P}_{N_u,bd}(\tau) =& \Xi_b \sum_{\kappa=1}^{N_u} (-1)^{\kappa+1} \binom{N_u}{\kappa} \sum_{j_1+j_2+\cdots+j_{\mathcal{M}_{bd}}=\kappa} \beta_{bd} \\ &\times \int_0^{r_b} e^{-\frac{\tau\alpha_{bd}\Omega_{bd}\sigma^2}{P_b G_B x^{-\eta}} - 2\pi\lambda_b \int_x^{r_b} \ell(1-v_{bd}(\ell))d\ell} f_{\ell_{bd}}(x)\,dx,\end{aligned} \tag{3.22}$$

where $j_1, j_2, \ldots, j_{\mathcal{M}_{bd}}$ are nonnegative integers, Ω_{bd} is defined as $\Omega_{bd} \triangleq j_1 + 2j_2 + \cdots + \mathcal{M}_{bd} j_{\mathcal{M}_{bd}}$, δ_{bd} is defined as

$$\delta_{bd} \triangleq \frac{\alpha_{bd}\tau}{P_b G_B \ell_{bd}^{-\eta}}, \tag{3.23}$$

β_{bd} is defined as

$$\begin{aligned}\beta_{bd} \triangleq& \binom{\kappa}{j_1}\binom{\kappa - j_1}{j_2} \cdots \binom{\kappa - j_1 - \cdots - j_{\mathcal{M}_{bd}-1}}{j_{\mathcal{M}_{bd}}} \binom{\mathcal{M}_{bd}}{1}^{j_1} \\ &\times \binom{\mathcal{M}_{bd}}{2}^{j_2} \cdots \binom{\mathcal{M}_{bd}}{\mathcal{M}_{bd}}^{j_{\mathcal{M}_{bd}}} (-1)^{j_1+j_2+\cdots+j_{\mathcal{M}_{bd}}+\Omega_{bd}},\end{aligned} \tag{3.24}$$

and $v_{bd}(\ell)$ is defined as

$$v_{bd}(\ell) \triangleq \frac{\theta_b}{2\pi} \prod_{q=1}^{\mathcal{M}_{bd}} \left(1 + \frac{q\delta_{bd} P_b G_B \ell^{-\eta}}{\mathcal{M}_{bd}}\right)^{-\mathcal{M}_{bd} j_q} + \left(1 - \frac{\theta_b}{2\pi}\right) \prod_{q=1}^{\mathcal{M}_{bd}} \left(1 + \frac{q\delta_{bd} P_b G_b \ell^{-\eta}}{\mathcal{M}_{bd}}\right)^{-\mathcal{M}_{bd} j_q}. \tag{3.25}$$

Proof: See Appendix A.1. ■

Theorem 3.2. Under the assumption that the CDF of $\hbar_{b_l}$ is approximated as $\left(1 - e^{-\alpha_{br}x}\right)^{\mathcal{M}_{br}}$, the SINR coverage probability of the BR link is derived as

$$\mathbb{P}_{br}(\tau) = \Xi_b \sum_{j=1}^{\mathcal{M}_{br}} (-1)^{j+1} \binom{\mathcal{M}_{br}}{j} \int_0^{r_b} e^{-\psi x^{\eta}\sigma^2 - 2\pi\lambda_b \int_x^{r_b} \ell\Omega_b d\ell} f_{\ell_{br}}(x)\,dx, \tag{3.26}$$

where $\psi \triangleq \frac{j\alpha_{br}\tau}{P_b G_B G_U}$ and Ω_b is defined as

$$\begin{aligned}\Omega_b \triangleq & \frac{\theta_u}{2\pi}\left(1-\frac{\theta_b}{2\pi}\right)\left[1-\left(1+\frac{\psi x^\eta P_b G_b G_U}{\mathcal{M}_{br}\ell^\eta}\right)^{-\mathcal{M}_{br}}\right] \\ & +\left(1-\frac{\theta_b}{2\pi}\right)\left(1-\frac{\theta_u}{2\pi}\right)\left[1-\left(1+\frac{\psi x^\eta P_b G_b G_u}{\mathcal{M}_{br}\ell^\eta}\right)^{-\mathcal{M}_{br}}\right] \\ & +\frac{\theta_b}{2\pi}\left(1-\frac{\theta_u}{2\pi}\right)\left[1-\left(1+\frac{\psi x^\eta P_b G_B G_u}{\mathcal{M}_{br}\ell^\eta}\right)^{-\mathcal{M}_{br}}\right] \\ & +\frac{\theta_b}{2\pi}\frac{\theta_u}{2\pi}\left[1-\left(1+\frac{\psi x^\eta P_b G_B G_U}{\mathcal{M}_{br}\ell^\eta}\right)^{-\mathcal{M}_{br}}\right]. \end{aligned} \tag{3.27}$$

Proof: See Appendix A.2. ■

Theorem 3.3. Under the assumption that $F_{\hbar_{nr}}(x)$ is approximated as $(1-e^{-\alpha_{rd}x})^{\mathcal{M}_{rd}}$, the SINR coverage probability of the RD link is derived as

$$\begin{aligned}\mathbb{P}_{N_u,rd}(\tau) = & \Xi_r \sum_{\kappa=1}^{N_u} (-1)^{\kappa+1} \binom{N_u}{\kappa} \sum_{j_1+j_2+\cdots+j_{\mathcal{M}_{rd}}=\kappa} \beta_{rd} \\ & \times \int_0^{r_u} e^{-\frac{\tau\alpha_{rd}\Omega_{rd}\sigma^2}{P_u G_U x^{-\eta}} - 2\pi\lambda_i \int_0^{r_u} \ell(1-v_{rd}(\ell))d\ell} f_{\ell_{rd}}(x)\, dx, \end{aligned} \tag{3.28}$$

where $j_1, j_2, \ldots, j_{\mathcal{M}_{rd}}$ are nonnegative integers, Ω_{rd} is defined as

$$\Omega_{rd} \triangleq j_1 + 2j_2 + \cdots + \mathcal{M}_{rd} j_{\mathcal{M}_{rd}}, \tag{3.29}$$

δ_{rd} is defined as

$$\delta_{rd} \triangleq \frac{\alpha_{rd}\tau}{P_u G_U \ell_{rd}^{-\eta}}, \tag{3.30}$$

β_{rd} is defined as

$$\begin{aligned}\beta_{rd} \triangleq & \binom{\kappa}{j_1}\binom{\kappa-j_1}{j_2}\cdots\binom{\kappa-j_1-\cdots-j_{\mathcal{M}_{rd}-1}}{j_{\mathcal{M}_{rd}}}\binom{\mathcal{M}_{rd}}{1}^{j_1} \\ & \times \binom{\mathcal{M}_{rd}}{2}^{j_2}\cdots\binom{\mathcal{M}_{rd}}{\mathcal{M}_{rd}}^{j_{\mathcal{M}_{rd}}} (-1)^{j_1+j_2+\cdots+j_{\mathcal{M}_{rd}}+\Omega_{rd}}, \end{aligned} \tag{3.31}$$

and $v_{rd}(\ell)$ is defined as

$$v_{rd}(\ell) \triangleq \frac{\theta_u}{2\pi}\prod_{q=1}^{\mathcal{M}_{rd}}\left(1+\frac{q\delta_{rd}P_u G_U \ell^{-\eta}}{\mathcal{M}_{rd}}\right)^{-\mathcal{M}_{rd}j_q} + \left(1-\frac{\theta_u}{2\pi}\right)\prod_{q=1}^{\mathcal{M}_{rd}}\left(1+\frac{q\delta_{rd}P_u G_u \ell^{-\eta}}{\mathcal{M}_{rd}}\right)^{-\mathcal{M}_{rd}j_q}. \tag{3.32}$$

Table 3.1: Simulation Parameters Used in Subchapter 3.3

Parameter	Value
BS transmit power, P_b	35 dBm
UE transmit power, P_u	25 dBm
Noise power, σ^2	0 dBm
Number of antennas at each BS, N_b	10
Number of antennas at each UE, N_u	4
Path loss exponent, η	2.4
Nakagami-$\mathcal{M}$ fading parameters, $\mathcal{M}_{bd}$, $\mathcal{M}_{br}$, $\mathcal{M}_{rd}$	2
BS LoS ball radius, r_b	100 m
UE LoS ball radius, r_u	20 m
BS LoS probability, ρ_b	0.9
UE LoS probability, ρ_u	0.63
LoS BS density, λ_b	$0.0002/\mathrm{m}^2$
LoS relay UE density, λ_r	$0.002/\mathrm{m}^2$
Multiplexing factor, ρ	0.9

Proof: See Appendix A.3. ■

We now examine the case where spatial correlation is ignored. In this case, the SINRs received at the N_u antennas of the destination UE are assumed to be independent with each other. Thus, when spatial correlation is ignored, the SINR coverage probability of the direct link, $\hat{\mathbb{P}}_{N_u,bd}(\tau)$, is given by

$$\hat{\mathbb{P}}_{N_u,bd}(\tau) = 1 - (1 - \mathbb{P}_{1,bd}(\tau))^{N_u}, \tag{3.33}$$

and the SINR coverage probability of the RD link, $\hat{\mathbb{P}}_{N_u,rd}(\tau)$, is given by

$$\hat{\mathbb{P}}_{N_u,rd}(\tau) = 1 - (1 - \mathbb{P}_{1,rd}(\tau))^{N_u}. \tag{3.34}$$

Therefore, when spatial correlation is ignored, the probability that the SINR at the N_u-antenna typical destination UE is higher than τ, $\hat{\mathbb{P}}_{N_u}(\tau)$, is obtained as

$$\hat{\mathbb{P}}_{N_u}(\tau) = 1 - \left[1 - \hat{\mathbb{P}}_{N_u,bd}(\tau)\right]\left[1 - \mathbb{P}_{br}(\tau)\hat{\mathbb{P}}_{N_u,rd}(\tau)\right]. \tag{3.35}$$

The difference between (3.21) and (3.35) indicates that spatial correlation imposes a profound impact on the SINR coverage probability of the N_u-antenna destination UE. This will be further examined in Subchapter 3.3.

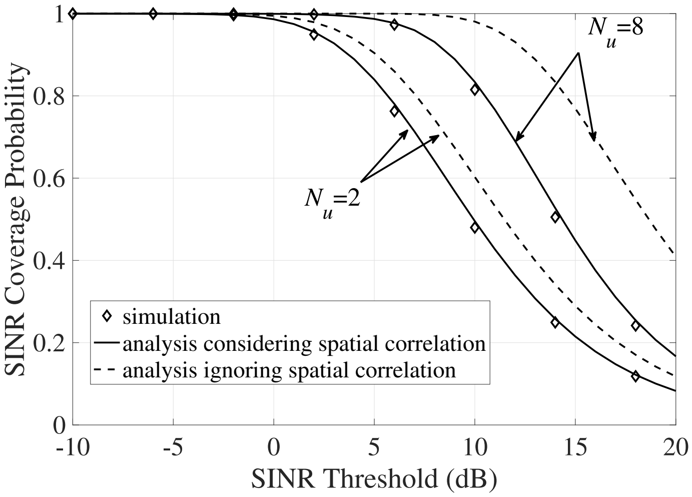


Figure 3.2: The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the SINR threshold, τ, for different values of N_u, when $\eta = 2$, $P_b = 30$ dBm, $P_u = 20$ dBm, and $N_b = 16$.

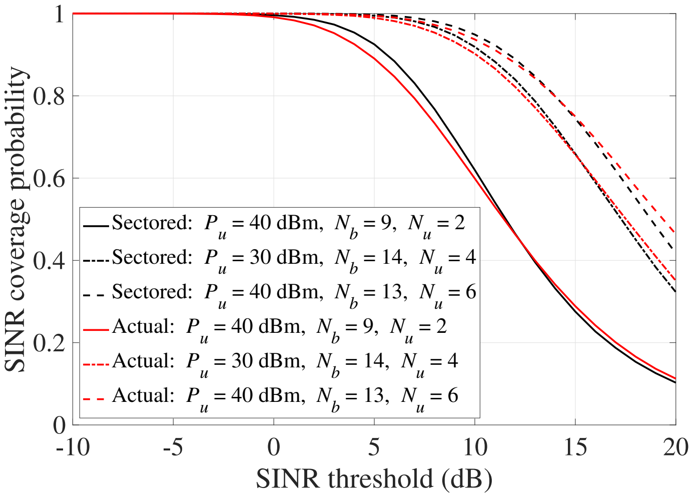


Figure 3.3: The SINR coverage probability based on the sectored gain model and the SINR coverage probability based on the actual gain, versus the SINR threshold, τ, for different values of P_u, N_b, and N_u.

3.3 Numerical Results and Discussions

In this subchapter, we present numerical results to evaluate the impact of spatial correlation and network parameters on the SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$. The number of samples in the Monte Carlo simulation in this Chapter is 10^6. The values of the parameters used in this subchapter are summarized in Table 3.1, unless otherwise specified.

First, we examine the impact of spatial correlation on $\mathbb{P}_{N_u}(\tau)$. We observe from Fig. 3.2

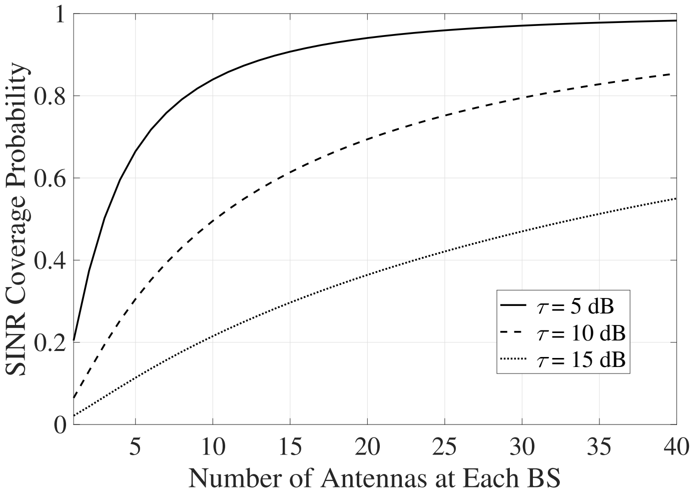


Figure 3.4: The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the number of antennas at each BS, N_b, for different values of τ, when $N_u = 2$.

that our analysis given by (3.21) matches the simulations very well, which confirms that (3.21) is an accurate approximation. Moreover, we observe that the SINR coverage probability based on the sectored gain model closely approximates the SINR coverage probability based on the actual gain in Fig. 3.3. Furthermore, we observe that ignoring spatial correlation leads to severe overestimation of the SINR coverage probability, which is evident when comparing the analysis ignoring spatial correlation, given by (3.35), with Monte Carlo simulation points. Specifically, when $\tau = 14$ dB and $N_u = 8$, $\hat{\mathbb{P}}_{N_u}(\tau)$ is 0.83 while the simulation shows that the actual SINR coverage probability is 0.51. Indeed, spatial correlation significantly reduces the probability of successful reception at the destination UE, thus cannot be ignored.

Second, we examine the impact of N_b and N_u on $\mathbb{P}_{N_u}(\tau)$. We observe from Fig. 3.4 and Fig. 3.5 that $\mathbb{P}_{N_u}(\tau)$ significantly increases when N_b increases. For example, we observe from Fig. 3.5 that when $\tau = 10$ dB and N_b increases from 4 to 8 and then to 16, $\mathbb{P}_{N_u}(\tau)$ increases from 0.25 to 0.43 and then to 0.63. This observation is expected because the beamwidth of BSs, given by $\theta_b = \frac{102\pi}{180N_b}$, decreases as N_b increases and BSs with narrower beams cause less interference. Moreover, we observe from Fig. 3.6 and Fig. 3.7 that $\mathbb{P}_{N_u}(\tau)$ significantly increases when N_u increases. For example, we observe from Fig. 3.7 that when $\tau = 10$ dB and N_u increases from 1 to 6 and then to 10, $\mathbb{P}_{N_u}(\tau)$ increases from 0.36 to 0.77 and then to 0.88. This observation is expected because the beamwidth of interfering UEs, given by $\theta_u = \frac{102\pi}{180N_u}$, decreases as N_u increases and interfering UEs with narrower beams cause less interference.

In order to further examine the analysis error caused by ignoring spatial correlation, we investigate the minimum number of antennas needed at the destination UE to achieve a given SINR coverage probability target, ξ, for a given τ. Mathematically, it is expressed as $N_{\min} \triangleq$

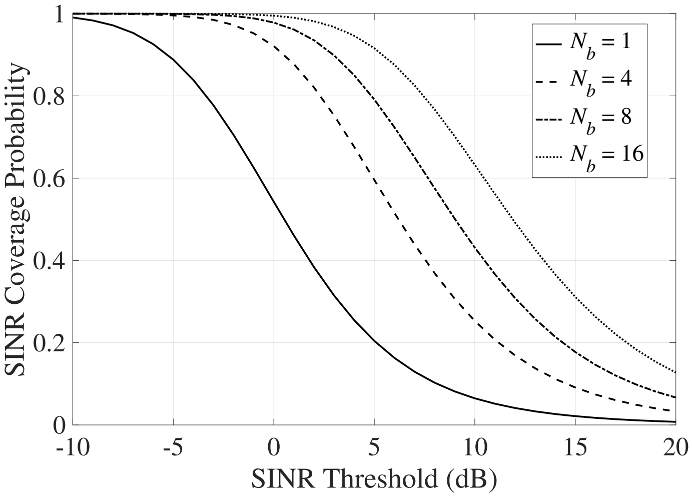


Figure 3.5: The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the SINR threshold, τ, for different values of N_b, when $N_u = 2$.

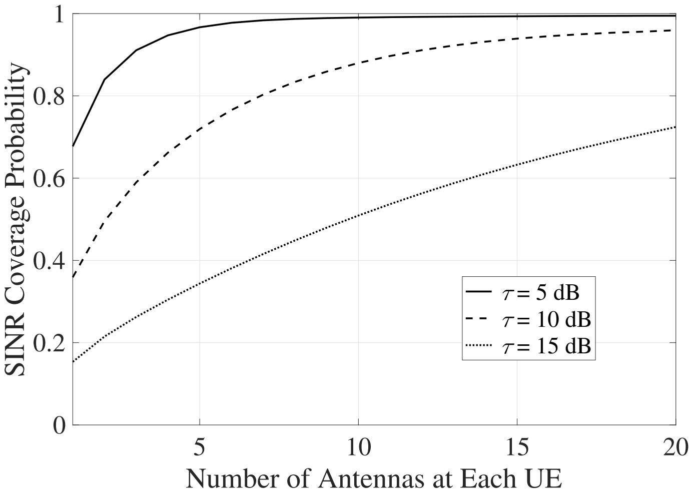


Figure 3.6: The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the number of antennas at each UE, N_u, for different values of τ.

Table 3.2: $N_{\min}$ versus $\hat{N}_{\min}$ for Different Values of ξ

ξ	60%	70%	80%	90%
$N_{\min}$	4	5	7	12
$\hat{N}_{\min}$	2	3	4	5

$\min\{N_u : \mathbb{P}_{N_u}(\tau) > \xi\}$ when spatial correlation is considered. When spatial correlation is

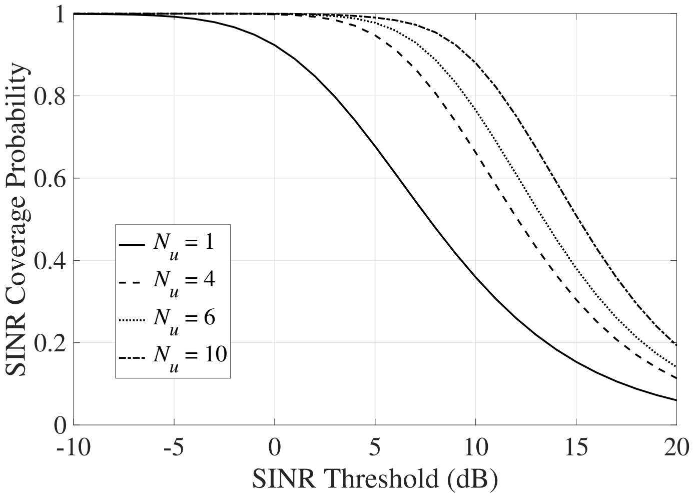


Figure 3.7: The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the SINR threshold, τ, for different values of N_u.

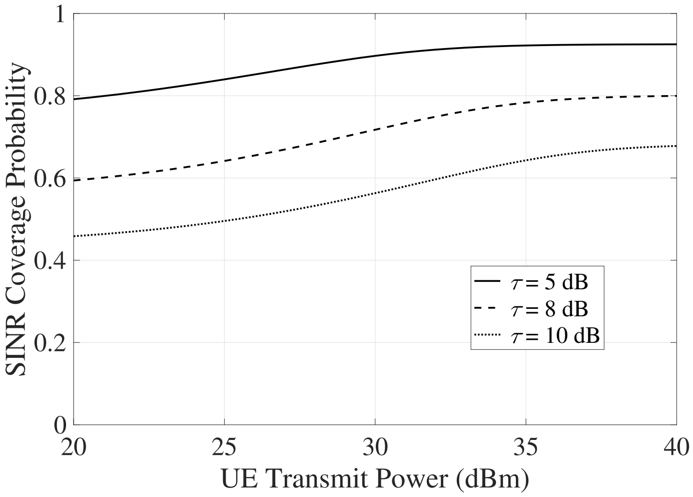


Figure 3.8: The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the UE transmit power, P_u, for different values of τ, when $N_u = 2$.

ignored, the expression is $\hat{N}_{\min} \triangleq \min\left\{N_u : \hat{\mathbb{P}}_{N_u}(\tau) > \xi\right\}$. Table 3.2 shows $N_{\min}$ and $\hat{N}_{\min}$ for different ξ when $\tau = 10$ dB. This table confirms the underestimation of the minimum number of antennas needed when spatial correlation is ignored. For example, when $\xi = 90\%$, if spatial correlation is ignored, the minimum number of antennas needed is 5. However, the actual minimum number of antennas needed is 12. Again, this example shows that ignoring spatial correlation is not acceptable when designing relay assisted mmWave cellular networks.

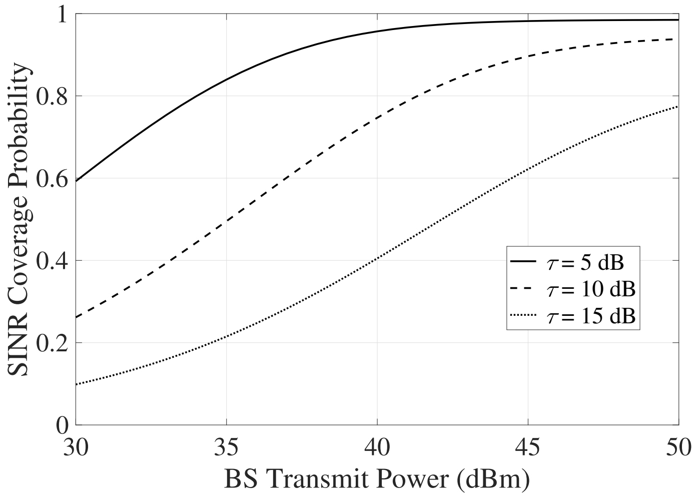


Figure 3.9: The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the BS transmit power, P_b, for different values of τ, when $N_u = 2$.

Third, we examine the impact of P_u and P_b on $\mathbb{P}_{N_u}(\tau)$. We observe from Fig. 3.8 that $\mathbb{P}_{N_u}(\tau)$ increases as P_u increases. Specifically, when $\tau = 10$ dB and P_u increases from 20 dBm to 30 dBm and then to 40 dBm, $\mathbb{P}_{N_u}(\tau)$ increases from 0.46 to 0.56 and then to 0.68. This is due to the fact that $\gamma_{rd,n}$ increases as P_u increases. Moreover, we observe from Fig. 3.9 that $\mathbb{P}_{N_u}(\tau)$ increases as P_b increases. Specifically, when $\tau = 10$ dB and P_b increases from 30 dBm to 40 dBm and then to 50 dBm, $\mathbb{P}_{N_u}(\tau)$ increases from 0.26 to 0.75 and then to 0.94. This is due to the fact that both $\gamma_{bd,n}$ and γ_{br} increase as P_b increases.

Fourth, we examine the impact of λ_b on $\mathbb{P}_{N_u}(\tau)$. We observe from Fig. 3.10 that $\mathbb{P}_{N_u}(\tau)$ first increases then decreases as λ_b increases. This is due to the fact that the impact of deploying more BSs is twofold. First, the associated BS is closer, which increases the desired signal power. Second, there are more interfering BSs, which increases the interference power. When λ_b increases from $10^{-4}/\text{m}^2$, the first impact dominates the second. After λ_b exceeds a certain value, which is the optimal BS density, λ_b^*, the second impact dominates the first. Thus, we highlight that λ_b^* that maximizes $\mathbb{P}_{N_u}(\tau)$ can be numerically found with the aid of our analysis. Specifically, when $N_b = 4$, 8, and 16, the optimal BS densities are $\lambda_b^* = 1.26 \times 10^{-3}/\text{m}^2$, $1.58 \times 10^{-3}/\text{m}^2$, and $1.78 \times 10^{-3}/\text{m}^2$, respectively.

Finally, we compare the SINR coverage probability without relays with $\mathbb{P}_{N_u}(\tau)$. We remark that the SINR coverage probability without relays is $\mathbb{P}_{N_u,bd}(\tau)$, since without relays, the relay mode does not exist. We observe from Fig. 3.11 and Fig. 3.12 that $\mathbb{P}_{N_u}(\tau)$ is significantly higher than $\mathbb{P}_{N_u,bd}(\tau)$. Specifically, we observe from Fig. 3.11 that when $N_u = 8$ and $P_b = 35$ dBm, $\mathbb{P}_{N_u}(\tau)$ is 0.83 while $\mathbb{P}_{N_u,bd}(\tau)$ is 0.59. This confirms that introducing relays into a mmWave cellular network vastly improves the SINR coverage probability. Moreover,

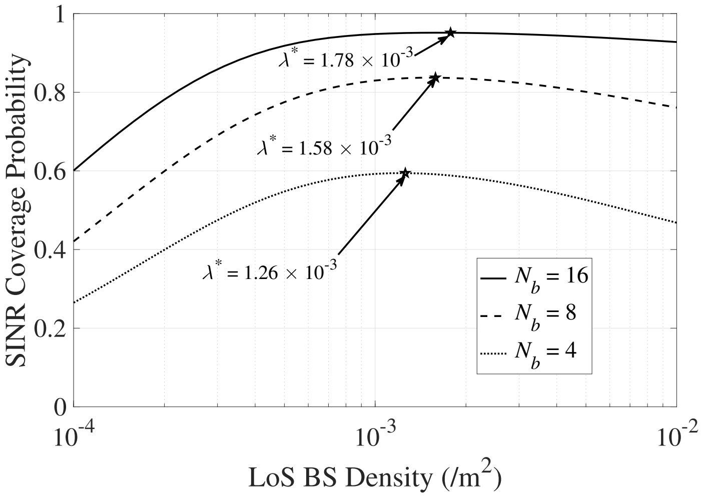


Figure 3.10: The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the LoS BS density, λ_b, for different values of N_b, when $\tau = 10$ dB.

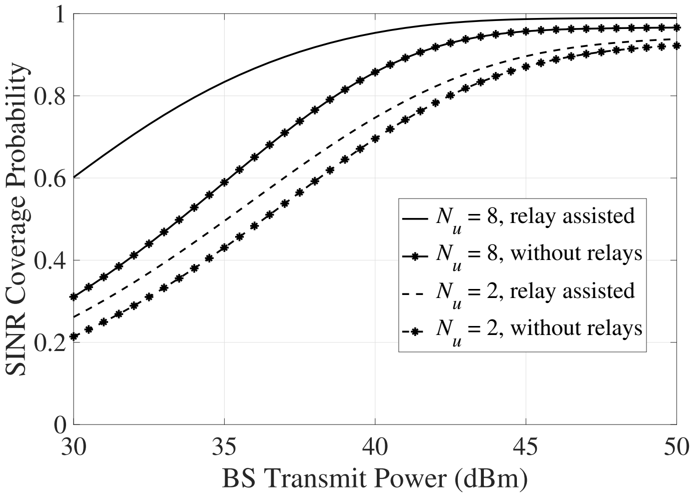


Figure 3.11: The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the BS transmit power, P_b, for different values of N_u, when $\tau = 10$ dB.

we observe from Fig. 3.11 that the gain brought by relays, i.e., the gap between $\mathbb{P}_{N_u}(\tau)$ and $\mathbb{P}_{N_u,bd}(\tau)$, becomes larger when N_u increases from 2 to 8, especially when P_b is not high, e.g., $P_b = 30 - 40$ dBm. This is due to the fact that when N_u increases, the directional gain of relay UEs, given by $G_U = N_u$, increases. When G_U increases, γ_{br} and $\gamma_{rd,n}$ increase. Thus, $\mathbb{P}_{br}(\tau)$ and $\mathbb{P}_{N_u,rd}(\tau)$ increase as N_u increases. Hence, based on (3.21), the gain brought by relays increases as N_u increases. Furthermore, we observe from Fig. 3.12 that $\mathbb{P}_{N_u,bd}(\tau)$ is constant

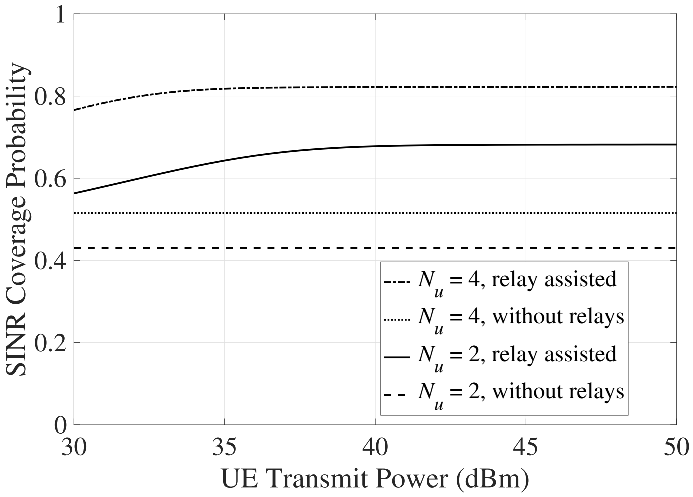


Figure 3.12: The SINR coverage probability, $\mathbb{P}_{N_u}(\tau)$, versus the UE transmit power, P_u, for different values of N_u, when $\tau = 10$ dB.

when P_u increases, while $\mathbb{P}_{N_u}(\tau)$ increases when P_u increases. This is due to the fact that without relays, the relay mode does not exist, i.e., P_u has no impact on $\mathbb{P}_{N_u,bd}(\tau)$.

3.4 Summary

In this chapter, we proposed a new analytical framework for the relay assisted mmWave cellular network which adopts selection combining at destination UEs to avoid the misalignment problem caused by traditional directional combining. Our results showed that ignoring spatial correlation to simplify the analysis is unacceptable, since it leads to severe overestimation of the SINR coverage probability. Moreover, we examined the impact of network parameters such as the BS density on the SINR coverage probability and found that the optimal BS density which maximizes the SINR coverage probability can be determined by using our analysis. Furthermore, our results showed that introducing relays into a mmWave cellular network vastly improves the SINR coverage performance.

Chapter 4

Conclusions and Future Research Directions

In this chapter, a summary of the conclusions drawn from this thesis is presented followed by a list of promising future research directions.

4.1 Conclusions

This thesis focused on the analysis and design of mmWave cellular networks. Specifically, two typical kinds of mmWave cellular networks were considered: 1) mmWave cellular HetNets and 2) relay assisted mmWave cellular networks.

In Chapter 2, we examined the impact of load balancing on the rate coverage probability in mmWave cellular HetNets. Specifically, we considered a two-tier mmWave HetNet where the high-power BSs in the macro-tier and the low-power BSs in the micro-tier co-exist to serve UEs. We assumed that a bias factor is used in the network to offload UEs to low-power BSs and characterized the impact of the bias factor on the system-level rate coverage probability. This characterization was completed in two steps. First, we analyzed the loads of high-power BSs and low-power BSs. Second, we analyzed the rate coverage probability of the network. In order to obtain tractable analytical results on the rate coverage probability, we modeled the considered network using the stochastic geometry based approach described in Chapter 1.1. Using numerical results, we demonstrated the accuracy of our analytical expressions and showed how the rate coverage probability is affected by the bias factor in various scenarios. Moreover, we found that the adoption of the bias factor significantly improves the rate coverage probability. Particularly, the optimal bias factor which maximizes the rate coverage probability can be numerically found using our analysis. Furthermore, we comprehensively examined the impact of network parameters such as BS beamwidths on the rate coverage probability. We obtained several insights into the design of mmWave cellular networks. For example, we found that the rate coverage probability improves when BS beams become narrower.

In Chapter 3, we evaluated the SINR coverage performance of a relay assisted mmWave cellular network which consists of multiple BSs, multiple relay UEs, and multiple destination UEs. In the considered relay assisted mmWave cellular network, the BS transmits data either directly to the destination UE, or indirectly via the relay UE when the direct transmission fails. We proposed a new analytical framework to derive the SINR coverage probability of the network. In this framework, we considered: 1) generalized Nakagami-$\mathcal{M}$ fading, 2) interference, 3) selection combining at destination UEs, and 4) spatial correlation. In order to obtain tractable analytical results on the system-level SINR coverage probability, we modeled the considered network using the stochastic geometry based approach described in Chapter 1.1. Using numerical results, we demonstrated the correctness of the derived SINR coverage probability. Moreover, we found that ignoring spatial correlation leads to severe overestimation of the coverage probability, as well as severe underestimation of the minimum number of antennas needed at destination UEs to achieve an SINR coverage probability target. Furthermore, we examined the impact of network parameters such as the BS density on the SINR coverage probability and found that the optimal BS density which maximizes the SINR coverage probability can be determined by using our analysis. In the end, we found that introducing relays into a mmWave cellular network vastly improves the SINR coverage probability and the gain brought by relays increases when the number of antennas at each UE increases.

4.2 Future Research Directions

In this subchapter, a number of future research directions are identified and listed as follows:

- Integrated mmWave/sub-6 GHz cellular networks: In Chapter 2 and Chapter 3, we considered mmWave cellular networks. A promising future research direction is to consider integrated mmWave/sub-6 GHz cellular networks where mmWave communications and sub-6 GHz communications co-exist [48, 137]. As mentioned in Chapter 1.1, mmWave communications has extremely high penetration loss, which makes mmWave communications extremely sensitive to blockage and brings serious challenges to the reliability of mmWave cellular networks. To achieve high reliability, mmWave BSs need to be extremely densely deployed to increase the LoS probability [48, 137]. However, deploying dense BSs leads to extremely high financial cost for cellular operators. An alternative method to improve the reliability of mmWave cellular networks is to adopt integrated mmWave/sub-6 GHz cellular networks, due to the fact that sub-6 GHz communications is robust to blockage [48, 137]. Specifically, in integrated mmWave/sub-6 GHz cellular networks, sub-6 GHz communications is used for two major purposes: 1) transmitting control signals and 2) transmitting data when mmWave communications is not available. We highlight that the modeling approaches we adopted in Chapter 2 and Chapter 3 can be used to model integrated mmWave/sub-6 GHz cellular networks with a few adjustments. In addition, the derivation methods we adopted in Chapter 2 and Chapter 3 can be used to derive the rate/SINR coverage probability of integrated mmWave/sub-6 GHz cellular networks.

- Mobility support in mmWave cellular networks: In Chapter 2 and Chapter 3, we considered stationary UEs. A promising future research direction is to consider moving UEs [138, 139, 140]. As mentioned in Chapter 1.1, mobility support is more difficult in mmWave cellular networks than in traditional sub-6 GHz cellular networks. First, the received signal strength at a moving UE in mmWave cellular networks can be highly fluctuating due to the huge amount of blockers in the communication environment. Second, mmWave BSs have small coverage areas. Thus, a moving UE can frequently move into and out of the coverage areas of mmWave BSs, which triggers frequent handoffs. We highlight that our study on optimal bias factors in cell association provides valuable insights into the design of handoff algorithms in mmWave cellular networks.
- Ultra-low latency mmWave cellular networks: In Chapter 2 and Chapter 3, we considered the coverage performance of mmWave cellular networks. A promising future research direction is to consider the latency performance of mmWave cellular networks [141, 142, 143, 144]. Many of the anticipated applications of future cellular networks have extreme latency constraints. For example, the applications such as immersive virtual reality, augmented reality, telesurgery, and real-time cloud/fog computing require end-to-end latencies to be less than 10 ms and the applications such as the collision avoidance service for self-driving cars require end-to-end latencies to be less than 1 ms [141]. We highlight that pursuing ultra-low latency is typically under the assumption that the SINR/rate constraint is met. Our studies in Chapter 2 and Chapter 3 provide detailed SINR/rate performance analysis in mmWave cellular networks.
- Transport layer design of mmWave cellular networks: In Chapter 2 and Chapter 3, we focused on the physical layer analysis and design of mmWave cellular networks. A promising future research direction is to consider the transport layer design of mmWave cellular networks [145, 146]. As mentioned in Chapter 1.1, LoS mmWave channels have high channel capacity, since mmWave channels have high bandwidths. However, mmWave channels can be highly intermittent due to the vulnerability of mmWave communications to blockage. Such intermittent mmWave channels present unique challenges in the transport layer. For example, fast transport layer feedback will be needed so that the congestion control algorithm can quickly adapt to changes in the channel capacity [146]. We highlight that our studies on the SINR performance of mmWave cellular networks provide valuable insights into the mmWave channel capacity.

Appendix A

This appendix contains the proofs needed in Chapter 3.

A.1 Proof of Theorem 3.1

We define ε_n as the event that $\gamma_{bd,n}$ is higher than τ, where $n \in \{1,2,\ldots,N_u\}$. As per the rules of selection combining, $\mathbb{P}_{N_u,bd}(\tau)$ is written as

$$\begin{aligned}\mathbb{P}_{N_u,bd}(\tau) &= \Pr\left(\bigcup_{n=1}^{N_u} \varepsilon_n\right) \\ &= \sum_{\kappa=1}^{N_u} (-1)^{\kappa+1} \sum_{1\leq\zeta_1<\zeta_2<\cdots<\zeta_\kappa\leq N_u} \Pr\left(\varepsilon_{\zeta_1} \cap \varepsilon_{\zeta_2} \cap \cdots \cap \varepsilon_{\zeta_\kappa}\right). \qquad \text{(A.1)}\end{aligned}$$

We note that in the considered network, the probability of the intersection of κ events, $\Pr\left(\varepsilon_{\zeta_1} \cap \varepsilon_{\zeta_2} \cap \cdots \cap \varepsilon_{\zeta_\kappa}\right)$, is the same, regardless of which κ out of N_u events are chosen. Thus, we rewrite (A.1) as

$$\mathbb{P}_{N_u,bd}(\tau) = \sum_{\kappa=1}^{N_u} (-1)^{\kappa+1} \binom{N_u}{\kappa} \Pr\left(\bigcap_{n=1}^{\kappa} \varepsilon_n\right). \qquad \text{(A.2)}$$

Based on (3.12), we derive $\Pr\left(\bigcap_{n=1}^{\kappa} \varepsilon_n\right)$ as

$$\begin{aligned}\Pr\left(\bigcap_{n=1}^{\kappa} \varepsilon_n\right) &= \Xi_b \Pr\left(\bigcap_{n=1}^{\kappa} \hbar_{nb_0} > \frac{\tau\left(\sum_{b_k\in\tilde{\Phi}_b} P_b G_{b_k} \ell_{b_k d}^{-\eta} \hbar_{nb_k} + \sigma^2\right)}{P_b G_B \ell_{bd}^{-\eta}}\right) \\ &\overset{(m)}{=} \Xi_b \mathbb{E}_{\ell_{bd},\tilde{\Phi}_b,G_{b_k},\vartheta_{bd}}\left[\prod_{n=1}^{\kappa}\sum_{j=1}^{\mathcal{M}_{bd}}\binom{\mathcal{M}_{bd}}{j}(-1)^{j+1} e^{-j\delta_{bd}\left(\sum_{b_k\in\tilde{\Phi}_b} P_b G_{b_k}\ell_{b_k d}^{-\eta}\hbar_{nb_k}+\sigma^2\right)}\right].\end{aligned} \tag{A.3}$$

Here, step (m) is achieved by following the binomial theorem and making the assumption that $F_{\hbar_{nb_l}}(x)$ is approximated as $\left(1-e^{-\alpha_{bd}x}\right)^{\mathcal{M}_{bd}}$, while ϑ_{bd} is defined as

$$\vartheta_{bd} \triangleq \sum_{q=1}^{\mathcal{M}_{bd}} \sum_{p=1}^{j_q} q\hbar_{I_{q,p},b_k}, \tag{A.4}$$

where $I_{q,p} \in \{1,\ldots,\kappa\}$ with $q \in \{1,\ldots,\mathcal{M}_{bd}\}$ and $p \in \{1,\ldots,j_q\}$ and the values of $I_{q,p}$ differ from each other. By defining $\Psi \triangleq \sum_{b_k\in\tilde{\Phi}_b} \delta_{bd} P_b G_{b_k} \ell_{b_k d}^{-\eta} \vartheta_{bd}$, we further derive (A.12) as

$$\begin{aligned}\Pr\left(\bigcap_{n=1}^{\kappa} \varepsilon_n\right) &= \Xi_b \sum_{j_1+j_2+\cdots+j_{\mathcal{M}_{bd}}=\kappa} \mathbb{E}_{\ell_{bd}}\left[e^{-\delta_{bd}\sigma^2\Omega_{bd}} \sum_{I_{q,p}} \mathbb{E}_{\tilde{\Phi}_b,G_{b_k},\vartheta_{bd}}\left[e^{-\Psi}\right]\right] \\ &\quad \times \binom{\mathcal{M}_{bd}}{1}^{j_1}\binom{\mathcal{M}_{bd}}{2}^{j_2}\cdots\binom{\mathcal{M}_{bd}}{\mathcal{M}_{bd}}^{j_{\mathcal{M}_{bd}}}(-1)^{j_1+j_2+\cdots+j_{\mathcal{M}_{bd}}+\Omega_{bd}} \\ &\overset{(n)}{=} \Xi_b \sum_{j_1+j_2+\cdots+j_{\mathcal{M}_{bd}}=\kappa} \beta_{bd}\mathbb{E}_{\ell_{bd}}\left[e^{-\delta_{bd}\sigma^2\Omega_{bd}}\mathbb{E}_{\tilde{\Phi}_b,G_{b_k},\vartheta_{bd}}\left[e^{-\Psi}\right]\right].\end{aligned} \tag{A.5}$$

To achieve step (n), we find that given $M_{\hbar_{nb_l}}(s)$, the MGF of ϑ_{bd} is given by

$$M_{\vartheta_{bd}}(s) = \prod_{q=1}^{\mathcal{M}_{bd}}\left(1-\frac{qs}{\mathcal{M}_{bd}}\right)^{-\mathcal{M}_{bd}j_q}. \tag{A.6}$$

This indicates that given $j_1, j_2, \ldots, j_{\mathcal{M}_{bd}}$, the MGF of ϑ_{bd} does not change with the values of $I_{q,p}$. We also find that given $j_1, j_2, \ldots, j_{\mathcal{M}_{bd}}$, $\mathbb{E}_{\tilde{\Phi}_b,G_{b_k},\vartheta_{bd}}\left[e^{-\Psi}\right]$ does not change with the values of $I_{q,p}$. Thus, step (n) is achieved due to

$$\sum_{I_{q,p}}\mathbb{E}_{\tilde{\Phi}_b,G_{b_k},\vartheta_{bd}}\left[e^{-\Psi}\right] = \binom{k}{j_1}\binom{k-j_1}{j_2}\cdots\binom{k-j_1-\cdots-j_{\mathcal{M}_{bd}-1}}{j_{\mathcal{M}_{bd}}}\mathbb{E}_{\tilde{\Phi}_b,G_{b_k},\vartheta_{bd}}\left[e^{-\Psi}\right]. \tag{A.7}$$

Using (A.6), we derive $\mathbb{E}_{\tilde{\Phi}_b,G_{b_k},\vartheta_{bd}}\left[e^{-\Psi}\right]$ as

$$\begin{aligned}
\mathbb{E}_{\tilde{\Phi}_b,G_{b_k},\vartheta_{bd}}\left[e^{-\Psi}\right] &= \mathbb{E}_{\tilde{\Phi}_b,G_b,\vartheta_{bd}}\left[\prod_{\tilde{\Phi}_b} e^{-\delta_{bd}P_bG_{b_k}\ell_{b_kd}^{-\eta}\vartheta_{bd}}\right] \\
&\overset{(o)}{=} \mathbb{E}_{\tilde{\Phi}_b,G_{b_k}}\left[\prod_{\tilde{\Phi}_b}\prod_{q=1}^{\mathcal{M}_{bd}}\left(1+\frac{q\delta_{bd}P_bG_{b_k}\ell_{b_kd}^{-\eta}}{\mathcal{M}_{bd}}\right)^{-\mathcal{M}_{bd}j_q}\right] \\
&= \mathbb{E}_{\tilde{\Phi}_b}\left[\prod_{\tilde{\Phi}_b} v_{bd}\left(\ell_{b_kd}\right)\right] \\
&= e^{-2\pi\lambda_b\int_{\ell_{bd}}^{r_b}\ell(1-v_{bd}(\ell))d\ell}. \qquad \text{(A.8)}
\end{aligned}$$

Here, step (o) is achieved by using (A.6). Finally, by substituting (A.5) and (A.8) into (A.2), we obtain the desired result in (3.22) and complete the proof.

A.2 Proof of Theorem 3.2

Based on (3.15), we derive $\mathbb{P}_{br}\left(\tau\right)$ as

$$\begin{aligned}
\mathbb{P}_{br}\left(\tau\right) &= \Xi_b\mathbb{E}_{\ell_{br},I_b^*}\left[\Pr\left(\hbar_{b_0} > \frac{\tau\left(I_b^*+\sigma^2\right)}{P_bG_BG_U\ell_{br}^{-\eta}}\right)\right] \\
&\overset{(p)}{=} \Xi_b\mathbb{E}_{\ell_{br},I_b^*}\left[\sum_{j=1}^{\mathcal{M}_{br}}(-1)^{j+1}\binom{\mathcal{M}_{br}}{j}e^{-\frac{j\alpha_{br}\tau\left(I_b^*+\sigma^2\right)}{P_bG_BG_U\ell_{br}^{-\eta}}}\right] \\
&= \Xi_b\sum_{j=1}^{\mathcal{M}_{br}}(-1)^{j+1}\binom{\mathcal{M}_{br}}{j}\mathbb{E}_{\ell_{br},I_b^*}\left[e^{-\frac{\psi\left(I_b^*+\sigma^2\right)}{\ell_{br}^{-\eta}}}\right] \\
&= \Xi_b\sum_{j=1}^{\mathcal{M}_{br}}(-1)^{j+1}\binom{\mathcal{M}_{br}}{j}\mathbb{E}_{\ell_{br}}\left[e^{-\frac{\psi\sigma^2}{\ell_{br}^{-\eta}}}\mathbb{E}_{I_b^*}\left[e^{-\frac{\psi I_b^*}{\ell_{br}^{-\eta}}}\right]\right] \\
&= \Xi_b\sum_{j=1}^{\mathcal{M}_{br}}(-1)^{j+1}\binom{\mathcal{M}_{br}}{j}\int_0^{r_b}e^{-\psi x^{\eta}\sigma^2}\mathbb{E}_{I_b^*}\left[e^{-\frac{\psi I_b^*}{x^{-\eta}}}\right]f_{\ell_{br}}\left(x\right)dx, \qquad \text{(A.9)}
\end{aligned}$$

where $I_b^* \triangleq \sum\limits_{b_k\in\tilde{\Phi}_b} P_bG_{b_kr}\ell_{b_kr}^{-\eta}\hbar_{b_k}$. Here, step (p) is achieved by following the binomial theorem and making the assumption that the CDF of $\hbar_{b_l}$ is approximated as $\left(1-e^{-\alpha_{br}x}\right)^{\mathcal{M}_{br}}$. We further derive $\mathbb{E}_{I_b^*}\left[e^{-\frac{\psi I_b^*}{x^{-\eta}}}\right]$ as

$$\begin{aligned}
\mathbb{E}_{I_b^*}\left[e^{-\frac{\psi I_b^*}{x^{-\eta}}}\right] &= \mathbb{E}_{I_b^*}\left[e^{-\psi x^{\eta} I_b^*}\right] \\
&\overset{(q)}{=} e^{-\lambda_b \int_x^{r_b} \ell \int_0^{2\pi}\left(1-\mathbb{E}_{\hbar_{b_k}}\left[e^{-\hbar_{b_k}\psi x^{\eta} P_b G_{b_k} r^{\ell^{-\eta}}}\right]\right) d\theta d\ell} \\
&\overset{(r)}{=} e^{-\lambda_b \int_x^{r_b} \ell \int_0^{2\pi}\left(1-\left(1+\frac{\psi x^{\eta} P_b G_{b_k} r^{\ell^{-\eta}}}{\mathcal{M}_{br}}\right)^{-\mathcal{M}_{br}}\right) d\theta d\ell} \\
&\overset{(s)}{=} e^{-2\pi\lambda_b \int_x^{r_b} \ell \Omega_b d\ell}.
\end{aligned} \tag{A.10}$$

Here, step (q) is achieved as per [127, Cor. (2.3.2)], step (r) is achieved by using (3.16), and step (s) is achieved as per (3.9). Finally, we substitute (A.10) into (A.9) to obtain the desired result in (3.26) and thus, the proof is completed.

A.3 Proof of Theorem 3.3

We define υ_n as the event that $\gamma_{rd,n}$ is higher than τ, where $n \in \{1, 2, \ldots, N_u\}$. Following the procedure presented in the proof of Theorem 3.1, $\mathbb{P}_{N_u,rd}(\tau)$ is written as

$$\begin{aligned}
\mathbb{P}_{N_u,rd}(\tau) &= \Pr\left(\bigcup_{n=1}^{N_u} \upsilon_n\right) \\
&= \sum_{\kappa=1}^{N_u} (-1)^{\kappa+1} \sum_{1\le\zeta_1<\zeta_2<\cdots<\zeta_\kappa\le N_u} \Pr\left(\upsilon_{\zeta_1} \cap \upsilon_{\zeta_2} \cap \cdots \cap \upsilon_{\zeta_\kappa}\right) \\
&= \sum_{\kappa=1}^{N_u} (-1)^{\kappa+1} \binom{N_u}{\kappa} \Pr\left(\bigcap_{n=1}^{\kappa} \upsilon_n\right).
\end{aligned} \tag{A.11}$$

Based on (3.18), we derive $\Pr\left(\bigcap_{n=1}^{\kappa} \upsilon_n\right)$ as

$$\begin{aligned}
\Pr\left(\bigcap_{n=1}^{\kappa} \upsilon_n\right) =& \Xi_r \Pr\left(\bigcap_{n=1}^{\kappa} \hbar_{nr} > \frac{\tau\left(\sum_{i_k\in\Phi_i} P_u G_e \ell_{i_k d}^{-\eta} \hbar_{ni_k} + \sigma^2\right)}{P_u G_U \ell_{rd}^{-\eta}}\right) \\
\overset{(t)}{=}& \Xi_r \mathbb{E}_{\ell_{rd},\Phi_i,G_e,\vartheta_{rd}}\left[\prod_{n=1}^{\kappa} \sum_{j=1}^{\mathcal{M}_{rd}} \binom{\mathcal{M}_{rd}}{j} (-1)^{j+1} e^{-j\delta_{rd}\left(\sum_{i_k\in\Phi_i} P_u G_e \ell_{i_k d}^{-\eta} \hbar_{ni_k} + \sigma^2\right)}\right].
\end{aligned} \tag{A.12}$$

Here, step (t) is achieved by following the binomial theorem and making the assumption that $F_{\hbar_{nr}}(x)$ is approximated as $\left(1-e^{-\alpha_{rd}x}\right)^{\mathcal{M}_{rd}}$, while ϑ_{rd} is defined as $\vartheta_{rd} \triangleq \sum_{q=1}^{\mathcal{M}_{rd}} \sum_{p=1}^{j_q} q\hbar_{I_{q,p},i_k}$, where $I_{q,p} \in \{1,\ldots,\kappa\}$ with $q \in \{1,\ldots,\mathcal{M}_{rd}\}$ and $p \in \{1,\ldots,j_q\}$ and the values of $I_{q,p}$ differ from each other. By defining $\Psi_{rd} \triangleq \sum_{i_k \in \Phi_i} \delta_{rd} P_u G_e \ell_{i_k d}^{-\eta} \vartheta_{rd}$, we further derive (A.12) as

$$\begin{aligned}\Pr\left(\bigcap_{n=1}^{\kappa} \upsilon_n\right) &= \Xi_r \sum_{j_1+j_2+\cdots+j_{\mathcal{M}_{rd}}=\kappa} \mathbb{E}_{\ell_{rd}}\left[e^{-\delta_{rd}\sigma^2\Omega_{rd}} \sum_{I_{q,p}} \mathbb{E}_{\Phi_i,G_e,\vartheta_{rd}}\left[e^{-\Psi_{rd}}\right]\right] \\ &\quad \times \binom{\mathcal{M}_{rd}}{1}^{j_1} \binom{\mathcal{M}_{rd}}{2}^{j_2} \cdots \binom{\mathcal{M}_{rd}}{\mathcal{M}_{rd}}^{j_{\mathcal{M}_{rd}}} (-1)^{j_1+j_2+\cdots+j_{\mathcal{M}_{rd}}+\Omega_{rd}} \\ &\overset{(u)}{=} \Xi_r \sum_{j_1+j_2+\cdots+j_{\mathcal{M}_{rd}}=\kappa} \beta_{rd} \mathbb{E}_{\ell_{rd}}\left[e^{-\delta_{rd}\sigma^2\Omega_{rd}} \mathbb{E}_{\Phi_i,G_e,\vartheta_{rd}}\left[e^{-\Psi_{rd}}\right]\right]. \end{aligned} \tag{A.13}$$

To achieve step (u), we find that given $M_{\hbar_{ni_k}}(s)$, the MGF of ϑ_{rd} is given by

$$M_{\vartheta_{rd}}(s) = \prod_{q=1}^{\mathcal{M}_{rd}} \left(1-\frac{qs}{\mathcal{M}_{rd}}\right)^{-\mathcal{M}_{rd} j_q}. \tag{A.14}$$

This indicates that given $j_1, j_2, \ldots, j_{\mathcal{M}_{rd}}$, the MGF of ϑ_{rd} does not change with the values of $I_{q,p}$. We also find that given $j_1, j_2, \ldots, j_{\mathcal{M}_{rd}}$, $\mathbb{E}_{\Phi_i,G_e,\vartheta_{rd}}\left[e^{-\Psi_{rd}}\right]$ does not change with the values of $I_{q,p}$. Thus, step (u) is achieved due to

$$\sum_{I_{q,p}} \mathbb{E}_{\Phi_i,G_e,\vartheta_{rd}}\left[e^{-\Psi_{rd}}\right] = \binom{k}{j_1}\binom{k-j_1}{j_2} \cdots \binom{k-j_1-\cdots-j_{\mathcal{M}_{rd}-1}}{j_{\mathcal{M}_{rd}}} \mathbb{E}_{\Phi_i,G_e,\vartheta_{rd}}\left[e^{-\Psi_{rd}}\right]. \tag{A.15}$$

Using (A.14), we derive $\mathbb{E}_{\Phi_i,G_e,\vartheta_{rd}}\left[e^{-\Psi_{rd}}\right]$ as

$$\begin{aligned}\mathbb{E}_{\Phi_i,G_e,\vartheta_{rd}}\left[e^{-\Psi_{rd}}\right] &= \mathbb{E}_{\Phi_i,G_e,\vartheta_{rd}}\left[\prod_{\Phi_i} e^{-\delta_{rd}P_u G_e \ell_{i_k d}^{-\eta} \vartheta_{rd}}\right] \\ &\overset{(v)}{=} \mathbb{E}_{\Phi_i,G_e}\left[\prod_{\Phi_i}\prod_{q=1}^{\mathcal{M}_{rd}} \left(1+\frac{q\delta_{rd}P_u G_e \ell_{i_k d}^{-\eta}}{\mathcal{M}_{rd}}\right)^{-\mathcal{M}_{rd} j_q}\right] \\ &= \mathbb{E}_{\Phi_i}\left[\prod_{\Phi_i} v_{rd}\left(\ell_{i_k d}\right)\right] \\ &= e^{-2\pi\lambda_b \int_{\ell_{rd}}^{r_b} \ell(1-v_{rd}(\ell))d\ell}. \end{aligned} \tag{A.16}$$

Here, step (v) is achieved by using (A.14). Finally, by substituting (A.13) and (A.16) into (A.11), we obtain the desired result in (3.28) and complete the proof.

Bibliography

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, “What will 5G be?” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014. (cited on pages 1 and 2)

[2] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, “An overview of signal processing techniques for millimeter wave MIMO systems,” *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016. (cited on pages 1, 2, 10, and 12)

[3] S. Rangan, T. S. Rappaport, and E. Erkip, “Millimeter-wave cellular wireless networks: Potentials and challenges,” *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014. (cited on pages 1 and 2)

[4] Z. Pi and F. Khan, “An introduction to millimeter-wave mobile broadband systems,” *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, Jun. 2011. (cited on pages 1, 7, 22, 27, and 48)

[5] A. Ghosh, T. A. Thomas, M. C. Cudak, R. Ratasuk, P. Moorut, F. W. Vook, T. S. Rappaport, G. R. MacCartney, S. Sun, and S. Nie, “Millimeter-wave enhanced local area systems: A high-data-rate approach for future wireless networks,” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1152–1163, Jun. 2014. (cited on page 1)

[6] P. Smulders, “Exploiting the 60 GHz band for local wireless multimedia access: Prospects and future directions,” *IEEE Commun. Mag.*, vol. 40, no. 1, pp. 140–147, Jan. 2002. (cited on page 1)

[7] R. C. Daniels and R. W. Heath, “60 GHz wireless communications: Emerging requirements and design recommendations,” *IEEE Veh. Technol. Mag.*, vol. 2, no. 3, pp. 41–50, Sep. 2007. (cited on page 1)

[8] T. S. Rappaport, W. Roh, and K. Cheun, “Mobile’s millimeter-wave makeover,” *IEEE Spectr.*, vol. 51, no. 9, pp. 34–58, Sep. 2014. (cited on page 1)

[9] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, “Millimeter wave mobile communications for 5G

cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013. (cited on pages 1, 2, 3, 7, and 18)

[10] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmWave) for 5G: Opportunities and challenges," *Wireless Networks*, vol. 21, no. 8, pp. 2657–2676, Nov. 2015. [Online]. Available: https://doi.org/10.1007/s11276-015-0942-z (cited on page 1)

[11] "Cisco visual networking index: Global mobile data traffic forecast update, 2017–2022 white paper," Cisco, Tech. Rep., Feb. 2019. (cited on page 1)

[12] X. An, C. S. Sum, R. V. Prasad, J. Wang, Z. Lan, J. Wang, R. Hekmat, H. Harada, and I. Niemegeers, "Beam switching support to resolve link-blockage problem in 60 GHz WPANs," in *Proc. IEEE PIMRC*, Tokyo, Japan, Sep. 2009, pp. 390–394. (cited on page 2)

[13] X. Zhang, S. Zhou, X. Wang, Z. Niu, X. Lin, D. Zhu, and M. Lei, "Improving network throughput in 60GHz WLANs via multi-AP diversity," in *Proc. IEEE ICC*, Ottawa, Canada, Jun. 2012, pp. 4803–4807. (cited on page 2)

[14] A. V. Alejos, M. G. Sanchez, and I. Cuinas, "Measurement and analysis of propagation mechanisms at 40 GHz: Viability of site shielding forced by obstacles," *IEEE Trans. Veh. Tech.*, vol. 57, no. 6, pp. 3369–3380, Nov. 2008. (cited on pages 2, 18, and 27)

[15] A. K. M. Isa, A. Nix, and G. Hilton, "Impact of diffraction and attenuation for material characterisation in millimetre wave bands," in *Proc. LAPC*, Loughborough, UK, Nov. 2015, pp. 1–4. (cited on pages 2 and 18)

[16] G. R. MacCartney, S. Deng, S. Sun, and T. S. Rappaport, "Millimeter-wave human blockage at 73 GHz with a simple double knife-edge diffraction model and extension for directional antennas," in *Proc. IEEE VTC Fall*, Montréal, Canada, Sep. 2016, pp. 1–6. (cited on pages 2 and 18)

[17] T. H. Barratt, E. Mellios, P. Cain, A. R. Nix, and M. A. Beach, "Measured and modelled corner diffraction at millimetre wave frequencies," in *Proc. IEEE PIMRC*, Valencia, Spain, Sep. 2016, pp. 1–5. (cited on pages 2 and 18)

[18] C. R. Anderson and T. S. Rappaport, "In-building wideband partition loss measurements at 2.5 and 60 GHz," *IEEE Trans. Wireless Commun.*, vol. 3, no. 3, pp. 922–928, May 2004. (cited on page 2)

[19] J. Lu, D. Steinbach, P. Cabrol, P. Pietraski, and R. V. Pragada, "Propagation characterization of an office building in the 60 GHz band," in *Proc. EuCAP*, Hague, Netherlands, Apr. 2014, pp. 809–813. (cited on page 2)

[20] M. E. Rasekh, A. A. Shishegar, and F. Farzaneh, “A study of the effect of diffraction and rough surface scatter modeling on ray tracing results in an urban environment at 60 GHz,” in *Proc. MMWaTT*, Tehran, Iran, Dec. 2009, pp. 27–31. (cited on page 2)

[21] M. Jacob, S. Priebe, R. Dickhoff, T. Kleine-Ostmann, T. Schrader, and T. Kurner, “Diffraction in mm and sub-mm wave indoor propagation channels,” *IEEE Trans. Microw. Theory Techn.*, vol. 60, no. 3, pp. 833–844, Mar. 2012. (cited on page 2)

[22] H. Zhang, S. Venkateswaran, and U. Madhow, “Channel modeling and MIMO capacity for outdoor millimeter wave links,” in *Proc. IEEE WCNC*, Sydney, Australia, Apr. 2010, pp. 1–6. (cited on pages 2 and 3)

[23] N. Moraitis and P. Constantinou, “Indoor channel measurements and characterization at 60 GHz for wireless local area network applications,” *IEEE Trans. Antennas Propag.*, vol. 52, no. 12, pp. 3180–3189, Dec. 2004. (cited on pages 2 and 3)

[24] T. Zwick, T. J. Beukema, and H. Nam, “Wideband channel sounder with measurements and model for the 60 GHz indoor radio channel,” *IEEE Trans. Veh. Technol.*, vol. 54, no. 4, pp. 1266–1277, Jul. 2005. (cited on page 2)

[25] J. G. Andrews, F. Baccelli, and R. K. Ganti, “A tractable approach to coverage and rate in cellular networks,” *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122–3134, Nov. 2011. (cited on page 3)

[26] H. ElSawy, E. Hossain, and M. Haenggi, “Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: A survey,” *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 996–1019, Mar. 2013. (cited on page 3)

[27] T. X. Brown, “Cellular performance bounds via shotgun cellular systems,” *IEEE J. Sel. Areas Commun.*, vol. 18, no. 11, pp. 2443–2455, Nov. 2000. (cited on page 3)

[28] C. Skouroumounis, C. Psomas, and I. Krikidis, “Low complexity base station cooperation in cellular networks with blockages,” in *Proc. IEEE WCNC*, Doha, Qatar, Apr. 2016, pp. 1–6. (cited on page 3)

[29] 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. (cited on pages 3, 23, and 26)

[30] 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. (cited on page 3)

[31] H. Xu, V. Kukshya, and T. S. Rappaport, "Spatial and temporal characteristics of 60-GHz indoor channels," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 3, pp. 620–630, Apr. 2002. (cited on pages 4 and 7)

[32] "Multiple input multiple output (MIMO) antennae in UTRA," 3GPP, Tech. Rep., Mar. 2007. (cited on page 5)

[33] G. Calcev, D. Chizhik, B. Goransson, S. Howard, H. Huang, A. Kogiantis, A. F. Molisch, A. L. Moustakas, D. Reed, and H. Xu, "A wideband spatial channel model for system-wide simulations," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 389–403, Mar. 2007. (cited on page 5)

[34] L. Liu, C. Oestges, J. Poutanen, K. Haneda, P. Vainikainen, F. Quitin, F. Tufvesson, and P. D. Doncker, "The COST 2100 MIMO channel model," *IEEE Wireless Commun.*, vol. 19, no. 6, pp. 92–99, Dec. 2012. (cited on page 5)

[35] M. K. Samimi and T. S. Rappaport, "3-D millimeter-wave statistical channel model for 5G wireless system design," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 7, pp. 2207–2225, Jul. 2016. (cited on pages 5 and 6)

[36] D. S. Baum, J. Hansen, J. Salo, G. Del Galdo, M. Milojevic, and P. Kyösti, "An interim channel model for beyond-3G systems: Extending the 3GPP spatial channel model (SCM)," in *Proc. IEEE VTC Spring*, Stockholm, Sweden, May 2005, pp. 3132–3136. (cited on page 6)

[37] 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. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 454–469, Apr. 2016. (cited on page 6)

[38] "METIS channel models," Mobile and wireless communications Enablers for the Twenty-twenty Information Society (METIS), Tech. Rep., Feb. 2015. (cited on page 6)

[39] J. Brady, N. Behdad, and A. M. Sayeed, "Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3814–3827, Jul. 2013. (cited on page 7)

[40] A. A. M. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE J. Sel. Areas Commun.*, vol. 5, no. 2, pp. 128–137, Feb. 1987. (cited on page 7)

[41] E. Ben-Dor, T. S. Rappaport, Y. Qiao, and S. J. Lauffenburger, "Millimeter-wave 60 GHz outdoor and vehicle AOA propagation measurements using a broadband channel

sounder," in *Proc. IEEE Globecom*, Houston, TX, Dec. 2011, pp. 1–6. (cited on page 7)

[42] X. Wu, C. Wang, J. Sun, J. Huang, R. Feng, Y. Yang, and X. Ge, "60-GHz millimeter-wave channel measurements and modeling for indoor office environments," *IEEE Trans. Antennas Propag.*, vol. 65, no. 4, pp. 1912–1924, Apr. 2017. (cited on page 7)

[43] J. Huang, C. Wang, R. Feng, J. Sun, W. Zhang, and Y. Yang, "Multi-frequency mmWave massive MIMO channel measurements and characterization for 5G wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1591–1605, Jul. 2017. (cited on page 7)

[44] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective," *IEEE Wireless Commun.*, vol. 21, no. 3, pp. 118–127, Jun. 2014. (cited on page 7)

[45] H. S. Jo, Y. J. Sang, P. Xia, and J. G. Andrews, "Heterogeneous cellular networks with flexible cell association: A comprehensive downlink SINR analysis," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3484–3495, Oct. 2012. (cited on pages 7, 21, and 29)

[46] D. Lopez-Perez, I. Guvenc, G. de la Roche, M. Kountouris, T. Q. S. Quek, and J. Zhang, "Enhanced intercell interference coordination challenges in heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 22–30, Jun. 2011. (cited on pages 7 and 9)

[47] T. E. Bogale and L. B. Le, "Massive MIMO and mmWave for 5G wireless HetNet: Potential benefits and challenges," *IEEE Veh. Technol. Mag.*, vol. 11, no. 1, pp. 64–75, Mar. 2016. (cited on pages 7 and 8)

[48] H. Elshaer, M. N. Kulkarni, F. Boccardi, J. G. Andrews, and M. Dohler, "Downlink and uplink cell association with traditional macrocells and millimeter wave small cells," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6244–6258, Sep. 2016. (cited on pages 7, 8, 26, 27, 49, and 66)

[49] E. Turgut and M. C. Gursoy, "Coverage in heterogeneous downlink millimeter wave cellular networks," *IEEE Trans. Commun.*, vol. 65, no. 10, pp. 4463–4477, Oct. 2017. (cited on pages 7, 8, 9, 21, 26, 27, 28, 40, 49, and 53)

[50] S. Singh, X. Zhang, and J. G. Andrews, "Uplink rate distribution in heterogeneous cellular networks with power control and load balancing," in *Proc. IEEE ICC Workshops*, London, UK, Jun. 2015, pp. 1275–1280. (cited on pages 7, 21, and 40)

[51] M. Cierny, H. Wang, R. Wichman, Z. Ding, and C. Wijting, "On number of almost blank subframes in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 5061–5073, Oct. 2013. (cited on page 7)

[52] J. G. Andrews, “Seven ways that HetNets are a cellular paradigm shift,” *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 136–144, Mar. 2013. (cited on pages 7, 8, and 9)

[53] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, “Femtocells: Past, present, and future,” *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 497–508, Apr. 2012. (cited on page 7)

[54] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, “Femtocell networks: A survey,” *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 59–67, Sep. 2008. (cited on pages 7 and 8)

[55] T. Z. Oo, N. H. Tran, W. Saad, J. Son, and C. S. Hong, “Traffic offloading via Markov approximation in heterogeneous cellular networks,” in *Proc. IEEE/IFIP NOMS*, Istanbul, Turkey, Apr. 2016, pp. 52–60. (cited on page 7)

[56] T. Z. Oo, N. H. Tran, W. Saad, D. Niyato, Z. Han, and C. S. Hong, “Offloading in HetNet: A coordination of interference mitigation, user association, and resource allocation,” *IEEE Trans. Mobile Comput.*, vol. 16, no. 8, pp. 2276–2291, Aug. 2017. (cited on page 7)

[57] J. G. Andrews, S. Singh, Q. Ye, X. Lin, and H. S. Dhillon, “An overview of load balancing in HetNets: Old myths and open problems,” *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 18–25, Apr. 2014. (cited on pages 7, 9, and 21)

[58] S. Singh and J. G. Andrews, “Joint resource partitioning and offloading in heterogeneous cellular networks,” *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 888–901, Feb. 2014. (cited on pages 7, 9, 21, 26, 29, 40, and 48)

[59] G. Athanasiou, P. C. Weeraddana, C. Fischione, and L. Tassiulas, “Optimizing client association for load balancing and fairness in millimeter-wave wireless networks,” *IEEE/ACM Trans. Netw.*, vol. 23, no. 3, pp. 836–850, Jun. 2015. (cited on page 7)

[60] T. K. Vu, M. Bennis, S. Samarakoon, M. Debbah, and M. Latva-aho, “Joint load balancing and interference mitigation in 5G heterogeneous networks,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6032–6046, Sep. 2017. (cited on page 7)

[61] S. Collonge, G. Zaharia, and G. E. Zein, “Influence of the human activity on wide-band characteristics of the 60 GHz indoor radio channel,” *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 2396–2406, Nov. 2004. (cited on pages 8 and 13)

[62] H. Elshaer, F. Boccardi, M. Dohler, and R. Irmer, “Downlink and uplink decoupling: A disruptive architectural design for 5G networks,” in *Proc. IEEE Globecom*, Austin, TX, Dec. 2014, pp. 1798–1803. (cited on page 8)

[63] Y. Zhou, F. R. Yu, J. Chen, and Y. Kuo, "Communications, caching, and computing for next generation HetNets," *IEEE Wireless Commun.*, vol. 25, no. 4, pp. 104–111, Aug. 2018. (cited on page 10)

[64] K. Poularakis and L. Tassiulas, "On the complexity of optimal content placement in hierarchical caching networks," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2092–2103, May 2016. (cited on page 10)

[65] L. Pu, L. Jiao, X. Chen, L. Wang, Q. Xie, and J. Xu, "Online resource allocation, content placement and request routing for cost-efficient edge caching in cloud radio access networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, pp. 1751–1767, Aug. 2018. (cited on page 10)

[66] A. Antonopoulos, E. Kartsakli, and C. Verikoukis, "Game theoretic D2D content dissemination in 4G cellular networks," *IEEE Commun. Mag.*, vol. 52, no. 6, pp. 125–132, Jun. 2014. (cited on page 10)

[67] G. Gao, W. Zhang, Y. Wen, Z. Wang, and W. Zhu, "Towards cost-efficient video transcoding in media cloud: Insights learned from user viewing patterns," *IEEE Trans. Multimedia*, vol. 17, no. 8, pp. 1286–1296, Aug. 2015. (cited on page 10)

[68] W. Zhang, Y. Wen, and H. Chen, "Toward transcoding as a service: Energy-efficient offloading policy for green mobile cloud," *IEEE Netw.*, vol. 28, no. 6, pp. 67–73, Nov. 2014. (cited on page 10)

[69] W. Hong, K. Baek, Y. Lee, Y. Kim, and S. Ko, "Study and prototyping of practically large-scale mmWave antenna systems for 5G cellular devices," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 63–69, Sep. 2014. (cited on pages 10 and 18)

[70] P. Shettigar and S. Pavan, "A 15 mW 3.6 GS/s CT–$\Delta\Sigma$ ADC with 36 MHz bandwidth and 83 dB DR in 90 nm CMOS," in *Proc. IEEE Int. Solid-State Circuits Conf.*, San Francisco, CA, Feb. 2012, pp. 156–158. (cited on page 10)

[71] B. Sung, D. Jo, I. Jang, D. Lee, Y. You, Y. Lee, H. Park, and S. Ryu, "26.4 A 21fJ/conv-step 9 ENOB 1.6GS/S 2$\times$ time-interleaved FATI SAR ADC with background offset and timing-skew calibration in 45nm CMOS," in *Proc. IEEE Int. Solid-State Circuits Conf.*, San Francisco, CA, Feb. 2015, pp. 1–3. (cited on page 10)

[72] Y. Dong, R. Schreier, W. Yang, S. Korrapati, and A. Sheikholeslami, "29.2 A 235 mW CT 0–3 MASH ADC achieving-167 dBFS/Hz NSD with 53 MHz BW," in *Proc. IEEE Int. Solid-State Circuits Conf.*, San Francisco, CA, Feb. 2014, pp. 480–481. (cited on page 10)

[73] E. Janssen, K. Doris, A. Zanikopoulos, A. Murroni, G. v. d. Weide, Y. Lin, L. Alvado, F. Darthenay, and Y. Fregeais, “An 11b 3.6 GS/s time-interleaved SAR ADC in 65nm CMOS,” in *Proc. IEEE Int. Solid-State Circuits Conf.*, San Francisco, CA, Feb. 2013, pp. 464–465. (cited on page 10)

[74] M. Miyahara, I. Mano, M. Nakayama, K. Okada, and A. Matsuzawa, “22.6 A 2.2 GS/s 7b 27.4mW time-based folding-flash ADC with resistively averaged voltage-to-time amplifiers,” in *Proc. IEEE Int. Solid-State Circuits Conf.*, San Francisco, CA, Feb. 2014, pp. 388–389. (cited on page 10)

[75] M. N. Kulkarni, A. Ghosh, and J. G. Andrews, “A comparison of MIMO techniques in downlink millimeter wave cellular networks with hybrid beamforming,” *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1952–1967, May 2016. (cited on pages 11, 15, and 16)

[76] X. Yu, J. Shen, J. Zhang, and K. B. Letaief, “Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems,” *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485–500, Apr. 2016. (cited on pages 11, 15, and 16)

[77] X. Gao, L. Dai, S. Han, C. I, and R. W. Heath, “Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays,” *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, Apr. 2016. (cited on pages 11, 15, and 16)

[78] A. Alkhateeb and R. W. Heath, “Frequency selective hybrid precoding for limited feedback millimeter wave systems,” *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1801–1818, May 2016. (cited on pages 11, 15, and 16)

[79] S. Han, C. 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. (cited on pages 11, 15, and 16)

[80] 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. (cited on pages 11, 15, and 16)

[81] Y. M. Tsang and A. S. Y. Poon, “Detecting human blockage and device movement in mmWave communication system,” in *Proc. IEEE Globecom*, Houston, TX, Dec. 2011, pp. 1–6. (cited on pages 13, 14, and 23)

[82] L. Dai, B. Wang, Y. Yuan, S. Han, C. I, and Z. Wang, “Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends,” *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015. (cited on page 16)

[83] S. A. R. Naqvi and S. A. Hassan, “Combining NOMA and mmWave technology for cellular communication,” in *Proc. IEEE VTC Fall*, Montréal, Canada, Sep. 2016, pp. 1–5. (cited on page 16)

[84] J. Men, J. Ge, and C. Zhang, “Performance analysis of nonorthogonal multiple access for relaying networks over Nakagami-*m* fading channels,” *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1200–1208, Feb. 2017. (cited on page 16)

[85] M. Zeng, W. Hao, O. A. Dobre, and H. V. Poor, “Energy-efficient power allocation in uplink mmWave massive MIMO with NOMA,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3000–3004, Mar. 2019. (cited on page 17)

[86] Z. Xiao, L. Zhu, J. Choi, P. Xia, and X. Xia, “Joint power allocation and beamforming for non-orthogonal multiple access (NOMA) in 5G millimeter wave communications,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2961–2974, May 2018. (cited on page 17)

[87] L. Zhu, J. Zhang, Z. Xiao, X. Cao, D. O. Wu, and X. Xia, “Joint power control and beamforming for uplink non-orthogonal multiple access in 5G millimeter-wave communications,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 9, pp. 6177–6189, Sep. 2018. (cited on page 17)

[88] J. Cui, Y. Liu, Z. Ding, P. Fan, and A. Nallanathan, “Optimal user scheduling and power allocation for millimeter wave NOMA systems,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1502–1517, Mar. 2018. (cited on page 17)

[89] Z. Wei, L. Zhao, J. Guo, D. W. K. Ng, and J. Yuan, “Multi-beam NOMA for hybrid mmWave systems,” *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1705–1719, Feb. 2019. (cited on page 17)

[90] L. Dai, B. Wang, M. Peng, and S. Chen, “Hybrid precoding-based millimeter-wave massive MIMO-NOMA with simultaneous wireless information and power transfer,” *IEEE J. Sel. Areas Commun.*, vol. 37, no. 1, pp. 131–141, Jan. 2019. (cited on page 17)

[91] Z. Ding, L. Dai, R. Schober, and H. Vincent Poor, “NOMA meets finite resolution analog beamforming in massive MIMO and millimeter-wave networks,” *IEEE Commun. Lett.*, vol. 21, no. 8, pp. 1879–1882, Aug. 2017. (cited on page 17)

[92] N. Rupasinghe, Y. Yapıcı, İ. Güvenç, and Y. Kakishima, “Non-orthogonal multiple access for mmWave drone networks with limited feedback,” *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 762–777, Jan. 2019. (cited on page 17)

[93] Y. Azar, G. N. Wong, K. Wang, R. Mayzus, J. K. Schulz, H. Zhao, F. Gutierrez, D. Hwang, and T. S. Rappaport, “28 GHz propagation measurements for outdoor cellular communications using steerable beam antennas in New York city,” in *Proc. IEEE ICC*, Budapest, Hungary, Jun. 2013, pp. 5143–5147. (cited on page 18)

[94] H. Zhao, R. Mayzus, S. Sun, M. Samimi, J. K. Schulz, Y. Azar, K. Wang, G. N. Wong, F. Gutierrez, and T. S. Rappaport, “28 GHz millimeter wave cellular communication measurements for reflection and penetration loss in and around buildings in New York city,” in *Proc. IEEE ICC*, Budapest, Hungary, Jun. 2013, pp. 5163–5167. (cited on page 18)

[95] “FCC concludes first high-band 5G airwaves auctions,” FCC, Tech. Rep., May 2019. (cited on page 18)

[96] A. Narayanan, J. Carpenter, E. Ramadan, Q. Liu, Y. Liu, F. Qian, and Z.-L. Zhang. A first measurement study of commercial mmWave 5G performance on smartphones. 2019. [Online]. Available: arXiv:1909.07532 (cited on page 19)

[97] “Spectrum released for the 5G rollout,” Australian Department of Communications and the Arts, Tech. Rep., Oct. 2019. (cited on page 19)

[98] “On harmonisation of the 24,25-27,5 GHz frequency band for terrestrial systems capable of providing wireless broadband electronic communications services in the union,” European Union, Tech. Rep., May 2019. (cited on page 19)

[99] “China to establish 5G millimeter wave spectrum use plan,” The Ministry of Industry and Information Technology of the People’s Republic of China, Tech. Rep., Jun. 2019. (cited on page 19)

[100] “5G live networks: SK Telecom,” Ericsson, Tech. Rep., Apr. 2019. (cited on page 20)

[101] A. Lobinger, S. Stefanski, T. Jansen, and I. Balan, “Coordinating handover parameter optimization and load balancing in LTE self-optimizing networks,” in *Proc. IEEE VTC Spring*, Budapest, Hungary, May 2011, pp. 1–5. (cited on page 20)

[102] R. Kwan, R. Arnott, R. Paterson, R. Trivisonno, and M. Kubota, “On mobility load balancing for LTE systems,” in *Proc. IEEE VTC Fall*, Ottawa, Canada, Sep. 2010, pp. 1–5. (cited on page 20)

[103] N. O. Tuncel and M. Koca, “Joint mobility load balancing and inter-cell interference coordination for self-organizing OFDMA networks,” in *Proc. IEEE VTC Spring*, Glasgow, Scotland, May 2015, pp. 1–5. (cited on page 20)

[104] M. M. Hasan, S. Kwon, and J. Na, “Adaptive mobility load balancing algorithm for LTE small-cell networks,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2205–2217, Jan. 2018. (cited on page 20)

[105] S. M. Shahid, Y. T. Seyoum, S. H. Won, and S. Kwon, “Load balancing for 5G integrated satellite-terrestrial networks,” *IEEE Access*, vol. 8, pp. 132 144–132 156, Jul. 2020. (cited on page 20)

[106] A. Lobinger, S. Stefanski, T. Jansen, and I. Balan, “Load balancing in downlink LTE self-optimizing networks,” in *Proc. IEEE VTC Spring*, Taipei, Taiwan, May 2010, pp. 1–5. (cited on page 20)

[107] H. Hu, J. Zhang, X. Zheng, Y. Yang, and P. Wu, “Self-configuration and self-optimization for LTE networks,” *IEEE Commun. Mag.*, vol. 48, no. 2, pp. 94–100, Jan. 2010. (cited on page 20)

[108] B. Ma, B. Yang, Y. Zhu, and J. Zhang, “Context-aware proactive 5G load balancing and optimization for urban areas,” *IEEE Access*, vol. 8, pp. 8405–8417, Jan. 2020. (cited on page 20)

[109] H. Farooq, A. Asghar, and A. Imran, “Mobility prediction based proactive dynamic network orchestration for load balancing with QoS constraint (OPERA),” *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3370–3383, Jan. 2020. (cited on pages 20 and 21)

[110] V. V. Chetlur and H. S. Dhillon, “Coverage and rate analysis of downlink cellular vehicle-to-everything (C-V2X) communication,” *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1738–1753, Dec. 2020. (cited on page 21)

[111] Y. Zhou, V. W. S. Wong, and R. Schober, “Coverage and rate analysis of millimeter wave NOMA networks with beam misalignment,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8211–8227, Oct. 2018. (cited on page 21)

[112] M. Comisso and F. Babich, “Coverage analysis for 2D/3D millimeter wave peer-to-peer networks,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 7, pp. 3613–3627, May 2019. (cited on page 21)

[113] V. V. Chetlur and H. S. Dhillon, “Downlink coverage analysis for a finite 3-D wireless network of unmanned aerial vehicles,” *IEEE Trans. Commun.*, vol. 65, no. 10, pp. 4543–4558, Jul. 2017. (cited on page 21)

[114] M. G. Khoshkholgh, K. Navaie, K. G. Shin, and V. C. M. Leung, “Coverage analysis of multi-stream MIMO HetNets with MRC receivers,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, pp. 7816–7833, Sep. 2017. (cited on page 21)

[115] D. Maamari, N. Devroye, and D. Tuninetti, “Coverage in mmWave cellular networks with base station co-operation,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 2981–2994, Apr. 2016. (cited on page 21)

[116] S. Xu, N. Yang, and S. Yan, “Impact of load balancing on rate coverage performance in millimeter wave cellular heterogeneous networks,” in *Proc. IEEE ICC Workshops*, Kansas City, MO, May 2018, pp. 1–6. (cited on pages 22, 23, and 49)

[117] N. Wei, X. Lin, and Z. Zhang, “Optimal relay probing in millimeter-wave cellular systems with device-to-device relaying,” *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10 218–10 222, Dec. 2016. (cited on page 22)

[118] S. Wu, R. Atat, N. Mastronarde, and L. Liu, “Improving the coverage and spectral efficiency of millimeter-wave cellular networks using device-to-device relays,” *IEEE Trans. Commun.*, vol. 66, no. 5, pp. 2251–2265, May 2018. (cited on pages 22, 23, 51, and 52)

[119] G. Yang and M. Xiao, “Performance analysis of millimeter-wave relaying: Impacts of beamwidth and self-interference,” *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 589–600, Feb. 2018. (cited on page 22)

[120] S. Biswas, S. Vuppala, J. Xue, and T. Ratnarajah, “An analysis on relay assisted millimeter wave networks,” in *Proc. IEEE ICC*, Kuala Lumpur, Malaysia, May 2016, pp. 1–6. (cited on page 22)

[121] Y. Cai, Y. Xu, Q. Shi, B. Champagne, and L. Hanzo, “Robust joint hybrid transceiver design for millimeter wave full-duplex MIMO relay systems,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 1199–1215, Feb. 2019. (cited on page 22)

[122] C. Skouroumounis, C. Psomas, and I. Krikidis, “Low-complexity base station cooperation for mmWave heterogeneous cellular networks,” in *Proc. IEEE Globecom*, Washington, DC, Dec. 2016, pp. 1–6. (cited on page 23)

[123] X. Yu, J. Zhang, M. Haenggi, and K. B. Letaief, “Coverage analysis for millimeter wave networks: The impact of directional antenna arrays,” *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1498–1512, Jul. 2017. (cited on pages 23, 26, 27, 28, 49, 50, and 53)

[124] S. Xu, N. Yang, B. He, and H. Jafarkhaniand, “Coverage analysis of relay assisted millimeter wave cellular networks with spatial correlation,” in *Proc. IEEE WCNC*, Seoul, South Korea, Apr. 2020. (cited on page 23)

[125] X. Zhou, J. Guo, S. Durrani, and M. D. Renzo, “Power beacon-assisted millimeter wave ad hoc networks,” *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 830–844, Feb. 2018. (cited on pages 26 and 40)

[126] S. Singh, M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "Tractable model for rate in self-backhauled millimeter wave cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 10, pp. 2196–2211, Oct. 2015. (cited on pages 26, 27, and 49)

[127] F. Baccelli and B. Błaszczyszyn, *Stochastic Geometry and Wireless Networks: Volume I Theory*. Hanover, MA: Now Publishers, 2009. (cited on pages 27, 34, 35, 49, 51, and 72)

[128] N. Yang, M. Elkashlan, and J. Yuan, "Outage probability of multiuser relay networks in Nakagami-*m* fading channels," *IEEE Trans. Veh. Tech.*, vol. 59, no. 5, pp. 2120–2132, Jun. 2010. (cited on page 28)

[129] M. Elkashlan, P. L. Yeoh, N. Yang, T. Q. Duong, and C. Leung, "A comparison of two MIMO relaying protocols in Nakagami-*m* fading," *IEEE Trans. Veh. Tech.*, vol. 61, no. 3, pp. 1416–1422, Mar. 2012. (cited on page 28)

[130] "Enhancements for very high throughput in the 60 GHz band," *IEEE Standard 802.11ad*, Dec. 2012. (cited on page 29)

[131] "Millimeter-wave-based alternative physical layer extension," *IEEE Standard 802.15.3c*, Oct. 2009. (cited on page 29)

[132] "High rate 60 GHz PHY, MAC and PALs," *ECMA-387*, Dec. 2010. (cited on page 29)

[133] D. Stoyan, W. Kendall, and J. Mecke, *Stochastic Geometry and Its Applications*. Chichester, New York: Wiley, 1987. (cited on page 29)

[134] A. Thornburg and R. W. Heath, "Capacity and coverage in clustered LOS mmWave ad hoc networks," in *Proc. IEEE Globecom*, Washington, DC, Dec. 2016, pp. 1–6. (cited on page 49)

[135] A. Thornburg, T. Bai, and R. W. Heath, "Performance analysis of outdoor mmWave ad hoc networks," *IEEE Trans. Signal Process.*, vol. 64, no. 15, pp. 4065–4079, Aug. 2016. (cited on pages 53 and 54)

[136] M. Haenggi, "Diversity loss due to interference correlation," *IEEE Commun. Lett.*, vol. 16, no. 10, pp. 1600–1603, Oct. 2012. (cited on page 54)

[137] J. Deng, O. Tirkkonen, R. Freij-Hollanti, T. Chen, and N. Nikaein, "Resource allocation and interference management for opportunistic relaying in integrated mmWave/sub-6 GHz 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 94–101, Jun. 2017. (cited on page 66)

[138] A. Alkhateeb, S. Alex, P. Varkey, Y. Li, Q. Qu, and D. Tujkovic, "Deep learning coordinated beamforming for highly-mobile millimeter wave systems," *IEEE Access*, vol. 6, pp. 37 328–37 348, Jun. 2018. (cited on page 67)

[139] S. Hur, H. Yu, J. Park, W. Roh, C. U. Bas, R. Wang, and A. F. Molisch, "Feasibility of mobility for millimeter-wave systems based on channel measurements," *IEEE Commun. Mag.*, vol. 56, no. 7, pp. 56–63, Jul. 2018. (cited on page 67)

[140] N. Gonzalez-Prelcic, A. Ali, V. Va, and R. W. Heath, "Millimeter-wave communication with out-of-band information," *IEEE Commun. Mag.*, vol. 55, no. 12, pp. 140–146, Dec. 2017. (cited on page 67)

[141] R. Ford, M. Zhang, M. Mezzavilla, S. Dutta, S. Rangan, and M. Zorzi, "Achieving ultra-low latency in 5G millimeter wave cellular networks," *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 196–203, Mar. 2017. (cited on page 67)

[142] G. Yang, M. Xiao, J. Gross, H. Al-Zubaidy, and Y. Huang, "Delay and backlog analysis for 60 GHz wireless networks," in *Proc. IEEE Globecom*, Washington, DC, Dec. 2016, pp. 1–7. (cited on page 67)

[143] G. Yang, M. Xiao, and H. V. Poor, "Low-latency millimeter-wave communications: Traffic dispersion or network densification?" *IEEE Trans. Commun.*, pp. 1–1, Mar. 2018. (cited on page 67)

[144] G. Yang, M. Xiao, M. Alam, and Y. Huang, "Low-latency heterogeneous networks with millimeter-wave communications," *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 124–129, Jun. 2018. (cited on page 67)

[145] M. Zhang, M. Mezzavilla, R. Ford, S. Rangan, S. Panwar, E. Mellios, D. Kong, A. Nix, and M. Zorzi, "Transport layer performance in 5G mmWave cellular," in *Proc. IEEE INFOCOM Workshops*, San Francisco, CA, Apr. 2016, pp. 730–735. (cited on page 67)

[146] M. Zhang, M. Polese, M. Mezzavilla, J. Zhu, S. Rangan, S. Panwar, and M. Zorzi, "Will TCP work in mmWave 5G cellular networks?" *IEEE Commun. Mag.*, vol. 57, no. 1, pp. 65–71, Jan. 2019. (cited on page 67)