A Stochastic Geometry Approach to the Analysis and Optimization of Cellular Networks

Jian Song

université PARIS-SACLAY

NNT : 2019SACLS545

Thèse de doctorat

A Stochastic Geometry Approach to the Analysis and Optimization of Cellular Networks

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud,
au laboratoire des signaux et systèmes (L2S)

École doctorale n°580 Sciences et technologies de l'information et de la communication (STIC)
Spécialité de doctorat : Réseaux, Information et Communications

Thèse présentée et soutenue à Gif-Sur-Yvette, le 19 décembre 2019, par

JIAN SONG

Composition du Jury :

M. Jalel Ben-Othman Professeur, Université Paris 13, France	Président
M. Mustapha Benjillali Maître de conférences, INPT, Maroc	Rapporteur
M. Laurent Clavier Professeur, Institut Mines-Telecom, France	Rapporteur
M. Jean-Marie Gorce Professeur, INSA-Lyon, France	Examinateur
Mme. Lina Mroueh Maître de conférences, ISEP, France	Examinateur
Mme. Valeria Loscri Chargé de recherche, Inria Lille-Nord Europe, France	Examinateur
Mme. Maryline Helard Professeur, IETR, France	Examinateur
M. Marco Di Renzo Directeur de recherche, CNRS, France	Directeur de thèse

université PARIS-SACLAY
ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre : Analyse et Optimisation des Réseaux Cellulaires par la Géométrie Stochastique

Mots clés : Réseaux cellulaires, géométrie stochastique, processus de Poisson, optimisation, surfaces intelligentes reconfigurables.

Résumé : Cette thèse porte principalement sur la modélisation, l'évaluation des performances et l'optimisation au niveau système des réseaux cellulaires de nouvelle génération à l'aide de la géométrie stochastique. En plus, la technologie émergente des surfaces intelligentes reconfigurables (RISs) est étudiée pour l'application aux futurs réseaux sans fil. En particulier, reposant sur un modèle d'abstraction basé sur la loi de Poisson pour la distribution spatiale des nœuds et des points d'accès, cette thèse développe un ensemble de nouveaux cadres analytiques pour le calcul d'importantes métriques de performance, telles que la probabilité de couverture et l'efficacité spectrale potentielle, qui peuvent être utilisés pour l'analyse et l'optimisation au niveau système. Plus spécifiquement, une nouvelle méthodologie d'analyse pour l'analyse de réseaux cellulaires tridimensionnels est introduite et utilisée pour l'optimisation du système. Un nouveau problème d'allocation de ressources est formulé et résolu en combinant pour la première fois géométrie stochastique et programmation non linéaire mixte en nombres entiers. L'impact du déploiement de surfaces réfléchissantes intelligentes sur un réseau sans fil est quantifié à l'aide de processus ponctuels, et les avantages potentiels des RISs contre le relais sont étudiés à l'aide de simulations numériques.

Title : A Stochastic Geometry Approach to the Analysis and Optimization of Cellular Networks

Keywords : Cellular Networks, Stochastic Geometry, Poisson Point Process, Optimization, Reconfigurable Intelligent Surfaces.

Abstract : The main focus of this thesis is on modeling, performance evaluation and system-level optimization of next-generation cellular networks by using stochastic geometry. In addition, the emerging technology of Reconfigurable Intelligent Surfaces (RISs) is investigated for application to future wireless networks. In particular, relying on a Poisson-based abstraction model for the spatial distribution of nodes and access points, this thesis develops a set of new analytical frameworks for the computation of important performance metrics, such as the coverage probability and potential spectral efficiency, which can be used for system-level analysis and optimization. More specifically, a new analytical methodology for the analysis of three-dimensional cellular networks is introduced and employed for system optimization. A novel resource allocation problem is formulated and solved by jointly combining for the first time stochastic geometry and mixed-integer non-linear programming. The impact of deploying intelligent reflecting surfaces throughout a wireless network is quantified with the aid of line point processes, and the potential benefits of RISs against relaying are investigated with the aid of numerical simulations.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France

To my family...

Acknowledgements

Undertaking this PhD has been a truly life-changing experience for me, and it would not have been possible to do without the support and guidance that I received from many people.
First of all, I am extremely grateful to my research guide, my PhD supervisor, Prof. Marco Di Renzo, for his valuable guidance, scholarly inputs and consistent encouragement that I received throughout the research work. This feat was possible only because of the unconditional support provided by him. I always consider it as a great opportunity to do my PhD under his guidance and to learn from his research expertise. Thank you for all your help and support.
Many thanks also to Dr. Majid Safari and Dr. Shenjie Huang who encouraged and convinced me during our many discussions in the University of Edinburgh that I should pursue my doctoral degree and who made it possible for me to obtain the MSc degree with distinction in the University of Edinburgh. Thank you for telling me that will power, motivation and hard work are the main ingredients for success and are as important as intelligence.
I greatly appreciate all of the jury members: Prof. Laurent Clavier, Dr. Mustapha Benjillali, Prof. Maryline Helard, Prof. Jean-Marie Gorce, Dr. Lina Mroueh, Dr. Loscrì Valeria and Prof. Jalel Ben Othman for their constructive comments/questions and in particular, the insightful suggestions for the future work.
Some colleagues and faculty members of the lab have been very kind enough to extend their help at various phases of this research, whenever I approached them, and I do hereby acknowledge all of them. I thank all my colleagues in the group: Tu Lam Thanh, Shanshan Wang, Xiaojun Xi, Xuewen Qian, Viet-Dung Nguyen, Jiang Liu, Fadil Danufane, Farshad Shams, Kishor Chandra Joshi, Romain Fara, Arzhang Shahbazi and Abdelhamed Mohamed. My thanks also go out to the support I received from administrative staffs in the lab: Stephanie Douesnard, Maryvonne Giron, Delphine Maucherat, Lucie Romelot, Boris Lambert, Thomas Cuidu and Christian David.
This PhD study would not have been possible without all my Chinese friends in France who opened their homes to me during my time at CentraleSupélec and who were always so helpful in numerous ways. I am also very grateful to all the ESRs of the H2020-ETN-5Gaura project and all my research collaborators.
Finally, I am indebted to all my family members and relatives for their love and support these years. Last but not least, I devote this thesis to my parents, Xiaohong Yue and Yuqiang Song, who encouraged and helped me at every stage of my personal and academic life, and longed to see this achievement come true.

Gif sur Yvette, 20 December 2019 Jian Song

Abstract

The main focus of this thesis is on modeling, performance evaluation and system-level optimization of next-generation cellular networks by using stochastic geometry. In addition, the emerging technology of Reconfigurable Intelligent Surfaces (RISs) is investigated for application to future wireless networks. In particular, relying on a Poisson-based abstraction model for the spatial distribution of nodes and access points, this thesis develops a set of new analytical frameworks for the computation of important performance metrics, such as the coverage probability and potential spectral efficiency, which can be used for system-level analysis and optimization. More specifically, a new analytical methodology for the analysis of three-dimensional cellular networks is introduced and employed for system optimization. A novel resource allocation problem is formulated and solved by jointly combining for the first time stochastic geometry and mixed-integer non-linear programming. The impact of deploying intelligent reflecting surfaces throughout a wireless network is quantified with the aid of line point processes, and the potential benefits of RISs against relaying are investigated with the aid of numerical simulations.

Key words: Cellular Networks, Stochastic Geometry, Poisson Point Process, Optimization, Reconfigurable Intelligent Surfaces.

Résumé

Cette thèse porte principalement sur la modélisation, l'évaluation des performances et l'optimisation au niveau système des réseaux cellulaires de nouvelle génération à l'aide de la géométrie stochastique. En plus, la technologie émergente des surfaces intelligentes reconfigurables (RISs) est étudiée pour l'application aux futurs réseaux sans fil. En particulier, reposant sur un modèle d'abstraction basé sur la loi de Poisson pour la distribution spatiale des nœuds et des points d'accès, cette thèse développe un ensemble de nouveaux cadres analytiques pour le calcul d'importantes métriques de performance, telles que la probabilité de couverture et l'efficacité spectrale potentielle, qui peuvent être utilisés pour l'analyse et l'optimisation au niveau système. Plus spécifiquement, une nouvelle méthodologie d'analyse pour l'analyse de réseaux cellulaires tridimensionnels est introduite et utilisée pour l'optimisation du système. Un nouveau problème d'allocation de ressources est formulé et résolu en combinant pour la première fois géométrie stochastique et programmation non linéaire mixte en nombres entiers. L'impact du déploiement de surfaces réfléchissantes intelligentes sur un réseau sans fil est quantifié à l'aide de processus ponctuels, et les avantages potentiels des RISs contre le relais sont étudiés à l'aide de simulations numériques.

Mots clefs : Réseaux cellulaires, géométrie stochastique, processus de Poisson, optimisation, surfaces intelligentes reconfigurables.

Contents

Acronyms

ADC Analog-to-Digital Converter

AF Amplify-and-Forward

AI Artificial Intelligence

ANN Artificial Neural Networks

AP Access Point

AP Area Spectral Efficiency

BS Base Station

CCDF Complementary Cumulative Distribution Function

CDF Cumulative Distribution Function

CF Characteristic Function

CSI Channel State Information

DAC Digital-to-Analog Converter

DF Decode-and-Forward

eNB E-UTRAN Node B

FD Full Duplexing

GBD Generalized Benders Decomposition

GPU Graphics Processing Unit

HCN Heterogeneous Cellular Network

HD Half Duplexing

IRS Intelligent Reflecting Surface

KPI Key Performance Indicator

LTE Long Term Evolution

LTE-A Long Term Evolution Advanced

LOS Line of Sight

MGF Moment Generating Function

mmWave Milimeter Wave

MIMO Multiple Input Multiple Output

MRC Maximal Ratio Combining

MRT Maximum Ratio Transmission

MT Mobile Terminal

MILP Mixed Integer Linear Programming

MINLP Mixed Integer Non-Linear Programming

NB-IoT Narrow Band Internet of Things

NFV Network Function Virtualization

NLOS Non Line of Sight

NR New Radio

PDF Probability Distribution Function

PGFL Probability Generating Functional

PMF Probability Mass Function

PP Point Process

PPP Poisson Point Process

PSE Power Spectral Efficiency

RAN Radio Access Network

RAT Radio Access Technology

RF Radio Frequency

RIS Reconfigurable Intelligent Surface

Rx Receiver

SDN Software Defined Networking

SLA Service Level Agreement

SNR Signal-to-Noise Ratio

SINR Signal-to-Interference-plus-Noise Ratio

SIR Signal-to-Interference Ratio

Tx Transmitter

UAV Unmanned Aerial Vehicle

UDN Ultra Dense Network

3GPP The 3rd Generation Partnership Project

5G 5th generation wireless systems

6G 6th generation wireless systems

Notation

The following notation is used throughout this thesis.

j	the imaginary unit $j = \sqrt{-1}$
$\mathbf{x}$	a vector
$\mathbf{x}_l$	the lth entry of $\mathbf{x}$
$\|\mathbf{x}\|$	the norm of $\mathbf{x}$
$\mathbf{X}$	a matrix
$\|\mathbf{X}\|$	the Frobenius norm of $\mathbf{X}$
$(\cdot)^T$	the transpose operator
$(\cdot)^H$	the Hermitian operator
$H(\cdot)$	the Heaviside function
$\mathbf{1}(\cdot)$	the indicator function
$\min\{\cdot,\cdot\}$	the minimum function
$\max\{\cdot,\cdot\}$	the maximum function
$\mathbb{E}\{\cdot\}$	the expectation operator
$\Pr\{\cdot\}$	the probability measure function
$\exp\{\cdot\}$	the exponential function
$\log(\cdot)$	the logarithm function
$f_X(\cdot)$	the Probability Density Function (PDF) of X
$F_X(\cdot)$	the Cumulative Distribution Function (CDF) of X $F_X(\xi) = \Pr\{X \leq \xi\}$
$\bar{F}_X(\cdot)$	the Complementary CDF (CCDF) of X $\bar{F}_X(\xi) = 1 - F_X(\xi)$
$M_X(\cdot)$	the Moment Generating Function (MGF) of X: if X is real, $M_X(s) = \mathbb{E}_X\{\exp(-sX)\}$
${}_pF_q\left(a_1,\ldots,a_p;b_1,\ldots,b_q;\cdot\right)$	the generalized hypergeometric function

List of Figures

List of Tables

1 Introduction

This chapter begins with *Section 1.1* which introduces overview of the potential enablers in the next generation cellular netwprks alongside with the corresponding research challenges. *Section 1.2* highlights the major contributions in this thesis work and the organization of the thesis. *Section 1.3* provides lists of publications produced during my Ph.D. candidature.

1.1 Overview - Next Generation Cellular Networks

The 3rd Generation Partnership Project (3GPP) proposed the series of specifications, Long Term Evolution (LTE), which was evolved to satisfy the requirement of the current 4G cellular networks in 2009 [1]. Some of the next generation (5G) requirements may also be accomplished under specific deployment scenarios with aid of such an evolution [2]. However, they are not able to address all the 5G traffic demands and use cases that are characterized by ultra-high data rates, ultra-low latency and support for a massive number of connections [3]. To be more precise, LTE operates within a maximum 20MHz band per carrier and thus the limited achievable data rate is provided. Meanwhile, its rigid frame structure makes it challenging to reduce the round-trip latency below 1ms, and the designation behind LTE does not account for energy efficiency because the pilot signals are always on. Moreover, the demand and challenge brought by the Narrow Band Internet of Things (NB-IoT), which is applied to serve a wide range of massive devices, is difficult to fulfill under the current operation of LTE.

In order to overcome the limitations of the current deployed 4G cellular networks, 3GPP defines a new Radio Access Technology (RAT), 3GPP New Radio (NR) [4], to propose the novel technologies and approaches towards the 5G requirements and use cases [5]. In particular, the deployment of small cell based Ultra-Dense Networks (UDNs) will deliver expansive coverage, enhanced throughput, and traffic burden relief in Macro-cells. Furthermore, NR exploits new spectrum, millimeter wave (mmWave) frequency band, to increase the user experienced rate, the spectral efficiency, and to support the new technologies such as massive Multiple Input and Multiple Output (MIMO). From the point view of the core network, the novel schemes have been introduced to offer network slicing and virtualization in order to meet multiple Service Level Agreements (SLAs) among the different tenants of users, especially on the application of designing radio resource management mechanism in RAN.

Those aforementioned technologies of 5G will be discussed as follows associated with its general background on motivations and challenges. In addition, the brief introduction over the 5G issues and a promising mechanism beyond 5G (6G) will be explained as well. Other candidate enablers such as massive MIMO are out of scope of this thesis.

1.1.1 Ultra-Dense Networks (UDNs)

As one of the most effective ways to convey the ever-increasing data rate and to enhance the user experience, UDN [6] has emerged to be prominent solution over the last few years by namely adding more Base Stations (BSs) and Access Points (APs) thus accessing more spatial reuse of spectrum, and releasing the high traffic demands. This infrastructure densification trends are set to continue into 5G and beyond. UDNs can be defined as those networks whose deployed BSs density is much higher than that of the active Mobile Terminals (MTs) [7]. Another definition of UDN can be stated quantitatively where the density of a candidate network is measured as $\geqslant 10^3 \text{cells/km}^2$ [8].

Compared with the conventional cellular networks, UDNs hold the following fundamental features [7]. Firstly, apart from the current deployed Macro BSs, the network elements or access nodes in UDN scenario are operated mainly with low-power and small coverage, i.e., small cells (Picocells and Femtocells). Thus, the inter-site distance in this scenario would be in the range of few meters and the network topology is more likely heterogeneous. Under such coverage environments, the MTs would be very close to the BSs. As a result, the second feature of UDN is high Line-Of-Sight (LOS) transmission probability. In addition, due to the existence of high density of small cells, many BSs would be inactive especially under the light traffic load. This motivates the proposition of the idle mode concept. Another significant feature of UDN is the drastic interference between the adjacent cells considering the close proximity of small cells to each other.

Based on the above features, several associated research challenges emerge. Since the network becomes heterogeneous, usual spatial models, lattice model or hexagonal model, employed for accounting for network spatial topology fail to capture their characteristics. Even though conventional spatial models may be applied in some limited areas whose network topology is regular, their lack of analytical tractability will hinder their efficiency. To fulfill the near LOS transmission conditions under UDN circumstances, the need for different propagation models becomes a necessity. In particular, the widely adopted power-decaying pathloss model with single slope may not characterize the short range transmission of UDNs. Furthermore, the current fully loaded network model assumes that BSs always transmit signals towards MTs, thus overestimating interference. In this sense, a suitable load model by taking into account the BSs density and the dense deployment of BSs is necessary to examine to impact of ultra-densification.

1.1.2 Millimeter Wave

Apart from the ultra-densification, to exploit the never-addressed millimeter wave (mmWave) spectrum beyond sub-6GHz, ranging from 30 – 300 GHz, constitutes another vital means to supporting the unprecedented requirement of 5G [9], [10], [11]. This is because applying such frequency bands as carriers allows for larger bandwidth allocation, which could directly be translated into high data rate. Furthermore, by increasing the channel bandwidth for mobile radio will be able to support much better internet-based access and applications with minimal latency since the data capacity is greatly boosted while the latency for digital traffic is considerably decreased. In addition, mmWave frequencies within 30 – 300 GHz, which corresponds to 1 – 10 mm, can support new spatial processing techniques such as massive MIMO due to the much smaller wavelength. Meanwhile, mmWave systems are operated with highly directional antenna which enable users to be isolated and thus providing reduced interference. Up to now, on the other hand, the mmWave frequency bands were not considered useful for dynamic commercial applications in wireless communications such as cellular system even though it has been investigated for large range satellite and terrestrial communication systems. This is mainly from the new introduced challenges in realizing the

propagation properties and the channel impairment of high frequency bands [12].

Concerning the mmWave propagation for the next generation cellular networks, the first challenge comes from several important components in the channel modeling, especially the large-scale attenuation in the free space omnidirectional pathloss, which grows with the square of the carrier frequency according to Friis transmission law. A more significant concern with respect to the mmWave channel is that the transmitted signals are extremely susceptible to shadowing. In dense urban cities, for instance, outdoor materials such as bricks and human bodies may attenuate signals considerably due to shadowing fading. Statistically, this phenomenon is modeled to be log-normal distributed random variable in mmWave channels, which further causes the higher complexity in analytical formulation and computation. Also, incorporating mmWave communication into UDNs would necessitate highly intermittent connectivity and communication needs to be adaptable due to the rapid channel fluctuation caused by high frequencies.

In addition to above stated challenges in terms of the wireless channels, due to the limited range of mmWave signals, majority of the candidate cellular networks of mmWave system are deployed as small cells and ultra-densification as investigated in the previous sections. Accordingly, coverage in the dense urban environment will encounter NLOS much more frequently under the impact of large-scale blockages. Such blockage effects, resulting from reflection, scattering, diffraction, refraction and absorption, deeply impair the link reliability of wireless communications. However, these characteristics are either neglected or over-simplified in modeling the current cellular networks. Although random shape theory can be used to identify random located buildings in the bi-dimensional area, which allows us to evaluate the system performance in a tractable analytical approach, the supplied probabilistic model introduces new complexity in analyzing the influence of building height. Additionally, except for the investigation over the direct propagation path of the blockage effects, modeling some other properties such as reflections is more challenging and almost no paper have addressed this issue.

1.1.3 Network Slicing

Advanced service requirements from vertical industries necessitate a novel design of 5G mobile networks, in order to increase the revenues of telecom providers and network operators [13]. To address this challenge, such as network programmability and virtualization, is required. Network programmability brings about the benefits of automation and reactiveness of software modules, which allows one to (re)configure mobile networks dynamically while in operation. Network virtualization overcomes the limitations of monolithic network infrastructures, by abstracting the concept of "network function" and providing flexibility in composing, placing and managing these functions. In this context, the novel definition of network slicing in mobile networks, mainly based on Software Defined Networking (SDN) and Network Function Virtualization (NFV), encompasses such new requirements and constitutes

an enabler for potential economical benefits [14]. New vertical industries, e.g., automotive, smart-grid, e-health, public safety, factories, etc., are entering into the telecom market and are disrupting the traditional business models of telecom operators. They are forcing infrastructure providers to open their networks to tenants, a solution that provides incentives for monetizing the availability of isolated and secure (virtualized) network slices.

This new disruptive concept has spurred research interest in both academic and industrial communities. Its realization, however, requires the solution of a number of technical challenges that, for the time being, are not completely addressed and solved either in academia or industry. In the future, telecom providers envision an increasing demand for end-to-end network slices, which involve heterogeneous SLAs comprising different key performance indicators (KPIs), such as throughput, latency and reliability. The possibility of having different SLAs per slice require appropriate automated admission control and resource allocation protocols for designing efficient work management systems. In partucular, the resource management of the Radio Access Network (RAN) is one of the most challenging aspects that needs to be dealt with. Many challenges need to be solved for efficiently slicing the RAN. The available resources of the RAN air interface can be sliced at multiple levels: in frequency, time and power domains. Slicing the network at this level of granularity requires to account for the cellular network topology, the other cell interference and the radio channel conditions experienced by users of every single tenant. However, this challenging issue is never addressed and solved.

1.1.4 Meta-surfaces - Wireless Networks beyond 5G (6G)

The last two decades have witnessed the extremely fast revolution over the wireless communications and mobile cellular networks from 1G to 4G. Many addressed challenges and issues might be overcome once the 5G technologies will be operated by 2020. Nevertheless, there are still some open issues associated with 5G networks which have not been fully identified and that are attracting further research efforts [15]. First of all, with the advent of 5G, as a primary means to increase the network capacity, the site densification process will be witnessed. This common under-going approach, however, will pose economical issues that may slow down the substantially spatial and temporal 5G deployment. Meanwhile, the underutilized spectrum is not unlimited, and exploring new frequency bands will always create more obstacles in antenna design, digital signal processing and energy consumption.

In summary, the tremendous increase of traffic demands for the cellular revolution consists of using more power and emitting more radio waves from the point view of wireless network operators. Normally, this is accomplished by exploiting new frequency bands for transmission, and by densifying the deployment of networks, which is coincident with the current endeavor in developing the next generation wireless technologies where there are always more energy and spectral efficiency than the previous one, and the power consumption and the emission of radio waves are always growing with respect to the previous one. Accordingly, the constraints

originate from the concept that more data requires more power and more emission of radio waves might be released from alternative solutions to overcome the bottleneck of the wireless networks beyond 5G [16]. The potential approach is provided alongside with two fundamental questions: 1) What if the wireless environment can be controlled to use the energy-constrained devices to sense and measure the data without transmitting new radio waves? 2) What if a remote software-operated controller equipped with predictive capabilities can be applied to customize the propagation of radio waves in order to enhance the coverage without increasing the power consumption?

To answer the above two questions and further address the challenges and expectations of the future wireless networks in 6G that allow more than people, mobile devices and objects to communicate, the creation of the smart radio environment would serve as the potential candidate enabler. More precisely, future wireless networks need an environment that itself can be turned into a reconfigurable space in transferring and processing information. As a consequence, the concept of meta-surfaces, which are capable of modifying the propagation of radio waves in fully customizable way, is proposed as the basic aspect of the smart radio environments to equip the low-energy devices and to perform the low-complex algorithms. In particular, some highlighted issues in 5G networks could be solved with aid of the reconfigurable meta-surfaces. For instance, the high interference generated by other cells in UDNs can be adaptively controlled and mitigated. Also, the NLOS transmitted high blockage-sensitive mmWave signals can be reflected by a reconfigurable meta-surface coated object, the received signal strength at the mobile users is enhanced accordingly.

The above mentioned key applications relies on the fact that such materials are attached to the environmental objects as reflectors which are distributed in space according to very complex spatial patterns. However, modeling, evaluating and optimizing the performance of the smart radio environments would require the knowledge of: 1) the distribution of the locations of the meta-surfaces in large-scale networks, 2) the wave manipulations applied by the meta-surfaces depending on their spatial locations and on the radio waves impinging upon them, and 3) the spatial locations and wave manipulations applied by other randomly distributed meta-surfaces. The lack of tractable models in communication theory and wireless networks further make the analytical analysis more challenging.

1.2 Thesis Overview and Major Contributions

In the present thesis, motivated by the above stated research challenges for the upcoming 5G and beyond 5G, we investigate the system-level performance evaluation and optimization of future candidate cellular networks by using tools from stochastic geometry, and highlight the performance gain in terms of achievable rate in meta-surface aided wireless communications.

The major contributions of this thesis are the following:

i) this thesis studies the system-level optimization as a function of the antenna height and density of base stations in 3D cellular networks based on a tractable analytical framework (chapter 2);

ii) this thesis develops a efficient algorithm based on Generalized Benders Decomposition (GBD) approach to solve Mixed-Integer Non-Linear Programing (MINLP) problems for optimal RAN resource allocation that is inspired by the concept of network slicing (chapter 3);

iii) this thesis proposes a novel and tractable modeling approach of the reflection probability in wireless networks in the presence of metasurface-coated environmental objects (chapter 4);

iv) this thesis compares the differences, similarities and end-to-end rate performance of multi-antenna relaying and Reconfigurable Intelligent Surfaces (RISs) aided communications (chapter 5).

1.3 Publications

Journal Papers The following is a list of publications in refereed journals produced from the research outcomes of this thesis. These journal papers [101, 102, 103, 104] are used as the basis for this thesis.

($J1$) M. Di Renzo and J. Song, "Reflection probability in wireless networks with metasurface-coated environmental objects: an approach based on random spatial processes", *EURASIP Journal on Wireless Communications and Networking*, April, 2019. [101]
Abstract: An emerging and promising vision of wireless networks consists of coating the environmental objects with reconfigurable metasurfaces that are capable of modifying the radio waves impinging upon them according to the generalized law of reflection. By relying on tools from point processes, stochastic geometry, and random spatial processes, we model the environmental objects with a modified random line process of fixed length and with random orientations and locations. Based on the proposed modeling approach, we develop the first analytical framework that provides one with the probability that a randomly distributed object that is coated with a reconfigurable metasurface acts as a reflector for a given pair of transmitter and receiver. In contrast to the conventional network setup where the environmental objects are not coated with reconfigurable metasurfaces, we prove that the probability that the typical random object acts as a reflector is independent of the length of the object itself. The proposed analytical approach is validated against Monte Carlo simulations, and numerical illus-

trations are given and discussed.

(*J*2) K. Ntontin, M. Di Renzo, J. Song, F. Lazarakis, J. de Rosny, D.-T. Phan-Huy, O. Simeone, R. Zhang, M. Debbah, G. Lerosey, M. Fink, S. Tretyakov and S. Shamai, "Reconfigurable Intelligent Surfaces vs. Relaying: Differences, Similarities, and Performance Comparison", *IEEE Communications Magazines*, Submitted, August, 2019. [102]

Abstract: Reconfigurable intelligent surfaces (RISs) have the potential of realizing the emerging concept of smart radio environments by leveraging the unique properties of meta-surfaces. In this article, we discuss the need and potential applications of RISs in wireless networks that operate at high-frequency bands, e.g., millimeter wave (30-100 GHz) and sub-millimeter wave (greater than 100 GHz) frequencies; we overview the technology enabler of reconfigurable meta-surfaces; we elaborate on the working principle of RISs when used as low-complexity transmitters and anomalous mirrors; and we describe promising use cases that encompass signal, interference, security, and scattering engineering. When used as anomalous mirrors, the RISs resemble multiple-antenna relay stations. Therefore, we elaborate on the key differences and similarities between RISs and relaying, and, notably, illustrate numerical results that highlight the spectral efficiency gains of RISs whose transverse size is large as compared with the wavelength of the radio wave. Finally, we discuss the key open issues that need to be addressed in order to unlock the potential of RISs, and, in particular, to unveil their ultimate performance limits and to optimize their deployment in future wireless networks.

(*J*3) J. Song, M. Di Renzo and D. Lopez-Perez, "System-Level Modeling and Optimization of the Coverage in 3D Poisson Cellular Networks - A Stochastic Geometry Approach", *IEEE Transactions on Wireless Communications*, in Submission, 2019. [103]

Abstract: In this paper, we propose a new tractable approach to model the 3D cellular networks where all the Base Stations (BSs) are deployed by taking into account their antenna height. In addition, the LOS/NLOS effect is quantified by a novel 3D blockage model which is verified to be tractable. Its rational relies on generalizing the ball model by incorporating the height of buildings and BSs with aid of the random shape theory. To the end, by leveraging the mathematical tool behind stochastic geometry, we develop a tractable analytical framework of coverage probability to evaluate the performance of 3D cellular networks. Beyond most of the current research outcomes, in particular, a closed-form upper bound of coverage probability is obtained under the typical assumptions, which provides a substantial approach for system-level optimization. It is disclosed that, under some sufficient conditions, there exists an optimal height of BSs to maximize coverage probability and the explicit expression of the optimal height can be obtained, which is analyzed under the impact of different system parameters. The exact expression of optimal BSs' density is also revealed under the assumption of fully

loaded networks, and the insightful performance trends are studied. The performance of the proposed methodologies are validated through numerical simulations.

(*J*4) J. Song and M. Di Renzo, "Optimization of Resource Allocation in Multi-Tenants Radio Access Network Slicing: A Generalized Benders Decomposition Algorithm", *IEEE Wireless Communications Letters*, in Submission, 2019. [104]
Abstract: Network virtualization and softwarization techniques are key-enablers of the novel Network Slicing concept behind 5G. Network slicing introduces new business models such as allowing telecom providers to lease a virtualized slice of their infrastructure to tenants such as industry verticals, e.g. automotive, ehealth, factories, etc. However, this new paradigm poses a major challenge when applied to Radio Access Networks (RAN): How to maximize the overall system performance while maintaining the diverse service level agreements (SLAs) requested by the infrastructure tenants?

In this paper, based on the concept of stochastic geometry, we propose a new analytical framework to model realistic RAN that leverages the business opportunities provided by network slicing. We design the radio resource allocation strategy through a mixed integer nonlinear programming (MINLP) optimization to best allocate transmit power and bandwidth (i.e., a slice of radio access resources) to the users of each infrastructure tenant by considering the per slice SLA guaranteed average rate condition. Under the linear separable property of the problem, we apply the efficient Generalized Benders Decomposition (GBD) where the global convergence is achieved. Numerical results are illustrated to validate the proposed solution, and the performance is compared against brute-force search and Greedy algorithm.

Conference Papers The following is a list of publications in refereed conference proceedings that originated from the main findings of this thesis. The conference papers [9, 105] contain material not presented in this thesis.

(*C*1) J. Song, L. T. Tu and M. Di Renzo, "On the Feasibility of Interference Alignment in Ultra-Dense Millimeter-Wave Cellular Networks", *IEEE Asilomar Conference on Signals, Systems and Computers*, pp. 1176–1180, Pacific Grove, CA, USA, Nov. 2016. [9]
Abstract: Recent studies have demonstrated that millimeterwave cellular networks may operate either in the noise- or in the interference-limited regime, depending on several parameters, which include the density of base stations, the density and size of obstacles/blockages, the antenna beamwidth, and the transmission bandwidth. The objective of the present paper is to exploit tools from stochastic geometry for obtaining

a mathematically tractable framework that allows us to identify the operating conditions under which millimeter-wave cellular networks operate in the inference-limited regime and to assess the potential advantages of interference alignment, by taking into account the overhead cost of base station cooperation due to the estimation of channel state information.

(*C*2) J. Song and M. Di Renzo, "On the Stochastic Geometry Modeling and Optimization of Multi-Tenant Network Slicing", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Invited, Sep. 2018. [105]

Abstract: Network virtualization and softwarization techniques are key-enablers of the novel Network Slicing concept. Network slicing introduces new business models such as allowing telecom providers to lease a virtualized slice of their infrastructure to tenants such as industry verticals, e.g. automotive, ehealth, factories, etc. In this paper, regarding this concept, we propose a new analytical framework, based on stochastic geometry theory, to model realistic RANs that leverage the business opportunities offered by network slicing. Moreover, we formulate a 2-Tenant optimization problem in order to optimally allocate transmit power and bandwidth (i.e., a slice of radio access resources) to the users of each infrastructure tenant. Numerical results are illustrated to validate our proposed solution in terms of potential spectral efficiency.

2 System-Level Modeling and Optimization of the Coverage in 3D Poission Cellular Networks – A Stochastic Geometry Approach

Chapter 2. System-Level Modeling and Optimization of the Coverage in 3D Poission Cellular Networks – A Stochastic Geometry Approach

In this chapter, we propose a new tractable approach to model the 3D cellular networks where all the Base Stations (BSs) are deployed by taking into account their antenna height. In addition, the LOS/NLOS effect is quantified by a novel 3D blockage model which is verified to be tractable. Its rational relies on generalizing the ball model by incorporating the height of buildings and BSs with aid of the random shape theory. To the end, by leveraging the mathematical tool behind stochastic geometry, we develop a tractable analytical framework of coverage probability to evaluate the performance of 3D cellular networks. Beyond most of the current research outcomes, in particular, a closed-form upper bound of coverage probability is obtained under the typical assumptions, which provides a substantial approach for system-level optimization. It is disclosed that, under the sufficient conditions, there exists an optimal height of BSs to maximize coverage probability and the explicit expression of the optimal height can be obtained, which is analyzed under the impact of different system parameters. The exact expression of optimal BSs' density is also revealed under the assumption of fully loaded networks, and the insightful performance trends are studied. The performance of the proposed methodologies are validated through numerical simulations.

2.1 Introduction

As an emerging mathematical tool, stochastic geometry has been widely applied for the system-level modeling, performance evaluation and optimization in the last few years. The principle of such tractable methodology is based on the following guidelines which are investigated originally in [17]: 1) All the Base Stations (BSs) or other network elements are modeled as points of a Poisson Point Process (PPP); 2) Path-losses behave as a fundamental role in wireless channels, and they are formulated as a simplistic distance-dependent function according to the power-decaying law; 3) The performance of typical Mobile Terminal (MT) located at the origin is analyzed and the BSs in cellular network is assumed to be fully loaded. The results illustrated in the above comprehensive work provide the general instructions for the theoretical studies. On the other hand, there is still a huge gap to fill from the philosophical research to practical application in modeling and optimizing realistic cellular network architectures. This gap mainly comes from the following aspects which are highlighted by the next generation communications and network topologies.

- From the empirical data, the practical Heterogeneous Cellular Network (HCNs) deployments are likely to exhibit different levels of correlations among the locations of BSs such as repulsion, spatial inhibition and spatial aggregation [18]. The Utilization of the completely random point processes such as homogeneous PPP seems to lose the accuracy to overcome such limitations. The non-PPP based models, however, are considered to be mathematical intractable for system-level analysis.

- The blockage effects severely impair the performance in urban cellular networks, especially at millimeter-wave (mmWave) frequency bands, which are expected to provide higher data rate in the 5G cellular system [19]. Recent experimental investigations have demonstrated a high sensitivity of the mmWave signals to blockage effects exists in dense urban cities. The bottom line of modeling such phenomenon is to develop a novel distance-dependent function in order to quantify significant Line-Of-Sight (LOS) and None-Line-Of-Sight (NLOS) effect caused by large-scale buildings.

- To model the elements of cellular networks in a bi-dimensional Euclidean space is not coincident with the practical network deployment and wireless channels. The more realistic assumptions in modeling cellular networks by taking into account impact of BSs antenna height and buildings height in the 3D candidate space, nevertheless, restricts the analytical formulations of performance metrics due to the general intractability when new system parameters are introduced.

- For simplicity, the load modeling in cellular networks is normally neglected for the convenience of mathematical analysis and the BSs are assumed to be fully-loaded. In this scenario, the performance of the typical MT is assessed. As a consequence, there exists an overestimation issue for the capacity due to the perfect geometry of BSs and weak interference from other tiers are neglected [20].

The above highlighted issues constitute the main research challenges in modeling 3D cellular networks, especially within the concept of Ultra Dense Networks (UDNs) [21]. In the following, we review the key investigation over the development of blockage models and the analysis over the generalized 3D cellular networks.

2.1.1 Related Work

Several available papers have provided comprehensive guidelines in modeling the blockage effects caused by buildings. In particular, the 3GPP standards [22] suggest modeling it by differentiating the LOS and NLOS links using a stochastic model to describe the probability of link states. This complex analytical model, however, is demonstrated to be intractable for system-level analysis. To overcome this issue, a tractable blockage model is proposed from random shape theory [23], where the centers of objects form a PPP, and each object is allowed to have independent shape, size, and orientation according to certain distributions. Relying on the random shape theory, this model is expressed by a negative exponential function to generalize the size and density of buildings. The resulting analytical frameworks, however, are formulated with high mathematical complexity for further insightful performance evaluation. To simplify the mathematical derivation in the system-level analysis, a LOS ball model is introduced in [24] as an approximation approach. Moreover, the generalized ball model is developed to increase the accuracy in [25] and the intensity matching based numerical optimization method is proposed to determine the model parameters. In summary, the above mentioned blockage models can only be applied in 2D case studies, which is not suitable for analyzing the actual 3D cellular systems. On the other hand, the current available 3D blockage model with the random height of buildings in [26] is verified to be mathematical intractable.

In terms of the analysis over the impact of BSs' height, several unfavorable research outcomes can be found in the fast few years. The results in [27] reveal that the promising solution to overcome the Area Spectral Efficiency (ASE) Crash is to lower the BSs' height in UDNs. Such findings, on the other hand, are obtained based on the simplistic system models where the networks are assumed to be fully loaded. The pessimistic results above, once again, are validated against Monte-Carlo simulations in [28], where the blockages with random placed buildings are considered. In particular, it is revealed that the maximum ASE is proportional to the inverse of the square of BSs' height. However, the proposed mathematical framework is based on 2D simplified blockage ball model and the association probability is independent of BSs' height.

Despite the negative findings, there are also some optimistic results from [29] and [30] where the optimal BSs' height exists. In [29], the framework of coverage probability is proposed by approximating the intractable 3D blockage model from [26] with an exponential function where tuned parameters are computed numerically. However, the rational behind such approximation does not agree with the random shape theory in [23], [24]. In [30], the tractable frameworks of coverage probability and ASE are obtained to indicate the optimal BSs' eleva-

tion angle. Nevertheless, due to the impact of integration inside the objective function, the optimization is only conducted numerically via. Based on the generalized system models, where both non-zero BSs' height and finite density of MTs are taken into account, the authors in [31] examine the asymptotic regime for coverage probability when density of BSs is infinity, which leads to a new SINR invariance law. Such desirable property is shown to be invalid in [32] for any nonzero height of BSs, and the conditions for SINR invariance law to be hold are summarized for different network models.

2.1.2 Research Contribution and Novelty

Despite providing new and interesting results, there are several shortcomings from the previous case studies. The findings are either too pessimistic or too optimistic in terms of the system performance. Thus, in the depicted context, the specific novel contributions made by this chapter are as follows:

- We introduce a tractable 3D blockage model with a modified 3D LOS ball, which is generalized with the impact of BSs height and buildings height. This model is verified analytically by the random shape theory by incorporating height of blockages [23].
- A tractable mathematical framework of coverage probability is proposed for performance evaluation under the generalized 3D system model with aid of stochastic geometry. The asymptotic regimes are studied based on this exact framework.
- We approximate the exact analytical expressions of coverage probability as a closed-form upper bound. The proposed upper bound is leveraged for the analysis beyond the asymptotic case studies with the exact formulations. Notably, the approximated closed-form expressions of the optimal BSs' height and deployment density are obtained for further investigation.
- We state the sufficient conditions for the optimal BSs' height to exist and the numerical results of optimization are obtained and verified through brute-force search. In the optimized settings, it is noticed that the SIR invariance property still holds.

The remainder of the present chapter is organized as follows. In Section 2.2, the system model is summarized. In Section 2.3, the proposed mathematical framework and closed-form upper bound of coverage probability are introduced, and several asymptotic regimes are analyzed. In Section 2.4, the system-level optimization over the BSs' height and density is studied. In Section 2.5, the analytical results are generalized by taking into account the new definition of coverage probability. In Section 2.6, numerical illustrations are shown to validate the mathematical findings and to analyze the impact of different system parameters. Finally, Section 2.7 concludes this chapter.

Notation: The main symbols and functions used in this chapter are reported in Table 2.1.

Table 2.1 – Summary of main symbols and functions used throughout the chapter.

Symbol/Function	Definition
$\mathbb{E}\{\cdot\}$, $\Pr\{\cdot\}$	Expectation operator, probability measure
λ_{BS}, λ_{MT}, $\lambda_{\text{buildings}}$	Density of base stations, mobile terminals, blockages
Ψ_{BS}, Ψ_{MT}, $\Psi_{\text{BS}}^{(\text{I})}$	PPP of base stations, mobile terminals, interfering base stations
$\text{BS}^{(0)}$, $\text{BS}^{(i)}$, $\text{BS}^{(n)}$	Serving, interfering, generic base station
h, $\tilde{h}$	height of BSs and blockages
D, $D_h = D\max\left\{\frac{h}{\tilde{h}},1\right\}$, $q_s^{[A,B)}$	2D, 3D LOS ball radius, LOS probability from $[A,B)$
P_{tx}, σ_{N}^2, $I_{\text{agg}}(\cdot)$	Transmit power of base stations, noise variance, aggregate-interference
$r^{(n)}$, $g_s^{(n)}$	Distance, fading power gain of a generic link in state s
$L_s\left(r^{(n)}\right)$, $L_s^{(n)}$, $L^{(0)}$	Path-loss, shorthand of path-loss in state s, path-loss of intended link
k_s, $\beta_s > 0$	Path-loss constant, slope (exponent) in state s
$\Lambda_s([0,x))$, $\Lambda_s^{(1)}([0,x))$	Path-loss intensity measure, its derivative in state s
γ_{th}, γ_{A}	Reliability threshold for decoding, cell association
$L\left(\frac{\lambda_{\text{MT}}}{\lambda_{\text{BS}}}\right) = 1-\left(1+\frac{\lambda_{\text{MT}}}{3.5\lambda_{\text{BS}}}\right)^{-3.5}$	Probability that a base station is in transmission mode
$f_X(\cdot)$, $M_X(\cdot)$	Probability density/mass, moment generating function of X
$\mathbf{1}(\cdot)$, ${}_2F_1(\cdot,\cdot,\cdot,\cdot)$, $H(\cdot)$, $\bar{H}(\cdot) = 1-H(\cdot)$	Indicator function, Gauss hypergeometric function, Heaviside function
$\max\{x,y\}$, $\min\{x,y\}$, $df(x)/dx$	Maximum, minimum between x and y, derivative of $f(x)$
$\Theta_s(z) = {}_2F_1\left(1,\frac{2}{\beta_s},1+\frac{2}{\beta_s},-z\right) \in [0,1]$	Shorthand
$\Gamma_s(z) = {}_2F_1\left(-\frac{2}{\beta_s},1,1-\frac{2}{\beta_s},-z\right) \in [1,\infty)$	Shorthand
$D_1 = \left(\frac{D^2}{\tilde{h}^2}+1\right)\mathbf{1}\left(h \geqslant \tilde{h}\right) + \left(\frac{D^2}{h^2}+1\right)\mathbf{1}\left(h < \tilde{h}\right)$	Shorthand
$\Delta, \Upsilon_1, \Upsilon_2, \Phi$, SINR	Shorthands for $\text{P}_{\text{cov}}^{\text{UB}}$, signal-to-interference-plus-noise-ratio
SIR, $\overline{\text{SNR}}$	Signal-to-interference-ratio, average signal-to-noise-ratio
P_{cov}, $\text{P}_{\text{cov}}^{\text{UB}}$, $\overline{\text{P}}_{\text{cov}}^{\text{UB}}$	Exact framework, upper-bound, approximated upper bound of coverage
$\text{P}_{\text{cov}}^{\text{G}}$, $\text{P}_{\text{cov}}^{\text{G}-\overline{\text{UB}}}$	Generalized framework, upper-bound of coverage (new definition)

2.2 System Model

2.2.1 Cellular Network Modeling

A downlink single-tier cellular network is considered. The BSs are modeled as points of a homogeneous PPP, denoted by Ψ_{BS}, of density λ_{BS}. The MTs are modeled as another homogeneous PPP, denoted by Ψ_{MT}, of density λ_{MT}. Ψ_{BS} and Ψ_{MT} are independent of each other. As a main component of 3D network, we assume all the BSs are elevated with the same height $h \geq 0$, whereas the MTs are located at the ground level for simplicity of exposition. The BSs and MTs are equipped with a single omnidirectional antenna. Each BS transmits with a constant power denoted by P_{tx}. The mathematical frameworks are developed for the typical MT, denoted by $\text{MT}^{(0)}$, that is located at the origin (Slivnyak theorem [33, Th. 1.4.5]). The BS serving $\text{MT}^{(0)}$ is denoted by $\text{BS}^{(0)}$. Throughout this chapter, the superscripts 0, i and n identify the intended link, a generic interfering link, and a generic BS-to-MT link, respectively. The set of interfering BSs is denoted by $\Psi_{\text{BS}}^{(\text{I})}$. Note that the analysis over more generalized network models by using non-PPP and random BSs' height will be investigated in the future work.

2.2.2 Blockage Modeling

To unify the analysis with 3D cellur systems, we adopt a new blockage model where the height of buildings is taken into account. The principle of such model can be justified from the concept of random shape theory in [23]. For simplicity but without loss of generality, the buildings are modeled as line segment process with the fixed height $\tilde{h} \geq 0$. The considered system model is depicted as follows in Fig. 2.1. Consider a building intersecting the link OX at the point which is at a horizontal distance y away from X. As shown in the figure, the building blocks the direct propagation path OX' only if its height $\tilde{h} > h_y$, where h_y can be computed as $h_y = \frac{(r-y)h}{r}$. r is the distance between the BS and the MT, y denotes the location of the building, which is a uniformly distributed random variable in $[0, r]$. Hence given the candidate building blocks OX, the probability it blocks OX' is:

$$\begin{aligned} \eta_h &= \Pr\{\tilde{h} > h_y\} = \mathbb{E}_y\{\Pr\{\tilde{h} > h_y | y\}\} = \frac{1}{r}\int_0^r \Pr\left\{\tilde{h} > \frac{(r-y)h}{r}\right\} dy = \frac{1}{r}\int_{r-\frac{\tilde{h}}{h}r}^{r} 1 dy \\ &\overset{(a)}{=} \frac{\tilde{h}}{h} 1\left(\tilde{h} \leqslant h\right) + 1 \cdot 1\left(\tilde{h} > h\right) \end{aligned} \tag{2.1}$$

where the (a) takes into account that the integral will be zero if $\tilde{h} > h$. If, however, $\tilde{h} < h$ then it is known that with probability equal to one the building will cross the line connecting the BS with the mobile terminal regardless of the location of the building in $[0, r]$. Theoretically, η_h accounts for the impact of the buildings and can be interpreted as the probability that a building is in between the BS and the MT then it blocks the LOS line that connects the BS and the MT. From the standpoint of practice, the obtained equation makes sense as it states that the blocking probability increases with the height of the building increases and decreases if the height of the BS increases. Also, it depends on their ratio.

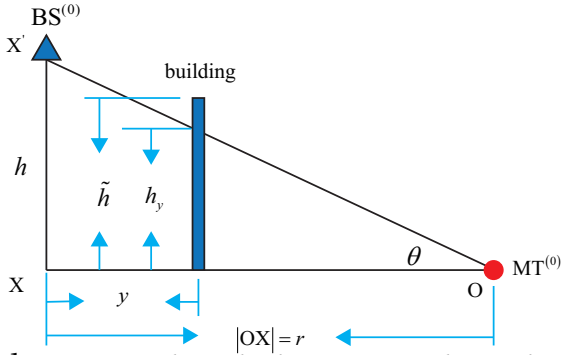

Figure 2.1 – Geometry relationship between a typical MT and its serving BS

From (2.1), it is noticed that η_h is only determined by the heights, thus, the LOS probability in 3D scenario also depends on the mean number of blockages between the BS and its associated MT with link length r, which is given by

$$\mathbb{E}\{K_h\} = \eta_h \mathbb{E}\{K\} = \eta_h \left(\frac{2\lambda_{\text{buildings}}\mathbb{E}\{L\}}{\pi}\right) r = \eta_h \mu_{\text{buildings}} r \tag{2.2}$$

where $\mathbb{E}\{K\}$ is the average number of blockages between OX, $\mu_{\text{buildings}} = \frac{2\lambda_{\text{buildings}}\mathbb{E}\{L\}}{\pi}$ depends on the density of blockages $\lambda_{\text{buildings}}$ and their average size $\mathbb{E}\{L\}$ [23]. Based on random shape theory and by the fact that the number of buildings that intersect the line connecting the BS with the MT is a Poisson random variable with mean $\mathbb{E}\{K_h\} = \eta_h \mu_{\text{buildings}} r$, the LOS probability can be formulated as follows:

$$p_{\text{LOS}}\left(r, h, \tilde{h}\right) = \exp\{-\mathbb{E}\{K_h\}\} = \exp\left\{-\left(\frac{\tilde{h}}{h}\mathbf{1}\left(\tilde{h} \leqslant h\right) + 1 \cdot \mathbf{1}\left(\tilde{h} > h\right)\right)\mu_{\text{buildings}} r\right\} \tag{2.3}$$

By using this result, we can generalize the LOS single-ball model by taking into account the height of the BSs and of the buildings. In particular, the main idea is to approximate the 3D LOS/NLOS model with an equivalent 2D LOS/NLOS model whose parameters depend not only on the horizontal distance r but also on the heights of the BSs and buildings. Based on the analytical expression of the LOS probability above, we evince the following:

Remark 2.1 *Given a horizontal distance between the BSs and the MTs, we observe that the LOS probability is the same as in the 2D case, i.e., the heights do not matter, if $\tilde{h} > h$. In other words, if the height of the BSs is less than the height of the buildings, then the 3D system boils down to a 2D system.*

Remark 2.2 *Given a horizontal distance between the BSs and the MTs, we observe that the LOS probability increases compared with the LOS probability of the 2D case if $h \geq \tilde{h}$. In other words, if the height of the BSs is greater than the height of the buildings, then the 3D system provides us with a better LOS probability compared with the 2D system. In other words, mobile terminals that, based on the 2D model, are located at longer distances from the BSs are still in LOS thanks to the condition $h \geq \tilde{h}$.*

These remarks allow us to generalize the concept of LOS ball to the 3D case, by taking as a reference the 3D case. In particular, based on [24], the 3D LOS ball can be written as follows:

$$\int_0^{+\infty} \exp\left\{-\left(\frac{\tilde{h}}{h}1\left(\tilde{h} \leq h\right) + 1 \cdot 1\left(\tilde{h} > h\right)\right)\mu_{\text{buildings}} r\right\} r\,dr = \int_0^{D_h} r\,dr \tag{2.4}$$

which yields

$$D_h = \frac{2}{\mu_{\text{buildings}}} \max\left\{\frac{h}{\tilde{h}}, 1\right\} = D \max\left\{\frac{h}{\tilde{h}}, 1\right\} \tag{2.5}$$

where $D = \frac{2}{\mu_{\text{buildings}}}$ is the LOS ball of the 2D model.

This result substantiates the adopted LOS model, by taking into account that the probability of the 2D LOS model is not one inside the ball and zero outsides but can be an arbitrary number

in [0,1]. Thus, the generalized LOS ball model for $s \in \{\text{LOS}, \text{NLOS}\}$ is given by

$$p_s\left(r, h, \tilde{h}\right) = q_s^{[0,D)} \mathbf{1}\left(r < D_h\right) + q_s^{[D,\infty)} \mathbf{1}\left(r > D_h\right) \tag{2.6}$$

where $\left\{q_s^{[0,D)}, q_s^{[D,\infty)}\right\}$ are the LOS probabilities of the underlying 2D model. Based on the previous studies in [25], by applying such model will provide the most tractable resulting mathematical frameworks for analysis.

2.2.3 Channel Modeling

A quasi-static channel model is assumed for each BS-to-MT link, where path-loss and fast-fading are considered for the different link states, i.e., LOS and NLOS. Shadowing is not explicitly taken into account because its net effect lies in modifying the density of the BSs [25]. All BS-to-MT links are assumed to be mutually independent and identically distributed (i.i.d.).

Path-Loss Consider a generic BS-to-MT link of length $r^{(n)}$ in state $s \in \{\text{LOS}, \text{NLOS}\}$. The path-loss is defined as $L_s\left(r^{(n)}\right) = k_s\left[\left(r^{(n)}\right)^2 + h^2\right]^{-\frac{\beta_s}{2}}$, where k_s and β_s are the path-loss constant and the path-loss slope (exponent).

Fast-Fading Consider a generic BS-to-MT link in state $s \in \{\text{LOS}, \text{NLOS}\}$. The power gain due to small-scale fading is assumed to follow an exponential distribution (i.e., Rayleigh fading is assumed) with mean Ω_s. Without loss of generality, $\Omega_{\text{LOS}} = \Omega_{\text{NLOS}} = 1$ is assumed. The power gain of a generic BS-to-MT link in state s is denoted by $g_s^{(n)}$.

2.2.4 Cell Association Criterion

A cell association criterion based on the minimum path-loss is assumed. Furthermore, we assume that the LOS or NLOS status of a BS-to-MT link is independent of the LOS or NLOS status of any other BS-to-MT link. In this sense, no spatial correlations among the BS-to-MT links is examined. Under this assumption, the homogeneous of PPP of the BSs, Ψ_{BS}, can be partioned into two independent homogeneous PPPs, $\Psi_{\text{BS},s}$, for $s \in \{\text{LOS,NLOS}\}$, such that $\cup_{s\in\{\text{LOS,NLOS}\}} \Psi_{\text{BS},s} = \Psi_{\text{BS}}$. Let $\text{BS}^{(n)}$ denote a generic BS of the network. The serving BS, $\text{BS}^{(0)}$, is obtained as follows:

$$\text{BS}^{(0)} = \underset{s\in\{\text{LOS,NLOS}\},\text{BS}^{(n)}\in\Psi_{\text{BS},s}}{\arg\min} \left\{L_s\left(r^{(n)}\right)\right\} = \underset{s\in\{\text{LOS,NLOS}\},\text{BS}^{(n)}\in\Psi_{\text{BS},s}}{\arg\min} \left\{L_s^{(n)}\right\} \tag{2.7}$$

where the short-hand notation $L_s^{(n)} = L_s\left(r^{(n)}\right)$ is used. As for the intended link, the following will hold: $L^{(0)} = \min\left\{L_{\text{LOS}}^{(0)}, L_{\text{NLOS}}^{(0)}\right\}$, where $L_s^{(0)}$ for $s \in \{\text{LOS,NLOS}\}$ is defined by $L^{(0)} = \min\limits_{\text{BS}^{(n)}\in\Psi_{\text{BS},s}} \left\{L_s^{(n)}\right\}$.

2.2.5 Load Modeling

Based on the cell association criterion in (2.7), several or no MTs can be associated to a generic BS. The latter case accounts for that the BSs will transmit zero power, i.e., $P_{tx} = 0$, and, thus, it does not generate inter-cell interference. However, as for the former case, the fundamental role of the MTs in cellular networks relies on the suitable load modeling. In the present chapter, let $N_{MT} \geq 1$ denote the number of MTs associated to a generic BS, we assume all the MTs are scheduled for transmission in the cell and that the resources (transmit power and bandwidth) from the BSs are equally shared among them. Another models are leveraged and compared in [34] for the optimization of energy efficiency. It is shown that they provide the same behavior in terms of the coverage and spectral efficiency.

Within the considered load model, the number of MTs that lie inside an arbitrary cell is N_{MT}. The probability that the BS is in idle mode $\mathbb{P}_{BS}^{(idle)}$, and in transmission mode, $\mathbb{P}_{BS}^{(tx)}$, can be formulated as follows [20]:

$$\begin{aligned} \mathbb{P}_{BS}^{(idle)} &= \Pr\{N_{MT} = 0\} = 1 - L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right) \\ \mathbb{P}_{BS}^{(tx)} &= \Pr\{N_{MT} \geqslant 1\} = 1 - \mathbb{P}_{BS}^{(idle)} = L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right) \end{aligned} \tag{2.8}$$

where we define $L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right) = 1 - \left(1 + \frac{\lambda_{MT}}{3.5\lambda_{BS}}\right)^{-3.5}$.

2.2.6 Problem Formulation

The performance metric of interest in the present chapter is coverage probability, denoted by P_{cov}. On the other hand, it is equivalent to evaluate the performance over Potential Spectral Efficiency (PSE) by adding a scaling factor [34]. Thus, we only focus on the analysis of coverage probability based on the following two definitions.

Definition 2.1 *The coverage probability is defined as the probability that a typical MT is able to achieve the Signal-to-Interference+Noise-Ratio (SINR), denoted by γ_{th}. State differently, the probability of coverage is the complementary cumulative distribution function (CCDF) of the SINR over the network, which is given by*

$$P_{cov} = \Pr\left\{SINR \geqslant \gamma_{th}\right\} \tag{2.9}$$

where γ_{th} is the mimimum SINR threshold for successful decoding. In particular, SINR can be formulated as follows:

$$SINR = \frac{P_{tx}\left|g^{(0)}\right|^2 \Big/ L^{(0)}}{\sigma_N^2 + P_{tx} I_{agg}\left(L^{(0)}\right)} \tag{2.10}$$

where σ_N^2 *is the noise power,* $I_{agg}\left(L^{(0)}\right) = \sum_{BS^{(i)} \in \Psi_{BS}^{(l)}} \frac{g^{(i)}}{L^{(i)}} 1\left(L^{(i)} > L^{(0)}\right)$ *is the aggregate interference.*

Definition 2.2 *Beyond the threshold for successfully decoding the infortion data,* γ_{th}*, we denote* γ_A *as the threshold for successfully detecting the serving BS,* $BS^{(0)}$*. The coverage probability is defined by [34]:*

$$P_{cov}\left(\gamma_{th}, \gamma_{\mathrm{A}}\right) = \begin{cases} \Pr\left\{\mathrm{SIR} \geq \gamma_{th}, \overline{\mathrm{SNR}} \geq \gamma_{\mathrm{A}}\right\} & \text{if} \quad MT^{(0)} \text{ is selected} \\ 0 & \text{if} \quad MT^{(0)} \text{ is not selected} \end{cases} \tag{2.11}$$

*where the Signal-to-Interference-Ratio (*SIR*) and the average Signal-to-Noise-Ratio (*$\overline{\mathrm{SNR}}$*) can be formulated, for the network model under analysis, as follows:*

$$SIR = \frac{P_{tx}\left|g^{(0)}\right|^2 \Big/ L^{(0)}}{P_{tx} I_{agg}\left(L^{(0)}\right)} \qquad \overline{\mathrm{SNR}} = \frac{\mathrm{P_{tx}}/L^{(0)}}{\sigma_{\mathrm{N}}^2} \tag{2.12}$$

Note that in the remaining part, the general mathematical frameworks of coverage probability is proposed and analyzed by using the conventional definition in (2.9). The application of the new definition in (2.11) will be generalized in Section 2.5. We release the notation $\mathrm{P_{cov}^{G}} \rightarrow \mathrm{P_{cov}}\left(\gamma_{\mathrm{th}}, \gamma_{\mathrm{A}}\right)$ to denote the generalized expression of coverage probability.

2.3 Mathematical Framework of Coverage Probability

2.3.1 Exact Mathematical Formulation of Coverage Probability

In this section, we develop the general mathematical framework of coverage probability based on the above reported system model. The obtained tractable expression, predominantly, relies on the fundamental theorems of stochastic geometry and PPP in [33]. The mathematical framework of coverage probability, as a consequence, is fomulated as follows.

$$\mathrm{P_{cov}} = \Pr\left\{\mathrm{SINR} \geqslant \gamma_{\mathrm{th}}\right\} = \int_0^{\infty} \exp\left(-\frac{\gamma_{\mathrm{th}}\sigma_{\mathrm{N}}^2}{\Omega \mathrm{P_{tx}}} x\right) M_{I_{\mathrm{agg}}}\left(\frac{\gamma_{\mathrm{th}} x}{\Omega} | x\right) f_{L^{(0)}}(x)\, dx \tag{2.13}$$

where $f_{L^{(0)}}(x) = \Lambda^{(1)}([0,x)) \exp\left(-\Lambda\left([0,x)\right)\right)$ denotes the PDF of minimum path-loss with the intensity measure defined by $\Lambda\left([0,x)\right) = \sum_{s \in \{\mathrm{LOS,NLOS}\}} \Lambda_s\left([0,x)\right)$ and its derivative defined by $\Lambda^{(1)}\left([0,x)\right) = \sum_{s \in \{\mathrm{LOS,NLOS}\}} \Lambda_s^{(1)}\left([0,x)\right)$. $M_{I_{\mathrm{agg}}}\left(p|x\right) = \prod_{s \in \{\mathrm{LOS,NLOS}\}} M_{I_{\mathrm{agg}}}^{s}\left(p|x\right)$ denotes the MGF of aggregagte interference with $p = \frac{\gamma_{\mathrm{th}} x}{\Omega}$. For an arbitrary link state $s \in$ {LOS,NLOS}, the subfunctions are defined in Table 2.2.

The framework of the coverage probability in (2.13) is exact under the working assumptions and is easy to compute since just one-fold integral is present. However, the system behavior is

Table 2.2 – Auxiliary functions used in (2.13).

Function definition
$\Lambda_s([0,x)) = \pi\lambda q_s^{[0,D)}\left(\left(\frac{x}{k_s}\right)^{\frac{2}{\beta_s}} - h^2\right)\bar{H}\left(x - k_s\left(D_h^2+h^2\right)^{\frac{\beta_s}{2}}\right)H\left(x-k_s h^{\beta_s}\right)$
$+\left\{\pi\lambda q_s^{[0,D)} D_h^2 + \pi\lambda q_s^{[D,\infty)}\left(\left(\frac{x}{k_s}\right)^{\frac{2}{\beta_s}} - h^2 - D_h^2\right)\right\}H\left(x - k_s\left(D_h^2+h^2\right)^{\frac{\beta_s}{2}}\right)H\left(x-k_s h^{\beta_s}\right)$
$\Lambda_s^{(1)}([0,x)) = \frac{2\pi\lambda}{\beta_s} q_s^{[0,D)}\left(\frac{1}{k_s}\right)^{\frac{2}{\beta_s}} x^{\frac{2}{\beta_s}-1}\bar{H}\left(x - k_s\left(D_h^2+h^2\right)^{\frac{\beta_s}{2}}\right)H\left(x-k_s h^{\beta_s}\right)$
$+\frac{2\pi\lambda}{\beta_s} q_s^{[D,\infty)}\left(\frac{1}{k_s}\right)^{\frac{2}{\beta_s}} x^{\frac{2}{\beta_s}-1} H\left(x - k_s\left(D_h^2+h^2\right)^{\frac{\beta_s}{2}}\right)H\left(x-k_s h^{\beta_s}\right)$
$M_{I_{\text{agg}}}^{s}\left(p\,\middle|\,x\right)$
$= \exp\left\{\begin{array}{l} \pi\lambda_{\text{BS}}\mathbb{P}_{\text{BS}}^{(\text{tx})} q_s^{[0,D)} h^2 \Theta_s\left(\frac{k_s h^{\beta_s}}{p\Omega_s}\right)\bar{H}\left(x-k_s h^{\beta_s}\right) \\ -\pi\lambda_{\text{BS}}\mathbb{P}_{\text{BS}}^{(\text{tx})} q_s^{[0,D)}\left(D_h^2+h^2\right)\Theta_s\left(\frac{k_s\left(D_h^2+h^2\right)^{\frac{\beta_s}{2}}}{p\Omega_s}\right)\bar{H}\left(x-k_s\left(D_h^2+h^2\right)^{\frac{\beta_s}{2}}\right) \\ +\pi\lambda_{\text{BS}}\mathbb{P}_{\text{BS}}^{(\text{tx})} q_s^{[0,D)}\left(\frac{x}{k_s}\right)^{\frac{2}{\beta_s}}\Theta_s\left(\frac{x}{p\Omega_s}\right)H\left(x-k_s h^{\beta_s}\right)\bar{H}\left(x-k_s\left(D_h^2+h^2\right)^{\frac{\beta_s}{2}}\right) \\ +\pi\lambda_{\text{BS}}\mathbb{P}_{\text{BS}}^{(\text{tx})} q_s^{[D,\infty)}\left(D_h^2+h^2\right)\left[1-\Gamma_s\left(\frac{p\Omega_s}{k_s\left(D_h^2+h^2\right)^{\frac{\beta_s}{2}}}\right)\right]\bar{H}\left(x-k_s\left(D_h^2+h^2\right)^{\frac{\beta_s}{2}}\right) \\ +\pi\lambda_{\text{BS}}\mathbb{P}_{\text{BS}}^{(\text{tx})} q_s^{[D,\infty)}\left(\frac{x}{k_s}\right)^{\frac{2}{\beta_s}}\left[1-\Gamma_s\left(\frac{p\Omega_s}{x}\right)\right]H\left(x-k_s\left(D_h^2+h^2\right)^{\frac{\beta_s}{2}}\right) \end{array}\right\}$
$\Theta_s(z) = {}_2F_1\left(1, \frac{2}{\beta_s}, 1+\frac{2}{\beta_s}, -z\right)$
$\Gamma_s(z) = {}_2F_1\left(-\frac{2}{\beta_s}, 1, 1-\frac{2}{\beta_s}, -z\right)$

not easy to perceive. Here we consider a simplified model, the pure LOS ball model, which allows us to further simplify the frameworks but still retaining the same properties of the original one. Thus, the following assumptions are adopted:

- Assumption I: $\sigma_{\text{N}}^2 = 0$ (Interferenc-limited).
- Assumption II: $q_{\text{LOS}}^{[0,D)} = 1, q_{\text{LOS}}^{[D,\infty)} = 0$ and $q_{\text{NLOS}}^{[0,D)} = 0, q_{\text{NLOS}}^{[D,\infty)} = 1$ (Pure LOS ball model).

Proposition 2.1 *With the above key assumptions, the coverage probability of SINR reduces to that of SIR, and the framework can be simplified under the pure LOS ball model. In general, the resulting framework is given by:*

$$P_{cov} = \int_0^{\infty} M_{I_{agg}}\left(\frac{\gamma_{th} x}{\Omega}\,\middle|\,x\right) f_{L^{(0)}}(x)\,dx = P_{cov}^{1} + P_{cov}^{2} \tag{2.14}$$

where P^1_{cov} and P^2_{cov} are formulated as follows.

$$P^1_{cov} = \int_{k_{LOS}h^{\beta_{LOS}}}^{k_{LOS}\left(D_h^2+h^2\right)^{\frac{\beta_{LOS}}{2}}} \exp\left\{\begin{array}{l} -\pi\lambda_{BS}L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\left(D_h^2+h^2\right)\Theta_{LOS}\left(\frac{k_{LOS}\left(D_h^2+h^2\right)^{\frac{\beta_{LOS}}{2}}}{\gamma_{th}x}\right) \\ +\pi\lambda_{BS}L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\left(\frac{x}{k_{LOS}}\right)^{\frac{2}{\beta_{LOS}}}\Theta_{LOS}\left(\frac{1}{\gamma_{th}}\right) \\ +\pi\lambda_{BS}L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\left(D_h^2+h^2\right)\left[1-\Gamma_{NLOS}\left(\frac{\gamma_{th}x}{k_{NLOS}\left(D_h^2+h^2\right)^{\frac{\beta_{NLOS}}{2}}}\right)\right] \end{array}\right\}$$
$$\times\left\{\frac{2\pi\lambda_{BS}}{\beta_{LOS}}\left(\frac{1}{k_{LOS}}\right)^{\frac{2}{\beta_{LOS}}}x^{\frac{2}{\beta_{LOS}}-1}\right\}\exp\left\{-\pi\lambda_{BS}\left(\left(\frac{x}{k_{LOS}}\right)^{\frac{2}{\beta_{LOS}}}-h^2\right)\right\}dx \tag{2.15}$$

and

$$P^2_{cov} = \frac{\exp\left\{\pi\lambda_{BS}h^2-\pi\lambda_{BS}\left(D_h^2+h^2\right)\left(1-L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\left[1-\Gamma_{NLOS}\left(\gamma_{th}\right)\right]\right)\right\}}{1-L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\left[1-\Gamma_{NLOS}\left(\gamma_{th}\right)\right]} \tag{2.16}$$

Proof*: Equations* (2.14)-(2.16) *are provided by directly applying the above assumptions to* (2.13)*. Given the analytical expression of the PDF of the smallest path-loss, the integral of the coverage can be split in two parts: $\left[k_{LOS}h^{\beta_{LOS}}, k_{LOS}\left(D_h^2+h^2\right)^{\frac{\beta_{LOS}}{2}}\right]$ and $\left[k_{NLOS}\left(D_h^2+h^2\right)^{\frac{\beta_{NLOS}}{2}},\infty\right)$. This is because we assumed, this is always true, that $k_{NLOS}\left(D_h^2+h^2\right)^{\frac{\beta_{NLOS}}{2}} \geqslant k_{LOS}\left(D_h^2+h^2\right)^{\frac{\beta_{LOS}}{2}}$. The integral in P^2_{cov} is solved based on: $\int_M^\infty x^{\frac{n}{\beta}-1}\exp\left\{-x^{\frac{n}{\beta}}K\right\}dx = \frac{\beta}{nK}\exp\left\{-M^{\frac{n}{\beta}}K\right\}$.* □

Remark 2.3 *The first integral P^1_{cov}, on the other hand, cannot be computed in closed-form but it can re-written by using the following change of variables: $y = \frac{x}{k_{LOS}\left(D_h^2+h^2\right)^{\frac{\beta_{LOS}}{2}}}$, which will lead to the following expression:*

$$P^1_{cov} = \frac{2\pi\lambda_{BS}}{\beta_{LOS}}\left(D_h^2+h^2\right)\exp\left\{\pi\lambda_{BS}h^2\right\}\int_{\left(1+\frac{D_h^2}{h^2}\right)^{-\frac{\beta_{LOS}}{2}}}^1 y^{\frac{2}{\beta_{LOS}}-1}$$
$$\times\exp\left\{-\pi\lambda_{BS}\left(D_h^2+h^2\right)\left[\begin{array}{l} y^{\frac{2}{\beta_{LOS}}}\left[1-L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\Theta_{LOS}\left(\frac{1}{\gamma_{th}}\right)\right]+L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\Theta_{LOS}\left(\frac{1}{\gamma_{th}y}\right) \\ +L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\left[\Gamma_{NLOS}\left(\gamma_{th}\frac{k_{LOS}\left(D_h^2+h^2\right)^{\frac{\beta_{LOS}}{2}}}{k_{NLOS}\left(D_h^2+h^2\right)^{\frac{\beta_{NLOS}}{2}}}y\right)-1\right] \end{array}\right]\right\}dy \tag{2.17}$$

It is verified once gain that the integral in $\mathrm{P}^1_{\mathrm{cov}}$ cannot be computed into a closed-form even though we changed variables with the above approach. Note that we specially refer the

simplified expressions in Proposition 2.1 as the exact framework of coverage probability.

2.3.2 Asymptotic Case Studies (Exact Framework)

In this section, based on the obtained tractable mathematical framework of coverage probability, we consider several asymptotic case studies under the impact of key system parameters. To gain the most of the insight for performance evaluation, the framework in Proposition 2.1 constitutes the departing point of the investigation.

Case Study 1 The LOS ball is small ($D \to 0$) and the height of BSs and buildings are non-zero ($h > 0$ and $\tilde{h} > 0$).

In this case, the following results hold for $\mathrm{P}_{\mathrm{cov}} = \mathrm{P}^1_{\mathrm{cov}} + \mathrm{P}^2_{\mathrm{cov}}$, where

$$\mathrm{P}^1_{\mathrm{cov}} \to 0,\ \mathrm{P}^2_{\mathrm{cov}} \to \frac{\exp\left\{\pi\lambda_{\mathrm{BS}} h^2 L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[1-\Gamma_{\mathrm{NLOS}}\left(\gamma_{\mathrm{th}}\right)\right]\right\}}{1-L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[1-\Gamma_{\mathrm{NLOS}}\left(\gamma_{\mathrm{th}}\right)\right]} \tag{2.18}$$

Proof: The results are obtained by directly replacing $D \to 0 \Rightarrow D_h^2 \to 0$, which specifically implies that the upper limit of the integral in (2.15) becomes the same as the lower limit, i.e., $k_{\mathrm{LOS}} h^{\beta_{\mathrm{LOS}}}$, thus the integral is solved to be $\mathrm{P}^1_{\mathrm{cov}} \to 0$. The second results for $\mathrm{P}^2_{\mathrm{cov}}$ is achieved immediately. □

Remark 2.4 *In this case study, the following comments hold true:*

- *As expected, coverage probability depends only on the NLOS links.*
- *It coincides with the findings in [28], which states that the coverage monotonically decreases with the BS's height. In fact,* $1-\Gamma_{NLOS}\left(\gamma_{th}\right) = 1 - {}_2F_1\left(-\frac{2}{\beta_{NLOS}}, 1, 1-\frac{2}{\beta_{NLOS}}, -\gamma_{th}\right) \leqslant 0$.
- *The expression above is correct for the 2D case when* $h = 0$, *and in this case it simplifies to:* $P^2_{cov} = \frac{1}{1-L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\left[1-\Gamma_{NLOS}(\gamma_{th})\right]} = \frac{1}{1-L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\left[1-{}_2F_1\left(-\frac{2}{\beta_{NLOS}},1,1-\frac{2}{\beta_{NLOS}},-\gamma_{th}\right)\right]}$, *which is consistent with the original paper from [17], since in this case we have a 2D system model where all the links are in NLOS and we obtain the usual* ${}_2F_1(\cdot)$ *expression of the coverage probability.*
- *This results highlights that for small values of the LOS ball, the term* P^2_{cov} *dominates the overall coverage with respect to* P^1_{cov}.

Case Study 2 The LOS ball is large ($D \to \infty$) and the height of BSs and buildings are non-zero ($h > 0$ and $\tilde{h} > 0$).

In this case, the following results hold for $\mathrm{P}_{\mathrm{cov}} = \mathrm{P}^1_{\mathrm{cov}} + \mathrm{P}^2_{\mathrm{cov}}$, where

$$\mathrm{P}^1_{\mathrm{cov}} \to \frac{\exp\left\{\pi\lambda_{\mathrm{BS}}h^2 L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[1-\Gamma_{\mathrm{LOS}}\left(\gamma_{\mathrm{th}}\right)\right]\right\}}{1-L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[1-\Gamma_{\mathrm{LOS}}\left(\gamma_{\mathrm{th}}\right)\right]}, \mathrm{P}^2_{\mathrm{cov}} \to 0 \tag{2.19}$$

Proof: See Appendix 2.8.1. □

Remark 2.5 *From this result, we conclude that the result is consistent in this limiting regime and that the same comments as in the Case Study 1 can be made. If the LOS ball is small, in fact, P^2_{cov} dominates the performance. For typical values of the LOS ball, however, it is realized that P^1_{cov} is the dominant term. It is useful to identify some conditions on the LOS ball that makes P^2_{cov} negligible compared with P^1_{cov}.*

Case Study 3 $k_{\mathrm{LOS}} = k_{\mathrm{NLOS}} = k$ and $\beta_{\mathrm{LOS}} = \beta_{\mathrm{NLOS}} = \beta$, and the LOS ball, and the height of BSs and buildings are non-zero ($D > 0$, $h > 0$ and $\tilde{h} > 0$).

In this case, the following results hold:

$$\mathrm{P}_{\mathrm{cov}} = \mathrm{P}^1_{\mathrm{cov}} + \mathrm{P}^2_{\mathrm{cov}} \to \frac{\exp\left\{\pi\lambda_{\mathrm{BS}}h^2 L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left(1-\Gamma\left(\gamma_{\mathrm{th}}\right)\right)\right\}}{1-L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left(1-\Gamma\left(\gamma_{\mathrm{th}}\right)\right)} \tag{2.20}$$

Proof: The proof follows from substituting $k_{\mathrm{LOS}} = k_{\mathrm{NLOS}} = k$ and $\beta_{\mathrm{LOS}} = \beta_{\mathrm{NLOS}} = \beta$ into (2.16) and (2.17) and computing the integral by invoking (2.51). □

Remark 2.6 *This results is coincident with Case Study 1 and Case Study 2 since the path-loss model is the same, which enable the system model and LOS ball blockage reduce to the normal scenario when all the links are either in LOS or NLOS. Thus, the same comments as for the first two case studies apply in this case as well. This case study, in particular, was studied in [28].*

Case Study 4 The height of the building is very small ($\tilde{h} \to 0$) and the height of the BSs is non-zero ($h > 0$).

In this case, the following results hold for $\mathrm{P}_{\mathrm{cov}} = \mathrm{P}^1_{\mathrm{cov}} + \mathrm{P}^2_{\mathrm{cov}}$, where

$$\mathrm{P}^1_{\mathrm{cov}} \to \frac{\exp\left\{\pi\lambda_{\mathrm{BS}}h^2 L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[1-\Gamma_{\mathrm{LOS}}\left(\gamma_{\mathrm{th}}\right)\right]\right\}}{1-L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[1-\Gamma_{\mathrm{LOS}}\left(\gamma_{\mathrm{th}}\right)\right]}, \mathrm{P}^2_{\mathrm{cov}} \to 0 \tag{2.21}$$

Proof: The proof follows from $\tilde{h} \to 0 \Rightarrow D_h^2 \to \infty$, which is the same as Case Study 2. □

This is expected because all the links will be in LOS in this case study. The same conclusions as in Case Study 2 can be drawn.

Case Study 5 The height of the building is very large ($\tilde{h} \to \infty$) and the height of the BSs is non-zero ($h > 0$) but finite.

In this case, the following results hold for $\mathrm{P}_{\mathrm{cov}} = \mathrm{P}^1_{\mathrm{cov}} + \mathrm{P}^2_{\mathrm{cov}}$, where

$$\mathrm{P}^1_{\mathrm{cov}} \to \frac{2\pi\lambda_{\mathrm{BS}}}{\beta_{\mathrm{LOS}}}\left(D^2+h^2\right)\exp\left\{\pi\lambda_{\mathrm{BS}}h^2\right\}\int_{\left(1+\frac{D^2}{h^2}\right)^{-\frac{\beta_{\mathrm{LOS}}}{2}}}^{1} y^{\frac{2}{\beta_{\mathrm{LOS}}}-1}$$
$$\times\exp\left\{-\pi\lambda_{\mathrm{BS}}\left(D^2+h^2\right)\left[\begin{array}{l} y^{\frac{2}{\beta_{\mathrm{LOS}}}}\left[1-L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)\right]+L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}y}\right) \\ +L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[\Gamma_{\mathrm{NLOS}}\left(\gamma_{\mathrm{th}}\frac{k_{\mathrm{LOS}}\left(D^2+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}}{k_{\mathrm{NLOS}}\left(D^2+h^2\right)^{\frac{\beta_{\mathrm{NLOS}}}{2}}}y\right)-1\right]\end{array}\right]\right\}dy \quad (2.22)$$
$$\mathrm{P}^2_{\mathrm{cov}} \to \frac{\exp\left\{\pi\lambda_{\mathrm{BS}}h^2-\pi\lambda_{\mathrm{BS}}\left(D^2+h^2\right)\left(1-L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[1-\Gamma_{\mathrm{NLOS}}\left(\gamma_{\mathrm{th}}\right)\right]\right)\right\}}{1-L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[1-\Gamma_{\mathrm{NLOS}}\left(\gamma_{\mathrm{th}}\right)\right]}$$

Proof: The proof follows from $\tilde{h} \to \infty \Rightarrow D_h^2 \to D^2$. □

Remark 2.7 *This integral cannot be computed in general, which motivates us to propose the closed-form expression of coverage probability for analysis in the following sections. This is, however, not an intuitive result and needs to be proved. If the height of the buildings is higher than the height of the BSs, then it means that the LOS/NLOS conditions of the links are determined only by presence or not of buildings in between BSs and MTs. The height of the BSs make the intended link worse, but it also reduce the amount of interference. Hence, this is a scenario to study with deep attention.*

Case Study 6 The height of the BSs is very small ($h \to 0$) and the height of the buildings is non-zero ($\tilde{h} > 0$) or even zero ($\tilde{h} = 0$).

In this case, the following results hold for $\mathrm{P}_{\mathrm{cov}} = \mathrm{P}^1_{\mathrm{cov}} + \mathrm{P}^2_{\mathrm{cov}}$, where

$$\begin{aligned}
\mathrm{P}^1_{\mathrm{cov}} &\to \frac{2\pi\lambda_{\mathrm{BS}}}{\beta_{\mathrm{LOS}}} D^2 \\
&\times \int_0^1 y^{\frac{2}{\beta_{\mathrm{LOS}}}-1} \exp\left\{ -\pi\lambda_{\mathrm{BS}} D^2 \left[\begin{array}{l} y^{\frac{2}{\beta_{\mathrm{LOS}}}} \left[1 - L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) \Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right) \right] + L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) \Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}} y}\right) \\ + L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) \left[\Gamma_{\mathrm{NLOS}}\left(\gamma_{\mathrm{th}} \frac{k_{\mathrm{LOS}} D^{\beta_{\mathrm{LOS}}}}{k_{\mathrm{NLOS}} D^{\beta_{\mathrm{NLOS}}}} y \right) - 1 \right] \end{array} \right] \right\} dy \\
\mathrm{P}^2_{\mathrm{cov}} &\to \frac{\exp\left\{ -\pi\lambda_{\mathrm{BS}} D^2 \left(1 - L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) \left[1 - \Gamma_{\mathrm{NLOS}}\left(\gamma_{\mathrm{th}}\right) \right] \right) \right\}}{1 - L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) \left[1 - \Gamma_{\mathrm{NLOS}}\left(\gamma_{\mathrm{th}}\right) \right]}
\end{aligned} \tag{2.23}$$

Proof: The proof follows from $h \to 0 \Rightarrow D_h^2 \to D^2$. □

This integral cannot be computed in general. In addition, this case study is equivalent to the 2D Poisson cellular network model, which has been investigated extensively in [25].

Case Study 7 The height of the BSs is very large ($h \to \infty$) and the height of the buildings is non-zero ($\tilde{h} > 0$) or even zero ($\tilde{h} = 0$) but finite.

In this case, the following results hold:

$$\mathrm{P}_{\mathrm{cov}} = \mathrm{P}^1_{\mathrm{cov}} + \mathrm{P}^2_{\mathrm{cov}} \to 0 \tag{2.24}$$

Proof: The proof follows from $h \to \infty \Rightarrow D_h^2 \to \infty$, which implies that both the upper limit and lower limit of $\mathrm{P}^1_{\mathrm{cov}}$ and $\mathrm{P}^2_{\mathrm{cov}}$ become infinity. Thus, the integrals are computed to be 0. □

Note that other asymptotic case studies of interest (notably as a function of the density of the BSs) cannot be handled by the exact framework with integrals in Proposition 2.1. In order to study the scenarios that cannot be clarified by the case studies described above, we introduce a bound, which is then proposed for the practical case studies where the heights of the BSs and buildings are non-zero and finite, as well as the LOS ball is non-zero and finite. All the other case studies can be retrieved either from other papers or from the case studies above.

2.3.3 Closed-Form Upper Bound of Coverage Probability

Despite the tractability of the exact framework $\mathrm{P}_{\mathrm{cov}}$ with one-fold integral in Proposition 2.1, the significant limitations can be discovered in scrutinizing the asymptotic case studies under the influence of density of BSs. As a consequence, in order to overcome this issue, by approximating $\mathrm{P}^1_{\mathrm{cov}}$ in (2.17), we propose a closed-form upper bound of coverage probability based on the following procedures.

Step 1 The term $L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[\Gamma_{\mathrm{NLOS}}\left(\gamma_{\mathrm{th}}\frac{k_{\mathrm{LOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}}{k_{\mathrm{NLOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{NLOS}}}{2}}}y\right)-1\right]$ can be ignored as it is negligible for typical system parameters.

This approximation holds for typical system parameters since i) $\frac{k_{\mathrm{LOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}}{k_{\mathrm{NLOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{NLOS}}}{2}}} \ll 1$; ii) $\Gamma(z) = {}_2F_1\left(-\frac{2}{\beta},1,1-\frac{2}{\beta},-z\right) \to 1$ if $z \to 0$. We observe that the first condition is sufficient to ensure that $\frac{k_{\mathrm{LOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}}{k_{\mathrm{NLOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{NLOS}}}{2}}}y \ll 1$ since $y \leqslant 1$. In addition, the decoding threshold is not too large in general. By ignoring this term, the integral simplifies as follows:

$$\begin{aligned}\mathrm{P}_{\mathrm{cov}}^{1} &\leqslant \frac{2\pi\lambda_{\mathrm{BS}}}{\beta_{\mathrm{LOS}}}\left(D_h^2+h^2\right)\exp\left\{\pi\lambda_{\mathrm{BS}}h^2\right\} \\ &\times\int_{\left(1+\frac{D_h^2}{h^2}\right)^{-\frac{\beta_{\mathrm{LOS}}}{2}}}^{1} y^{\frac{2}{\beta_{\mathrm{LOS}}}-1}\exp\left\{-\pi\lambda_{\mathrm{BS}}\left(D_h^2+h^2\right)\begin{bmatrix} y^{\frac{2}{\beta_{\mathrm{LOS}}}}\left[1-L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)\right] \\ +L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}y}\right)\end{bmatrix}\right\}dy\end{aligned} \tag{2.25}$$

which can be easily proved to be an upper-bound.

Step 2 Polynomial approximation $\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}y}\right) = {}_2F_1\left(1,\frac{2}{\beta_{\mathrm{LOS}}},1+\frac{2}{\beta_{\mathrm{LOS}}},-\frac{1}{\gamma_{\mathrm{th}}y}\right) \approx K_1 + K_2 y^{\frac{2}{\beta_{\mathrm{LOS}}}}$.

Remark 2.8 *The proposed approximation is a lower-bound, i.e., $K_1+K_2y^{\frac{2}{\beta_{LOS}}} \leqslant f(y)$ for $0 \leqslant y \leqslant 1$, where K_1 and K_2 are constant values. This approximation yields that the resulting expression of coverage probability consititutes the upper bound of the exact one, which is given by*

$$\begin{aligned}&\int_{\left(1+\frac{D_h^2}{h^2}\right)^{-\frac{\beta_{LOS}}{2}}}^{1} y^{\frac{2}{\beta_{LOS}}-1}\exp\left\{-\pi\lambda_{BS}\left(D_h^2+h^2\right)\begin{Bmatrix} y^{\frac{2}{\beta_{LOS}}}\left[1-L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\Theta_{LOS}\left(\frac{1}{\gamma_{th}}\right)\right] \\ +L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\left(K_1+K_2y^{\frac{2}{\beta_{LOS}}}\right)\end{Bmatrix}\right\}dy \\ &\geqslant\int_{\left(1+\frac{D_h^2}{h^2}\right)^{-\frac{\beta_{LOS}}{2}}}^{1} y^{\frac{2}{\beta_{LOS}}-1}\exp\left\{-\pi\lambda_{BS}\left(D_h^2+h^2\right)\begin{Bmatrix} y^{\frac{2}{\beta_{LOS}}}\left[1-L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\Theta_{LOS}\left(\frac{1}{\gamma_{th}}\right)\right] \\ +L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\Theta_{LOS}\left(\frac{1}{\gamma_{th}y}\right)\end{Bmatrix}\right\}dy\end{aligned} \tag{2.26}$$

where all the sub-functions are defined in Table 2.3.

Proof: *The proof of $K_1 + K_2 y^{\frac{2}{\beta_{LOS}}} \leqslant f(y)$ directly follows from that by substituting y_1, y_2, f_1, f_2 inside K_1 and K_2 defined in Table 2.3 and by invoking the property of the Hypergeometric function $\Theta_{LOS}\left(\frac{1}{\gamma_{th}y}\right)$ that is increasing in y and takes values between in the range $[0,1]$.* □

Remark 2.9 *The upper-bound is applicable for all case studies where $K_1 \neq 0$ and $K_2 \neq 0$. This is true for all system setups, except the 2D cellular network model where the lower-limit of the*

Table 2.3 – Auxiliary functions used in Remark 2.8.

Function definition
$f(y) = \Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}} y}\right) = {}_2F_1\left(1, \frac{2}{\beta_{\text{LOS}}}, 1+\frac{2}{\beta_{\text{LOS}}}, -\frac{1}{\gamma_{\text{th}} y}\right)$
$y_1 = \left(\frac{D_h^2}{h^2}+1\right)^{-\frac{\beta_{\text{LOS}}}{2}}$, $y_2 = 1$
$f_1 = f(y_1) = \Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}} y_1}\right)$, $f_2 = f(y_2) = \Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}} y_2}\right)$
$K_1 = \frac{y_2^{\frac{2}{\beta_{\text{LOS}}}} f_1 - y_1^{\frac{2}{\beta_{\text{LOS}}}} f_2}{y_2^{\frac{2}{\beta_{\text{LOS}}}} - y_1^{\frac{2}{\beta_{\text{LOS}}}}}$, $K_2 = \frac{f_2 - f_1}{y_2^{\frac{2}{\beta_{\text{LOS}}}} - y_1^{\frac{2}{\beta_{\text{LOS}}}}}$

integral is equal to zero (see (2.23)*), which would yield* $K_1 = 0$*. The 2D cellular network model is, however, well-studied and can be found in [25].*

Proposition 2.2 *With the approximation adopted above, the integral inside the upper bound of coverage probability can be solved into a closed-form, which is given by:*

$$P_{cov}^{UB} = \frac{\exp\{\pi\lambda_{BS}h^2\}\exp\left(-\pi\lambda_{BS}\left(D_h^2+h^2\right)\Upsilon_1\right)}{\Delta} - \frac{\exp\{\pi\lambda_{BS}h^2\}\exp\left(-\pi\lambda_{BS}\left(D_h^2+h^2\right)\Upsilon_2\right)}{\Delta} + \frac{\exp\{\pi\lambda_{BS}h^2\}\exp\left(-\pi\lambda_{BS}\left(D_h^2+h^2\right)\Phi\right)}{\Phi} \tag{2.27}$$

where we define the following short-hands: $\Delta = K_2 L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right) + 1 - L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\Theta_{LOS}\left(\frac{1}{\gamma_{th}}\right)$, $\Phi = 1 - L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\left[1-\Gamma_{NLOS}(\gamma_{th})\right]$, $\Upsilon_1 = K_1 L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right) + \Delta y_1^{\frac{2}{\beta_{LOS}}}$, $\Upsilon_2 = K_1 L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right) + \Delta y_2^{\frac{2}{\beta_{LOS}}}$ *where* y_1, y_2, K_1, K_2 *are defined in Table 2.3.*

Even though this upper-bound is formulated in a closed-form, it is still not sufficiently simple to get insight for system optimization. Thus, we introduce a further approximation in Step 3 that is used in some cases in order to compute the optimal parameters that maximize the coverage probability in closed-form. We note, in addition, that the approximation is used only for the case study $h > \tilde{h}$, since for the other case study we are able to show that the coverage decreases with the height of the BSs in Theorem 2.2.

Before introducing this approximation to get insight for system optimization, we introduce some important corollaries that are used in the sequel to prove some important performance trends.

Corollary 2.1 $0 \leqslant \Upsilon_1 \leqslant 1$, $\Upsilon_2 = 1$*, and the exact expression of* Υ_1 *is formulated as follows:*

$$\Upsilon_1 = \Theta_{LOS}\left(\frac{1}{\gamma_{th} D_1^{-\frac{\beta_{LOS}}{2}}}\right) L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right) + D_1^{-1}\left[1 - L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\Theta_{LOS}\left(\frac{1}{\gamma_{th}}\right)\right] \tag{2.28}$$

where $D_1 = \frac{D_h^2}{h^2} + 1 \geqslant 1$.

Proof: See Appendix 2.8.2. □

Corollary 2.2 $1 - D_1\Upsilon_1 \leqslant 0$, $1 - D_1\Upsilon_2 \leqslant 0$, *and* $1 - D_1\Phi \leqslant 0$, *where* $D_1 = \frac{D_h^2}{h^2} + 1 \geqslant 1$.

Proof: See Appendix 2.8.3. □

Corollary 2.3 *Given* $D_1 = \frac{D_h^2}{h^2} + 1 \geqslant 1$, $g(D_1)$ *is a negative decreasing function and the following asymptotic results hold true for the later analysis, which are given by*

$$g(D_1) = \Theta_{LOS}\left(\frac{1}{\gamma_{th}}\right) - D_1\Theta_{LOS}\left(\frac{1}{\gamma_{th}D_1^{-\frac{\beta_{LOS}}{2}}}\right) \leqslant 0 \tag{2.29}$$

and

$$\lim_{D_1 \to 1} g(D_1) = 0, \ \lim_{D_1 \to \infty} g(D_1) \leqslant 0 \tag{2.30}$$

Proof*: See Appendix 2.8.4.* □

Corollary 2.4 *Given* $D_1 = \frac{D_h^2}{h^2} + 1 \geqslant 1$, *it is determined that* $0 \leqslant \Delta \leqslant 1$, *and the exact expression for* Δ *is formulated by:*

$$\Delta = 1 + L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\frac{g(D_1)}{D_1 - 1} = 1 + L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\left[\frac{\Theta_{LOS}\left(\frac{1}{\gamma_{th}}\right) - D_1\Theta_{LOS}\left(\frac{1}{\gamma_{th}D_1^{-\frac{\beta_{LOS}}{2}}}\right)}{D_1 - 1}\right] \tag{2.31}$$

Proof*: See Appendix 2.8.5.* □

Step 3 Based on the sufficient condition stated in Proposition 2.3, we propose an approximated upper bound of coverage probability which can be utilized in system-level optimization for the case study $h > \tilde{h}$.

Proposition 2.3 *Under the following condition:*

$$\pi\lambda_{BS}h^2 D_1 L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\left[\Gamma_{NLOS}(\gamma_{th}) - 1\right] \geqslant \ln\left(\frac{1}{\varepsilon}\right) \tag{2.32}$$

with $D_1 = \frac{D^2}{\tilde{h}^2} + 1$ *and* $\varepsilon = 0.1$*, the framework of coverage probability can be approximated by*

$$P_{cov} \approx \overline{P}_{cov}^{UB} = \frac{\exp\{\pi\lambda_{BS}h^2\}\exp\left(-\pi\lambda_{BS}\left(D_h^2+h^2\right)\Upsilon_1\right)}{\Delta} - \frac{\exp\{\pi\lambda_{BS}h^2\}\exp\left(-\pi\lambda_{BS}\left(D_h^2+h^2\right)\Upsilon_2\right)}{\Delta} \tag{2.33}$$

Proof*: See Appendix 2.8.6.* □

We highlight that the approximation is required only for the case study $h > \tilde{h}$. The reason is that for the other case study, we can prove that the coverage is monotonically decreasing with the height of the BSs in Theorem 2.2 and this approximation is not needed. Also, the case study $h > \tilde{h}$ is the case of interest.

Before analyzing the performance trends and insight for system design that can be gained from the proposed upper-bound of the coverage probability, we study the accuracy of the proposed bound in some important asymptotic regimes as a function of the height of the BSs and blockages as well as the density of the BSs.

2.3.4 Asymptotic Case Studies (Upper Bound)

The proposed bound allows us to compute the coverage probability in some relevant asymptotic regimes that is not possible to study, in general, with the exact expression of the coverage because of the impossibility of computing the integral.

More precisely, we focus our attention only on the analysis of the density of the BSs, since the impact of the antenna height was discussed in Section 2.3.2, where the two setups are adopted: $h \to \infty$ when $h \geqslant \tilde{h}$; $h \to 0$ when $h < \tilde{h}$. In particular, that the case study $h < \tilde{h}$ is not of much interest since the coverage monotonically decreases with the height of the BSs and thus it is not an interesting scenario to study. This is proved in Theorem 2.2.

Note that:

- It is not necessary to use the approximation in Step 3 to analyze the impact of BS densification.
- It is not necessary to investigate the case study with a very low density of BSs, since the candidate system model is noise-free and, thus, this case study does not make much sense.
- The extreme setups that are analyzed in Section 2.3.2 discussed above are out of scope, especially the case where the heights of the BSs and of the buildings are zero and the LOS ball is zero, since these are special cases already discussed.

Case Study 1 If $\lambda_{\mathrm{BS}} \to \infty$ and $L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) < 1$, then the coverage probability is as follows:

$$\mathrm{P}_{\mathrm{cov}}^{\mathrm{UB}} \to \exp\left\{\pi h^2 \lambda_{\mathrm{MT}} g\left(D_1\right)\right\} = \exp\left\{\pi h^2 \lambda_{\mathrm{MT}} \left[\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right) - D_1 \Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}} D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right)\right]\right\} \tag{2.34}$$

Proof: See Appendix 2.8.7. □

Remark 2.10 *It is known that* $D_1 = \frac{D_h^2}{h^2} + 1$, *which can be divided into two cases (*$h \neq 0$ *and* $\tilde{h} \neq 0$*):* $D_1 = \frac{D_h^2}{h^2} + 1 = \frac{D^2}{\tilde{h}^2} + 1$ *if* $h \geqslant \tilde{h}$; $D_1 = \frac{D_h^2}{h^2} + 1 = \frac{D^2}{h^2} + 1$ *if* $h < \tilde{h}$. *Thus, the following holds true:*

- *If* $h \geqslant \tilde{h}$, *then* D_1 *is independent of the height of the BSs and we conclude that, if the density of the BSs is very large, then the coverage probability decreases as the height of the BSs increases. Thus, in dense networks, it is necessary to lower the height of the BSs as indicated from other literatures.*
- *If* $h \geqslant \tilde{h}$, *then* D_1 *depends on the height of the buildings. We know that the negative function* $g\left(D_1\right) = \Theta_{LOS}\left(\frac{1}{\gamma_{th}}\right) - D_1 \Theta_{LOS}\left(\frac{1}{\gamma_{th} D_1^{-\frac{\beta_{LOS}}{2}}}\right)$ *is monotonically decreasing in* D_1. *This implies that the larger* D_1 *is, the smaller this function is. Thus, the coverage decreases as* D_1 *increases. In dense cellular networks, the larger the LOS ball is, the smaller the height of the buildings is, and the worse the coverage probability would be. The above statement can be easily drawn from the very large interference.*
- *If* $h < \tilde{h}$, *then* D_1 *is independent of the height of the buildings and so is the coverage probability.*
- *If* $h < \tilde{h}$, *then* $D_1 = \frac{D^2}{h^2} + 1$ *depends on the height of the BSs. The following function:* $-h^2 g\left(D_1\right) = -h^2\left[\Theta_{LOS}\left(\frac{1}{\gamma_{th}}\right) - D_1 \Theta_{LOS}\left(\frac{1}{\gamma_{th} D_1^{-\frac{\beta_{LOS}}{2}}}\right)\right]$ *is determined to be decreasing with* h, *which implies that the coverage probability decreases with* h.

In conclusion, it is revealed that, in very dense networks, the coverage probability decreases with the height of the BSs. Note that this result holds true if the density of the BSs goes to infinity, which cannot be overlooked.

Case Study 2 If $\lambda_{\mathrm{BS}} \to \infty$ and $L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) = 1$ (the density of MTs is much larger than the density of the BSs), then the coverage probability is as follows:

$$\mathrm{P}_{\mathrm{cov}}^{\mathrm{UB}} \to 0 \tag{2.35}$$

Proof: This result follows directly from Case Study 1 when the density of the MTs goes to infinity.

Case Study 3 If $\lambda_{\mathrm{BS}} \to \infty$ and $L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) < 1$, the exact framework of coverage probability is as follows:

$$\mathrm{P_{cov}} \to \exp\left\{\pi h^2 \lambda_{\mathrm{MT}} g\left(D_1\right)\right\} = \exp\left\{\pi h^2 \lambda_{\mathrm{MT}} \left[\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right) - D_1 \Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}} D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right)\right]\right\} \quad (2.36)$$

which coincides with the result from the upper-bound.

Proof: See Appendix 2.8.8. □

2.4 System-Level Optimization and Performance Trends Analysis

2.4.1 Optimal BSs' Height

In this section, we evaluate the performance of cellular networks as a function of the height of the BSs. The main result is stated in the following theorem. Note that the parameters must satisfy the constraint identified above for the approximated upper-bound to be accurate and hold true.

In Theorem 1 ($h \geqslant \tilde{h}$), it can be proved that, if the approximated bound is accurate, there exist an optimal height of the BSs that maximizes the coverage probability under the assumption that the resulting optimal height is greater than $\tilde{h}$. This provides some constraints on the parameters to users. If this condition is not fulfilled, then the coverage probability decreases monotonically with the height of the BSs. In Theorem 2 ($h \leqslant \tilde{h}$), it is proved that the coverage probability decreases monotonically with the height of the BSs.

It is very important to note that we do not know a priori if the maximum coverage under the assumptions of Theorem 2 is smaller than the maximum coverage under assumptions of Theorem 1. However, it is not necessary use it for replacing the exact framework that is quite simple to compute but just to get insight and to have an approximate estimate of the optimal setup of parameters for coverage optimization. Once the approximated optimal choice of the parameters is obtained, this can be plugged into the exact framework of the coverage and a reliable estimate for it can be found. Based on the results of both Theorem, however, the following can be given ($h_{\mathrm{Th1}}^{\mathrm{opt}}$ is the optimum under Theorem 1):

$$h^{\mathrm{opt}} = \begin{cases} h_{\mathrm{Th1}}^{\mathrm{opt}} & \text{if} \quad \mathrm{P_{cov}}\left(h_{\mathrm{Th1}}^{\mathrm{opt}}\right) > \mathrm{P_{cov}}\left(0\right) \\ 0 & \text{if} \quad \mathrm{P_{cov}}\left(h_{\mathrm{Th1}}^{\mathrm{opt}}\right) < \mathrm{P_{cov}}\left(0\right) \end{cases} \quad (2.37)$$

and we emphasize that the exact framework is used to estimate the coverage due to its general

tractability. This is a general approach: We always compute the exact framework and use the bounds only for system optimizations and to get insight for system design.

Theorem 2.1 *Assume $h \geqslant \tilde{h}$, there exists a unique optimal height of the BSs that maximizes the coverage probability. The optimal height of the BSs is equal to:*

$$h_{Th1}^{opt} = \sqrt{\frac{\log\left[D_1\Upsilon_2 - 1\right] - \log\left[D_1\Upsilon_1 - 1\right]}{\pi\lambda_{BS}D_1\left(\Upsilon_2 - \Upsilon_1\right)}} = \sqrt{\frac{\log\left(\frac{D_1-1}{\Upsilon_1 D_1 - 1}\right)}{\pi\lambda_{BS}D_1\left(1-\Upsilon_1\right)}} \tag{2.38}$$

where $D_1 = \frac{D^2}{\tilde{h}^2} + 1$, Υ_1 and Υ_2 are defined in (2.28).

Proof*: This results can be easily obtained by computing the derivative of the approximated upper bound in* (2.33) *with respect to h, i.e., $\frac{d\bar{P}_{cov}^{UB}(h)}{dh} = 0$.* □

Remark 2.11 *Assume now that the optimum height can be computed by using the upper-bound instead of the approximated upper-bound. In other words, we compute the derivative of* (2.27)*: $\frac{dP_{cov}^{UB}(h)}{dh} = 0$, it would be difficult to obtain the analytical solutions. However, only numerical methods can be used to compute the solutions and check whether they are consistent, i.e., are positive and greater than the height of the BSs. So, thanks to the upper bound we now know that an optimal solution may exist but it depends on the parameters. However, it is difficult to get insight from this, even simplified, upper-bound, since the solution can only be computed numerically. A closed-form solution from which some important design information can be obtained is the one that is applicable under our simplifying assumptions, i.e., by ignoring the third summand.*

Remark 2.12 *It is important to mention that when solving the equation $\frac{d\bar{P}_{cov}^{UB}(h)}{dh} = 0$, the solution $h = 0$ is not considered since we assume $h \geqslant \tilde{h} > 0$. These are, in addition, special cases that we analyzed in Section 2.3.2 and, thus, they are not considered here.*

Remark 2.13 *The optimal value of the BSs' height is verified to be acceptable due to the fact that $\log\left[D_1\Upsilon_2 - 1\right] \geqslant \log\left[D_1\Upsilon_1 - 1\right]$, which yields $\Upsilon_2 \geqslant \Upsilon_1$. This implies that the stationary point always exists and it is always well defined.*

Remark 2.14 *Let $\frac{d\bar{P}_{cov}^{UB}(h)}{dh} \geqslant 0$, which implies that $\frac{d\bar{P}_{cov}^{UB}(h)}{dh} \geqslant 0 \Rightarrow h_{Th1}^{opt} \leqslant \sqrt{\frac{\log\left(\frac{D_1-1}{\Upsilon_1 D_1 - 1}\right)}{\pi\lambda_{BS}D_1(1-\Upsilon_1)}}$. Thus, the unique stationary point is indeed a maximum because the function first increases and then decreases.*

Remark 2.15 *If $h \geqslant \tilde{h} > 0$, Υ_1, Υ_2, Δ, and Φ are independent of the BS's height.*

Proposition 2.4 *Due to the fact that the optimal height is obtained under the assumption* $h \geqslant \tilde{h} > 0$*, this would result in* $h_{Th1}^{opt} = \sqrt{\frac{\log\left(\frac{D_1-1}{\Upsilon_1 D_1-1}\right)}{\pi\lambda_{BS}D_1(1-\Upsilon_1)}} = \sqrt{\frac{\log\left(\frac{D^2/\tilde{h}^2}{\Upsilon_1\left(D^2/\tilde{h}^2+1\right)-1}\right)}{\pi\lambda_{BS}\left(D^2/\tilde{h}^2+1\right)(1-\Upsilon_1)}} \geqslant \tilde{h}$*, which further indicates that the following condition for BSs' density needs to be satisfied in order to ensure the optimum:*

$$\lambda_{BS} \leqslant \frac{1}{\pi\tilde{h}^2\left[D^2/\tilde{h}^2 + L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)g\left(D^2/\tilde{h}^2\right)\right]}\log\left(-\frac{D^2/\tilde{h}^2}{L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)g\left(D^2/\tilde{h}^2\right)}\right) \tag{2.39}$$

where $g\left(D^2/\tilde{h}^2\right) = \Theta_{LOS}\left(\frac{1}{\gamma_{th}}\right) - \left(\frac{D^2}{\tilde{h}^2}+1\right)\Theta_{LOS}\left(\frac{\left(\frac{D^2}{\tilde{h}^2}+1\right)^{\frac{\beta_{LOS}}{2}}}{\gamma_{th}}\right) \leqslant 0$*, which follows from Corollary 2.3.*

If the above condition is not satisfied, then there is no optimal height of the BSs and the coverage will decrease monotonically with the antenna height. This is an important result that shows that the existence of an optimum depends on many parameters and that general conclusions cannot be drawn for some parameters based on numerical tests. Also, we need to consider the conditions to make the upper-bound applicable.

Corollary 2.5 *Let us consider* $\lambda_{BS} = \frac{1}{\pi R_{cell}^2}$*, the condition denoted in* (2.39) *can be rewritten as* $\frac{R_{cell}^2}{\tilde{h}^2} \geqslant F\left(D^2/\tilde{h}^2\right) = \frac{D^2/\tilde{h}^2 + L(\lambda_{MT}/\lambda_{BS})g\left(D^2/\tilde{h}^2\right)}{\log\left(-\frac{D^2/\tilde{h}^2}{L(\lambda_{MT}/\lambda_{BS})g\left(D^2/\tilde{h}^2\right)}\right)}$*, then* $F\left(D^2/\tilde{h}^2\right)$ *is monotonically increasing with the ratio* $D^2/\tilde{h}^2$*, and reaches infinity.*

Proof*: See Appendix 2.8.9.* □

This means that the larger the ratio between the LOS ball and the height of the building is, the larger the ratio between the average cell size and the height of the building is. For a fixed value of $\tilde{h}$, this implies that the bigger the LOS ball is, then the larger the cell size is. In order words, the fewer the buildings are the less dense the cellular network needs to be. Note that this comment is strictly correct under the condition $L\left(\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}}\right) = 1$. The reason is that this function, in general, depends on the density of the BSs as well.

Based on this result, we conclude, that, as the LOS ball increases, then the network should be sparser for an optimal BS's height to exist. In particular, we observe that, if the LOS ball goes to infinity, then the average cell size should go to infinity. This means that for very large values of the LOS ball there is no optimal height of the BSs. This result is consistent with Case Study 2 in Section 2.3.2 ($D \to \infty$).

Corollary 2.6 *Let us consider* $\lambda_{BS} = \frac{1}{\pi R_{cell}^2}$*, the condition denoted in* (2.39) *can be rewritten as*

$R^2_{cell} \geqslant \tilde{h}^2 \times \frac{D^2/\tilde{h}^2 + L(\lambda_{MT}/\lambda_{BS})g\left(D^2/\tilde{h}^2\right)}{\log\left(-\frac{D^2/\tilde{h}^2}{L(\lambda_{MT}/\lambda_{BS})g\left(D^2/\tilde{h}^2\right)}\right)} = \tilde{h}^2 F\left(D^2/\tilde{h}^2\right)$, *then* $\tilde{h}^2 F\left(D^2/\tilde{h}^2\right)$ *is monotonically increasing with the height of buildings* $\tilde{h}^2$.

Proof*: See Appendix 2.8.10.* □

Based on this result, once again, it is concluded that as the heights of the buildings increases, the network should be sparser for an optimal BS's height to exist.

Theorem 2.2 *Let us assume* $h < \tilde{h}$*. The coverage probability monotonically decreases with the height of the BSs. In other words,* $h^{opt} = 0$.

Proof*: See Appendix 2.8.11.* □

Based on the above obtained results, we can study the impact of several system parameters on $h^{\text{opt}}_{\text{Th1}}$ in (2.38). They are summarized in the following corollaries. We note that the analysis is correct if it is assumed that the parameters belong to the range of values for which: i) the approximated upper-bound is accurate, and ii) the optimal height of the BSs exists. These two assumptions are implicit in the rest of the analysis and need to always be checked.

Corollary 2.7 *Let* $h \geqslant \tilde{h} > 0$*, the optimal height of the BSs (obtained by maximizing the coverage probability) monotonically decreases as the density of the BSs increases.*

Proof*: See Appendix 2.8.12.* □

Remark 2.16 *Note that when cellular network is assumed to be fully loaded, i.e.,* $L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right) = 1$*, it can be found* $h^{opt}_{Th1} \propto \sqrt{\frac{1}{\pi\lambda_{BS}}}$*, which indicates that the SIR invariance property holds.*

Corollary 2.8 *Let* $h \geqslant \tilde{h} > 0$*, the optimal height of the BSs (obtained by maximizing the coverage probability) monotonically decreases as the radius of the LOS ball (D) increases.*

Proof: See Appendix 2.8.13. □

Corollary 2.9 *Let* $h \geqslant \tilde{h} > 0$*, the optimal height of the BSs (obtained by maximizing the coverage probability) monotonically increases as the height of the buildings* $(\tilde{h})$ *increases.*

Proof: This corollary directly follows from the proof of Corollary 2.8. In fact, $h^{\text{opt}}_{\text{Th1}}(D_1)$ is proved to increase with D_1. Now that $D_1 = \frac{D^2}{\tilde{h}^2} + 1$, this implies that $h^{\text{opt}}_{\text{Th1}}\left(\tilde{h}\right)$ increases with $\tilde{h}$.

Corollary 2.10 *Let $h \geqslant \tilde{h} > 0$, the optimal height of the BSs (obtained by maximizing the coverage probability) monotonically decreases as the density of the mobile terminals (λ_{MT}) increases.*

Proof: See Appendix 2.8.14. □

2.4.2 Optimal BSs' Density

In this section, we evaluate the performance of cellular networks as a function of the density of the BSs. The main result is stated in the following theorem. Similar to the case study where the height of the BSs was of interest, we apply the approximated upper bound in (2.33) for optimization. Therefore, the conditions under which this holds true need to be taken into account in the rest of the analysis.

Remark 2.17 *The first remark is that now, with the exception of $\Upsilon_2 = 1$, all the parameters depends on λ_{BS}. This makes the analysis very complicated. Evidently, the derivatives can be computed, but no much feasible results can be obtained after doing so. The only case study that be analyzed and from which some insights can be gained from it is when $L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right) \approx 1$. In this case, all the parameters are independent of λ_{BS}. Thus, the derivative can be computed and some insights can be gained. For this reason, only this case study is taken into account.*

Remark 2.18 *It is worth noting that in this case it is not necessary to consider the case studies $h \geqslant \tilde{h}$ and $h < \tilde{h}$ independently, since the derivative is computed only with respect to λ_{BS}. So, the study and analysis is unified.*

Theorem 2.3 *Let us assume $L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right) = 1$. There exists a unique deployment density of the BSs that maximizes the coverage probability. The optimal density of the BSs is equal to:*

$$\lambda_{BS}^{opt} = \frac{\log\left[\frac{D_h^2}{\left(D_h^2+h^2\right)\Upsilon_1 - h^2}\right]}{\pi\left(D_h^2+h^2\right)\left(1-\Upsilon_1\right)} \tag{2.40}$$

or, equivalently, the cell radius of the BSs is equal to:

$$R_{cell}^{opt} = \sqrt{\frac{\left(D_h^2+h^2\right)\left(1-\Upsilon_1\right)}{\log\left[\frac{D_h^2}{\left(D_h^2+h^2\right)\Upsilon_1 - h^2}\right]}} \tag{2.41}$$

where Υ_1 is defined in (2.28).

Proof*: This results can be easily obtained by computing the derivative of the approximated upper bound in* (2.33) *with respect to λ_{BS}, i.e., $\frac{d\bar{P}_{cov}^{UB}(\lambda_{BS})}{d\lambda_{BS}} = 0$.* □

Remark 2.19 *The obtained expression makes sense if* $\left(D_h^2 + h^2\right)\Upsilon_1 - h^2 \geqslant 0$*, in order for the logarithmic function to be well defined. From Corollary 2.2, we know that that following is true:* $1 - \left(\frac{D_h^2}{h^2} + 1\right)\Upsilon_1 \leqslant 0 \Rightarrow \left(D_h^2 + h^2\right)\Upsilon_1 - h^2 \geqslant 0$*. This implies that the optimum is always well defined.*

Based on the above obtained results, we analyze the impact of several system parameters on $R_{\text{cell}}^{\text{opt}}$ in (2.41). They are summarized in the following corollaries.

Corollary 2.11 *Let* $h \geqslant \tilde{h} > 0$*, the cell radius of the BSs (obtained by maximizing the coverage probability) monotonically increases as the height of the BSs (h) increases. More precisely, it linearly increases with h.*

Proof*: This statement can be proved as follows. If* $h \geqslant \tilde{h} > 0$*, according to the definition of* D_h*,* (2.41) *becomes:*

$$R_{cell}^{opt} = \sqrt{h^2\left(\frac{D^2}{\tilde{h}^2} + 1\right)(1 - \Upsilon_1) \Big/ \log\left[\frac{\frac{D^2}{\tilde{h}^2}}{\left(\frac{D^2}{\tilde{h}^2} + 1\right)\Upsilon_1 - 1}\right]}$$

The proof follows by direct inspection of the equation, since Υ_1 *does not depend on h.*

Corollary 2.12 *Let* $h < \tilde{h}$*, the cell radius of the BSs (obtained by maximizing the coverage probability) monotonically increases as the height of the BSs increases (h) increases.*

Proof*: See Appendix 2.8.15.* □

Corollary 2.13 *The cell radius of the BSs (obtained by maximizing the coverage probability) monotonically increases as the LOS ball (D) increases.*

Proof*: See Appendix 2.8.16.* □

Corollary 2.14 *The cell radius of the BSs (obtained by maximizing the coverage probability) monotonically decreases as the height of blockages* $(\tilde{h})$ *increases if* $h \geqslant \tilde{h}$ *and is independent of the height of blockages if* $h < \tilde{h}$*.*

Proof: This statement immediately follow from Corollary 2.13, since it is apparent that the behavior as a function of $\tilde{h}$ is opposite to that of D, if $h \geqslant \tilde{h}$, while does not appear if $h < \tilde{h}$.

2.5 Generalization

In this section, we generalize the proposed analytical framework by taking into account the impact of noise, which is concerned with the interference-limited assumption. This

assumption is reasonable if the density of the BSs is very large. It is, however, not acceptable if the density of the BSs is not very high. Under the interference-limited assumption, in particular, the results about the coverage probability for small densities of the BSs may not be very reliable (e.g., the coverage probability is not equal to zero if the density of the BSs tends to zero).

2.5.1 Exact Mathematical Formulation of Coverage Probability

Based on the Definition 2.2 in Section 2.2.6, as a result, the mathematical framework of coverage probability becomes:

$$\mathrm{P}_{\mathrm{cov}}^{\mathrm{G}} = \Pr\left\{\mathrm{SIR} \geqslant \gamma_{\mathrm{th}}, \overline{\mathrm{SNR}} \geqslant \gamma_{\mathrm{A}}\right\} = \int_0^{\frac{P}{\gamma_{\mathrm{A}}\sigma_{\mathrm{N}}^2}} M_{I_{\mathrm{agg}}}\left(\frac{\gamma_{\mathrm{th}} x}{\Omega} | x\right) f_{L^{(0)}}(x)\, dx \tag{2.42}$$

We note that the only difference between the original framework in (2.14) and the new analytical framework in (2.42) is the upper-limit of the integral that now does not go to infinity. According to this property, we can apply the same analytical framework of the previous sections under the pure-LOS ball model. In general, the resulting framework is obtained in Table 2.4, where the Heaviside functions originate from comparing the upper limit of integral, $\frac{\mathrm{P_{tx}}}{\gamma_{\mathrm{A}}\sigma_{\mathrm{N}}^2}$, with three essential terms, $k_{\mathrm{LOS}} h^{\beta_{\mathrm{LOS}}}$, $k_{\mathrm{LOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}$, and $k_{\mathrm{NLOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{NLOS}}}{2}}$, while computing the initial integral.

Table 2.4 – Auxiliary functions used in (2.42).

Function definition
$\mathrm{P}_{\mathrm{cov}}^{\mathrm{G}} = \mathrm{P}_{\mathrm{cov}}^{\mathrm{G1}} \times H\left(\frac{\mathrm{P_{tx}}}{\gamma_{\mathrm{A}}\sigma_{\mathrm{N}}^2} - k_{\mathrm{LOS}} h^{\beta_{\mathrm{LOS}}}\right) \bar{H}\left(\frac{\mathrm{P_{tx}}}{\gamma_{\mathrm{A}}\sigma_{\mathrm{N}}^2} - k_{\mathrm{LOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}\right)$
$+\mathrm{P}_{\mathrm{cov}}^{\mathrm{G2}} \times H\left(\frac{\mathrm{P_{tx}}}{\gamma_{\mathrm{A}}\sigma_{\mathrm{N}}^2} - k_{\mathrm{LOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}\right) \bar{H}\left(\frac{\mathrm{P_{tx}}}{\gamma_{\mathrm{A}}\sigma_{\mathrm{N}}^2} - k_{\mathrm{NLOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{NLOS}}}{2}}\right)$
$+\mathrm{P}_{\mathrm{cov}}^{\mathrm{G3}} \times H\left(\frac{\mathrm{P_{tx}}}{\gamma_{\mathrm{A}}\sigma_{\mathrm{N}}^2} - k_{\mathrm{NLOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{NLOS}}}{2}}\right)$
$\mathrm{P}_{\mathrm{cov}}^{\mathrm{G1}} = \frac{2\pi\lambda_{\mathrm{BS}}}{\beta_{\mathrm{LOS}}} \exp\left(\pi\lambda_{\mathrm{BS}} h^2\right)\left(D_h^2+h^2\right) \int_{\left(\frac{D_h^2}{h^2}+1\right)^{-\frac{\beta_{\mathrm{LOS}}}{2}}}^{\frac{\mathrm{P_{tx}}}{k_{\mathrm{LOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}\gamma_{\mathrm{A}}\sigma_{\mathrm{N}}^2}} y^{\frac{2}{\beta_{\mathrm{LOS}}}-1} \exp\left[\tau(y)\right] dy$
$\mathrm{P}_{\mathrm{cov}}^{\mathrm{G2}} = \frac{2\pi\lambda_{\mathrm{BS}}}{\beta_{\mathrm{LOS}}} \exp\left(\pi\lambda_{\mathrm{BS}} h^2\right)\left(D_h^2+h^2\right) \int_{\left(\frac{D_h^2}{h^2}+1\right)^{-\frac{\beta_{\mathrm{LOS}}}{2}}}^{1} y^{\frac{2}{\beta_{\mathrm{LOS}}}-1} \exp\left[\tau(y)\right] dy$
$\mathrm{P}_{\mathrm{cov}}^{\mathrm{G3}} = \mathrm{P}_{\mathrm{cov}}^{\mathrm{G2}} + \frac{\exp\left(\pi\lambda_{\mathrm{BS}} h^2\right)\exp\left\{-\left(D_h^2+h^2\right)\pi\lambda_{\mathrm{BS}}\Phi\right\} - \exp\left(\pi\lambda_{\mathrm{BS}} h^2\right)\exp\left\{-\pi\lambda_{\mathrm{BS}}\left(\frac{\mathrm{P_{tx}}}{\gamma_{\mathrm{A}}\sigma_{\mathrm{N}}^2 k_{\mathrm{NLOS}}}\right)^{\frac{2}{\beta_{\mathrm{NLOS}}}}\Phi\right\}}{\Phi}$
$\tau(y) = -\pi\lambda_{\mathrm{BS}} L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left(D_h^2+h^2\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}} y}\right) - \pi\lambda_{\mathrm{BS}} L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left(D_h^2+h^2\right) y^{\frac{2}{\beta_{\mathrm{LOS}}}}\left[1-\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)\right]$
$+\pi\lambda_{\mathrm{BS}} L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left(D_h^2+h^2\right)\left[1-\Gamma_{\mathrm{NLOS}}\left(\frac{\gamma_{\mathrm{th}} k_{\mathrm{LOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}} y}{k_{\mathrm{NLOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{NLOS}}}{2}}}\right)\right]$

Remark 2.20 *If $\sigma_N^2 \to 0$, the generalized framework of coverage probability reduces to the original one based on the conventional definition in* (2.17).

Remark 2.21 *If $\sigma_N^2 \to \infty$, the generalized framework of coverage probability tends to be zero, i.e., $P_{cov}^G \to 0$.*

Remark 2.22 *If $\lambda_{BS} \to 0$, the generalized framework of coverage probability tends to be zero, i.e., $P_{cov}^G \to 0$.*

Remark 2.23 *If $\lambda_{BS} \to \infty$, the generalized framework of coverage probability tends to be coincident with the results obtained in Case Study 3 of Section 2.3.4, which depends on the impact of λ_{MT}.*

2.5.2 Closed-Form Upper Bound of Coverage Probability

The general expression of the coverage probability in (2.42) depends on the ratio $\frac{\mathrm{P_{tx}}}{\gamma_{\mathrm{A}}\sigma_{\mathrm{N}}^2}$, and there are different integrals that need to be computed. For simplicity, we study only the most reasonable case, which is likely to occur in practice. The other case studies can be investigated by using similar steps. The most reasonable case study is when the following holds true: $H\left(\frac{\mathrm{P_{tx}}}{\gamma_{\mathrm{A}}\sigma_{\mathrm{N}}^2} - k_{\mathrm{NLOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{NLOS}}}{2}}\right) = 1$, which implies that

$$\frac{\mathrm{P_{tx}}}{\sigma_{\mathrm{N}}^2 k_{\mathrm{NLOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{NLOS}}}{2}}} > \gamma_{\mathrm{A}} \tag{2.43}$$

Note that the left-hand side is the received power at an horizontal distance that corresponds to the LOS ball and as a function of the BSs' height. It is very improbable that detection threshold is chosen in a way that the above inequality is not fulfilled, because it would imply that, for a given BS's height, that it would not be fulfilled for distances larger than the radius of the LOS ball. Since this situation arises quite often in practice, except for BSs' densities that lead to cell radius much smaller that the radius of the LOS ball. Hence, this is a reasonable assumption to consider in practice. By all means, the consistency of the results need to be assessed by checking if the obtained results (for example the optimal values of the BSs' density or height) are consistent with the assumption $\frac{\mathrm{P_{tx}}}{\sigma_{\mathrm{N}}^2 k_{\mathrm{NLOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{NLOS}}}{2}}} > \gamma_{\mathrm{A}}$, which is deterministic in any case.

Thus, in this case, the coverage probability would be $\mathrm{P_{cov}^{G3}} \times H\left(\frac{\mathrm{P_{tx}}}{\gamma_{\mathrm{A}}\sigma_{\mathrm{N}}^2} - k_{\mathrm{NLOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{NLOS}}}{2}}\right)$, and the corresponding upper bound can be obtained directly based on the previous results while the only change is the last term that contains the transmit power. In particular, the

resulting closed-form upper bound is formulated by:

$$\mathrm{P}_{\mathrm{cov}}^{\mathrm{G-UB}} = H\left(\frac{\mathrm{P_{tx}}}{\gamma_{\mathrm{A}}\sigma_{\mathrm{N}}^2} - k_{\mathrm{NLOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{NLOS}}}{2}}\right) \times \left\{ \left[\frac{\exp\left\{\pi\lambda_{\mathrm{BS}}h^2\right\}\exp\left(-\pi\lambda_{\mathrm{BS}}\left(D_h^2+h^2\right)\Upsilon_1\right)}{\Delta} - \frac{\exp\left\{\pi\lambda_{\mathrm{BS}}h^2\right\}\exp\left(-\pi\lambda_{\mathrm{BS}}\left(D_h^2+h^2\right)\Upsilon_2\right)}{\Delta} \right] + \frac{\exp\left(\pi\lambda_{\mathrm{BS}}h^2\right)\left[\exp\left\{-\left(D_h^2+h^2\right)\pi\lambda_{\mathrm{BS}}\Phi\right\} - \exp\left\{-\left(\frac{\mathrm{P_{tx}}}{\gamma_{\mathrm{A}}\sigma_{\mathrm{N}}^2}\right)^{\frac{2}{\beta_{\mathrm{NLOS}}}}\pi\lambda_{\mathrm{BS}}\left(\frac{1}{k_{\mathrm{NLOS}}}\right)^{\frac{2}{\beta_{\mathrm{NLOS}}}}\Phi\right\}\right]}{\Phi} \right\} \tag{2.44}$$

where the notation is always the same.

Remark 2.24 *The primary difference between the upper bound of SIR coverage in* (2.27) *and the genralized one based on the new definition in* (2.44) *occurs when the density of the BSs tends to be zero. Specifically, if* $\lambda_{BS} \to 0$*, the following difference can be highlighted:*

$$\lim_{\lambda_{BS}\to 0} P_{cov}^{UB} = \frac{1}{\Phi} \neq 0, \lim_{\lambda_{BS}\to 0} P_{cov}^{G-UB} = 0 \tag{2.45}$$

2.6 Numerical Results

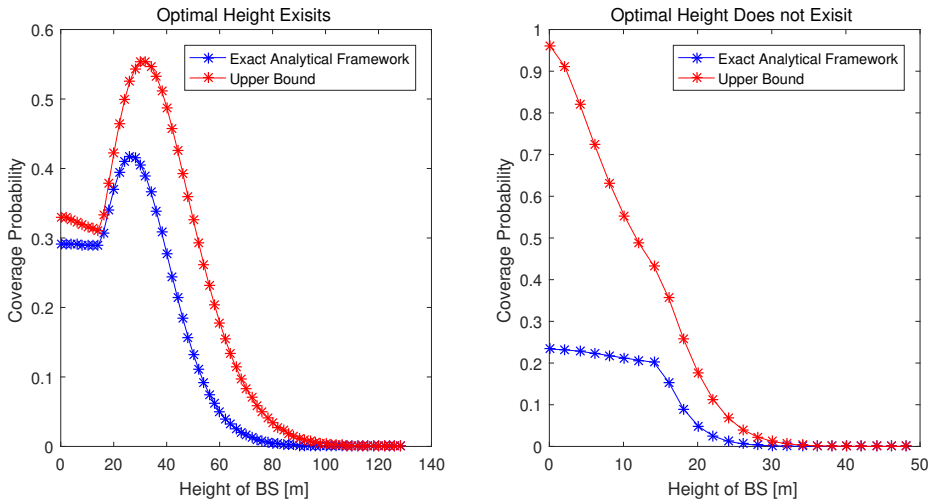

Figure 2.2 – Coverage probability vs height of BSs in meter for different case studies. Setup: $R_{\mathrm{cell}} = 150$ m (left) and $R_{\mathrm{cell}} = 50$ m (right).

In this section, we illustrate several numerical results that substantiate the applicability of the proposed approach for the modeling and analysis of the 3D cellular network deployments. In particular, the optimal height and density of BSs are hilighted. Unless otherwise stated, the simulation setup is summarized in Table 2.5.

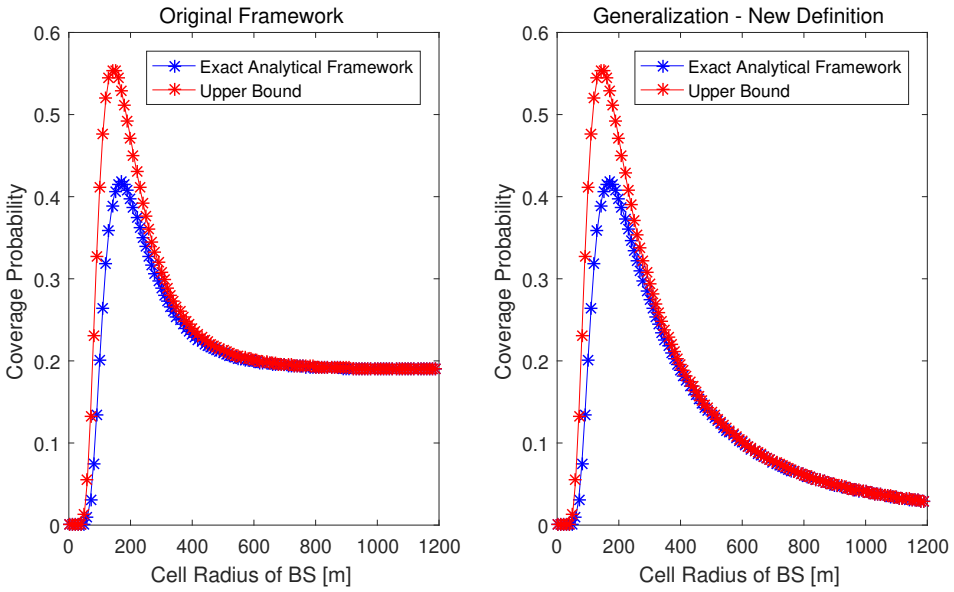

Figure 2.3 – Coverage probability vs density of BSs (R_{cell} is average cell radius in meter) for different case studies). Setup: Height of BSs, $h = 30$ m.

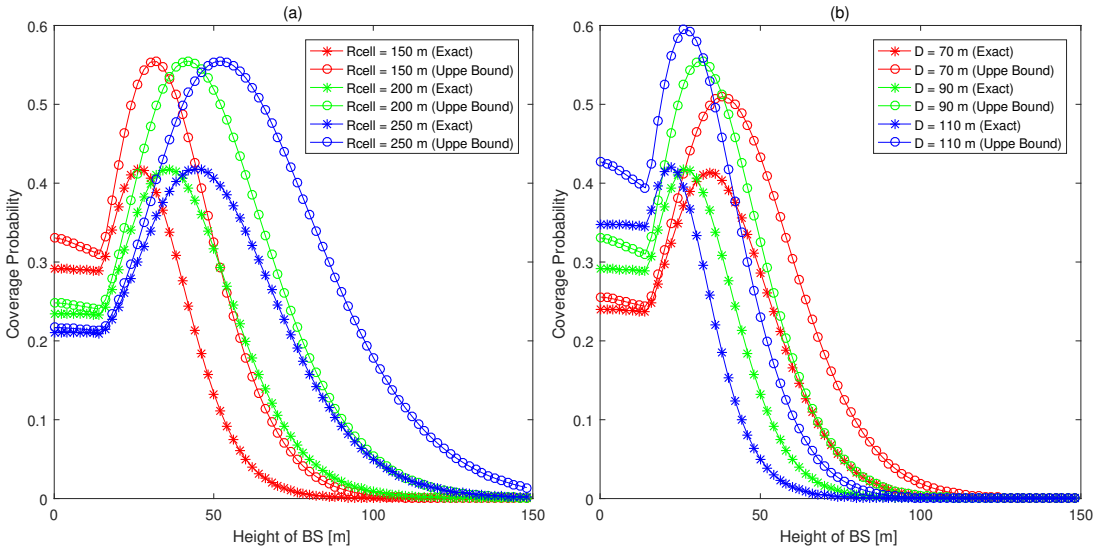

Figure 2.4 – Coverage probability vs height of BSs under the impact of different values of: (a) density of BS R_{cell}, (b) 2D LOS ball radius D.

In Figure 2.2, we illustrate the coverage probability versus the height of BSs. The exact analytical framework and the upper bound are formulated in Proposition 2.1 and Proposition 2.2, respectively. In particular, for the case that the optimal height of BSs exists, the cell radius of BSs is selected as $R_{\text{cell}} = 150$ m which indicates the condition in (2.39) of Proposition 2.4 is satisfied. For the case that the optimal height of BSs does not exists, on the other hand, the cell

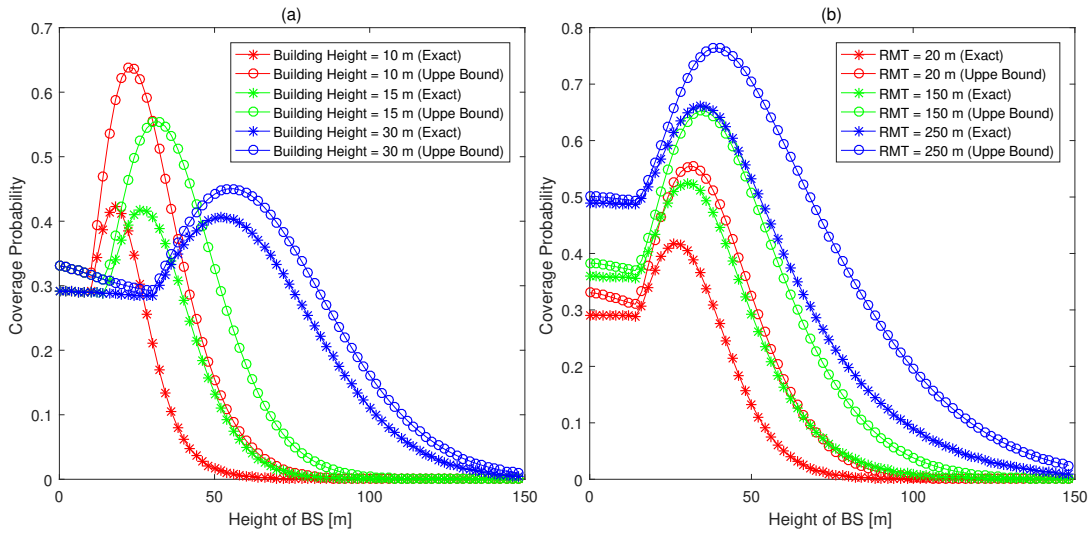

Figure 2.5 – Coverage probability vs height of BSs under the impact of different values of: (a) height of building $\tilde{h}$, (b) density of MT R_{MT}.

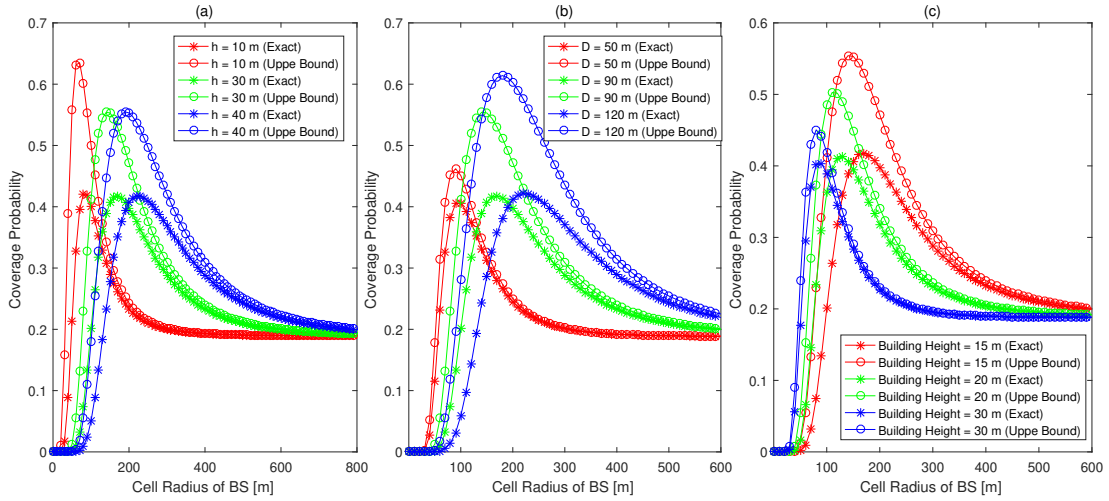

Figure 2.6 – Coverage probability vs density of BSs under the impact of different values of: (a) height of BS h, (b) 2D LOS ball radius D, (c) height of building $\tilde{h}$.

radius of BSs is selected as $R_{\mathrm{cell}} = 50$ m, which means the condition in (2.39) of Proposition 2.4 is not satisfied and coverage probability monotonically decreases as the height of BSs. We highlight that the optimal height of BSs exists to maximize the coverage probaility only when this sufficient condition is guaranteed, or otherwise, the optimal height becomes zero.

In Figure 2.3, we illustrate the coverage probability versus the density of BSs. The exact analytical framework and the upper bound of the original case are formulated in Proposition 2.1 and Proposition 2.2. As for the generalized framework based on the new definition, we

Table 2.5 – Setup of parameters (unless otherwise stated).

Parameter	Value
β_{LOS}	2.1
β_{NLOS}	3.9
$k_{\text{LOS}} = k_{\text{NLOS}} = \left(4\pi f_c/3\cdot 10^8\right)^2$	f_c = 2.5 GHz
Bandwidth	B_w = 200 KHz
Noise Power $\sigma^2_{\text{N}} = -174 + 10\log_{10}(B_w) + N_f$	N_f = 10 dBm
$\tilde{h}$	15 m
D	90 m
R_{cell}	150 m
$L\left(\frac{\lambda_{\text{MT}}}{\lambda_{\text{BS}}}\right)$	1 (Fully-loaded)
P_{tx}	30 dBm
$\Omega_{\text{LOS}} = \Omega_{\text{NLOS}}$	1
γ_{th}	10 dB

apply the exact formulation and its upper bound in (2.42) and (2.44), respectively. We notice that the optimal density of BSs exists to maximize the overall coverage probabilily for both cases. It is worth noting that here we highlight the necessity of introducing the generalized frameworks with the new definition due to the fact that its behavior trends goes to zero when the density of the BSs tends to be zero, which coincident with the practical assumption and the statement in Remark 2.24.

In Figure 2.4 and Figure 2.5, the behavior of coverage probability as a function of height of BS is illustrated under the impact of different system parameters. In particular, we observe the performance trends of the optimal height of BSs are identical to the comments in Corollary 2.7 - Corollary 2.10. In Figure 2.6, the behavior of coverage probability as a function of density of BS is illustrated under the impact of different system parameters. It is shown that the performance trends of the optimal height of BSs are identical to the comments in Corollary 2.11 - Corollary 2.14.

2.7 Conclusion

In this chapter, we have proposed a tractable mathematical framework of coverage probability to evaluate the performance of 3D cellular networks, and it has been shown that several coincident results can be found with other literatures based on the proposed framework. In addition, we have introduced a closed-form upper bound to gain more insights where the exact expression may not be able to provide. State differently, under the sufficient conditions, the upper bound would yield the closed-form expressions of the optimal BSs' height and density to optimize the coverage probability. We have analyzed the impact of key system parameters on the obtained expressions of the optimal height and density, which is verified to be consistent with the practical setup in engineering application. Moreover, we have generalized the mathematical framework of coverage based on the new definition to overcome the limitations caused by the interference-limited assumption and the impact of noise. All the performance trends and optimization findings have been substantiated with aid of numerical simulations.

We conclude that the obtained results for the system-level modeling and optimization of 3D cellular networks are effective and promising.

2.8 Appendices

2.8.1 Proof of Case Study 2 in Section 2.3.2

Based on the condition when $D \to \infty \Rightarrow D_h^2 \to \infty$, (2.15) becomes:

$$\mathrm{P}_{\mathrm{cov}}^{1} \overset{D_h^2 \to \infty}{=} \int_{k_{\mathrm{LOS}} h^{\beta_{\mathrm{LOS}}}}^{\infty} \exp\left\{ \begin{array}{l} -\pi\lambda_{\mathrm{BS}} L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left(D_h^2+h^2\right)\Theta_{\mathrm{LOS}}\left(\frac{k_{\mathrm{LOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}}{\gamma_{\mathrm{th}} x}\right) \\ +\pi\lambda_{\mathrm{BS}} L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left(\frac{x}{k_{\mathrm{LOS}}}\right)^{\frac{2}{\beta_{\mathrm{LOS}}}}\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right) \\ +\pi\lambda_{\mathrm{BS}} L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left(D_h^2+h^2\right)\left[1-\Gamma_{\mathrm{NLOS}}\left(\frac{\gamma_{\mathrm{th}} x}{k_{\mathrm{NLOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{NLOS}}}{2}}}\right)\right] \end{array} \right\} \times \left\{\frac{2\pi\lambda_{\mathrm{BS}}}{\beta_{\mathrm{LOS}}}\left(\frac{1}{k_{\mathrm{LOS}}}\right)^{\frac{2}{\beta_{\mathrm{LOS}}}} x^{\frac{2}{\beta_{\mathrm{LOS}}}-1}\right\} \exp\left\{-\pi\lambda_{\mathrm{BS}}\left(\left(\frac{x}{k_{\mathrm{LOS}}}\right)^{\frac{2}{\beta_{\mathrm{LOS}}}}-h^2\right)\right\} dx \tag{2.46}$$

In order to solve the integral, the following two primary results are leveraged:

$$\left(D_h^2+h^2\right)\Theta_{\mathrm{LOS}}\left(\frac{k_{\mathrm{LOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}}{\gamma_{\mathrm{th}} x}\right) \overset{D_h^2\to\infty}{\to} \left(\frac{\gamma_{\mathrm{th}} x}{k_{\mathrm{LOS}}}\right)^{\frac{2}{\beta_{\mathrm{LOS}}}} {}_2F_1\left(\frac{2}{\beta_{\mathrm{LOS}}}, \frac{2}{\beta_{\mathrm{LOS}}}, 1+\frac{2}{\beta_{\mathrm{LOS}}}, 1\right) \tag{2.47}$$

$$\left(D_h^2+h^2\right)\left[1-\Gamma_{\mathrm{NLOS}}\left(\frac{\gamma_{\mathrm{th}} x}{k_{\mathrm{NLOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{NLOS}}}{2}}}\right)\right] \overset{D_h^2\to\infty}{\to} 0 \tag{2.48}$$

The first result can be proved as follows. Let us define ($a>0$):

$$\mathscr{F}(D) = (D+a)\,\Theta\left(A(D+a)^{\frac{\beta}{2}}\right) \overset{(a)}{=} (D+a)\left[1+A(D+a)^{\frac{\beta}{2}}\right]^{-\frac{2}{\beta}} {}_2F_1\left(\frac{2}{\beta}, \frac{2}{\beta}, 1+\frac{2}{\beta}, \frac{A(D+a)^{\frac{\beta}{2}}}{1+A(D+a)^{\frac{\beta}{2}}}\right)$$

where (a) is hold by applying the equality ${}_2F_1(a,b,c,-z) = (1+z)^{-b}\,{}_2F_1\left(b, c-a, c, \frac{z}{z+1}\right)$.

Thus, the limit can be computed by:

$$\lim_{D\to\infty} \mathscr{F}(D) = A^{-\frac{2}{\beta}}\,{}_2F_1\left(\frac{2}{\beta}, \frac{2}{\beta}, 1+\frac{2}{\beta}, 1\right)$$

which is based on the following:

$$(D+a)\left[1+A(D+a)^{\frac{\beta}{2}}\right]^{-\frac{2}{\beta}} \overset{D\to\infty}{\to} (D+a)\left[A(D+a)^{\frac{\beta}{2}}\right]^{-\frac{2}{\beta}} = A^{-\frac{2}{\beta}}$$
$${}_2F_1\left(\frac{2}{\beta},\frac{2}{\beta},1+\frac{2}{\beta},\frac{A(D+a)^{\frac{\beta}{2}}}{1+A(D+a)^{\frac{\beta}{2}}}\right) \overset{D\to\infty}{\to} {}_2F_1\left(\frac{2}{\beta},\frac{2}{\beta},1+\frac{2}{\beta},1\right)$$

The second result can be proved as follows. Let us define ($a>0$):

$$\mathscr{F}(D) = (D+a)\left[1-\Gamma\left(A(D+a)^{-\frac{\beta}{2}}\right)\right] \overset{(a)}{=} (D+a) - \frac{D+a}{1+A(D+a)^{-\frac{\beta}{2}}}\,{}_2F_1\left(1,1,1-\frac{2}{\beta},\frac{1}{1+\frac{1}{A(D+a)^{-\frac{\beta}{2}}}}\right)$$

where (a) is hold by applying the equality ${}_2F_1(a,b,c,-z) = (1+z)^{-b}\,{}_2F_1\left(b,c-a,c,\frac{z}{z+1}\right)$.

Now since $\lim_{D\to\infty} {}_2F_1\left(1,1,1-\frac{2}{\beta},\frac{1}{1+\frac{1}{A(D+a)^{-\frac{\beta}{2}}}}\right) = {}_2F_1\left(1,1,1-\frac{2}{\beta},0\right) = 1$, which implies that

$$\lim_{D\to\infty} f(D) = \lim_{D\to\infty}\left[(D+a) - \frac{D+a}{1+A(D+a)^{-\frac{\beta}{2}}}\right] \overset{x=D+a}{=} \lim_{x\to\infty}\left[x - \frac{x}{1+Ax^{-\frac{\beta}{2}}}\right] = \lim_{x\to\infty}\left[\frac{Ax^{1-\frac{\beta}{2}}}{1+Ax^{-\frac{\beta}{2}}}\right] = 0$$

Hence, we can formulated $\mathrm{P}^1_{\mathrm{cov}}$ as follows based on the above crucial results:

$$\mathrm{P}^1_{\mathrm{cov}} \overset{D_h^2\to\infty}{=} \frac{2\pi\lambda_{\mathrm{BS}}}{\beta_{\mathrm{LOS}}}\left(\frac{1}{k_{\mathrm{LOS}}}\right)^{\frac{2}{\beta_{\mathrm{LOS}}}} \exp\left\{\pi\lambda_{\mathrm{BS}}h^2\right\}\int_{k_{\mathrm{LOS}}h^{\beta_{\mathrm{LOS}}}}^{\infty} x^{\frac{2}{\beta_{\mathrm{LOS}}}-1}\exp\left\{-\pi\lambda_{\mathrm{BS}}\left(\frac{1}{k_{\mathrm{LOS}}}\right)^{\frac{2}{\beta_{\mathrm{LOS}}}}\mathrm{T}x^{\frac{2}{\beta_{\mathrm{LOS}}}}\right\}dx$$
$$= \frac{1}{\mathrm{T}}\exp\left\{-\pi\lambda_{\mathrm{BS}}h^2(\mathrm{T}-1)\right\} \tag{2.49}$$

where $\mathrm{T} = 1+L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\gamma_{\mathrm{th}}^{\frac{2}{\beta_{\mathrm{LOS}}}}\,{}_2F_1\left(\frac{2}{\beta_{\mathrm{LOS}}},\frac{2}{\beta_{\mathrm{LOS}}},1+\frac{2}{\beta_{\mathrm{LOS}}},1\right) - L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right){}_2F_1\left(1,\frac{2}{\beta_{\mathrm{LOS}}},1+\frac{2}{\beta_{\mathrm{LOS}}},-\frac{1}{\gamma_{\mathrm{th}}}\right)$. Finally, the resulting expression is obtained as follows.

$$\mathrm{P}^1_{\mathrm{cov}} = \frac{1}{\mathrm{T}}\exp\left\{-\pi\lambda_{\mathrm{BS}}h^2(\mathrm{T}-1)\right\} \overset{(a)}{=} \frac{\exp\left\{\pi\lambda_{\mathrm{BS}}h^2L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[1-{}_2F_1\left(-\frac{2}{\beta_{\mathrm{LOS}}},1,1-\frac{2}{\beta_{\mathrm{LOS}}},-\gamma_{\mathrm{th}}\right)\right]\right\}}{1-L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[1-{}_2F_1\left(-\frac{2}{\beta_{\mathrm{LOS}}},1,1-\frac{2}{\beta_{\mathrm{LOS}}},-\gamma_{\mathrm{th}}\right)\right]} \tag{2.50}$$

where (a) is hold based on the following property:

$$z^{\frac{2}{\beta}}\,{}_2F_1\left(\frac{2}{\beta},\frac{2}{\beta},1+\frac{2}{\beta},1\right) = {}_2F_1\left(1,\frac{2}{\beta},1+\frac{2}{\beta},-\frac{1}{z}\right) + {}_2F_1\left(-\frac{2}{\beta},1,1-\frac{2}{\beta},-z\right) - 1 \tag{2.51}$$

2.8.2 Proof of Corollary 2.1

Firstly, $\Upsilon_1 \leqslant 1$ is proved as follows:

$$\begin{aligned}\Upsilon_1 &= K_1 L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) + \Delta y_1^{\frac{2}{\beta_{\mathrm{LOS}}}} \overset{(a)}{=} D_1^{-1}\left\{D_1 \Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}} D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right) L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) + 1 - L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)\right\} \\ &\overset{(b)}{\leqslant} D_1^{-1}\left\{L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)(D_1 - 1) + 1\right\} \overset{(c)}{\leqslant} D_1^{-1}\{1 \times (D_1 - 1) + 1\} = 1\end{aligned} \tag{2.52}$$

where (a) is hold by sustituting y_1, y_2, f_1, f_2,K_1,K_2 and Δ, (b) is obtained by applying the inequality: $\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}} D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right) \leqslant \Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)$, (c) is hold becasue $L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right) \leqslant 1$.

Similarly, we can prove $\Upsilon_1 \geqslant 0$ as follows:

$$\begin{aligned}\Upsilon_1 &= D_1^{-1}\left\{D_1 \Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}} D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right) L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) + 1 - L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)\right\} \\ &\overset{(a)}{\geqslant} D_1^{-1}\left\{D_1 \Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}} D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right) L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) + \Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)\left[1 - L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\right]\right\} \geqslant 0\end{aligned} \tag{2.53}$$

where the first inequality (a) holds true because: $1 \geqslant \Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)$.

By using the same approach, it can be easily proved that $\Upsilon_2 = 1$.

2.8.3 Proof of Corollary 2.2

The following is always true by substituting K_1, Δ inside Υ_1 with y_1 and f_1:

$$\begin{aligned}&1 - \left(\frac{D_h^2}{h^2} + 1\right)\Upsilon_1 = 1 - \left(\frac{D_h^2}{h^2} + 1\right)\left\{K_1 L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) + \Delta y_1^{\frac{2}{\beta_{\mathrm{LOS}}}}\right\} \\ &= 1 + \left(\frac{D_h^2}{h^2} + 1\right)\left\{y_1^{\frac{2}{\beta_{\mathrm{LOS}}}}\left[L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right) - 1\right] - f_1 L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\right\}\end{aligned}$$

In order to prove that the above expression is negative, it is sufficient to prove that:

$$L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right) - 1 \leqslant 0 \Rightarrow L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right) \leqslant 1$$

This is because $\left(\frac{D_h^2}{h^2} + 1\right) \geqslant 1$ and $f_1 L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) \geqslant 0$ and $y_1 \geqslant 0$. Since we know that $L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) \leqslant 1$ and

$\Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}}}\right) = {}_2F_1\left(1, \frac{2}{\beta_{\text{LOS}}}, 1+\frac{2}{\beta_{\text{LOS}}}, -\frac{1}{\gamma_{\text{th}}}\right) \leqslant 1$, which concludes the proof for $1-\left(\frac{D_h^2}{h^2}+1\right)\Upsilon_1 \leqslant 0$.

$1-\left(\frac{D_h^2}{h^2}+1\right)\Upsilon_2 \leqslant 0$ immediately follows from the fact that $\Upsilon_2 = 1$ and $\left(\frac{D_h^2}{h^2}+1\right) \geqslant 1$. Similarly, $1-\left(\frac{D_h^2}{h^2}+1\right)\Phi \leqslant 0$ follows because $\Gamma_{\text{NLOS}}\left(\gamma_{\text{th}}\right) = {}_2F_1\left(-\frac{2}{\beta_{\text{NLOS}}}, 1, 1-\frac{2}{\beta_{\text{NLOS}}}, -\gamma_{\text{th}}\right) \geqslant 1$, $L\left(\frac{\lambda_{\text{MT}}}{\lambda_{\text{BS}}}\right) \geqslant 0$, and $\left(\frac{D_h^2}{h^2}+1\right) \geqslant 1$.

2.8.4 Proof of Corollary 2.3

Denote $g(D_1) = \Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}}}\right) - D_1\Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}} D_1^{-\frac{\beta_{\text{LOS}}}{2}}}\right)$, this result can be proved by noting that: i) $\lim_{D_1\to 1} g(D_1) = 0$; ii) The first-order derivative of $g(D_1)$ is the following: $\frac{dg(D_1)}{D_1} = -\frac{1}{1+\frac{1}{\gamma_{\text{th}} D_1^{-\frac{\beta_{\text{LOS}}}{2}}}} <$ 0, which indicates that $g(D_1)$ is monotonically decreasing. These two properties imply that $g(D_1)$ is negative.

It is important to analyze $g(D_1)$ when D_1 is very large, i.e., $\lim_{D_1\to\infty} g(D_1)$. This can be done as follows. First, we rewrite $\bar{g}(D_1) = D_1\Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}} D_1^{-\frac{\beta_{\text{LOS}}}{2}}}\right)$ by using the Pfaff transformation ${}_2F_1(a,b,c,-z) = (1+z)^{-b}\,{}_2F_1\left(b, c-a, c, \frac{z}{z+1}\right)$ as follows:

$$\bar{g}(D_1) = D_1\Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}} D_1^{-\frac{\beta_{\text{LOS}}}{2}}}\right) = \frac{D_1}{\left(1+\frac{D_1^{\frac{\beta_{\text{LOS}}}{2}}}{\gamma_{\text{th}}}\right)^{\frac{2}{\beta_{\text{LOS}}}}}\,{}_2F_1\left(\frac{2}{\beta_{\text{LOS}}}, \frac{2}{\beta_{\text{LOS}}}, 1+\frac{2}{\beta_{\text{LOS}}}, \frac{1}{1+\frac{\gamma_{\text{th}}}{D_1^{\frac{\beta_{\text{LOS}}}{2}}}}\right)$$

Let us now examine both terms of this expression.

$$\lim_{D_1\to\infty} \frac{D_1}{\left(1+\frac{D_1^{\frac{\beta_{\text{LOS}}}{2}}}{\gamma_{\text{th}}}\right)^{\frac{2}{\beta_{\text{LOS}}}}} \overset{(a)}{=} \lim_{D_1\to\infty} \frac{dD_1/D_1}{d\left[\left(1+\frac{D_1^{\frac{\beta_{\text{LOS}}}{2}}}{\gamma_{\text{th}}}\right)^{\frac{2}{\beta_{\text{LOS}}}}\right]\Big/ D_1} = \lim_{D_1\to\infty} \frac{1}{\left(D_1^{-\frac{\beta_{\text{LOS}}}{2}}+\frac{1}{\gamma_{\text{th}}}\right)^{\frac{2}{\beta_{\text{LOS}}}-1}\frac{1}{\gamma_{\text{th}}}} = \gamma_{\text{th}}^{\frac{2}{\beta_{\text{LOS}}}}$$

$$\lim_{D_1\to\infty} {}_2F_1\left(\frac{2}{\beta_{\text{LOS}}}, \frac{2}{\beta_{\text{LOS}}}, 1+\frac{2}{\beta_{\text{LOS}}}, \frac{1}{1+\frac{\gamma_{\text{th}}}{D_1^{\frac{\beta_{\text{LOS}}}{2}}}}\right) = {}_2F_1\left(\frac{2}{\beta_{\text{LOS}}}, \frac{2}{\beta_{\text{LOS}}}, 1+\frac{2}{\beta_{\text{LOS}}}, 1\right)$$

where (a) is hold by applying the L'Hopital's rule.

Thus, in summary, the following is obtained:

$$\lim_{D_1\to\infty} g(D_1) = {}_2F_1\left(1, \frac{2}{\beta_{\text{LOS}}}, 1+\frac{2}{\beta_{\text{LOS}}}, -\frac{1}{\gamma_{\text{th}}}\right) - \gamma_{\text{th}}^{\frac{2}{\beta_{\text{LOS}}}}\,{}_2F_1\left(\frac{2}{\beta_{\text{LOS}}}, \frac{2}{\beta_{\text{LOS}}}, 1+\frac{2}{\beta_{\text{LOS}}}, 1\right) \leqslant 0 \quad (2.54)$$

2.8.5 Proof of Corollary 2.4

Based on the definition of Δ, it can be formulated as follows:

$$\Delta = K_2 L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)+1-L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)\overset{(a)}{=}1+L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[\frac{\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)-D_1\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right)}{D_1-1}\right]$$

where (a) is hold by sustituting K_2 with y_1, y_2, f_1, f_2 inside.

Thus, $\Delta \leqslant 1$ can be proved directly from the results obtained in Corollary 2.4, where $\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)-D_1\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right)\leqslant 0$.

$\Delta \geqslant 0$ can be proved as follows.

$$\Delta = 1-L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[\frac{D_1\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right)-\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)}{D_1-1}\right]\overset{(a)}{\geqslant}0$$

where (a) is hold because $L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\in[0,1]$ and $\frac{D_1\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right)-\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)}{D_1-1}\leqslant\frac{D_1\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)-\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)}{D_1-1}=\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)\leqslant 1$.

2.8.6 Proof of Proposition 2.3

In order to understand the essence of the proposed approximation for system optimization, let us assume $h>\tilde{h}$, and, then, $D_h^2=D^2\frac{h^2}{\tilde{h}^2}$, which implies $D_1=\frac{D^2}{\tilde{h}^2}+1$. Now that D_1 is independent of h, based on Corollary 2.1 and Corollary 2.4, it can be easily determined that all the following short-hands are independent of h: $0\leqslant\Upsilon_1\leqslant 1$, $\Upsilon_2=1$, $0\leqslant\Delta\leqslant 1$, and $\Phi=1+L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[\Gamma_{\mathrm{NLOS}}\left(\gamma_{\mathrm{th}}\right)-1\right]\geqslant 1$.

We denote the three terms of $\mathrm{P}_{\mathrm{cov}}^{\mathrm{UB}}$ in (2.27) as follows.

$$\mathrm{P}_{\mathrm{cov}}^{\mathrm{1a}}=\frac{\exp\left\{\pi\lambda_{\mathrm{BS}}h^2\right\}\exp\left(-\pi\lambda_{\mathrm{BS}}\left(D_h^2+h^2\right)\Upsilon_1\right)}{\Delta}=\frac{\exp\left(-\pi\lambda_{\mathrm{BS}}h^2\left[D_1\Upsilon_1-1\right]\right)}{\Delta}$$
$$\mathrm{P}_{\mathrm{cov}}^{\mathrm{1b}}=\frac{\exp\left\{\pi\lambda_{\mathrm{BS}}h^2\right\}\exp\left(-\pi\lambda_{\mathrm{BS}}\left(D_h^2+h^2\right)\Upsilon_2\right)}{\Delta}=\frac{\exp\left(-\pi\lambda_{\mathrm{BS}}h^2\left[D_1-1\right]\right)}{\Delta}$$
$$\mathrm{P}_{\mathrm{cov}}^{2}=\frac{\exp\left\{\pi\lambda_{\mathrm{BS}}h^2\right\}\exp\left(-\pi\lambda_{\mathrm{BS}}\left(D_h^2+h^2\right)\Phi\right)}{\Phi}=\frac{\exp\left(-\pi\lambda_{\mathrm{BS}}h^2\left[D_1\Phi-1\right]\right)}{\Phi}$$

It is realized that: 1) All the exponential terms are negative since the terms in the square brackets are all positive (or equal to zero); 2) $\mathrm{P}_{\mathrm{cov}}^{1\mathrm{a}} \geqslant \mathrm{P}_{\mathrm{cov}}^{1\mathrm{b}}$ since $0 \leqslant \Upsilon_1 \leqslant 1$; 3) $\mathrm{P}_{\mathrm{cov}}^{1\mathrm{b}} \geqslant \mathrm{P}_{\mathrm{cov}}^{2}$ since $1/\Delta \geqslant 1/\Psi$ and $\exp\left(-\pi\lambda_{\mathrm{BS}}h^2\left[D_1-1\right]\right) \geqslant \exp\left(-\pi\lambda_{\mathrm{BS}}h^2\left[D_1\Phi-1\right]\right)$, since $\Phi \geqslant 1$.

Thus, we conclude that $\mathrm{P}_{\mathrm{cov}}^{2}$ is the smallest terms among the three. In addition, it tends to be quite small for the typical setups that we are interested in and are not covered by the special cases discussed above. More precisely, the case studies of interest are those where the height of the BSs is a few meters and the LOS ball is a few tens of meters, as well as the cellular network is not too sparse but up to a regular/normal density. Under these operating conditions, we argue that $\mathrm{P}_{\mathrm{cov}}^{2}$ can usually be neglected as compared to the other two terms. So, we propose the following approximation:

$$\mathrm{P}_{\mathrm{cov}} \leqslant \overline{\mathrm{P}}_{\mathrm{cov}}^{\mathrm{UB}}\left(h > \tilde{h}\right) \approx \frac{\exp\left(-\pi\lambda_{\mathrm{BS}}h^2\left[D_1\Upsilon_1-1\right]\right)}{\Delta} - \frac{\exp\left(-\pi\lambda_{\mathrm{BS}}h^2\left[D_1-1\right]\right)}{\Delta}$$

Let us introduce the following notation: $\mathrm{P}_{\mathrm{cov}}^{1\mathrm{a}} = \frac{\widetilde{\mathrm{P}}_{\mathrm{cov}}^{1\mathrm{a}}}{\Delta} = \frac{\exp\left(-\pi\lambda_{\mathrm{BS}}h^2\left[D_1\Upsilon_1-1\right]\right)}{\Delta}$, $\mathrm{P}_{\mathrm{cov}}^{1\mathrm{b}} = \frac{\widetilde{\mathrm{P}}_{\mathrm{cov}}^{1\mathrm{b}}}{\Delta} = \frac{\exp\left(-\pi\lambda_{\mathrm{BS}}h^2\left[D_1-1\right]\right)}{\Delta}$, and $\mathrm{P}_{\mathrm{cov}}^{2} = \frac{\widetilde{\mathrm{P}}_{\mathrm{cov}}^{2}}{\Phi} = \frac{\exp\left(-\pi\lambda_{\mathrm{BS}}h^2\left[D_1\Psi-1\right]\right)}{\Phi}$.

More formally, the proposed approximation is applicable if the following inequality is satisfied:

$$\min\left\{\widetilde{\mathrm{P}}_{\mathrm{cov}}^{1\mathrm{a}}, \widetilde{\mathrm{P}}_{\mathrm{cov}}^{1\mathrm{b}}\right\} = \widetilde{\mathrm{P}}_{\mathrm{cov}}^{1\mathrm{b}} \gg \widetilde{\mathrm{P}}_{\mathrm{cov}}^{2} \Rightarrow \widetilde{\mathrm{P}}_{\mathrm{cov}}^{2} \leqslant \varepsilon\widetilde{\mathrm{P}}_{\mathrm{cov}}^{1\mathrm{b}}$$

where ε is a small non-zero positive constant. In practice, $\varepsilon = 0.1$ can be assumed to be a good choice, since all probabilities with "tilde" are less than one by definition (are negative exponential functions), and, thus, being ten times smaller means to be one order of magnitude less than a number that is less than one. Thus this is an acceptable condition for being negligible.

We note, in addition, that: 1) The denominators Δ and Φ have been ignored in order to get a more restrictive condition. In fact, we know that $1/\Delta \geqslant 1/\Phi$. If $\widetilde{\mathrm{P}}_{\mathrm{cov}}^{2} \leqslant \varepsilon\widetilde{\mathrm{P}}_{\mathrm{cov}}^{1\mathrm{b}}$, then, for sure, $\mathrm{P}_{\mathrm{cov}}^{2} \leqslant \varepsilon\mathrm{P}_{\mathrm{cov}}^{1\mathrm{b}}$; 2) It is better not to add Δ and Φ because it is true that, by definition $\widetilde{\mathrm{P}}_{\mathrm{cov}}^{2} \leqslant 1$ and $\widetilde{\mathrm{P}}_{\mathrm{cov}}^{1\mathrm{b}} \leqslant 1$, but is not necessarily true that $\mathrm{P}_{\mathrm{cov}}^{1\mathrm{b}} \leqslant 1$, since $\Delta \leqslant 1$. It is always true, on the other hand, that $\mathrm{P}_{\mathrm{cov}}^{2} \leqslant 1$, since $\Phi \geqslant 1$.

More precisely, the following will be hold:

$$\widetilde{\mathrm{P}}_{\mathrm{cov}}^{2} \leqslant \varepsilon\widetilde{\mathrm{P}}_{\mathrm{cov}}^{1\mathrm{b}} \Rightarrow \pi\lambda_{\mathrm{BS}}h^2 D_1\left[\Phi-1\right] \geqslant \ln\left(\frac{1}{\varepsilon}\right)$$

which implies that

$$\widetilde{\mathrm{P}}_{\mathrm{cov}}^{2} \leqslant \varepsilon\widetilde{\mathrm{P}}_{\mathrm{cov}}^{1\mathrm{b}} \Rightarrow \pi\lambda_{\mathrm{BS}}h^2\left(\frac{\hat{D}^2}{\tilde{h}^2}+1\right)L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[\Gamma_{\mathrm{NLOS}}\left(\gamma_{\mathrm{th}}\right)-1\right] \geqslant \ln\left(\frac{1}{\varepsilon}\right)$$

2.8.7 Proof of Case Study 1 in Section 2.3.4

Consider the following asmptotic results:

$$\lambda_{\mathrm{BS}} \to \infty \Rightarrow L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) = 1 - \left(1 + \frac{\lambda_{\mathrm{MT}}}{3.5\lambda_{\mathrm{BS}}}\right)^{-3.5} \to 0$$

$$\lambda_{\mathrm{BS}} \to \infty \Rightarrow \Delta\left(\lambda_{\mathrm{BS}}\right) = 1 + L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[\frac{D_1^{-1}\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right) - \Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}} D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right)}{1 - D_1^{-1}}\right] \to 1$$

$$\lambda_{\mathrm{BS}} \to \infty \Rightarrow \Upsilon_1\left(\lambda_{\mathrm{BS}}\right) = \Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}} D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right) L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) + D_1^{-1}\left[1 - L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)\right] \to \frac{1}{D_1}$$

$$\lambda_{\mathrm{BS}} \to \infty \Rightarrow \Phi\left(\lambda_{\mathrm{BS}}\right) = 1 - L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[1 - \Gamma_{\mathrm{NLOS}}\left(\gamma_{\mathrm{th}}\right)\right] \to 1$$

which implies that

$$\lambda_{\mathrm{BS}} \to \infty \Rightarrow \lambda_{\mathrm{BS}} L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) \to \lambda_{\mathrm{MT}}$$

$$\lambda_{\mathrm{BS}} \to \infty \Rightarrow \lambda_{\mathrm{BS}}\left[1 - D_1\Upsilon_1\left(\lambda_{\mathrm{BS}}\right)\right] = -D_1\lambda_{\mathrm{BS}} L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}} D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right) + \lambda_{\mathrm{BS}} L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)$$

$$\to -\lambda_{\mathrm{MT}} D_1 \Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}} D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right) + \lambda_{\mathrm{MT}}\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)$$

$$\lambda_{\mathrm{BS}} \to \infty \Rightarrow \lambda_{\mathrm{BS}}\left[1 - D_1\Phi\left(\lambda_{\mathrm{BS}}\right)\right] = \lambda_{\mathrm{BS}}\left[1 - D_1 - D_1 L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left[1 - \Gamma_{\mathrm{NLOS}}\left(\gamma_{\mathrm{th}}\right)\right]\right] \to -\infty$$

As a result, it can be obtained that

$$\lim_{\lambda_{\mathrm{BS}}\to\infty} \mathrm{P}_{\mathrm{cov}}^{\mathrm{UB}} = \lim_{\lambda_{\mathrm{BS}}\to\infty}\left(\frac{\exp\left(\pi\lambda_{\mathrm{BS}} h^2\left[1 - D_1\Upsilon_1\left(\lambda_{\mathrm{BS}}, h\right)\right]\right)}{\Delta\left(\lambda_{\mathrm{BS}}, h\right)} - \frac{\exp\left(\pi\lambda_{\mathrm{BS}} h^2\left[1 - D_1\right]\right)}{\Delta\left(\lambda_{\mathrm{BS}}, h\right)} + \frac{\exp\left(\pi\lambda_{\mathrm{BS}} h^2\left[1 - D_1\Phi\left(\lambda_{\mathrm{BS}}, h\right)\right]\right)}{\Phi\left(\lambda_{\mathrm{BS}}, h\right)}\right)$$

$$= \lim_{\lambda_{\mathrm{BS}}\to\infty}\left(\frac{\exp\left(\pi\lambda_{\mathrm{BS}} h^2\left[1 - D_1\Upsilon_1\left(\lambda_{\mathrm{BS}}, h\right)\right]\right)}{\Delta\left(\lambda_{\mathrm{BS}}, h\right)}\right) = \exp\left\{\pi h^2\lambda_{\mathrm{MT}}\left[\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right) - D_1\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}} D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right)\right]\right\}$$

which concludes the proof.

2.8.8 Proof of Case Study 3 in Section 2.3.4

In this case, it can be argued that in the very dense regime the following will be true:

- The serving BS will be in LOS.
- The interfering BSs that dominate the aggregate interference are in LOS.

which implies the following expression of coverage probability:

$$\mathrm{P_{cov}} = \int_0^\infty M_{I_{\mathrm{agg}}}\left(\frac{\gamma_{\mathrm{th}}x}{\Omega}|x\right) f_{L^{(0)}}(x)\,dx \approx \int_0^\infty M^{\mathrm{LOS}}_{I_{\mathrm{agg}}}\left(\frac{\gamma_{\mathrm{th}}x}{\Omega}|x\right) \Lambda^{(1)}_{\mathrm{LOS}}([0,x)) \exp\left(-\Lambda_{\mathrm{LOS}}([0,x))\right) dx$$

$$= \int_{k_{\mathrm{LOS}}h^{\beta_{\mathrm{LOS}}}}^{k_{\mathrm{LOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}} \exp\left\{\begin{array}{l} -\pi\lambda_{\mathrm{BS}}L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left(D_h^2+h^2\right)\Theta_{\mathrm{LOS}}\left(\frac{k_{\mathrm{LOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}}{\gamma_{\mathrm{th}}x}\right) \\ +\pi\lambda_{\mathrm{BS}}L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\left(\frac{x}{k_{\mathrm{LOS}}}\right)^{\frac{2}{\beta_{\mathrm{LOS}}}}\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right) \end{array}\right\}$$

$$\times \frac{2\pi\lambda_{\mathrm{BS}}}{\beta_{\mathrm{LOS}}}\left(\frac{1}{k_{\mathrm{LOS}}}\right)^{\frac{2}{\beta_{\mathrm{LOS}}}} x^{\frac{2}{\beta_{\mathrm{LOS}}}-1} \exp\left(-\pi\lambda_{\mathrm{BS}}\left(\left(\frac{x}{k_{\mathrm{LOS}}}\right)^{\frac{2}{\beta_{\mathrm{LOS}}}}-h^2\right)\right) dx$$

Thus, the following holds when $\lambda_{\mathrm{BS}} \to \infty$:

$$\lim_{\lambda_{\mathrm{BS}}\to\infty} \mathrm{P_{cov}} \overset{(a)}{=} \int_{k_{\mathrm{LOS}}h^{\beta_{\mathrm{LOS}}}}^{k_{\mathrm{LOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}} \exp\left\{-\pi\lambda_{\mathrm{MT}}\left(D_h^2+h^2\right)\Theta_{\mathrm{LOS}}\left(\frac{k_{\mathrm{LOS}}\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}}{\gamma_{\mathrm{th}}x}\right)+\pi\lambda_{\mathrm{MT}}\left(\frac{x}{k_{\mathrm{LOS}}}\right)^{\frac{2}{\beta_{\mathrm{LOS}}}}\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)\right\}$$

$$\times \frac{2\pi\lambda_{\mathrm{BS}}}{\beta_{\mathrm{LOS}}}\left(\frac{1}{k_{\mathrm{LOS}}}\right)^{\frac{2}{\beta_{\mathrm{LOS}}}} x^{\frac{2}{\beta_{\mathrm{LOS}}}-1} \exp\left(-\pi\lambda_{\mathrm{BS}}\left(\left(\frac{x}{k_{\mathrm{LOS}}}\right)^{\frac{2}{\beta_{\mathrm{LOS}}}}-h^2\right)\right) dx$$

$$= \int_0^{\sqrt{\pi}D_h} \exp\left\{-\pi\lambda_{\mathrm{MT}}\left(D_h^2+h^2\right)\Theta_{\mathrm{LOS}}\left(\frac{\left(D_h^2+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}}{\gamma_{\mathrm{th}}\left(\frac{\varsigma^2}{\pi}+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}}\right)+\pi\lambda_{\mathrm{MT}}\left(\frac{k_{\mathrm{LOS}}\left(\frac{\varsigma^2}{\pi}+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}}{k_{\mathrm{LOS}}}\right)^{\frac{2}{\beta_{\mathrm{LOS}}}}\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)\right\}$$

$$\times \left[2\lambda_{\mathrm{BS}}\varsigma\exp\left(-\lambda_{\mathrm{BS}}\varsigma^2\right)\right] d\varsigma \overset{(c)}{=} \exp\left\{\pi\lambda_{\mathrm{MT}}h^2\left[\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)-D_1\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right)\right]\right\}$$

where (a) is hold because $\lambda_{\mathrm{BS}} \to \infty \Rightarrow \lambda_{\mathrm{BS}}L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) \to \lambda_{\mathrm{MT}}$; (b) is hold by changing the variable: $\pi\left(\left(\frac{x}{k_{\mathrm{LOS}}}\right)^{\frac{2}{\beta_{\mathrm{LOS}}}}-h^2\right) = \varsigma^2 \Rightarrow x = k_{\mathrm{LOS}}\left(\frac{\varsigma^2}{\pi}+h^2\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}$; (c) is hold based on the property of function $\mathscr{T}(\varsigma) = 2\lambda_{\mathrm{BS}}\varsigma\exp\left(-\lambda_{\mathrm{BS}}\varsigma^2\right)$ when $\lambda_{\mathrm{BS}} \to \infty$, which is similar as Dirac's delta function: $\int \mathscr{H}(\varsigma)\mathscr{T}(\varsigma)\,d\varsigma = \mathscr{H}(0)$.

2.8.9 Proof of Corollary 2.5

Let us denote $x = D^2/\tilde{h}^2$, then the function can be re-written as follows: $F(x) = \frac{x+L(\lambda_{\rm MT}/\lambda_{\rm BS})g(x)}{\log\left(-\frac{x}{L(\lambda_{\rm MT}/\lambda_{\rm BS})g(x)}\right)}$, where $g(x) = \Theta_{\rm LOS}\left(\frac{1}{\gamma_{\rm th}}\right) - (x+1)\,\Theta_{\rm LOS}\left(\frac{(x+1)^{\frac{\beta_{\rm LOS}}{2}}}{\gamma_{\rm th}}\right) \leqslant 0$.

The derivative of $g(x)$, with respect to x, is as follows: $\frac{dg(x)}{dx} = -\frac{\gamma_{\rm th}}{(x+1)^{\frac{\beta_{\rm LOS}}{2}}+\gamma_{\rm th}} < 0$, The derivative of $F(x)$, with respect to x, is computed as follows:

$$\frac{dF(x)}{dx} = \frac{1}{\left[\log\left(-\frac{x}{L(\lambda_{\rm MT}/\lambda_{\rm BS})g(x)}\right)\right]^2}\left(\begin{array}{l}-1-\dfrac{L\left(\lambda_{\rm MT}/\lambda_{\rm BS}\right)g(x)}{x}+\log\left(-\dfrac{x}{L\left(\lambda_{\rm MT}/\lambda_{\rm BS}\right)g(x)}\right)\\ +\dfrac{x+L\left(\lambda_{\rm MT}/\lambda_{\rm BS}\right)g(x)\left[1+\log\left(-\frac{x}{L(\lambda_{\rm MT}/\lambda_{\rm BS})g(x)}\right)\right]}{g(x)}\dfrac{dg(x)}{dx}\end{array}\right)$$

Due to the fact that $g(x) < 0$, we denote $\tilde{g}(x) = -g(x) > 0$ in order to better formulate the function and test the signs. Then $\frac{dF(x)}{dx}$ becomes:

$$\frac{dF(x)}{dx} = \frac{1}{\left[\log\left(\frac{x}{L(\lambda_{\rm MT}/\lambda_{\rm BS})\tilde{g}(x)}\right)\right]^2}\left(\vartheta_1 + \vartheta_2\underbrace{\frac{dg(x)/x}{\tilde{g}(x)}}_{\leqslant 0}\right) \geqslant 0$$

which can be proved as follows.

$$\begin{aligned}\vartheta_1 &= -1+\frac{L\left(\lambda_{\rm MT}/\lambda_{\rm BS}\right)\tilde{g}(x)}{x}+\log\left(1+\frac{x-L\left(\lambda_{\rm MT}/\lambda_{\rm BS}\right)\tilde{g}(x)}{L\left(\lambda_{\rm MT}/\lambda_{\rm BS}\right)\tilde{g}(x)}\right)\\ &\overset{(a)}{>} -1+\frac{L\left(\lambda_{\rm MT}/\lambda_{\rm BS}\right)\tilde{g}(x)}{x}+\frac{\frac{x-L(\lambda_{\rm MT}/\lambda_{\rm BS})\tilde{g}(x)}{L(\lambda_{\rm MT}/\lambda_{\rm BS})\tilde{g}(x)}}{\frac{x-L(\lambda_{\rm MT}/\lambda_{\rm BS})\tilde{g}(x)}{L(\lambda_{\rm MT}/\lambda_{\rm BS})\tilde{g}(x)}+1} = -1+\frac{x}{x} = 0\end{aligned}$$

where (a) follows by applying the inequality $\log(1+x) \geqslant \frac{x}{x+1}$, $x > -1$.

$$\begin{aligned}\vartheta_2 &\overset{(a)}{<} L\left(\lambda_{\rm MT}/\lambda_{\rm BS}\right)\tilde{g}(x+1)\left[1+\frac{x-L\left(\lambda_{\rm MT}/\lambda_{\rm BS}\right)\tilde{g}(x+1)}{L\left(\lambda_{\rm MT}/\lambda_{\rm BS}\right)\tilde{g}(x+1)}\right]-x\\ &= L\left(\lambda_{\rm MT}/\lambda_{\rm BS}\right)\tilde{g}(x+1)\left[\frac{x}{L\left(\lambda_{\rm MT}/\lambda_{\rm BS}\right)\tilde{g}(x+1)}\right]-x = x-x = 0\end{aligned}$$

where (a) follows by applying the inequality $\log(1+x) < x$, $x > -1$.

Let us study in detail the extreme cases: i) $x \to 0$; and ii) $x \to \infty$. The case study $x \to 0$ cannot be examined, since we know that the upper-bound is not a good approximation in this case. In fact, the system performance depends on the NLOS term, as discussed in the Case Studies. Accordingly, we

consider only the case study $x \to \infty$.

$$\lim_{x\to\infty} F(x) = \lim_{x\to\infty} \frac{d\left[x + L\left(\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}}\right) g(x)\right]/dx}{d\log\left(-\frac{x}{L(\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}})g(x)}\right)\Big/dx} = \lim_{x\to\infty} \frac{1 - L\left(\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}}\right) \frac{\gamma_{\mathrm{th}}}{(x+1)^{\frac{\beta_{\mathrm{LOS}}}{2}} + \gamma_{\mathrm{th}}}}{\frac{1}{x} + \frac{\frac{\gamma_{\mathrm{th}}}{(x+1)^{\frac{\beta_{\mathrm{LOS}}}{2}} + \gamma_{\mathrm{th}}}}{\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right) - (x+1)\Theta_{\mathrm{LOS}}\left(\frac{(x+1)^{\frac{\beta_{\mathrm{LOS}}}{2}}}{\gamma_{\mathrm{th}}}\right)}} \overset{(a)}{=} \infty$$

where the equality in (a) comes from the following result:

$$\lim_{x\to\infty} \frac{\frac{\gamma_{\mathrm{th}}}{(x+1)^{\frac{\beta_{\mathrm{LOS}}}{2}} + \gamma_{\mathrm{th}}}}{\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right) - (x+1)\,\Theta_{\mathrm{LOS}}\left(\frac{(x+1)^{\frac{\beta_{\mathrm{LOS}}}{2}}}{\gamma_{\mathrm{th}}}\right)} = \lim_{x\to\infty} \frac{-\frac{(x+1)^{-1+\frac{\beta_{\mathrm{LOS}}}{2}} \beta_{\mathrm{LOS}} \gamma_{\mathrm{th}}}{2\left[(x+1)^{\frac{\beta_{\mathrm{LOS}}}{2}} + \gamma_{\mathrm{th}}\right]^2}}{-\frac{\gamma_{\mathrm{th}}}{(x+1)^{\frac{\beta_{\mathrm{LOS}}}{2}} + \gamma_{\mathrm{th}}}} = \lim_{x\to\infty} \frac{\beta_{\mathrm{LOS}}}{2\left[(x+1) + \frac{\gamma_{\mathrm{th}}}{(x+1)^{-1+\frac{\beta_{\mathrm{LOS}}}{2}}}\right]} = 0$$

which concludes the proof.

2.8.10 Proof of Corollary 2.6

For ease of writing, let us denote the variable $v = \tilde{h}^2$. So, we can focus on the following function: $\Xi(v) = vF(v)$, where $F(v) = \frac{D^2/v + L(\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}})g(v)}{\log\left(-\frac{D^2/v}{L\left(\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}}\right)g(v)}\right)}$, and $g(v) = \Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right) - \left(\frac{D^2}{v} + 1\right)\Theta_{\mathrm{LOS}}\left(\frac{\left(\frac{D^2}{v}+1\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}}{\gamma_{\mathrm{th}}}\right) \leqslant 0$.
Based on the previous proof, we know that $\frac{dF(v)}{dv} \leqslant 0$, and $\frac{dg(v)}{dv} \geqslant 0$. Now, we need to study the combined function $\Xi(v) = vF(v)$.

The derivative of $\Xi(v)$ can be computed by

$$\frac{d\Xi(v)}{dv} \overset{(a)}{=} \frac{\left[vL\left(\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}}\right)\tilde{g}(v)\left(1 + \log\left(\frac{D^2}{vL(\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}})\tilde{g}(v)}\right)\right) - D^2\right]\left[vdg(v)/dv - \tilde{g}(v)\right]}{v\tilde{g}(v)\left[\log\left(\frac{D^2}{vL(\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}})\tilde{g}(v)}\right)\right]^2}$$

where (a) is hold by denoting $g(v) = -\tilde{g}(v) > 0$ in order to better formulate the function and test the signs. Note that he denominator is always positive, so we only need to consider the numerator, which contains the following two terms.

The first term is determined to be negative as follows.

$$\begin{aligned} &vL\left(\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}}\right)\tilde{g}(v)\left(1 + \log\left(\frac{D^2}{vL\left(\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}}\right)\tilde{g}(v)}\right)\right) - D^2 \\ &\overset{(a)}{<} vL\left(\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}}\right)\tilde{g}(v)\left(1 + \frac{D^2 - vL\left(\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}}\right)\tilde{g}(v)}{vL\left(\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}}\right)\tilde{g}(v)}\right) - D^2 = D^2 - D^2 = 0 \end{aligned}$$

The second term is given by:

$$\zeta(v) = v\frac{dg(v)}{dv} - \tilde{g}(v) = \frac{D^2\gamma_{\text{th}}}{v\left[\left(\frac{D^2}{v}+1\right)^{\frac{\beta_{\text{LOS}}}{2}}+\gamma_{\text{th}}\right]} + \Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}}}\right) - \left(\frac{D^2}{v}+1\right)\Theta_{\text{LOS}}\left(\frac{\left(\frac{D^2}{v}+1\right)^{\frac{\beta_{\text{LOS}}}{2}}}{\gamma_{\text{th}}}\right)$$

which can be proved to be negative as well.

Firstly, We compute the derivative of $\zeta(v)$ as follows.

$$\frac{d\zeta(v)}{dv} = \frac{D^4\left(\frac{D^2}{v}+1\right)^{\frac{1}{2}(-2+\beta_{\text{LOS}})}\beta_{\text{LOS}}\gamma_{\text{th}}}{2v^3\left[\left(\frac{D^2}{v}+1\right)^{\frac{\beta_{\text{LOS}}}{2}}+\gamma_{\text{th}}\right]^2} > 0$$

which means the function will be monotonically increasing in $v \in [0,\infty)$.

Then we analyze the asymptotic results for $\lim_{v\to 0}\zeta(v) \leqslant 0$ as follows.

$$\lim_{v\to 0}\frac{D^2\gamma_{\text{th}}}{v\left[\left(\frac{D^2}{v}+1\right)^{\frac{\beta_{\text{LOS}}}{2}}+\gamma_{\text{th}}\right]} = \lim_{v\to 0}\frac{D^2\gamma_{\text{th}}}{v^{\frac{\beta_{\text{LOS}}}{2}}v^{1-\frac{\beta_{\text{LOS}}}{2}}\left(\frac{D^2}{v}+1\right)^{\frac{\beta_{\text{LOS}}}{2}}+v\gamma_{\text{th}}} = \frac{D^2\gamma_{\text{th}}}{\infty\times\left(D^2\right)^{\frac{\beta_{\text{LOS}}}{2}}} = 0$$

and

$$\begin{aligned}
&\lim_{v\to 0}\left\{-\left(\frac{D^2}{v}+1\right)\Theta_{\text{LOS}}\left(\frac{\left(\frac{D^2}{v}+1\right)^{\frac{\beta_{\text{LOS}}}{2}}}{\gamma_{\text{th}}}\right)\right\} \\
&\overset{(a)}{=}\lim_{v\to 0}\left\{-\left(\frac{D^2}{v}+1\right)\left[1+\frac{\left(\frac{D^2}{v}+1\right)^{\frac{\beta_{\text{LOS}}}{2}}}{\gamma_{\text{th}}}\right]^{-\frac{2}{\beta_{\text{LOS}}}}{}_2F_1\left(\frac{2}{\beta_{\text{LOS}}},\frac{2}{\beta_{\text{LOS}}},1+\frac{2}{\beta_{\text{LOS}}},\frac{\left(\frac{D^2}{v}+1\right)^{\frac{\beta_{\text{LOS}}}{2}}\Big/\gamma_{\text{th}}}{1+\left(\frac{D^2}{v}+1\right)^{\frac{\beta_{\text{LOS}}}{2}}\Big/\gamma_{\text{th}}}\right)\right\} \\
&= -\gamma_{\text{th}}^{\frac{2}{\beta_{\text{LOS}}}}{}_2F_1\left(\frac{2}{\beta_{\text{LOS}}},\frac{2}{\beta_{\text{LOS}}},1+\frac{2}{\beta_{\text{LOS}}},1\right) \leqslant 0
\end{aligned}$$

where (a) is hold based on the transformation: $(1+z)^{-b}{}_2F_1\left(b,c-a,c;\frac{z}{z+1}\right)$.

Thus, it can be proved that $\lim_{v\to 0}\zeta(v) = \Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}}}\right) - \gamma_{\text{th}}^{\frac{2}{\beta_{\text{LOS}}}}{}_2F_1\left(\frac{2}{\beta_{\text{LOS}}},\frac{2}{\beta_{\text{LOS}}},1+\frac{2}{\beta_{\text{LOS}}},1\right) \leqslant 0$, which follows from Corollary 2.3.

Then we analyze the asymptotic results for $\lim_{\upsilon\to\infty}\zeta(\upsilon)=0$ as follows.

$$\lim_{\upsilon\to\infty}\zeta(\upsilon)=\lim_{\upsilon\to\infty}\frac{D^2\gamma_{\text{th}}}{\upsilon\left[\left(\frac{D^2}{\upsilon}+1\right)^{\frac{\beta_{\text{LOS}}}{2}}+\gamma_{\text{th}}\right]}+\Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}}}\right)+\lim_{\upsilon\to\infty}\left\{-\left(\frac{D^2}{\upsilon}+1\right)\Theta_{\text{LOS}}\left(\frac{\left(\frac{D^2}{\upsilon}+1\right)^{\frac{\beta_{\text{LOS}}}{2}}}{\gamma_{\text{th}}}\right)\right\}$$
$$=\lim_{\upsilon\to\infty}\frac{D^2\gamma_{\text{th}}}{\upsilon\left[1+\gamma_{\text{th}}\right]}+\Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}}}\right)-\Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}}}\right)=0$$

In conclusion, there are: $\zeta(\upsilon)$ is monotonically increasing; $\lim_{\upsilon\to 0}\zeta(\upsilon)\leqslant 0$; and $\lim_{\upsilon\to\infty}\zeta(\upsilon)=0$. As a result, $\zeta(\upsilon)$ is a negative function, which implies that $\Xi(\upsilon)$ is a monotonically increasing since its first-order derivative is the product of two negative functions, and, thus, it is positive.

2.8.11 Proof of Theorem 2.2

This result can be proved from the upper-bound by setting $D_h=D$ as follows:

$$\mathrm{P}_{\text{cov}}^{\text{UB}}(h)=\frac{\exp\{\pi\lambda_{\text{BS}}h^2\}\exp\left(-\pi\lambda_{\text{BS}}\left(D^2+h^2\right)\Upsilon_1\right)}{\Delta(h)}-\frac{\exp\{\pi\lambda_{\text{BS}}h^2\}\exp\left(-\pi\lambda_{\text{BS}}\left(D^2+h^2\right)\Upsilon_2\right)}{\Delta(h)}$$
$$+\frac{\exp\{\pi\lambda_{\text{BS}}h^2\}\exp\left(-\pi\lambda_{\text{BS}}\left(D^2+h^2\right)\Phi\right)}{\Phi}$$
$$=\frac{\exp\left(-\pi\lambda_{\text{BS}}\left(D^2+h^2\right)\left(\Upsilon_1(h)-1\right)\right)-1}{\Delta(h)}\exp\left(-\pi\lambda_{\text{BS}}D^2\right)+\frac{\exp\left(-\pi\lambda_{\text{BS}}D^2\Phi\right)\exp\left(-\pi\lambda_{\text{BS}}h^2\left(\Phi-1\right)\right)}{\Phi}$$

where $\Upsilon_1(h)$, $\Upsilon_2=1$, and $\Delta(h)$ are defined in Corollary 1 and Corollary 4, which are dependent of the BS's height if $h<\tilde{h}$.

According to the definition, $\Phi=1-L\left(\frac{\lambda_{\text{MT}}}{\lambda_{\text{BS}}}\right)\left[1-\Gamma_{\text{NLOS}}\left(\gamma_{\text{th}}\right)\right]=1+L\left(\frac{\lambda_{\text{MT}}}{\lambda_{\text{BS}}}\right)\left[\Gamma_{\text{NLOS}}\left(\gamma_{\text{th}}\right)-1\right]\geqslant 1$, which is independent of h. Thus, $\exp\left(-\pi\lambda_{\text{BS}}h^2\left(\Phi-1\right)\right)$ can be determined to be monotonically decreasing with h, which implies that the last term is proved to monotonically decrease with h. In order to prove the theorem, we need to prove that:

$$f(h)=\frac{\exp\left(-\pi\lambda_{\text{BS}}\left(D^2+h^2\right)\left(\Upsilon_1(h)-1\right)\right)-1}{\Delta(h)}\exp\left(-\pi\lambda_{\text{BS}}D^2\right)$$

is decreasing with h. The derivative of $f(h)$ can be computed by

$$\frac{df(h)}{dh}=\frac{1}{\Delta^2(h)}\left(\begin{array}{l}\exp\{-\pi\lambda_{\text{BS}}D^2\}\frac{d\Delta(h)}{dh}-\exp\{\pi\lambda_{\text{BS}}\left[h^2-\left(D^2+h^2\right)\Upsilon_1(h)\right]\}\\ \times\left(\frac{d\Delta(h)}{dh}-\pi\lambda_{\text{BS}}\Delta(h)\left[2h\left(1-\Upsilon_1(h)\right)-\left(D^2+h^2\right)\frac{d\Upsilon_1(h)}{dh}\right]\right)\end{array}\right)$$
$$\overset{(a)}{=}\frac{2L\left(\frac{\lambda_{\text{MT}}}{\lambda_{\text{BS}}}\right)}{h\Delta^2(h)}\left[\frac{\gamma_{\text{th}}}{\left(\frac{D^2}{h^2}+1\right)^{\frac{\beta_{\text{LOS}}}{2}}+\gamma_{\text{th}}}+\frac{h^2m(h)}{D^2}\right]\left[\begin{array}{l}\exp\{-\pi\lambda_{\text{BS}}D^2\}+\exp\left\{\pi\lambda_{\text{BS}}L\left(\frac{\lambda_{\text{MT}}}{\lambda_{\text{BS}}}\right)h^2m(h)\right\}\\ \times\left(\pi\lambda_{\text{BS}}D^2+\pi\lambda_{\text{BS}}L\left(\frac{\lambda_{\text{MT}}}{\lambda_{\text{BS}}}\right)h^2m(h)-1\right)\end{array}\right]$$

where (a) is hold because of the following:

$$m(h) = \Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right) - \left(\frac{D^2}{h^2}+1\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}\left(\frac{D^2}{h^2}+1\right)^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right) \leqslant 0$$

$$\frac{d\Upsilon_1(h)}{dh} = \frac{2D^2}{\left(\frac{D^2}{h^2}+1\right)^2 h^3}\left[1 - \left(\frac{D^2}{h^2}+1\right)L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\frac{\gamma_{\mathrm{th}}}{\left(\frac{D^2}{h^2}+1\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}+\gamma_{\mathrm{th}}} - L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)m(h)\right]$$

$$\frac{d\Delta(h)}{dh} = \frac{2L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)}{h}\left[\frac{\gamma_{\mathrm{th}}}{\left(\frac{\hat{D}^2}{h^2}+1\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}+\gamma_{\mathrm{th}}} + \frac{h^2}{\hat{D}^2}m(h)\right]$$

We first analyze the following term:

$$\begin{aligned}
&\exp\left\{-\pi\lambda_{\mathrm{BS}}D^2\right\} + \exp\left\{\pi\lambda_{\mathrm{BS}}L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)h^2 m(h)\right\}\left(\pi\lambda_{\mathrm{BS}}D^2 + \pi\lambda_{\mathrm{BS}}L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)h^2 m(h) - 1\right)\\
&\overset{(a)}{>} \exp\left\{-\pi\lambda_{\mathrm{BS}}D^2\right\} + \exp\left\{\pi\lambda_{\mathrm{BS}}L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)h^2 m(h)\right\}\left(-\exp\left\{-\pi\lambda_{\mathrm{BS}}D^2 - \pi\lambda_{\mathrm{BS}}L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)h^2 m(h)\right\}\right)\\
&= \exp\left\{-\pi\lambda_{\mathrm{BS}}D^2\right\} - \exp\left\{-\pi\lambda_{\mathrm{BS}}D^2\right\} = 0
\end{aligned}$$

where (a) is true because $1-x<e^{-x} \Rightarrow x-1>-e^{-x}$.

Then, the proof would follow if we are able to prove the following:

$$\frac{\gamma_{\mathrm{th}}}{\left(\frac{D^2}{h^2}+1\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}+\gamma_{\mathrm{th}}} + \frac{h^2 m(h)}{D^2} = \frac{\gamma_{\mathrm{th}}}{\left(\frac{D^2}{h^2}+1\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}+\gamma_{\mathrm{th}}} + \frac{h^2}{D^2}\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right) - \left(1+\frac{h^2}{D^2}\right)\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}\left(\frac{D^2}{h^2}+1\right)^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right) \overset{(a)}{\leqslant} 0$$

where (a) is hold based on the same approach as in Appendix 2.8.10

2.8.12 Proof of Corollary 2.7

Denote $D_1 = \frac{D^2}{h^2}+1 \geqslant 1$, the optimal height of BSs $h_{\mathrm{Th1}}^{\mathrm{opt}}$ in (2.38) can be written as a function of λ_{BS}:

$$h_{\mathrm{Th1}}^{\mathrm{opt}}(\lambda_{\mathrm{BS}}) = \sqrt{\frac{\log\left(\frac{D_1-1}{\Upsilon_1(\lambda_{\mathrm{BS}})D_1-1}\right)}{\pi\lambda_{\mathrm{BS}}D_1(1-\Upsilon_1(\lambda_{\mathrm{BS}}))}}$$

As the equivalent method, we analyze the monotonicity of $h_{\mathrm{Th1}}^{\mathrm{opt}}(\lambda_{\mathrm{BS}})$ without the square root, which is given by

$$\frac{d}{d\lambda_{\mathrm{BS}}}\frac{\log\left(\frac{D_1-1}{\Upsilon_1(\lambda_{\mathrm{BS}})D_1-1}\right)}{\pi\lambda_{\mathrm{BS}}D_1(1-\Upsilon_1(\lambda_{\mathrm{BS}}))} = \frac{\psi(\lambda_{\mathrm{BS}})}{[\lambda_{\mathrm{BS}}(1-\Upsilon_1(\lambda_{\mathrm{BS}}))]^2}$$

where the denominator is always positive $[\lambda_{\mathrm{BS}}(1-\Upsilon_1(\lambda_{\mathrm{BS}}))]^2 \geqslant 0$, and the numerator $\psi(\lambda_{\mathrm{BS}})$ can be detrermined by:

$$\begin{aligned}
&\psi(\lambda_{\mathrm{BS}}) = (\Upsilon_1(\lambda_{\mathrm{BS}})-1)\log\left(\frac{D_1-1}{\Upsilon_1(\lambda_{\mathrm{BS}})D_1-1}\right) + \lambda_{\mathrm{BS}}\frac{d\Upsilon_1(\lambda_{\mathrm{BS}})}{d\lambda_{\mathrm{BS}}}\log\left(\frac{D_1-1}{\Upsilon_1(\lambda_{\mathrm{BS}})D_1-1}\right)\\
&+D_1\frac{\lambda_{\mathrm{BS}}(\Upsilon_1(\lambda_{\mathrm{BS}})-1)\left[d\Upsilon_1(\lambda_{\mathrm{BS}})/d\lambda_{\mathrm{BS}}\right]}{\Upsilon_1(\lambda_{\mathrm{BS}})D_1-1}\\
&\overset{(a)}{\leqslant}(\Upsilon_1(\lambda_{\mathrm{BS}})-1)\frac{D_1-1}{\Upsilon_1(\lambda_{\mathrm{BS}})D_1-1} + \lambda_{\mathrm{BS}}\frac{d\Upsilon_1(\lambda_{\mathrm{BS}})}{d\lambda_{\mathrm{BS}}}\frac{D_1-1}{\Upsilon_1(\lambda_{\mathrm{BS}})D_1-1} + D_1\frac{\lambda_{\mathrm{BS}}(\Upsilon_1(\lambda_{\mathrm{BS}})-1)}{\Upsilon_1(\lambda_{\mathrm{BS}})D_1-1}\frac{d\Upsilon_1(\lambda_{\mathrm{BS}})}{d\lambda_{\mathrm{BS}}}\\
&=(\Upsilon_1(\lambda_{\mathrm{BS}})-1)\frac{D_1-1}{\Upsilon_1(\lambda_{\mathrm{BS}})D_1-1} + \lambda_{\mathrm{BS}}\frac{d\Upsilon_1(\lambda_{\mathrm{BS}})}{d\lambda_{\mathrm{BS}}} \overset{(b)}{\leqslant} 0
\end{aligned}$$

where (a) is hold by applying $\frac{D_1-1}{\Upsilon_1(\lambda_{\mathrm{BS}})D_1-1} \geqslant 1$ and $\log\left(\frac{D_1-1}{\Upsilon_1(\lambda_{\mathrm{BS}})D_1-1}\right) \leqslant \frac{D_1-1}{\Upsilon_1(\lambda_{\mathrm{BS}})D_1-1}$; (b) is hold since $\Upsilon_1(\lambda_{\mathrm{BS}})-1 \leqslant 0$, $\frac{D_1-1}{\Upsilon_1(\lambda_{\mathrm{BS}})D_1-1} \geqslant 1 \geqslant 0$, and $\frac{d\Upsilon_1(\lambda_{\mathrm{BS}})}{d\lambda_{\mathrm{BS}}} \leqslant 0$ which is proved as follows:

$$\frac{d\Upsilon_1(\lambda_{\mathrm{BS}})}{d\lambda_{\mathrm{BS}}} = \left(\frac{D^2}{\tilde{h}^2}+1\right)^{-1}\left\{-\frac{dL\left(\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}}\right)}{d\lambda_{\mathrm{BS}}}g\left(\frac{D^2}{\tilde{h}^2}\right)\right\} \leqslant 0$$

where the inequalily follows from $g\left(\frac{D^2}{\tilde{h}^2}\right) = \Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right) - \left(\frac{D^2}{\tilde{h}^2}+1\right)\Theta_{\mathrm{LOS}}\left(\frac{\left(\frac{D^2}{\tilde{h}^2}+1\right)^{\frac{\beta_{\mathrm{LOS}}}{2}}}{\gamma_{\mathrm{th}}}\right) \leqslant 0$, and $\frac{dL(\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}})}{d\lambda_{\mathrm{BS}}} \leqslant 0$.

To summarize, it is determined that $\frac{d}{d\lambda_{\mathrm{BS}}}\frac{\log\left(\frac{D_1-1}{\Upsilon_1(\lambda_{\mathrm{BS}})D_1-1}\right)}{\pi\lambda_{\mathrm{BS}}D_1(1-\Upsilon_1(\lambda_{\mathrm{BS}}))} \leqslant 0$, which concludes the proof.

2.8.13 Proof of Corollary 2.8

Denote $D_1 = \frac{D^2}{\tilde{h}^2}+1 \geqslant 1$, the optimal height of BSs $h_{\mathrm{Th1}}^{\mathrm{opt}}$ in (2.38) can be written as a function of D_1:

$$h_{\mathrm{Th1}}^{\mathrm{opt}}(D_1) = \sqrt{\frac{\log\left(\frac{D_1-1}{\Upsilon_1(D_1)D_1-1}\right)}{\pi\lambda_{\mathrm{BS}}D_1(1-\Upsilon_1(D_1))}}$$

Similarly, we analyze the monotonicity of $h_{\mathrm{Th1}}^{\mathrm{opt}}(D_1)$ without the square root, which is given by

$$\begin{aligned}
&\frac{d}{dD_1}\frac{\log\left(\frac{D_1-1}{\Upsilon_1(D_1)D_1-1}\right)}{\pi\lambda_{\mathrm{BS}}D_1(1-\Upsilon_1(D_1))}\\
&=\frac{\frac{D_1(1-\Upsilon_1(D_1))}{D_1-1}\frac{-\frac{d\Upsilon_1(D_1)}{dD_1}D_1(D_1-1)-(1-\Upsilon_1(D_1))}{\Upsilon_1(D_1)D_1-1} - \log\left(\frac{D_1-1}{\Upsilon_1(D_1)D_1-1}\right)\left(1-\Upsilon_1(D_1)-\frac{d\Upsilon_1(D_1)}{dD_1}D_1\right)}{[D_1(1-\Upsilon_1(D_1))]^2}\\
&\overset{(a)}{<}\frac{\frac{D_1(1-\Upsilon_1(D_1))}{D_1-1}\frac{-\frac{d\Upsilon_1(D_1)}{dD_1}D_1(D_1-1)-(1-\Upsilon_1(D_1))}{\Upsilon_1(D_1)D_1-1} + \frac{(\Upsilon_1(D_1)-1)D_1}{D_1-1}\left(1-\Upsilon_1(D_1)-\frac{d\Upsilon_1(D_1)}{dD_1}D_1\right)}{[D_1(1-\Upsilon_1(D_1))]^2}\\
&=\frac{-D_1[1-\Upsilon_1(D_1)]^2\left[\frac{d\Upsilon_1(D_1)}{dD_1}D_1^2+\Upsilon_1(D_1)D_1\right]}{(D_1-1)(\Upsilon_1(D_1)D_1-1)[D_1(1-\Upsilon_1(D_1))]^2} \overset{(b)}{\leqslant} 0
\end{aligned}$$

where the inequality in (a) holds because $1-\Upsilon_1(D_1)-\frac{d\Upsilon_1(D_1)}{dD_1}D_1 = 1-L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\frac{1}{\frac{\beta_{\mathrm{LOS}}}{2}} > 0$ and $-\log\left(\frac{D_1-1}{\Upsilon_1(D_1)D_1-1}\right) = -\log\left(1-\frac{(\Upsilon_1(D_1)-1)D_1}{\Upsilon_1(D_1)D_1-1}\right) < \frac{(\Upsilon_1(D_1)-1)D_1}{D_1-1}$ by using $-\log(1-x) < \frac{x}{1-x}$, $x<1 \wedge x \neq 0$; (b) is hold since $\frac{d\Upsilon_1(D_1)}{dD_1}D_1^2 + \Upsilon_1(D_1)D_1 = \frac{D_1 L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\gamma_{\mathrm{th}}}{\frac{\beta_{\mathrm{LOS}}}{2}} > 0$, $D_1-1>0$ and $\Upsilon_1(D_1)D_1-1>0$.

The proof follows from the fact that $D_1 = \frac{D^2}{h^2}+1 \geqslant 1$, and, thus, the trend as a function of D and D_1 are the same.

2.8.14 Proof of Corollary 2.10

Denote $D_1 = \frac{D^2}{h^2}+1 \geqslant 1$, the optimal height of BSs $h_{\mathrm{Th1}}^{\mathrm{opt}}$ in (2.38) can be written as a function of λ_{MT}:

$$h_{\mathrm{Th1}}^{\mathrm{opt}}(\lambda_{\mathrm{MT}}) = \sqrt{\frac{\log\left(\frac{D_1-1}{\Upsilon_1(\lambda_{\mathrm{MT}})D_1-1}\right)}{\pi\lambda_{\mathrm{BS}}D_1(1-\Upsilon_1(\lambda_{\mathrm{MT}}))}}$$

Similarly, we analyze the monotonicity of $h_{\mathrm{Th1}}^{\mathrm{opt}}(\lambda_{\mathrm{MT}})$ without the square root, which is given by

$$\begin{aligned}\frac{d}{d\lambda_{\mathrm{MT}}}\frac{\log\left(\frac{D_1-1}{\Upsilon_1(\lambda_{\mathrm{MT}})D_1-1}\right)}{1-\Upsilon_1(\lambda_{\mathrm{MT}})} &= -\frac{\frac{d\Upsilon_1(\lambda_{\mathrm{MT}})}{d\lambda_{\mathrm{MT}}}D_1}{(\Upsilon_1(\lambda_{\mathrm{MT}})D_1-1)(1-\Upsilon_1(\lambda_{\mathrm{MT}}))} + \frac{\frac{d\Upsilon_1(\lambda_{\mathrm{MT}})}{d\lambda_{\mathrm{MT}}}\log\left(\frac{D_1-1}{\Upsilon_1(\lambda_{\mathrm{MT}})D_1-1}\right)}{(1-\Upsilon_1(\lambda_{\mathrm{MT}}))^2} \\ &\overset{(a)}{\leqslant} -\frac{\frac{d\Upsilon_1(\lambda_{\mathrm{MT}})}{d\lambda_{\mathrm{MT}}}D_1}{(\Upsilon_1(\lambda_{\mathrm{MT}})D_1-1)(1-\Upsilon_1(\lambda_{\mathrm{MT}}))} + \frac{\frac{d\Upsilon_1(\lambda_{\mathrm{MT}})}{d\lambda_{\mathrm{MT}}}\left[\frac{(D_1-1)}{(\Upsilon_1(\lambda_{\mathrm{MT}})D_1-1)}-1\right]}{(1-\Upsilon_1(\lambda_{\mathrm{MT}}))^2} = 0\end{aligned}$$

where the inequality in (a) follows from $\log\left(\frac{D_1-1}{\Upsilon_1(\lambda_{\mathrm{MT}})D_1-1}\right) \leqslant \frac{D_1-1}{\Upsilon_1(\lambda_{\mathrm{MT}})D_1-1}-1$.

This concludes the proof.

2.8.15 Proof of Corollary 2.12

This statement can be proved as follows. If $h<\tilde{h}$, we have the following:

$$R_{\mathrm{cell}}^{\mathrm{opt}}(h) = \sqrt{\frac{\left(D^2+h^2\right)(1-\Upsilon_1(h))}{\log\left[\frac{D^2}{(D^2+h^2)\Upsilon_1(h)-h^2}\right]}}$$

where $\Upsilon_1(h)$ and its derivative $\frac{d\Upsilon_1(h)}{dh}$ are given by:

$$\Upsilon_1(h) = \Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}\left(\frac{D^2}{h^2}+1\right)^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right) + \left(\frac{D^2}{h^2}+1\right)^{-1}\left[1-\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)\right]$$

and

$$\frac{d\Upsilon_1(h)}{dh} = \frac{2D^2}{\left(\frac{D^2}{h^2}+1\right)^2 h^3}\left\{1-\Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}}}\right)-\left(\frac{D^2}{h^2}+1\right)\left[\frac{\gamma_{\text{th}}}{\left(\frac{D^2}{h^2}+1\right)^{\frac{\beta_{\text{LOS}}}{2}}+\gamma_{\text{th}}}-\Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}}\left(\frac{D^2}{h^2}+1\right)^{-\frac{\beta_{\text{LOS}}}{2}}}\right)\right]\right\}$$

Based on the same approach, we analyze the monotonicity of $R_{\text{cell}}^{\text{opt}}(h)$ without the square root, which is given by

$$\frac{d}{dh}\frac{\left(D^2+h^2\right)\left(1-\Upsilon_1(h)\right)}{\log\left[\frac{D^2}{(D^2+h^2)\Upsilon_1(h)-h^2}\right]} = \frac{\rho_2\rho_3}{\left(\log\left[\frac{D^2}{(D^2+h^2)\Upsilon_1(h)-h^2}\right]\right)^2\rho_1}$$

where the following short-hands are defined:

$$\rho_1 = -\left[\left(D^2+h^2\right)\Upsilon_1(h)-h^2\right]$$

$$\rho_2 = -D^2-h^2-h^2\log\left[\frac{D^2}{\left(\hat{D}^2+h^2\right)\Upsilon_1(h)-h^2}\right]+\left(D^2+h^2\right)\left(1+\log\left[\frac{D^2}{\left(D^2+h^2\right)\Upsilon_1(h)-h^2}\right]\right)\Upsilon_1(h)$$

$$\rho_3 = 2h\left[\Upsilon_1(h)-1\right]+\left(D^2+h^2\right)\frac{d\Upsilon_1(h)}{dh}$$

The signs of these three terms are analyzed as follows individually.

$$\rho_1 = -\left[\left(D^2+h^2\right)\Upsilon_1(h)-h^2\right] = h^2\left[1-\left(\frac{D^2}{h^2}+1\right)\Upsilon_1(h)\right] \leqslant 0$$

where the inequality follows from Corollary 2.2.

$$\begin{aligned}\rho_2 &= -D^2-h^2-h^2\log\left[\frac{D^2}{\left(D^2+h^2\right)\Upsilon_1(h)-h^2}\right]+\left(D^2+h^2\right)\left(1+\log\left[\frac{D^2}{\left(D^2+h^2\right)\Upsilon_1(h)-h^2}\right]\right)\Upsilon_1(h)\\ &= \left(D^2+h^2\right)\left[\Upsilon_1(h)-1\right]+\left[\left(D^2+h^2\right)\Upsilon_1(h)-h^2\right]\log\left[\frac{D^2}{\left(D^2+h^2\right)\Upsilon_1(h)-h^2}\right]\\ &\overset{(a)}{<} \left(D^2+h^2\right)\left[\Upsilon_1(h)-1\right]+\left[\left(D^2+h^2\right)\Upsilon_1(h)-h^2\right]\left(\frac{D^2}{\left(D^2+h^2\right)\Upsilon_1(h)-h^2}-1\right)\\ &= \left(D^2+h^2\right)\left[\Upsilon_1(h)-1+1-\Upsilon_1(h)\right] = 0\end{aligned}$$

where (a) holds true from the inequality: $\log(x) < x-1$.

$$\begin{aligned}\rho_3 &= 2h\left[\Upsilon_1(h)-1\right]+\left(D^2+h^2\right)\frac{d\Upsilon_1(h)}{dh}\\ &\overset{(a)}{=} 2h\left\{\left(1+\frac{D^2}{h^2}\right)\Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}}\left(\frac{D^2}{h^2}+1\right)^{-\frac{\beta_{\text{LOS}}}{2}}}\right)-\Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}}}\right)-\frac{D^2}{h^2}\frac{\gamma_{\text{th}}}{\left(\frac{D^2}{h^2}+1\right)^{\frac{\beta_{\text{LOS}}}{2}}+\gamma_{\text{th}}}\right\}\overset{(b)}{\geqslant} 0\end{aligned}$$

where the equality (a) follows by substituting $\Upsilon_1(h)$ and $\frac{d\Upsilon_1(h)}{dh}$, the inequality (b) can be proved as follows.

Denote $D_1 = 1 + \frac{D^2}{h^2}$, the term inside parenthesis can be written as:

$$\phi(D_1) = D_1 \Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}} D_1^{-\frac{\beta_{\text{LOS}}}{2}}}\right) - \Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}}}\right) - (D_1 - 1)\frac{\gamma_{\text{th}}}{D_1^{\frac{\beta_{\text{LOS}}}{2}} + \gamma_{\text{th}}}$$

The derivative of $\phi(D_1)$ can be computed as follows:

$$\frac{d\phi(D_1)}{dD_1} = \frac{(D_1 - 1) D_1^{-1+\frac{\beta_{\text{LOS}}}{2}} \beta_{\text{LOS}} \gamma_{\text{th}}}{2\left(D_1^{\frac{\beta_{\text{LOS}}}{2}} + \gamma_{\text{th}}\right)^2} > 0$$

which implies that the function is monotonically increasing with D_1, and $\phi(D_1)$ is proved to be positive since $\lim_{D_1 \to 1} \phi(D_1) = 0$.

To summarize, it is determined that $\frac{d}{dh}\frac{(D^2+h^2)(1-\Upsilon_1(h))}{\log\left[\frac{D^2}{(D^2+h^2)\Upsilon_1(h)-h^2}\right]} = \frac{\rho_2\rho_3}{\left(\log\left[\frac{D^2}{(D^2+h^2)\Upsilon_1(h)-h^2}\right]\right)^2\rho_1} \geqslant 0$, which concludes the proof.

2.8.16 Proof of Corollary 2.13

This statement can be proved as follows. Firstly, note that we can directly consider $D_h = D\max\left\{\frac{h}{h}, 1\right\}$, since the height of the BSs and of the blockages are fixed. As the equivalent approach, we analyze $R_{\text{cell}}^{\text{opt}}$ as a function of $D_1 = \frac{D_h^2}{h^2} + 1$, which is given by

$$R_{\text{cell}}^{\text{opt}}(D_1) = \sqrt{\frac{h^2 D_1 (1 - \Upsilon_1(D_1))}{\log\left[\frac{h^2 D_1 - 1}{h^2 D_1 \Upsilon_1(D_1) - h^2}\right]}} = h\sqrt{\frac{D_1(1-\Upsilon_1(D_1))}{\log\left[\frac{D_1-1}{D_1\Upsilon_1(D_1)-1}\right]}}$$

where

$$\Upsilon_1(D_1) = \Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}} D_1^{-\frac{\beta_{\text{LOS}}}{2}}}\right) + D_1^{-1}\left[1 - \Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}}}\right)\right]$$

The function $\Upsilon_1(D_1)$ is verified to be monotonically decreasing as D_1, since both $\Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}} D_1^{-\frac{\beta_{\text{LOS}}}{2}}}\right)$ and $D_1^{-1}\left[1 - \Theta_{\text{LOS}}\left(\frac{1}{\gamma_{\text{th}}}\right)\right]$ are decreasing as D_1. This implies that $\frac{d\Upsilon_1(D_1)}{D_1} \leqslant 0$.

Similarly, the monotonicity of $R_{\text{cell}}^{\text{opt}}(D_1)$ can be investigated by only considering the term inside squre root, which is

$$\frac{d}{dD_1}\frac{D_1(1-\Upsilon_1(D_1))}{\log\left[\frac{D_1-1}{D_1\Upsilon_1(D_1)-1}\right]} = \frac{(1-\Upsilon_1(D_1)) - D_1\frac{d\Upsilon_1(D_1)}{D_1}}{\log\left[\frac{D_1-1}{D_1\Upsilon_1(D_1)-1}\right]} + \frac{D_1(1-\Upsilon_1(D_1))\left(\frac{(D_1-1)\left[\Upsilon_1(D_1)+D_1\frac{d\Upsilon_1(D_1)}{D_1}\right]}{D_1\Upsilon_1(D_1)-1} - 1\right)}{(D_1-1)\left(\log\left[\frac{D_1-1}{D_1\Upsilon_1(D_1)-1}\right]\right)^2}$$

The first term is determined to be positive, $\frac{(1-\Upsilon_1(D_1))-D_1\frac{d\Upsilon_1(D_1)}{D_1}}{\log\left[\frac{D_1-1}{D_1\Upsilon_1(D_1)-1}\right]} \geqslant 0$. It can be proved as follows: $\frac{d\Upsilon_1(D_1)}{D_1} \leqslant 0$, $0<\Upsilon_1(D_1)<1$, which indicate that the numerator is positive. In addition, $\frac{D_1-1}{D_1\Upsilon_1(D_1)-1}>1$ directly results in that the denominator is positive.

The second term is determined to be positive as well, which is illustrated by analyzing its numerator as follows since the denominator is always positive:

$$D_1(1-\Upsilon_1(D_1))\left(\frac{(D_1-1)\left[\Upsilon_1(D_1)+D_1\frac{d\Upsilon_1(D_1)}{D_1}\right]}{D_1\Upsilon_1(D_1)-1}-1\right)=\underbrace{\frac{D_1-1}{D_1\Upsilon_1(D_1)-1}}_{>1}\frac{\omega(D_1)}{D_1-1}\geqslant 0$$

where the inequality above is hold since $\omega(D_1)$, which is proved as follows.

$$\omega(D_1)=1-\Upsilon_1(D_1)+\left(D_1^2-D_1\right)\frac{d\Upsilon_1(D_1)}{D_1}=\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}}\right)-D_1\Theta_{\mathrm{LOS}}\left(\frac{1}{\gamma_{\mathrm{th}}D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}\right)+(D_1-1)\frac{1}{1+\frac{D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}{\gamma_{\mathrm{th}}}}$$

Based on the definition of D_1, i.e., $D_1\in[1,\infty)$, there is $\lim_{D_1\to 1}\omega(D_1)=0$, and its derivative can be determined by:

$$\frac{d\omega(D_1)}{dD_1}=\underbrace{\frac{1}{1+\frac{D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}{\gamma_{\mathrm{th}}}}-\frac{1}{1+\frac{D_1^{\frac{\beta_{\mathrm{LOS}}}{2}}}{\gamma_{\mathrm{th}}}}}_{>0}+\frac{(D_1-1)D_1^{-1-\frac{\beta_{\mathrm{LOS}}}{2}}\beta_{\mathrm{LOS}}}{2\gamma_{\mathrm{th}}\left(1+\frac{1}{1+\frac{D_1^{-\frac{\beta_{\mathrm{LOS}}}{2}}}{\gamma_{\mathrm{th}}}}\right)^2}>0$$

As a result, the derivative $\frac{d}{dD_1}\frac{D_1(1-\Upsilon_1(D_1))}{\log\left[\frac{D_1-1}{D_1\Upsilon_1(D_1)-1}\right]}\geqslant 0$ and the corollary is proved.

3 Resource Allocation in Multi-Tenants Radio Access Networks: Combining Stochastic Geometry with Generalized Benders Decomposition

Chapter 3. Resource Allocation in Multi-Tenants Radio Access Networks: Combining Stochastic Geometry with Generalized Benders Decomposition

Network virtualization and softwarization techniques are key-enablers of the novel Network Slicing concept behind 5G. Network slicing introduces new business models such as allowing telecom providers to lease a virtualized slice of their infrastructure to tenants such as industry verticals, e.g. automotive, ehealth, factories, etc. However, this new paradigm poses a major challenge when applied to Radio Access Networks (RAN): How to maximize the overall system performance while maintaining the diverse service level agreements (SLAs) requested by the infrastructure tenants?

In this chaper, based on the concept of stochastic geometry, we propose a new analytical framework to model realistic RAN that leverages the business opportunities provided by network slicing. We design the radio resource allocation strategy through a mixed integer nonlinear programming (MINLP) optimization to best allocate transmit power and bandwidth (i.e., a slice of radio access resources) to the users of each infrastructure tenant by considering the per slice SLA guaranteed average rate condition. Under the linear separable property of the problem, we apply the efficient Generalized Benders Decomposition (GBD) where the global convergence is achieved. Numerical results are illustrated to validate the proposed solution, and the performance is compared against brute-force search and Greedy algorithm.

3.1 Introduction

Advanced heterogeneous service requirements from vertical industries necessitate a novel design of 5th generation (5G) mobile networks, in order to increase the revenues of telecom providers and network operators. To address this challenge, the introduction of new technologies, such as network programmability and virtualization, is required. Network programmability delivers the benefits of automation and reactiveness of software modules, which allows one to (re)configure mobile networks dynamically while in operation. Network virtualization overcomes the limitations of monolithic network infrastructure, by invoking the concept of "network function" and providing flexibility in composing, placing and managing these functions. In the chapter, the novel definition of network slicing [35] encompasses such new requirements and constitutes an enabler for potential economical benefits. New verticle industies, such as automotive, e-health and factories, etc., are entering into the telecom market and are disrupting the traditional business models of telecom operators. They are forcing infrastructure providers to open their network to tenants, a solution that provides incentives for monetizing the availability of isolated and secure (virtualized) network slices [36], [37].

This new disruptive concept of network slicing, has gained extensive attentions from both academia and industrial communities. Its realization, on the other hand, requires the solution of a number of technical challenges that, for the time being, are not completely addressed and solved either in academia or industry [38]. In the future, a increasing demand for end-to-end network slices, which involve multiple advanced service level agreements (SLAs) comprising different key performance indicators (KPIs) [39], such as throughput, latency and reliability, is conceived by telecom providers. Having different SLAs throughout slices, however, require appropriate automated admission control and resource allocation protocols for designing efficient network management systems [40]. The resource management of the radio access network (RAN), in particular, is one of the most challenging prospects that need to be clarified, and many challenges need to be solved for efficiently slicing the RAN [41]. The available resources of the RAN air interface can be sliced at multiple levels: time, frequency and power domains. Slicing the network at such levels, on the other hand, necessity the cellular network topology, other-cell interference and the radio channel conditions experienced by the users of each tenant [42]. To solve this never addressed and challenging issue, we leverage the theory of stochastic geometry and spatial point process to model the network topology based on the tractable analytical formulations [43]. To the best of our knowledge, this is new work in designing the automated RAN slicing admission control and resource allocation scheme that provides rate guarantees in RAN, where the topology of cellular system and other-cell interference are taken into account.

The main research contributions of this chapter contain the following: i) A new analytical framework of spectral efficiency, which explicitly accounts for the interplay among the transmit power of the cellular base stations (BSs), the available spectrum and the deployment densities of BSs and mobile terminals (MTs) of each tenant, is proposed with aid of stochastic geometry. ii) A deep study of the analytical properties of the newly proposed utility function for slicing the RAN is accomplished. iii) A designation of RAN admission control and resource management optimization problem are formulated and an efficient algorithm, relying on the Generalized Benders Decomposition (GBD), is implemented, analyzed and compared against brute-force and greedy algorithm.

The remainder of the present chapter is organized as follows. In Section 3.2, the system model is summarized including the formulation, analysis of the utility function and the optimization problem. In Section 3.3, the algorithm of GBD with its subproblems are illustrated. In Section 3.4, numerical illustrations are shown to validate the applied algorithms against the brute-force search and the Greedy

Table 3.1 – Summary of main symbols and functions used throughout the chapter.

Symbol/Function	Definition
$\mathbb{E}\{\cdot\}$, $\Pr\{\cdot\}$	Expectation operator, probability measure
$i \in \mathscr{I}$, $\mathscr{I} = \{1,2,...,\mathrm{T}\}$, X_i	Generic tenant, set of available tenants, X of a generic tenant i
λ_{BS}, λ_{MT}	Density of base stations, mobile terminals
Ψ_{BS}, Ψ_{MT}, $\Psi_{\mathrm{BS}}^{(I)}$	PPP of base stations, mobile terminals, interfering base stations
$\mathrm{BS}^{(0)}$, $\mathrm{BS}^{(\eta)}$, $\mathrm{BS}^{(n)}$, $\mathrm{MT}^{(0)}$	Serving, interfering, generic base station, typical mobile terminal
N_{MT}, $\bar{N}_{\mathrm{MT}}$	Number of MTs and number of MTs without $\mathrm{MT}^{(0)}$ in the typical cell
P, B, P_{tot}, B_{tot}	Transmit power, bandwidth, total transmit power and total bandwidth
N_0, β, τ	Noise power spectral density, binary variable, minimum required rate
$r^{(n)}$, $h^{(n)}$	Distance, fading power gain of a generic link
$l\left(r^{(n)}\right)$, $L^{(n)}$, $L^{(0)}$	Path-loss, shorthand of path-loss, path-loss of intended link
k, $\alpha > 0$	Path-loss constant, slope (exponent)
UBD, LBD, ϵ	Upper bound, lower bound, convergence tolerence of GBD
γ_I, γ_{A}	Reliability threshold for decoding, cell association
$L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right) = 1 - \left(1 + \frac{\lambda_{\mathrm{MT}}}{3.5\lambda_{\mathrm{BS}}}\right)^{-3.5}$	Probability that a base station is in transmission mode
$f(\cdot)$, $\phi(\cdot)$, β_0, θ, $\mathbf{G}(\cdot)$	Objective functions and constraints vector
$\mathbf{1}(\cdot)$, ${}_2F_1(\cdot,\cdot,\cdot,\cdot)$, $\log(\cdot)$	Indicator function, Gauss hypergeometric function, logarithm function
$\boldsymbol{\mu}$, $\boldsymbol{\rho}$	Lagrange multipliers of feasible and infeasible primal problem
$L^*\left(\beta_i, \boldsymbol{\mu}^{(k_1)}\right)$, $L_*\left(\beta_i, \boldsymbol{\rho}^{(k_2)}\right)$	Support functions of the master problem
SIR, $\overline{\mathrm{SNR}}$	Signal-to-interference-ratio, average signal-to-noise-ratio
$\mathrm{PSE}(\cdot)$	Potential spectral efficiency

algorithm. Finally, Section 3.5 concludes this chapter.

Notation: The main symbols and functions used in this chapter are reported in Table 3.1.

3.2 System Model

Efficient and effective resource management schemes for the RAN cannot be developed without taking into account the network topology, which accounts for, in particular, the spatial distributions of BSs and MTs of each tenant throughout the cellular network. The number of BSs and MTs per unit area directly determines the inter-site distance and average traffic load of the network, which in turns affects the fact that how to best allocate the radio resources, such as the transmit power of BSs and available spectrum to share, among each tenant of MTs. The network topology affects the distributions of pathloss experienced for the intended link, other-cell interference and, thus, received signal-to-interference-plus-noise ratio (SINR). These crucial components are of major importance for designing the optimal resource management mechanism. In this chapter, we explicitly apply the theory of stochastic geometry to model the network topology, which allows on to obtain the system of tractable analytical frameworks of several key KPIs, due to its tractability. Towards the end, accordingly, we describe the cellular network model, channel model along with the tractable analytical frameworks in the following subsections.

3.2.1 Cellular Networks Modeling

A downlink cellular network is considered. The BSs are modeled as points of a homogeneous Poisson point process (PPP), denoted by Ψ_{BS}, of density λ_{BS}. Assume multiple tenants, denoted by a set $\mathscr{I} = \{1, 2, ..., \mathrm{T}\}$, are available in the network. MTs belong to a particular tenant, $i \in \mathscr{I}$, are modeled following a different independent homogeneous PPP, denoted by Ψ_{MT_i}, of density λ_{MT_i}. The BSs and MTs are equipped with a single omnidirectional antenna. Each BS is assumed to transmit with a constant power and P_{tot} denotes the total power budget. Each BS serves MTs belong to different tenants in a non-overlapping frequency band of total bandwidth B_{tot}. The percentage of transmit power and bandwidth used by MT_i are denoted by P_i and B_i for $i \in \mathscr{I}$. Due to the use of non-overlapping frequency bands among the MTs of different tenants, consequently, no inter-tenant interference is available. The other-cell interence is generated from all the other BSs that transmitting via the same spectrum under the consideration of full frequency reuse. Given a generic tenant, the analytical frameworks are developed for the typical MT, denoted by $\mathrm{MT}^{(0)}$, that is located at the origin (Slivnyak theorem [33]). The superscripts 0, η and n identify the intended link, a generic interfering link and a generic BS-to-MT link. The set of interfering BSs is denoted by $\Psi_{\mathrm{BS}}^{(I)}$. Note that the analytical frameworks are developed for the generic tenant, so we release the notation i in the following subsections.

3.2.2 Channel Modeling

For each BS-to-MT link, pathloss and fast-fading are considered. Shadowing is not taken into account as its effect lies in modifying the density of the BSs through the intensity matching approach [25]. All BS-to-MT links are assumed to by mutually independent and identically distributed (i.i.d.).

Pathloss Consider a generic BS-to-MT link of length $r^{(n)}$, for simplicity, the pathloss is expressed as an unbounded power decaying function $l\left(r^{(n)}\right) = k\left(r^{(n)}\right)^{\alpha}$, where k and α are the pathloss constant and the path loss exponent.

Fast-fading Consider a generic BS-to-MT link, Rayleigh fading is assumed where the power gain follows exponential distribution with mean Ω. Without loss of generality, $\Omega = 1$ is assumed. The power gain of a generic BS-to-MT link is denoted by $h^{(n)}$.

3.2.3 Cell Association Modeling

We assume the cell association criterion is based on the maximum average received power. Let $\mathrm{BS}^{(n)} \in \Psi_{\mathrm{BS}}$ denote a generic BS from the cellular network, the serving BS, $\mathrm{BS}^{(0)}$, is selected according to the following rule:

$$\mathrm{BS}^{(0)} = \arg \max_{\mathrm{BS}^{(n)} \in \Psi_{\mathrm{BS}}} \left\{1/l\left(r^{(n)}\right)\right\} = \arg \max_{\mathrm{BS}^{(n)} \in \Psi_{\mathrm{BS}}} \left\{1/L^{(n)}\right\} \tag{3.1}$$

where the short-hand $L^{(n)} = l\left(r^{(n)}\right)$ is used. For the intended link, there is $L^{(0)} = \min_{r^{(n)} \in \Psi_{\mathrm{BS}}} \left\{L^{(n)}\right\}$.

3.2.4 Load Modeling

According to cell association criterion in (3.1), several or no MTs belong to a generic tenant can be associated with a BS. Let us consider a generic tenant. When no MTs are inside the range of a BS cell radius, the transmit power of BS is zero, $P = 0$, and it will not generate inter-cell interference. For the former case, on the other hand, denote the number of MTs associated to a BS as $N_{\mathrm{MT}} \geqslant 1$, we assume all the MTs are scheduled for transmission in the cell and that the resources (transmit power and bandwidth) from the BS are equally splitted among B/N_{MT} and the power spectral density at the typical MT is equal to P/B.

3.2.5 Potential Spectral Efficiency

For a generic tenant whose MTs constitute a PPP with density λ_{MT}, and whose BSs allocate transmit power P and bandwidth B. Let $\bar{N}_{\mathrm{MT}}$ be the number of MTs that lie in the cell of the typical MT, $\mathrm{MT}^{(0)}$, with the exception of $\mathrm{MT}^{(0)}$. The PSE is formulated according to the definition of coverage probability in [34] as follows:

$$\mathrm{PSE}\left(P, B, \lambda_{\mathrm{MT}}\right) = \sum_{u=0}^{+\infty} \mathrm{PSE}\left(P, B, \lambda_{\mathrm{MT}} \,|\, u+1\right) \Pr\left\{\bar{N}_{\mathrm{MT}} = u\right\} \tag{3.2}$$

where $\mathrm{PSE}\left(P, B, \lambda_{\mathrm{MT}} \,|\, u+1\right)$ is the PSE by conditioning on the number of MTs in a generic cell where $N_{\mathrm{MT}} = \bar{N}_{\mathrm{MT}} + 1$, $\Pr\left\{\bar{N}_{\mathrm{MT}} = u\right\}$ is the probability mass function in the considered system model, which explicitly defines the probability that, given a MT in a cell, there are other u MTs in it.

Relying on the definition of coverage probability in [34], let us define γ_I as the reliable threshold for successfully decoding a data packet, and γ_A as the reliable threshold for detecting the presence of the serving BS during the cell association phase. With aid of stochastic geometry, $\mathrm{PSE}\left(P, B, \lambda_{\mathrm{MT}} \,|\, u+1\right)$ and $\Pr\left\{\bar{N}_{\mathrm{MT}} = u\right\}$ are formulated as follows [20], [34]:

$$\begin{aligned} &\mathrm{PSE}\left(P, B, \lambda_{\mathrm{MT}} \,|\, u+1\right) = \lambda_{\mathrm{MT}} \frac{B}{u+1} \log_2\left(1+\gamma_I\right) \\ &\times \Pr\left\{\mathrm{SIR}\left(u+1\right) \geqslant \gamma_I, \overline{\mathrm{SNR}}\left(u+1\right) \geqslant \gamma_A\right\} \end{aligned} \tag{3.3}$$

$$\Pr\left\{\bar{N}_{\mathrm{MT}} = u\right\} = \frac{3.5^{3.5}\,\Gamma\left(u+4.5\right)\left(\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}}\right)^{u}}{\Gamma\left(3.5\right)\Gamma\left(u+1\right)\left(3.5+\lambda_{\mathrm{MT}}/\lambda_{\mathrm{BS}}\right)^{u+4.5}} \tag{3.4}$$

where $\Gamma\left(\cdot\right)$ is the gamma function, $\Pr\left\{\mathrm{SIR}\left(u+1\right) \geqslant \gamma_I, \overline{\mathrm{SNR}}\left(u+1\right) \geqslant \gamma_A\right\}$ denotes the coverage probability, $\mathrm{SIR}\left(u+1\right)$ and $\overline{\mathrm{SNR}}\left(u+1\right)$ denote the signal-to-interference ratio (SIR) and the average signal-to-noise ratio (SNR), given the number of MTs inside a generic cell, during the information decoding and the cell association detection phase, respectively. They are defined as follows:

$$\mathrm{SIR}\left(u+1\right) = \frac{\left[P/\left(u+1\right)\right] h^{(0)}/L^{(0)}}{\sum\limits_{\mathrm{BS}^{(\eta)} \in \Psi_{\mathrm{BS}}^{(I)}} \left[P/\left(u+1\right)\right] h^{(\eta)}/L^{(\eta)} \mathbf{1}\left(L^{(\eta)} > L^{(0)}\right)} \tag{3.5}$$

$$\overline{\mathrm{SNR}}(u+1) = \frac{[P/(u+1)]/L^{(0)}}{N_0[B/(u+1)]} \tag{3.6}$$

where $h^{(0)}$ and $h^{(\eta)}$ are the fading power gains of the serving and interfering BSs of a generic MT, respectively. $L^{(0)}$ and $L^{(\eta)}$ are the pathlosses of serving and interfering BSs, respectively. N_0 is the noise power spectral density. The indicator function $\mathbf{1}\left(L^{(\eta)} > L^{(0)}\right)$ accounts for the cell association criterion. It is worth noting that $\overline{\mathrm{SNR}}(u+1)$ is averaged SNR with respect to the fast fading in order to avoid frequent handovers due to the channel variations. Under this new definition, a closed-form expression of PSE can be obtained, which explicitly depends on hte transmit power of BSs. The PSE is illustrated in the following proposition.

Proposition 3.1 *The exact analytical expression of PSE is formulated as follows:*

$$\begin{aligned} PSE(P,B,\lambda_{MT}) &= B\log_2\left(1+\gamma_I\right)\frac{\lambda_{BS}L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)}{1+L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\Upsilon\left(\gamma_I,\alpha\right)} \\ &\times\left[1-\exp\left(-\pi\lambda_{BS}\left(\xi_A\frac{P}{B}\right)^{2/\alpha}\left(1+L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right)\Upsilon\left(\gamma_I,\alpha\right)\right)\right)\right] \end{aligned} \tag{3.7}$$

where $\xi_A = \left(k\gamma_A N_0\right)^{-1}$, *and:*

$$L\left(\frac{\lambda_{MT}}{\lambda_{BS}}\right) = 1-\left(1+\frac{1}{3.5}\frac{\lambda_{MT}}{\lambda_{BS}}\right)^{-3.5} \geqslant 0 \tag{3.8}$$

$$\Upsilon\left(\gamma_I,\alpha\right) = {}_2F_1\left(-\frac{2}{\alpha},1,1-\frac{2}{\alpha},-\gamma_I\right)-1 \geqslant 0 \tag{3.9}$$

where ${}_2F_1(\cdot)$ *denoting the Gauss hypergeometric function.*

Proof*: The proof follows from [34].* □

Remark 3.1 *It is worth noting that the framework based on the new definition is more realistic since it accounts for the non-zero values of* γ_A*, which results in that the MT cannot detect an arbitrary weak signal. If* $\gamma_A = 0$*, the PSE becomes independent of transmit power,* P*, and the proposed framework simplifies to the conventional formulation [17]. The resulting PSE, in addition, will be a linear function of bandwidth,* B*. By considering the MTs have the finite sensitity in detecting the presence of the BSs, i.e.,* $\gamma_A \neq 0$*, the more accurate analytical formulation, which depends on the fundamental parameter pair* (P,B)*, are obtained for formulating the optimization problem in multi-tenants network slicing.*

The property of the proposed analytical framework of PSE as a function of (P,B) is reported in the following proposition.

Proposition 3.2 *The utility function* $PSE(P,B)$ *is joint concave in transmit power* P *and bandwidth* B*.*

Proof*: See Appendix A.* □

Based on the introduced closed-from formulation of PSE and its fundamental property in transmit power and bandwidth, we design the optimal resource management scheme through an optimization problem in the following subsection.

3.2.6 Problem Formulation

We consider a RAN with multiple access points among the cellular BSs, such as LTE eNBs, mmWave access points and femto cells, where multiple tenants $\mathscr{I} = \{1, 2, ..., \mathrm{T}\}$ are available in the network [44]. The generic infrastructure tenant $i \in \mathscr{I}$ is willing to pay for managing a "slice" of the resources of the RAN given a certain SLA is satisfied. We consider a single network operator with its physical resources available to all the T tenants, MTs belongs to a tenant i access the RAN and share common resources of the same air interface with distinctive previleges. The BSs of cellular netork, in particular, are assumed to be operated in an open access mode for all the MTs among each tenant. Each BS, nevertheless, serves MTs of different tenants in a non-overlapping frequency band of total bandwidth and by using part of the total transmit power budget. A dedidcated part of the spectrum and of the power, which constitutes an isolated and distinctive slice of the available resources of the RAN, therefore, are requested for MTs of different tenants from the network operator. Note that the requests of the tenants from the network operator, in particular, are not related to the specific spectrum share or transmit power are equally assigned to them. The requests of tenants, on the other hand, are authorized in order to get a minimum average spectral efficiency (expressed in bit/sec/m^2), which allows them to satisfy the specific SLA of their own users (i.e., MTs) with a given density, regardless of the existence of the other tenants in the network.

In this chapter, we formulate the optimization problem based on the above assumptions and identify the optimal transmit power and spectrum to be allocated to each tenant when a sertain SLA per slice is guaranteed such that sum spectral efficiency is maximized over all the tenants. To be more precise, we introduce a binary variable $\beta_i = \{0, 1\}$ for $i \in \mathscr{I}$, in particular, as an indicator variable to control that the physical resources are only assigned among each tenant when the SLA constraint is satisfied. The optimizaiton problem for multi-tenants, $i \in \mathscr{I}$, is formulated as follows.

Problem I

$$\max_{\substack{P_i, B_i, \beta_i, \\ i \in \mathscr{I}}} \quad f\left(P_i, B_i, \beta_i\right) = \sum_{i=1}^{\mathrm{T}} \log\left[\left(1+\beta_i\right)\left(1+\mathrm{PSE}\left(P_i, B_i\right)\right)\right] \tag{3.10}$$

$$\text{s.t.} \quad \mathrm{PSE}\left(P_i, B_i\right) \geqslant \beta_i \lambda_{\mathrm{MT}_i} \tau_i \tag{3.11}$$

$$\sum_{i=1}^{\mathrm{T}} B_i \leqslant B_{\mathrm{tot}} \tag{3.12}$$

$$\sum_{i=1}^{\mathrm{T}} P_i \leqslant P_{\mathrm{tot}} \tag{3.13}$$

$$0 \leqslant B_i \leqslant \beta_i B_{\mathrm{tot}} \tag{3.14}$$

$$0 \leqslant P_i \leqslant \beta_i P_{\mathrm{tot}} \tag{3.15}$$

$$\beta_i \in \{0, 1\} \tag{3.16}$$

The objective function $f\left(P_i, B_i, \beta_i\right)$ in (3.10) is formulated as a function of the continuous variables, i.e., the available resources (P_i, B_i) to be assigned for each tenant $i \in \mathscr{I}$, and the corresponding binary

variables β_i. In the present chapter, relying on the assumptions stated previously, the infrastructure tenants are interested in getting a minimum requested spectral efficiency given the density of MTs of each tenant, λ_{MT_i}, in order to satisfy the per slice SLA, which results in that the problem is subjected to a sets of T nonlinear constraints, expressed in (3.11). $\lambda_{\mathrm{MT}_i}\tau_i$ denotes the minimum requested PSE for a tenant of MTs with given density (τ_i therefore denotes the minimum required average rate given density of MTs). The linear constraints in (3.12) and (3.13) are formulated to ensure that the assigned dedicated resources of power and bandwidth satisfy their total budgets. Constraints (3.14), (3.15) and (3.16), furthermore, define the upper and lower bounds of the continuous variables, and integer variables, and it guarantees that no resources will be allocated when $\beta_i = 0$.

Remark 3.2 *Follow the property of PSE(P,B) in Proposition 3.2, if the binary variables are fixed,* ***Problem I*** *is a convex optimization problem due to the fact that the objective function in* (3.10)*, the nonlinear constraint in* (3.11) *along with other constraints are convex [45].*

To be more precise, if $\beta_i = 0$, for $i \in \mathscr{I}$, the logarithm and binary variable guarantee that the function $\log\left[\left(1+\beta_i\right)(1+PSE(P_i,B_i))\right] = 0$ and the nonlinear constraints in (3.11) *always hold true. If $\beta_i = 1$, in addition, the* ***Problem I*** *boils down to a typical convex optimization problem.*

Remark 3.3 *The maximization of objective function in* (3.10) *is equivalent to maximize the sum of PSE, which is given by* (3.17) *as follows.*

$$\mathscr{F}\left(P_i,B_i,\beta_i\right) = \sum_{i=1}^{\mathrm{T}} \beta_i PSE(P_i,B_i) \tag{3.17}$$

The formulation with logarithm in the objective function (3.10)*, on the other hand, ensures fairness among the tenants. As a result, the objective function can be rewritten as:*

$$f\left(P_i,B_i,\beta_i\right) = \sum_{i=1}^{\mathrm{T}} \left[\log\left(1+\beta_i\right) + \log\left(1+PSE(P_i,B_i)\right)\right] \tag{3.18}$$

Moreover, this property of separability also holds for all the constraints.

Based on these properties, in the next section, we propose and design the efficient algorithm to solve it.

3.3 Generalized Benders Decomposition Approach

Within the existence of both the integer variables and continuous variables, the **Problem I** constitutes a mixed integer nonlinear programming (MINLP) [46], which is a NP-hard combinatorial problem and no polynomial time algorithm has yet been discovered for solving it. Relying on i) the convex property of **Problem I** when the binary variables are fixed as stated in Remark 3.2, and ii) the binary variables are separable from the continuous variables in both objective function and constraints in Remark 3.3, which allow us to apply efficiently the Generalized Benders Decomposition (GBD) algorithm [47] to solve the original MINLP. In particular, the property of separability between two types of variables constitutes the Property P of GBD, which results in that the optimization over (P_i,B_i) can be obtained independently as β_i, and thus in the reduction of complexity.

The rational of GBD is to decompose the original MINLP into a primal problem, which corresponds to the original problem with fixed complicating variables, and a master problem derived through nonlinear duality theory [47] using the Lagrange multipliers obtained from the primal problem. The algorithm, accordingly, is operated by solving them iteratively and the global optimum can be reached through the finite convergence in Theorem 3.1.

Theorem 3.1 *Finite Convergence: The GBD algorithm terminates in a finite number of iterations for any given tolerance* $\varepsilon \geqslant 0$.

Proof*: The proof follows from [Theorem 2.5, [47]].* □

The GBD algorithm, based on the finite convergence property, is demonstrated as follows in Algorithm 1, where all the subproblems are defined as follows.

Algorithm 1 Generalized Benders Decomposition

1: Select an initial value for $\bar{\beta}_i$ which makes the primal problem in **Problem III** feasible, and obtain its solution $\left(P_i^*, B_i^*\right)$, along with the multiliers associated with the constraints, $\boldsymbol{\mu}^*$. Set $K_1 = 1$, $K_2 = 0$, $\boldsymbol{\mu}^{(1)} = \boldsymbol{\mu}^*$, the upper bound $UBD = \phi\left(P_i^*, B_i^*, \bar{\beta}_i\right)$, and determine $L^*\left(\beta_i, \boldsymbol{\mu}^{(1)}\right)$.
2: Solve the relaxed master problem in **Problem IV** and obtain its global solution $\left(\beta_0^*, \beta_i^*\right)$. Set the lower bound $LBD = \beta_0^*$.
3: **while** $UBD - LBD \geqslant \varepsilon$ **do**
4: Solve the primal problem with fixed $\bar{\beta}_i = \beta_i^*$.
5: **if** The primal problem is feasible **then**
6: Obtain its solution and multipliers $\left(P_i^*, B_i^*, \mu^*\right)$. Set $K_1 = K_1 + 1$, $\boldsymbol{\mu}^{(K_1)} = \boldsymbol{\mu}^*$ and determine $L^*\left(\beta_i, \boldsymbol{\mu}^{(K_1)}\right)$. If $\phi\left(P_i^*, B_i^*, \bar{\beta}_i\right) < UBD$, set $UBD = \phi\left(P_i^*, B_i^*, \bar{\beta}_i\right)$. Return to Step 2.
7: **else**
8: Solve the problem **Problem V**, and obtain its solution $\left(P_i^*, B_i^*, \boldsymbol{\rho}^*\right)$. Set $K_2 = K_2 + 1$, $\boldsymbol{\rho}^{(K_2)} = \boldsymbol{\rho}^*$, determine $L_*\left(\beta_i, \boldsymbol{\rho}^{(K_2)}\right)$ and return to Step 2.
9: **end if**
10: **end while**

We write the original MINLP in **Problem I** as the following standard minimization form for the convenience of presenting the subproblems in GBD.

Problem II:

$$\min_{\substack{P_i, B_i, \beta_i \\ i \in \mathscr{I}}} \quad \phi\left(P_i, B_i, \beta_i\right) = -f\left(P_i, B_i, \beta_i\right) \tag{3.19}$$

$$\text{s.t.} \quad \mathbf{G}\left(P_i, B_i, \beta_i\right) \leqslant 0 \tag{3.20}$$

$$\beta_i \in \{0, 1\} \tag{3.21}$$

where $\mathbf{G}\left(P_i, B_i, \bar{\beta}_i\right) = \begin{bmatrix} G_1 & G_2 & \cdots & G_7 \end{bmatrix}^T$ is a vector contains all the constraint functions of the primal problem with $G_1 = g\left(P_i, B_i, \beta_i\right)$, $G_2 = \sum_{i=1}^{\mathrm{T}} P_i - P_{\mathrm{tot}}$, $G_3 = \sum_{i=1}^{\mathrm{T}} B_i - B_{\mathrm{tot}}$, $G_4 = B_i - \beta_i B_{\mathrm{tot}}$, $G_5 = P_i - \beta_i P_{\mathrm{tot}}$, $G_6 = -B_i$ and $G_7 = -P_i$.

The primal problem and master problem, based on the standard form of the original MINLP in **Problem II**, are formulated as follows.

Problem III: Primal Problem

$$\min_{\substack{P_i,B_i \\ i\in\mathscr{I}}} \quad \phi\left(P_i,B_i,\bar{\beta}_i\right) \tag{3.22}$$

$$\text{s.t.} \quad \mathbf{G}\left(P_i,B_i,\bar{\beta}_i\right)\leqslant 0 \tag{3.23}$$

Problem IV: Master Problem

$$\min_{\substack{\beta_0,\beta_i \\ i\in\mathscr{I}}} \quad \beta_0 \tag{3.24}$$

$$\text{s.t.} \quad \beta_0 \geqslant L^*\left(\beta_i,\boldsymbol{\mu}^{(k_1)}\right), k_1=1,...,K_1 \tag{3.25}$$

$$L_*\left(\beta_i,\boldsymbol{\rho}^{(k_2)}\right)\leqslant 0, k_2=1,...,K_2 \tag{3.26}$$

The above formulation of the primal problem results from fixing the binary variables to a particular 0-1 combination, denoted by $\bar{\beta}_i$. Since the optimal solution (if exists) to the **Problem III** is also a feasible solution to the **Problem II**, the optimal value $\phi\left(P_i^*,B_i^*,\bar{\beta}_i\right)$ constitutes the upper bound to the original problem. The optimal solution of the master problem, which originates from the nonlinear duality theory, futhermore, provides the lower bound to the original problem.

Remark 3.4 *The objective function of the original problems,* ***Problem I*** *and* ***Problem II****, along with the constraints are linear separable in β_i and (P_i,B_i), which indicates that property P is satisfied [47] and the minimum over the continuous variables is obtained independently of of the binary variables. Thus, the support functions $L^*\left(\beta_i,\boldsymbol{\mu}^{(k_1)}\right)$ and $L_*\left(\beta_i,\boldsymbol{\rho}^{(k_2)}\right)$ in the master problem are explicitly determined.*

Proposition 3.3 *Follow the property P of GBD in Remark 3.4, the explicit expression of $L^*\left(\beta_i,\boldsymbol{\mu}^{(k_1)}\right)$ and $L_*\left(\beta_i,\boldsymbol{\rho}^{(k_2)}\right)$ in master problem are formulated in* (3.27) *and* (3.28) *as follows.*

$$\begin{aligned} L^*\left(\beta_i,\mu^{(k_1)}\right) &= \sum_{i=1}^{\mathrm{T}}\beta_i\left(\mu_i^{(k_1)}\lambda_{MT_i}\tau_i-\log(2)-\mu_{0Bi}^{(k_1)}B_{tot}-\mu_{0Pi}^{(k_1)}P_{tot}\right)-\mu_B^{(k_1)}B_{tot}-\mu_P^{(k_1)}P_{tot} \\ &+ \min_{P_i,B_i} \left\{\begin{array}{l} -\sum_{i=1}^{\mathrm{T}}\log\left[1+PSE(P_i,B_i)\right]+\sum_{i=1}^{T}\left(\mu_{0Bi}^{(k_1)}-\mu_{00Bi}^{(k_1)}\right)B_i+\sum_{i=1}^{T}\left(\mu_{0Pi}^{(k_1)}-\mu_{00Pi}^{(k_1)}\right)P_i \\ +\mu_B^{(k_1)}\sum_{i=1}^{\mathrm{T}}B_i+\mu_P^{(k_1)}\sum_{i=1}^{\mathrm{T}}P_i-\sum_{i=1}^{\mathrm{T}}\mu_i^{(k_1)}PSE(P_i,B_i) \end{array}\right\} \end{aligned} \tag{3.27}$$

$$\begin{aligned} L_*\left(\beta_i,\rho^{(k_2)}\right) &= \sum_{i=1}^{\mathrm{T}}\beta_i\left(\rho_i^{(k_2)}\lambda_{MT_i}\tau_i-\rho_{0Bi}^{(k_2)}B_{tot}-\rho_{0Pi}^{(k_2)}P_{tot}\right)-\rho_B^{(k_2)}B_{tot}-\rho_P^{(k_2)}P_{tot} \\ &+ \min_{P_i,B_i} \left\{\begin{array}{l} \sum_{i=1}^{T}\left(\rho_{0Bi}^{(k_2)}-\rho_{00Bi}^{(k_2)}\right)B_i+\sum_{i=1}^{T}\left(\rho_{0Pi}^{(k_2)}-\rho_{00Pi}^{(k_2)}\right)P_i \\ +\rho_B^{(k_2)}\sum_{i=1}^{\mathrm{T}}B_i+\rho_P^{(k_2)}\sum_{i=1}^{\mathrm{T}}P_i-\sum_{i=1}^{\mathrm{T}}\rho_i^{(k_2)}PSE(P_i,B_i) \end{array}\right\} \end{aligned} \tag{3.28}$$

where the vectors are given by the order of the constraints $\mathbf{G}\left(P_i, B_i, \bar{\beta}_i\right)$ *as follows.*

$$\mu^{(k_1)} = \left[\begin{array}{ccccccc} \mu_i^{(k_1)} & \mu_B^{(k_1)} & \mu_P^{(k_1)} & \mu_{0Bi}^{(k_1)} & \mu_{0Pi}^{(k_1)} & \mu_{00Bi}^{(k_1)} & \mu_{00Pi}^{(k_1)} \end{array} \right]^T \tag{3.29}$$

$$\rho^{(k_2)} = \left[\begin{array}{ccccccc} \rho_i^{(k_2)} & \rho_B^{(k_2)} & \rho_P^{(k_2)} & \rho_{0Bi}^{(k_2)} & \rho_{0Pi}^{(k_2)} & \rho_{00Bi}^{(k_2)} & \rho_{00Pi}^{(k_2)} \end{array} \right]^T \tag{3.30}$$

Proof*: See Appendix B.* □

Remark 3.5 *The functions* $L^*\left(\beta_i, \boldsymbol{\mu}^{(k_1)}\right)$ *and* $L_*\left(\beta_i, \boldsymbol{\rho}^{(k_2)}\right)$ *are formulated within two sets of inner optimization problems in* (P_i, B_i)*, which are independent of the optimization over* β_i*. This, accordingly, results in that the master problem becomes mixed integer linear programming (MILP) in* β_i*. A branch of algorithms are available to solve it with relatively lower complexity compared with the original MINLP.*

Remark 3.6 *It is worth noting that the two inner optimization problems in* (3.27) *and* (3.28) *are obtained by invoking the minimization of Lagrange functions from the convex primal problem, which results, accordingly, in the inner optimizations are convex problems [45].*

Note that not all the choices of binary variables provide the feasible solutions to primal problem. There are two cases for the primal problem, i.e., feasible and infeasible, which deliver the Lagrange multipliers for the master problem. The vector of multipliers $\boldsymbol{\mu}$ is obtained when the primal problem is solved to be feasible. The other vector of multipliers $\boldsymbol{\rho}$, however, is obtained by solving the following problem when the primal problem is infeasible [46], [48], [49].

Problem V: Determination of $\boldsymbol{\rho}$

$$\min_{\substack{P_i, B_i, \theta \\ i \in \mathscr{I}}} \quad \theta \tag{3.31}$$

$$\text{s.t.} \quad \mathbf{G}\left(P_i, B_i, \bar{\beta}_i\right) - \theta \mathbf{1} \leqslant 0 \tag{3.32}$$

Remark 3.7 *Due to the fact that constraints in* $\mathbf{G}\left(P_i, B_i, \bar{\beta}_i\right) - \theta \mathbf{1}$ *are convex, and the objective function is a constant, the optimization in* ***Problem V*** *is a convex problem accordingly.*

3.4 Numerical Results

In this section, we illustrate several numerical simulations that substantiate the applied GBD algorithm. The accuracy of GBD is verified against the brute-force search for the binary variables, which is stated in Algorithm 2. In addition, we implement a dynamic programming scheme called Greedy algorithm in Algorithm 3 as a comparison for the performance evaluation.

The following simulation setup is considered: i) density of BSs, measured in cell radius $R_{\text{cell}} = 500$m, ii) decoding thresholds $\gamma_I = \gamma_A = 10$dB, iii) carrier frequency is 2.1GHz, iv) pathloss exponent $\alpha = 3.5$, v) total bandwidth $B_{\text{tot}} = 100$MHz, total power $P_{\text{tot}} = 30$dBm, vi) noise power spectral density

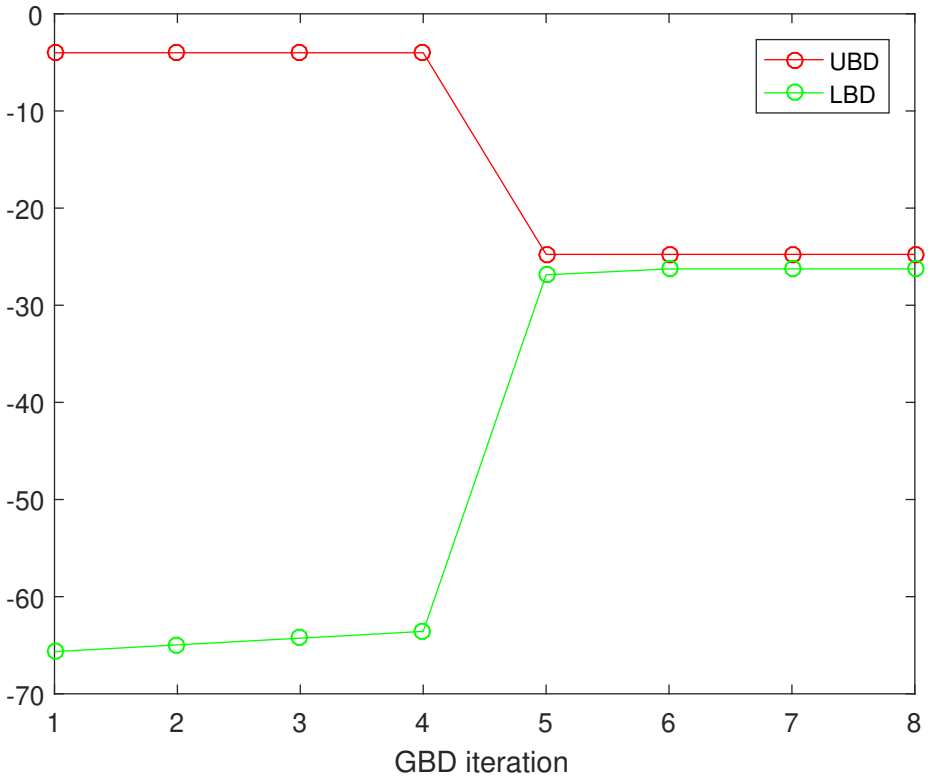

Figure 3.1 – Finite convergence of GBD. Number of tenants T = 20.

$N_0 = -174\text{dBm}/\text{Hz}$. We apply the OPTI toolbox [50] intergrated with Matlab to solve the optimization problems, and the corresponding solvers and the associated algorithms are summarized in Table 3.2. Except for the master problem, which is formulated as MILP and solved by using SCIP solver and spatial branch and bound algorithm [51], all the other subproblems are presented as convex optimizations, and thus, they are determined through IPOPT with typical primal-dual interior point method [52].

Table 3.2 – Solvers and Algorithms of OPTI toolbox [50] for subproblems of GBD

	Solver	Algorithm
Problem III	IPOPT	Primal-dual interior point
Inner Optimizations ((3.27), (3.28))	IPOPT	Primal-dual interior point
Problem V	IPOPT	Primal-dual interior point
Problem IV	SCIP	Spatial branch and bound

Fig. 3.1 shows the finite convergence of GBD algorithm as illustrated in Theorem 3.1. We select the parameters for large values of number of tenants T = 20 in order to better declare convergence. The random input parameters for minimum PSE, or to be more precise, the minimum required average rates (bit/s) and densities $(1/\pi R_{\text{MT}_i}^2)$ of all the tenants are given by: $\tau_1 = 1498.65$, $\tau_2 = 6596.05$, $\tau_3 = 5185.95$, $\tau_4 = 9729.75$, $\tau_5 = 6489.91$, $\tau_6 = 8003.31$, $\tau_7 = 4537.98$, $\tau_8 = 4323.92$, $\tau_9 = 25.12$, $\tau_{10} = 11.68$, $\tau_{11} = 17.27$, $\tau_{12} = 0.62$, $\tau_{13} = 39.36$, $\tau_{14} = 6.69$, $\tau_{15} = 4.25$, $\tau_{16} = 14.90$, $\tau_{17} = 7.92$, $\tau_{18} = 19.59$, $\tau_{19} = 13.58$, $\tau_{20} = 38.07$, and $R_{\text{MT}_1} = 27.28$, $R_{\text{MT}_2} = 1.70$, $R_{\text{MT}_3} = 2.86$, $R_{\text{MT}_4} = 20.87$, $R_{\text{MT}_5} = 3.87$, $R_{\text{MT}_6} = 32.73$, $R_{\text{MT}_7} = 32.70$,

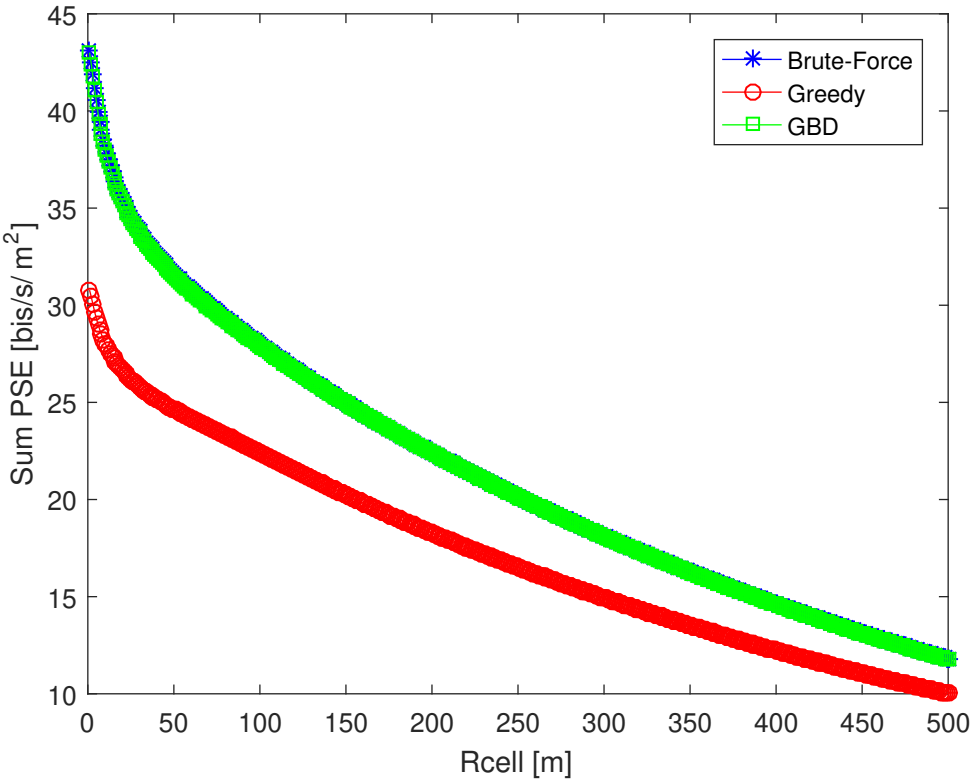

Figure 3.2 – Sum PSE $\mathscr{F}\left(P_i^*, B_i^*, \beta_i^*\right)$ in (3.17) as a function of density of BSs for different algorithms. Number of tenants T = 8.

$R_{\mathrm{MT}_8} = 28.90$, $R_{\mathrm{MT}_9} = 825.31$, $R_{\mathrm{MT}_{10}} = 83.47$, $R_{\mathrm{MT}_{11}} = 133.17$, $R_{\mathrm{MT}_{12}} = 173.39$, $R_{\mathrm{MT}_{13}} = 390.94$, $R_{\mathrm{MT}_{14}} = 831.38$, $R_{\mathrm{MT}_{15}} = 803.36$, $R_{\mathrm{MT}_{16}} = 60.47$, $R_{\mathrm{MT}_{17}} = 399.26$, $R_{\mathrm{MT}_{18}} = 526.88$, $R_{\mathrm{MT}_{19}} = 416.80$, $R_{\mathrm{MT}_{20}} = 656.86$. It is shown that the GBD algorithm convergences after 6 iterations based on the current setup.

Fig 3.2 shows the performance sum PSE in (3.17) as a function of density of BSs ($\lambda_{\mathrm{BS}} = 1/\pi R_{\mathrm{cell}}^2$) for three algorithms. We select the paramters of the number of tenants T = 8 in order to obtain the solution of brute-force in an accepted time slots. The optimal values $\left(P_i^*, B_i^*, \beta_i^*\right)$ for three algorithms are obtained as the input of (3.17). The random input parameters for the minimum required average rates (bit/s) and densities ($1/\pi R_{\mathrm{MT}_i}^2$) of all the tenants are given by: $\tau_1 = 1307.38$, $\tau_2 = 1913.87$, $\tau_3 = 1871.46$, $\tau_4 = 915.77$, $\tau_5 = 480.96$, $\tau_6 = 29.75$, $\tau_7 = 4.24$, $\tau_8 = 27.26$, and $R_{\mathrm{MT}_1} = 8.59$, $R_{\mathrm{MT}_2} = 2.50$, $R_{\mathrm{MT}_3} = 0.49$, $R_{\mathrm{MT}_4} = 5.807$, $R_{\mathrm{MT}_5} = 6.35$, $R_{\mathrm{MT}_6} = 458.34$, $R_{\mathrm{MT}_7} = 455.60$, $R_{\mathrm{MT}_8} = 444.39$. It is shown that there is a good agreement of the GBD algorithm and brute-force search, which indicates the optimal convergence of GBD. In addition, it is illustrated that the applied GBD approach outperforms that of Greedy algorithm and the high performance gain can be reached for the BSs deployed in a higher density.

In Figure 3.3, we study the complexity of the applied GBD approach against the brute-force search and the Greedy algorithm. To be more precise, the computational time of three algorithms is compared as a function of the number of tenants. It is clear that the computational time of three algorithms increases as the number of optimization variables increases. Moreover, the applied GBD approach remains the similar level of lower complexity as the Greedy algorithm, and even achieve the best performance

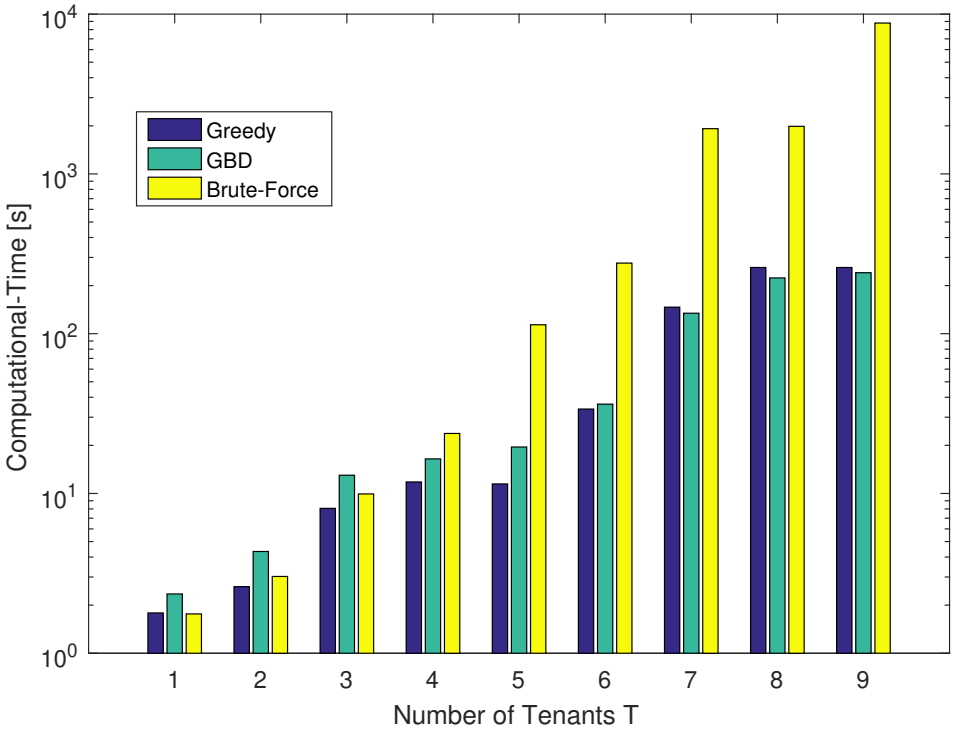

Figure 3.3 – Computational analysis of three algorithms.

in terms of computational time for the large values of T. The brute-force search, as a verification benckmark, on the other hand, constitutes the worst candidate algorithm in complexity regardless of its optimal solutions for larger variables.

3.5 Conclusion

In this chapter, we have designed an optimal resource allocation mechanism for application to network slicing in RANs. In particular, we have taken into account the network slice requests with diverse SLAs in terms of the minimum required spectral efficiency per tenant. To analytically formulate the problem, we have exploited the theory of stochastic geometry, which allowed us to model the network topology in a tractable yet sufficiently realistic manner. The proposed utility function, PSE, have been identified as concave with respect to the transmit power and bandwidth. The optimization problem has been formulated as a MINLP and solved by using a Generalized Benders Decomposition approach where the global optimality has been retained compared with the brute-force search approach. The numerical results have further verified the accuracy of the adopted algorithm, and the visible performance gain of sum PSE can be achieved by comparing with the Greedy algorithm. In conclusion, we evince that the applied GBD algorithm is capable of achieving the optimal performance with a relatively lower computational complexity in solving the formulated MINLP.

Algorithm 2 Brute-Force Search

1: Generate all the possible combinations of $\beta_i \in \{0,1\}$ for $i = 1,2,...,\mathrm{T}$ as $\beta_i^{(j)}$.
2: **for** $j = 1 : 2^{\mathrm{T}}$ **do**
3: Set $\beta_i = \beta_i^{(j)}$ for all $i = 1,2,...,\mathrm{T}$, solve the primal problem in **Problem II** and obtain the solutions $\left(P_i^{(j)}, B_i^{(j)}\right)$ and $\phi\left(P_i^{(j)}, B_i^{(j)}, \beta_i^{(j)}\right)$.
4: **end for**
5: Find $\hat{j} = \arg\min_j \left\{\phi\left(P_i^{(j)}, B_i^{(j)}, \beta_i^{(j)}\right)\right\}$, and return the optimal solutions $\left(P_i^*, B_i^*, \beta_i^*\right) = \left(P_i^{(\hat{j})}, B_i^{(\hat{j})}, \beta_i^{(\hat{j})}\right)$ for all i.

Algorithm 3 Greedy Algorithm

1: Initialize $\beta_i = 1$ for all $i = 1,2,...,\mathrm{T}$.
2: **for** $i = 1 : \mathrm{T}$ **do**
3: Solve the primal problem in **Problem II** with β_i as input.
4: **if** The primal problem is infeasible **then**
5: Set $\beta_{\hat{i}} = 0$ where $\hat{i} = \arg\max_i \{\tau_i\}$.
6: **else**
7: Return the current solutions as the optima $\left(P_i^*, B_i^*, \beta_i^*\right) = \left(P_i, B_i, \beta_i\right)$.
8: **end if**
9: **end for**

3.6 Appendices

3.6.1 Appendix A – Proof of Proposition 3.2

For simplicity, we denote the function of PSE in (3.7) as follows:

$$F(x,y) = ax\left(1 - \exp\left(-c\left(\frac{y}{x}\right)^b\right)\right) \tag{3.33}$$

where x and y denote B and P, respectively, and $a = \log_2\left(1+\gamma_I\right)\frac{\lambda_{\mathrm{BS}} L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)}{1+L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\Upsilon(\gamma_I,\alpha)}$, $b = 2/\alpha < 1$, and $c = \pi\lambda_{\mathrm{BS}}\xi_A^b\left(1 + L\left(\frac{\lambda_{\mathrm{MT}}}{\lambda_{\mathrm{BS}}}\right)\Upsilon\left(\gamma_I,\alpha\right)\right)$.

The Hessian matrix of $F(x,y)$ is given by

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x \partial y} \\ \frac{\partial^2 F}{\partial y \partial x} & \frac{\partial^2 F}{\partial y^2} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{3.34}$$

with the elements expressed by

$$\begin{aligned}
A_{11} &= abc(y)^b x^{-b-1}\exp\left(-c\left(\frac{y}{x}\right)^b\right)\left[b-1-bc\left(\frac{y}{x}\right)^b\right]<0\\
A_{22} &= abc\left(\frac{y}{x}\right)^{b-1} y^{-1}\exp\left(-c\left(\frac{y}{x}\right)^b\right)\left[(b-1)-bc\left(\frac{y}{x}\right)^b\right]<0\\
A_{12} &= A_{21} = abc\left(\frac{y}{x}\right)^b y^{-1}\exp\left(-c\left(\frac{y}{x}\right)^b\right)\left[1-b+bc\left(\frac{y}{x}\right)^b\right]>0
\end{aligned} \tag{3.35}$$

Thus, given a non-zero vector $\mathbf{v} = \begin{bmatrix} v_1 & v_2 \end{bmatrix}^T$, the Hessian matrix is determined to be negative-semidefinite as follows:

$$\begin{aligned}
&\mathbf{v}^T\mathbf{H}\mathbf{v} = A_{11}v_1^2 + A_{22}v_2^2 + 2A_{12}v_1v_2\\
&= \left(abc\left(\frac{y}{x}\right)^b\frac{1}{y}\exp\left(-c\left(\frac{y}{x}\right)^b\right)\left[1-b+bc\left(\frac{y}{x}\right)^b\right]\right)\\
&\times\left[2v_1v_2-\left(v_1^2\frac{y}{x}+\frac{x}{y}v_2^2\right)\right]\leqslant 0
\end{aligned} \tag{3.36}$$

where the inequality is hold since $2v_1v_2-\left(v_1^2\frac{y}{x}+\frac{x}{y}v_2^2\right)\leqslant 2v_1v_2-2\sqrt{v_1^2\frac{y}{x}\times\frac{x}{y}v_2^2}=0$.

This implies that the function $F(x,y)$ is joint concave in x and y.

3.6.2 Appendix B – Proof of Proposition 3.3

The master problem is obtained by using the nonlinear duality theory [47], $L^*\left(\beta_i,\boldsymbol{\mu}^{(k_1)}\right)$ and $L_*\left(\beta_i,\boldsymbol{\rho}^{(k_2)}\right)$ are formulated based on the minimum of their Langrange functions as follows.

$$L^*\left(\beta_i,\boldsymbol{\mu}^{(k_1)}\right) = \min_{P_i,B_i}\mathscr{L}_1\left(P_i,B_i,\beta_i,\boldsymbol{\mu}^{(k_1)}\right) \tag{3.37}$$

$$L_*\left(\beta_i,\boldsymbol{\rho}^{(k_2)}\right) = \min_{P_i,B_i}\mathscr{L}_2\left(P_i,B_i,\beta_i,\boldsymbol{\rho}^{(k_2)}\right) \tag{3.38}$$

where the Lagrange functions are given by

$$\begin{aligned}
&\mathscr{L}_1\left(P_i,B_i,\beta_i,\mu^{(k_1)}\right) = f\left(P_i,B_i,\beta_i\right)+\left(\mu^{(k_1)}\right)^T\mathbf{G}\left(P_i,B_i,\bar{\beta}_i\right)\\
&= f\left(P_i,B_i,\beta_i\right)+\sum_{i=1}^{\mathrm{T}}\mu_i^{(k_1)}g\left(P_i,B_i,\beta_i\right)-\sum_{i=1}^{T}\mu_{00Bi}^{(k_1)}B_i-\sum_{i=1}^{T}\mu_{00Pi}^{(k_1)}P_i\\
&+\mu_B^{(k_1)}\left(\sum_{i=1}^{\mathrm{T}}B_i-B_{\mathrm{tot}}\right)+\mu_P^{(k_1)}\left(\sum_{i=1}^{\mathrm{T}}P_i-P_{\mathrm{tot}}\right)\\
&+\sum_{i=1}^{T}\mu_{0Bi}^{(k_1)}\left(B_i-\beta_iB_{\mathrm{tot}}\right)+\sum_{i=1}^{T}\mu_{0Pi}^{(k_1)}\left(P_i-\beta_iP_{\mathrm{tot}}\right)
\end{aligned} \tag{3.39}$$

$$
\begin{aligned}
&\mathscr{L}_2\left(P_i, B_i, \beta_i, \rho^{(k_2)}\right) = \left(\rho^{(k_2)}\right)^T \mathbf{G}\left(P_i, B_i, \bar{\beta}_i\right) \\
&= \sum_{i=1}^{\mathrm{T}} \rho_i^{(k_2)} g\left(P_i, B_i, \beta_i\right) - \sum_{i=1}^{T} \rho_{00Bi}^{(k_2)} B_i - \sum_{i=1}^{T} \rho_{00Pi}^{(k_2)} P_i \\
&+ \rho_B^{(k_2)} \left(\sum_{i=1}^{\mathrm{T}} B_i - B_{\mathrm{tot}}\right) + \rho_P^{(k_2)} \left(\sum_{i=1}^{\mathrm{T}} P_i - P_{\mathrm{tot}}\right) \\
&+ \sum_{i=1}^{T} \rho_{0Bi}^{(k_2)} \left(B_i - \beta_i B_{\mathrm{tot}}\right) + \sum_{i=1}^{T} \rho_{0Pi}^{(k_2)} \left(P_i - \beta_i P_{\mathrm{tot}}\right)
\end{aligned}
\tag{3.40}
$$

Therefore, (3.27) and (3.28) are obtained by calling the property P of GBD.

4 Wireless Networks with Metasurface-Coated Environmental Objects: An Approach Based on Random Spatial Processes

Chapter 4. Wireless Networks with Metasurface-Coated Environmental Objects: An Approach Based on Random Spatial Processes

An emerging and promising vision of wireless networks consists of coating the environmental objects with reconfigurable metasurfaces that are capable of modifying the radio waves impinging upon them according to the generalized law of reflection. By relying on tools from point processes, stochastic geometry, and random spatial processes, we model the environmental objects with a modified random line process of fixed length, and with random orientations and locations. Based on the proposed modeling approach, we develop the first analytical framework that provides one with the probability that a randomly distributed object that is coated with a reconfigurable metasurface acts as a reflector for a given pair of transmitter and receiver. In contrast to the conventional network setup where the environmental objects are not coated with reconfigurable metasurfaces, we prove that the probability that the typical random object acts as a reflector is independent of the length of the object itself. The proposed analytical approach is validated against Monte Carlo simulations, and numerical illustrations are given and discussed.

4.1 Introduction

Future wireless networks will be more than allowing people, mobile devices, and objects to communicate [53]. Future wireless networks will be turned into a distributed intelligent wireless communications, sensing, and computing platform, which, besides communications, will be capable of sensing the environment to provide context-awareness capabilities, of locally storing and processing data to enable its time critical and energy efficient delivery, of accurately localizing people and objects in harsh propagation environments. Future wireless networks will have to fulfill the challenging requirement of interconnecting the physical and digital worlds in a seamless and sustainable manner [54], [55].

To fulfill these challenging requirements, it is apparent that it is not sufficient anymore to rely solely on wireless networks whose logical operation is software-controlled and optimized [56]. The wireless environment itself needs to be turned into a software-reconfigurable entity [57], whose operation is optimized to enable uninterrupted connectivity. Future wireless networks need a smart radio environment, i.e., a wireless environment that is turned into a reconfigurable space that plays an active role in transferring and processing information.

Different solutions towards realizing this wireless future are currently emerging [58]-[66]. Among them, the use of reconfigurable metasurfaces constitutes a promising and enabling solution to fulfill the challenging requirements of future wireless networks [67]. Metasurfaces are thin metamaterial layers that are capable of modifying the propagation of the radio waves in fully customizable ways [68], thus owing the potential of making the transfer and processing of information more reliable [69]. Also, they constitute a suitable distributed platform to perform low-energy and low-complexity sensing [70], storage [71], and analog computing [72]. For this reason, they are particularly useful for improving the performance of non-line-of-sight transmission, e.g., to appropriately customize the impact of multipath propagation.

In [66], in particular, the authors have put forth a network scenario where every environmental object is coated with reconfigurable metasurfaces, whose response to the radio waves is programmed in software by capitalizing on the enabling technology and hardware platform currently being developed in [73]. Current research efforts towards realizing this vision are, however, limited to implement hardware testbeds, e.g., reflect-arrays and metasurfaces, and on realizing point-to-point experimental tests [58]-[66]. To the best of the authors knowledge, notably, there exists no analytical framework that investigates the performance of large-scale wireless networks in the presence of reconfigurable metasurfaces. In the present chapter, motivated by these considerations, we develop the first analytical approach that allows one to study the probability that a random object coated with a reconfigurable metasurface acts as a reflector according to the generalized laws of reflection [68]. To this end, we capitalize on the mathematical tool of random spatial processes [23], [74].

Random spatial processes are considered to be the most suitable analytical tool to shed light on the ultimate performance limits of innovative technologies when applied in wireless networks, and to guide the design of optimal algorithms and protocols for attaining such ultimate limits. Several recent results on the application of random spatial processes in wireless networks can be found in [17], [75]-[80]. Despite the many results available, however, fundamental issues remain open [34]. In the current literature, in particular, the environmental objects are modeled as entities that can only attenuate the

Table 4.1 – Main symbols and functions used throughout the chapter.

Symbol/Function	Definition
$\Pr\{A\}$	Probability of Event A
$\Pr\left\{\overline{A}\right\}$	Probability of complement of Event A
$H(\cdot)$, $\bar{H}(\cdot)$	Heaviside function, complementary Heaviside function
$\left(x_{\text{Tx}}, y_{\text{Tx}}\right)$	Location of the transmitter
$\left(x_{\text{Rx}}, y_{\text{Rx}}\right)$	Location of the receiver
$\left(x_{\text{object}}, y_{\text{object}}\right)$	Location of the center of the typical object
$\left(x_{\text{end1}}, y_{\text{end1}}\right)$, $\left(x_{\text{end2}}, y_{\text{end2}}\right)$	Coordinates of the end points of the typical object
L	Length of the typical object
R_{net}	Radius of the network

signals, by making the links either line-of-sight or non-line-of-sight, e.g., [76], [25]. Modeling anything else is acknowledged to be difficult. Just in [81], the authors have recently investigated the impact of reflections, but only based on conventional Snell's laws. This work highlights the analytical complexity, the relevance, and the non-trivial performance trade-offs: The authors emphasize that the obtained trends highly depend on the fact that the total distance of the reflected paths is almost always two times larger than the distance of the direct paths. This occurs because the angles of incidence and reflection are the same based on Snell's law. In the presence of reconfigurable metasurfaces, on the other hand, the random objects can optimize the reflected signals in anomalous directions beyond Snell's law. The corresponding achievable performance and the associated optimal setups are unknown.

Motivated by these considerations, we develop an analytical framework that allows one to quantify the probability that a random object coated with reconfigurable metasurfaces acts as a reflector for a given pair of transmitter and receiver. Even though reconfigurable metasurfaces can be used to control and customize the refractions from environmental objects, in the present chapter we focus our attention on controlling and customizing only the reflections of signals, since refractions may be subject to severe signal's attenuation. Our proposed approach, in particular, is based on modeling the environmental objects with a modified random line process of fixed length, and with random orientations and locations. In contrast to the conventional network setup where the environmental objects are not coated with reconfigurable metasurfaces, we prove that the probability that the typical random object acts as a reflector is independent of the length of the object itself. The proposed analytical approach is validated against Monte Carlo simulations, and numerical illustrations are given and discussed. In the present chapter, we limit ourselves to analyze a 2D network scenarios, but our approach can be applied to 3D network topologies as well. This non-trivial generalization is postponed to a future research work.

The remainder of the present chapter is organized as follows. In Section 3, the system model is introduced. In Section 4, the problem is formulated in mathematical terms. In Section 5, the analytical framework of the reflection probability is described. In Section 6, numerical results are illustrated, and the proposed approach is validated against Monte Carlo simulations. Finally, Section 7 concludes the chapter.

Notation: The main symbols and functions used in this chapter are reported in Table 4.1.

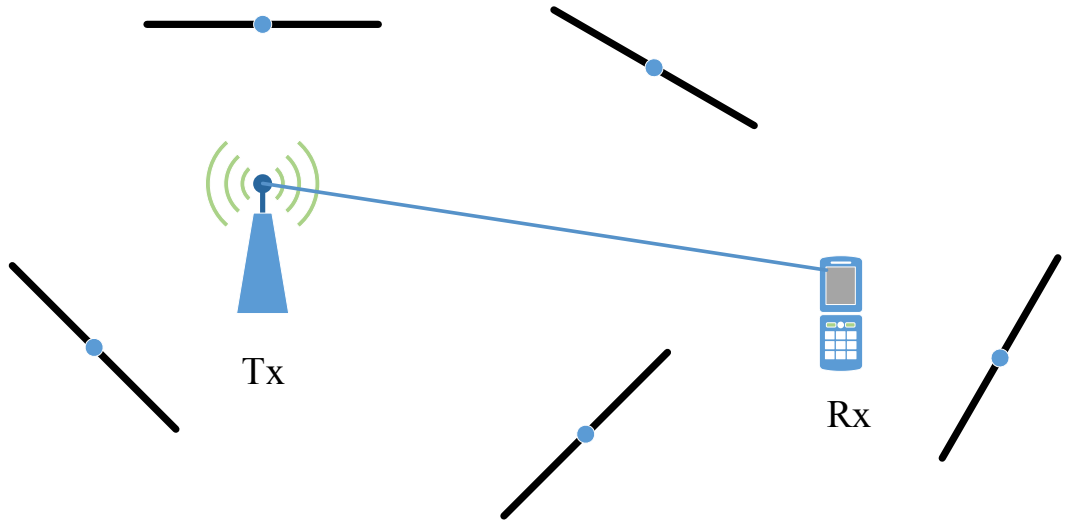

Figure 4.1 – Probe transmitter (Tx) and receiver (Rx) in the presence of randomly distributed environmental objects.

4.2 System Model

We consider a wireless network on a bi-dimensional plane, where the transmitters and receivers are distributed independently of each other. Without loss of generality, the location of the transmitter and receiver of interest, i.e., the probe transmitter and receiver, are denoted by $(x_{\mathrm{Tx}}, y_{\mathrm{Tx}})$ and $(x_{\mathrm{Rx}}, y_{\mathrm{Rx}})$, respectively.

Besides the transmitter and receiver, we assume that environmental objects, e.g., buildings in a urban outdoor scenario, are randomly distributed in the same region. An example of the network model is depicted in Fig. 4.1. More precisely, the environmental objects are assumed to follow a Boolean model of line segments with the following properties [23]:

- The center points of the objects form a homogeneous Poisson point process.
- The orientation of the objects are independent and identically distributed in $[0, 2\pi]$.
- The lengths of the objects are fixed and all equal to L.
- The random orientation and the center points of the objects are independent of each other.

We consider a generic environmental object, i.e., the typical object, and denote its center by $(x_{\mathrm{object}}, y_{\mathrm{object}})$, and the coordinates of its two end points by $(x_{\mathrm{end1}}, y_{\mathrm{end1}})$ and $(x_{\mathrm{end2}}, y_{\mathrm{end2}})$.

4.3 Problem Formulation

The objective of the present chapter is to compute the probability that a randomly distributed object can act as a reflector for the pair of transmitter and receiver located in $(x_{\mathrm{Tx}}, y_{\mathrm{Tx}})$ and $(x_{\mathrm{Rx}}, y_{\mathrm{Rx}})$, respec-

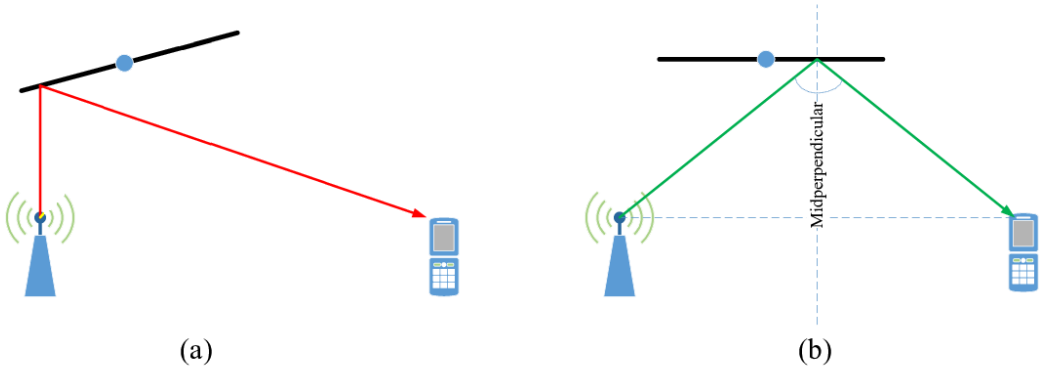

Figure 4.2 – Scenario I (a): In the presence of reconfigurable metasurfaces, the angle of incidence and reflection are not necessarily the same. Scenario II (b): According to Snell's law of reflection, the angle of incidence and reflection are the same.

tively. We analyze two case studies:

- **Scenario I**: The first scenario corresponds to the case study where the typical object is coated with a reconfigurable metasurface, which can optimize the angle of reflection regardless of the angle of incidence [68].
- **Scenario II**: The second scenario corresponds to the case study where the typical object is not coated with a reconfigurable metasurface. This is the state-of-the-art scenario, where the angle of reflection needs to be equal to the angle of incidence according to Snell's law of reflection [68].

The aim of the present chapter is to develop a mathematical theory to compute the probability that the typical object can act as a reflector, i.e., the reflection probability, and to quantify the gain of adding reconfigurable metasurfaces in wireless networks. For analytical tractability, we assume that the reconfigurable metasurfaces are capable of producing any angle of reflection for any given location of transmitter and receiver, for any angle of incidence, and for any length. This yields the best-case performance bound compared with conventional Snell's law of reflection. The analysis, in addition, is conducted by relying on ray tracing arguments, in order to highlight the potential gains of using reconfigurable metasurfaces. The two case studies are sketched in Fig. 4.2. Generalizations of the proposed analytical framework are left to future research works.

4.3.1 Scenario I: Reflections in the Presence of Reconfigurable Metasurfaces

In the presence of reconfigurable metasurfaces, an arbitrary angle of reflection can be obtained for any angle of incidence. This implies that the typical object acts as a reflector for a transmitter and receiver if they are both located on the same side of the infinite line passing through the end points $(x_{\text{end1}}, y_{\text{end1}})$ and $(x_{\text{end2}}, y_{\text{end2}})$ of the typical object.

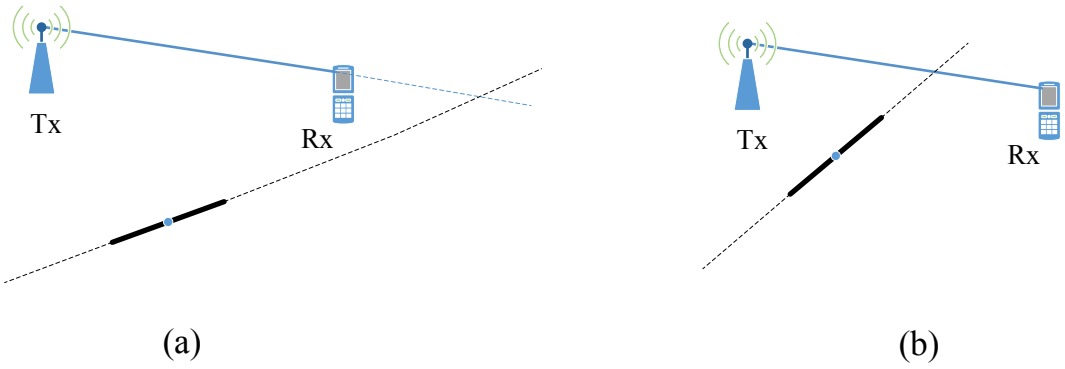

Figure 4.3 – Illustration of Event 1 based on Approach 1. (a) Event 1 holds true if the intersection point between the infinite lines falls outside the Tx-Rx line segment. (b) Otherwise, Event 1 does not hold true.

For ease of exposition, we introduce the following event.

Event 4.1 *The probe transmitter, Tx, and receiver, Rx, are located on the same side of the infinite line passing through the end points* $(x_{\text{end1}}, y_{\text{end1}})$ *and* $(x_{\text{end2}}, y_{\text{end2}})$ *of the typical object.*

Therefore, the typical object acts as a reflector if Event 4.1 holds true. Our objective is to formulate the probability of Event 4.1, i.e., to compute $\Pr\{\text{Event 1}\}$. This latter probability can be formulated in two different but equivalent ways.

Approach 1

Let us consider the infinite line that connects the locations of transmitter and receiver, and the infinite line that connects the two end points of the typical object. Event 1 holds true if the intersection point, denoted by (x^*, y^*), of these two infinite lines falls outside the line segment that connects that transmitter and the receiver. An illustration is given in Fig. 4.3.

In mathematical terms, therefore, $\Pr\{\text{Event 1}\}$ can be formulated as follows:

$$\Pr\{\text{Event 1}\} = 1 - \Pr\left\{\overline{\text{Event 1}}\right\} \tag{4.1}$$

where $\overline{\text{Event 1}}$ is the complement of Event 1, and $\Pr\left\{\overline{\text{Event 1}}\right\}$ denotes the probability that the intersection point (x^*, y^*) is on the Tx-Rx line segment:

$$\Pr\left\{\overline{\text{Event 1}}\right\} = \Pr\left\{\begin{matrix}\min(x_{\text{Tx}}, x_{\text{Rx}}) \leqslant x^* \leqslant \max(x_{\text{Tx}}, x_{\text{Rx}}) \\ \cap\ \min(y_{\text{Tx}}, y_{\text{Rx}}) \leqslant y^* \leqslant \max(y_{\text{Tx}}, y_{\text{Rx}})\end{matrix}\right\} \tag{4.2}$$

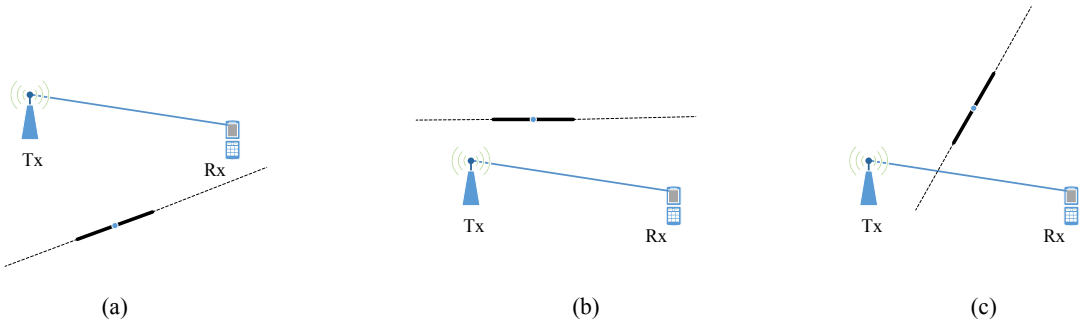

Figure 4.4 – Illustration of Event 1 based on Approach 2. Event 1 holds true if both Tx and Rx are above (a) or below (b) the infinite line corresponding to the typical object. Event 1 does not hold true if Tx and Rx are not on the same side of the line (c).

Approach 2

The probability that Event 4.1 holds true can be formulated also in terms of the positions of the transmitter and receiver with respect to the infinite line that connects the end points of the typical object. In particular, Event 4.1 holds true if both the transmitter and receiver are on the same side of the infinite line, i.e., either they are above or they are below the infinite line. This interpretation can be viewed as a problem of classifying points with respect to a line. This interpretation is depicted in Fig. 4.4.

In mathematical terms, therefore, $\Pr\{\text{Event 1}\}$ can be formulated as follows:

$$\Pr\{\text{Event 1}\} = \Pr\left\{\begin{matrix}\{[\text{Tx is above the line}] \ \cap \ [\text{Rx is above the line}]\} \\ \cup \ \{[\text{Tx is below the line}] \ \cap \ [\text{Rx is below the line}]\}\end{matrix}\right\} \tag{4.3}$$

4.3.2 Scenario II: Reflections in the Absence of Reconfigurable Metasurfaces

In the absence of reconfigurable metasurfaces, the typical object acts as a reflector, for a given transmitter and receiver, only if the angles of reflection and incidence are the same. This is agreement with Snell's law of reflection, and imposes some geometric constraints among the locations of the typical object, the transmitter, and the receiver. In order to compute the corresponding probability of occurrence, we introduce the following event.

Event 4.2 *The mid-perpendicular of the line segment that connects the transmitter and receiver intersects the line segment that represents the typical object.*

Let (x_*, y_*) denote the intersection between the mid-perpendicular of the line segment that connects the transmitter and receiver, and the line segment that represents the typical object. According to Snell's law of reflection, for some given locations of the transmitter and receiver, the typical object acts as a reflector if the mid-perpendicular of the line segment that connects the transmitter and receiver

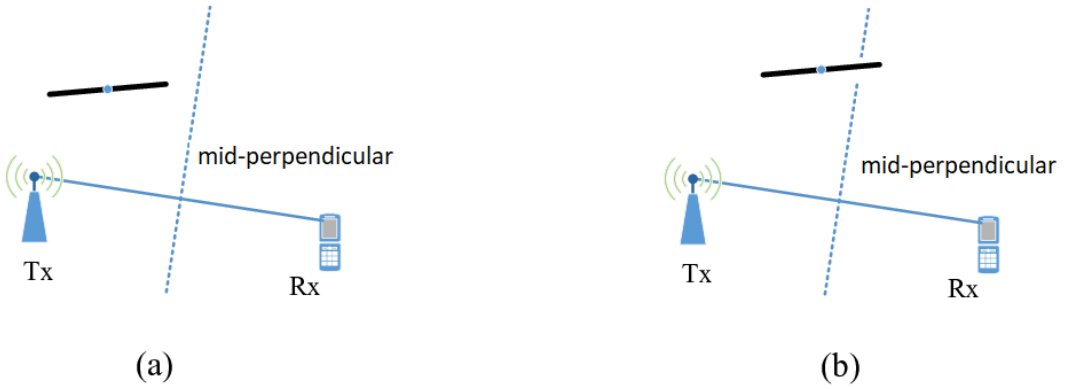

Figure 4.5 – Illustration of Event 3. In (a), Event 4.1 and Event 4.2 hold true: The typical object acts as a reflector. In (b), Event 4.1 holds true but Event 4.2 does not hold true: The typical object cannot be a reflector.

intersects the line segment that represents the typical object (i.e., Event 4.2), and, at the same time, the transmitted and receiver are located on the same side of the infinite line passing through the end points of the typical object (i.e., Event 4.1).

In mathematical terms, the probability of occurrence of Event 4.2 can be formulated as follows:

$$\Pr\{\text{Event 2}\} = \Pr\left\{\begin{matrix}\min(x_{\text{end1}}, x_{\text{end2}}) \leqslant x_* \leqslant \max(x_{\text{end1}}, x_{\text{end2}}) \\ \cap\ \min(y_{\text{end1}}, y_{\text{end2}}) \leqslant y_* \leqslant \max(y_{\text{end1}}, y_{\text{end2}})\end{matrix}\right\} \tag{4.4}$$

Based on Snell's law of reflection, therefore, the typical object acts a reflector if the following event holds true.

Event 4.3 *The transmitter and receiver are located on the same side of the infinite line passing through the end points of the typical object, and the mid-perpendicular of the line segment that connects the transmitter and receiver intersects the line segment that represents the typical object.*

An illustration is given in Fig. 4.5. In mathematical terms, the probability of occurrence of Event 4.1 can be formulated as follows:

$$\Pr\{\text{Event 3}\} = \Pr\{\text{Event 1}\ \cap\ \text{Event 2}\} \tag{4.5}$$

4.4 Analytical Formulation of the Reflection Probability

In this section, we introduce analytical expressions of the probability of occurrence of the three events introduced in the previous sections, and, therefore, characterize the probability that the typical object

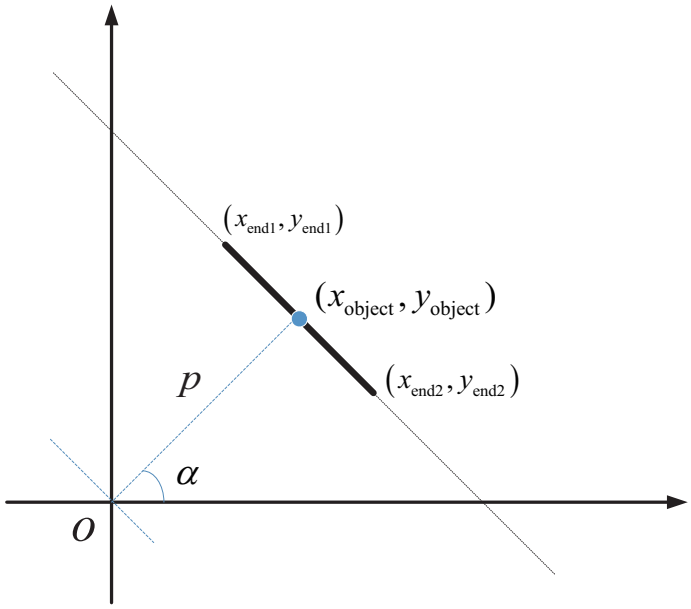

Figure 4.6 – Geometric representation of the typical object in polar coordinates.

acts as a reflector in the presence and in the absence of reconfigurable metasurfaces. First, we begin with some preliminary results.

4.4.1 Preliminary Results

Lemma 4.1 *Let* $(x_{\text{Tx}}, y_{\text{Tx}})$ *and* $(x_{\text{Rx}}, y_{\text{Rx}})$ *be the locations of the probe transmitter and receiver, respectively. The infinite line passing through them can be formulated as follows:*

$$y = mx + z \tag{4.6}$$

where $m = \frac{y_{\text{Tx}} - y_{\text{Rx}}}{x_{\text{Tx}} - x_{\text{Rx}}}$, *and* $z = y_{\text{Rx}} - m x_{\text{Rx}}$.

Proof: It follows by definition of line passing through two points. □

Lemma 4.2 *Let us consider the typical object of length L depicted in Fig. 4.6. The distance between the center of the line segment and the origin is* $p = R_{\text{net}}\sqrt{u}$, *where u is a uniformly distributed random variable in* $[0,1]$, *and* R_{net} *is largest size of the region of interest. Let* α *be the angle between the perpendicular line to the line segment, which passes through the center of the object, and the horizontal axis. The infinite line passing through the end points of the object can be formulated as follows:*

$$x\cos\alpha + y\sin\alpha = p \tag{4.7}$$

where $\alpha \in [0, 2\pi]$.

In addition, the center of the line segment can be written as $\left(x_{\text{object}}, y_{\text{object}}\right) = \left(p\cos\alpha, p\sin\alpha\right)$*, and its end points* $\left(x_{\text{end1}}, y_{\text{end1}}\right)$ *and* $\left(x_{\text{end2}}, y_{\text{end2}}\right)$ *can be formulated as follows:*

$$\begin{aligned} x_{\text{end1}} &= x_{\text{object}} - \frac{L}{2}\sin\alpha,\ y_{\text{end1}} = y_{\text{object}} + \frac{L}{2}\cos\alpha \\ x_{\text{end2}} &= x_{\text{object}} + \frac{L}{2}\sin\alpha,\ y_{\text{end2}} = y_{\text{object}} - \frac{L}{2}\cos\alpha \end{aligned} \tag{4.8}$$

Proof: The proof follows by noting that the centers of the line segments (the objects) are distributed according to a Poisson point process with random orientations, which implies $p = R_{\text{net}}\sqrt{u}$ *and* $\alpha \in [0, 2\pi]$*. The rest follows from geometric considerations.* □

Lemma 4.3 *The mid-perpendicular of the infinite line in* (4.6) *is as follows:*

$$y = m_p x + z_p \tag{4.9}$$

where $m_p = -\frac{1}{m}$*, and* $z_p = \frac{1}{2m}\left(x_{\text{Tx}} + x_{\text{Rx}}\right) + \frac{1}{2}\left(y_{\text{Tx}} + y_{\text{Rx}}\right)$.

Proof: See Appendix A. □

Lemma 4.4 *The intersection point between the infinite line that connects the transmitter and the receiver, and the infinite line that connects the end points of the line segment representing the typical object can be formulated as follows:*

$$\begin{aligned} x^* &= \frac{p - z\sin\alpha}{m\sin\alpha + \cos\alpha} \\ y^* &= mx^* + z \end{aligned} \tag{4.10}$$

The intersection point between the (infinite) mid-perpendicular line to the line segment that connects the transmitter and the receiver, and the infinite line that connects the end points of the line segment representing the typical object can be formulated as follows:

$$\begin{aligned} x_* &= \frac{p - z_p\sin\alpha}{m_p\sin\alpha + \cos\alpha} \\ y_* &= m_p x_* + z_p \end{aligned} \tag{4.11}$$

Proof: Equation (4.10) *follows by solving the system of equations in* (4.6) *and* (4.7)*. Equation* (4.11) *follows by solving the system of equations in* (4.7) *and* (4.9)*.* □

Lemma 4.5 *Let a generic infinite line formulated as:* $ax + by + c = 0$*. The following holds true:*

Table 4.2 – Auxiliary functions used in Theorem 4.1.

Function definition
$f(\alpha,\xi)=\frac{1}{R_{\text{net}}}\left([m\sin\alpha+\cos\alpha]\,\xi+z\sin\alpha\right)$
$g(\alpha,\omega)=\frac{1}{R_{\text{net}}}\left(\frac{[m\sin\alpha+\cos\alpha][\omega-z]}{m}+z\sin\alpha\right)$
$\Theta\left(\alpha\middle\vert\begin{array}{cc}\mu_1&\mu_2\\ \mu_3&\mu_4\end{array}\right)=\left[\left(\min\{\mu_1,\mu_2,1\}\right)^2-\left(\max\{\mu_3,\mu_4,0\}\right)^2\right]H\left(\min\{\mu_1,\mu_2,1\}-\max\{\mu_3,\mu_4,0\}\right)$
$\theta_1(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}})=\Theta\left(\alpha\middle\vert\begin{array}{cc}f(\alpha,\max(x_{\text{Tx}},x_{\text{Rx}}))&g(\alpha,\max(y_{\text{Tx}},y_{\text{Rx}}))\\ f(\alpha,\min(x_{\text{Tx}},x_{\text{Rx}}))&g(\alpha,\min(y_{\text{Tx}},y_{\text{Rx}}))\end{array}\right)\times H(m)$
$\theta_2(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}})=\Theta\left(\alpha\middle\vert\begin{array}{cc}f(\alpha,\max(x_{\text{Tx}},x_{\text{Rx}}))&g(\alpha,\min(y_{\text{Tx}},y_{\text{Rx}}))\\ f(\alpha,\min(x_{\text{Tx}},x_{\text{Rx}}))&g(\alpha,\max(y_{\text{Tx}},y_{\text{Rx}}))\end{array}\right)\times \bar{H}(m)$
$\theta_3(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}})=\Theta\left(\alpha\middle\vert\begin{array}{cc}f(\alpha,\min(x_{\text{Tx}},x_{\text{Rx}}))&g(\alpha,\min(y_{\text{Tx}},y_{\text{Rx}}))\\ f(\alpha,\max(x_{\text{Tx}},x_{\text{Rx}}))&g(\alpha,\max(y_{\text{Tx}},y_{\text{Rx}}))\end{array}\right)\times H(m)$
$\theta_4(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}})=\Theta\left(\alpha\middle\vert\begin{array}{cc}f(\alpha,\min(x_{\text{Tx}},x_{\text{Rx}}))&g(\alpha,\max(y_{\text{Tx}},y_{\text{Rx}}))\\ f(\alpha,\max(x_{\text{Tx}},x_{\text{Rx}}))&g(\alpha,\min(y_{\text{Tx}},y_{\text{Rx}}))\end{array}\right)\times \bar{H}(m)$

- *The point* (x_1,y_1) *is above the line if* $ax_1+by_1+c>0$ *and* $b>0$*, or if* $ax_1+by_1+c<0$ *and* $b<0$.
- *The point* (x_1,y_1) *is below the line if* $ax_1+by_1+c<0$ *and* $b>0$*, or if* $ax_1+by_1+c>0$ *and* $b<0$.

Proof: See Appendix B. □

4.4.2 Scenario I: Reflection Probability in the Presence of Reconfigurable Metasurfaces

Theorem 4.1 and Theorem 4.2 provide one with analytical expressions of the probability that the typical object acts as a reflector if it is coated with reconfigurable metasurfaces. Theorem 4.1 is computed based on Approach 1, and Theorem 4.2 based on the Approach 2.

Theorem 4.1 *Based on Approach 1, the probability of occurrence of Event 4.1 is as follows:*

$$\begin{aligned}&\Pr\{\text{Event1}\}=1-\Pr\left\{\overline{\text{Event1}}\right\}\\&=1-\frac{1}{2\pi}\left\{\begin{array}{l}\displaystyle\int_0^{\delta_1}\theta_1(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}})d\alpha+\int_{\delta_2}^{2\pi}\theta_1(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}})d\alpha\\ \displaystyle+\int_0^{\delta_1}\theta_2(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}})d\alpha+\int_{\delta_2}^{2\pi}\theta_2(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}})d\alpha\\ \displaystyle+\int_{\delta_1}^{\delta_2}\theta_3(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}})d\alpha++\int_{\delta_1}^{\delta_2}\theta_4(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}})d\alpha\end{array}\right\}\end{aligned}\tag{4.12}$$

where the integral limits are defined as $\delta_1=2\tan^{-1}\left(m+\sqrt{1+m^2}\right)$*, and* $\delta_2=2\pi+2\tan^{-1}\left(m-\sqrt{1+m^2}\right)$*, and the auxiliary functions are given in Table 4.2.*

Proof: See Appendix C. □

Theorem 4.2 *Based on Approach 2, the probability of occurrence of Event 4.1 is as follows:*

$$\begin{aligned}&\Pr\{\text{Event1}\}\\&=\frac{1}{2\pi}\int_0^{2\pi}\rho_1\left(\alpha,x_{\text{Tx}},y_{\text{Tx}},x_{\text{Rx}},y_{\text{Rx}}\right)d\alpha+\frac{1}{2\pi}\int_0^{2\pi}\rho_2\left(\alpha,x_{\text{Tx}},y_{\text{Tx}},x_{\text{Rx}},y_{\text{Rx}}\right)d\alpha\end{aligned}\tag{4.13}$$

where the integrand functions are defined as:

$$\begin{aligned}&\rho_1\left(\alpha,x_{\text{Tx}},y_{\text{Tx}},x_{\text{Rx}},y_{\text{Rx}}\right)=\left[\min\left\{\frac{x_{\text{Tx}}\cos\alpha+y_{\text{Tx}}\sin\alpha}{R_{\text{net}}},\frac{x_{\text{Rx}}\cos\alpha+y_{\text{Rx}}\sin\alpha}{R_{\text{net}}},1\right\}\right]^2\\&\times H\left(\min\left\{\frac{x_{\text{Tx}}\cos\alpha+y_{\text{Tx}}\sin\alpha}{R_{\text{net}}},\frac{x_{\text{Rx}}\cos\alpha+y_{\text{Rx}}\sin\alpha}{R_{\text{net}}},1\right\}\right)\end{aligned}\tag{4.14}$$

and

$$\begin{aligned}&\rho_2\left(\alpha,x_{\text{Tx}},y_{\text{Tx}},x_{\text{Rx}},y_{\text{Rx}}\right)=\left[1-\left(\max\left\{\frac{x_{\text{Tx}}\cos\alpha+y_{\text{Tx}}\sin\alpha}{R_{\text{net}}},\frac{x_{\text{Rx}}\cos\alpha+y_{\text{Rx}}\sin\alpha}{R_{\text{net}}},0\right\}\right)^2\right]\\&\times H\left(1-\max\left\{\frac{x_{\text{Tx}}\cos\alpha+y_{\text{Tx}}\sin\alpha}{R_{\text{net}}},\frac{x_{\text{Rx}}\cos\alpha+y_{\text{Rx}}\sin\alpha}{R_{\text{net}}},0\right\}\right)\end{aligned}\tag{4.15}$$

Proof: See Appendix D. □

Remark 4.1 *Theorem 4.1 and Theorem 4.2 are two analytical formulations of the same event. In the sequel, we show that they coincide.* □

4.4.3 Scenario II: Reflection Probability in the Absence of Reconfigurable Metasurfaces

The probability of occurrence of Event 4.3 is not easy to compute. The reason is that Event 4.3 is formulated in terms of the intersection of Event 4.1 and Event 4.2, which are not independent. In order to avoid the analytical complexity that originates from the correlation between Event 4.1 and Event 4.2, we propose a upper-bound to compute the probability of occurrence of Event 4.3. Before stating the main result, we introduce the following proposition that provides one with the probability of occurrence of Event 4.2.

Proposition 4.1 *The probability of occurrence of Event 4.2 can be formulated as follows.*

$$\begin{aligned}&\Pr\{\text{Event2}\}=\Pr\left\{\begin{matrix}\min\left(x_{\text{end1}},x_{\text{end2}}\right)\leqslant x_*\leqslant\max\left(x_{\text{end1}},x_{\text{end2}}\right)\\\cap\min\left(y_{\text{end1}},y_{\text{end2}}\right)\leqslant y_*\leqslant\max\left(y_{\text{end1}},y_{\text{end2}}\right)\end{matrix}\right\}\\&=\frac{1}{2\pi}\left\{\int_{\frac{3\pi}{2}}^{2\pi}\Gamma_1\left(\alpha\right)d\alpha+\int_{\pi}^{\frac{3\pi}{2}}\Gamma_2\left(\alpha\right)d\alpha+\int_0^{\frac{\pi}{2}}\Gamma_3\left(\alpha\right)d\alpha+\int_{\frac{\pi}{2}}^{\pi}\Gamma_4\left(\alpha\right)d\alpha\right\}\end{aligned}\tag{4.16}$$

where $\Gamma_1\left(\alpha\right)=\Gamma_1^a\left(\alpha\right)+\Gamma_1^b\left(\alpha\right)+\Gamma_1^c\left(\alpha\right)+\Gamma_1^d\left(\alpha\right)$, $\Gamma_2\left(\alpha\right)=\Gamma_2^a\left(\alpha\right)+\Gamma_2^b\left(\alpha\right)+\Gamma_2^c\left(\alpha\right)+\Gamma_2^d\left(\alpha\right)$, $\Gamma_3\left(\alpha\right)=\Gamma_3^a\left(\alpha\right)+\Gamma_3^b\left(\alpha\right)+\Gamma_3^c\left(\alpha\right)+\Gamma_3^d\left(\alpha\right)$, *and* $\Gamma_4\left(\alpha\right)=\Gamma_4^a\left(\alpha\right)+\Gamma_4^b\left(\alpha\right)+\Gamma_4^c\left(\alpha\right)+\Gamma_4^d\left(\alpha\right)$, *which are all defined in Table 4.3.*

Table 4.3 – Auxiliary functions used in Proposition 4.1.

Function definition
$F(\alpha,t)=\frac{1}{R_{\text{net}}}\left(t+\frac{z_p\sin\alpha}{m_p\sin\alpha+\cos\alpha}\right)\left(\frac{1}{m_p\sin\alpha+\cos\alpha}-\cos\alpha\right)^{-1}$
$G(\alpha,v)=\frac{1}{R_{\text{net}}}\left(v+\frac{m_p z_p\sin\alpha}{m_p\sin\alpha+\cos\alpha}-z_p\right)\left(\frac{m_p}{m_p\sin\alpha+\cos\alpha}-\sin\alpha\right)^{-1}$
$\Gamma_1^a(\alpha)=\Theta\left(\alpha\left\vert\begin{matrix}F\left(\alpha,-\frac{L}{2}\sin\alpha\right) & G\left(\alpha,\frac{L}{2}\cos\alpha\right)\\ F\left(\alpha,\frac{L}{2}\sin\alpha\right) & G\left(\alpha,-\frac{L}{2}\cos\alpha\right)\end{matrix}\right.\right)H\left(\frac{1}{m_p\sin\alpha+\cos\alpha}-\cos\alpha\right)H\left(\frac{m_p}{m_p\sin\alpha+\cos\alpha}-\sin\alpha\right)$
$\Gamma_1^b(\alpha)=\Theta\left(\alpha\left\vert\begin{matrix}F\left(\alpha,-\frac{L}{2}\sin\alpha\right) & G\left(\alpha,-\frac{L}{2}\cos\alpha\right)\\ F\left(\alpha,\frac{L}{2}\sin\alpha\right) & G\left(\alpha,\frac{L}{2}\cos\alpha\right)\end{matrix}\right.\right)H\left(\frac{1}{m_p\sin\alpha+\cos\alpha}-\cos\alpha\right)\bar{H}\left(\frac{m_p}{m_p\sin\alpha+\cos\alpha}-\sin\alpha\right)$
$\Gamma_1^c(\alpha)=\Theta\left(\alpha\left\vert\begin{matrix}F\left(\alpha,\frac{L}{2}\sin\alpha\right) & G\left(\alpha,\frac{L}{2}\cos\alpha\right)\\ F\left(\alpha,-\frac{L}{2}\sin\alpha\right) & G\left(\alpha,-\frac{L}{2}\cos\alpha\right)\end{matrix}\right.\right)\bar{H}\left(\frac{1}{m_p\sin\alpha+\cos\alpha}-\cos\alpha\right)H\left(\frac{m_p}{m_p\sin\alpha+\cos\alpha}-\sin\alpha\right)$
$\Gamma_1^d(\alpha)=\Theta\left(\alpha\left\vert\begin{matrix}F\left(\alpha,\frac{L}{2}\sin\alpha\right) & G\left(\alpha,-\frac{L}{2}\cos\alpha\right)\\ F\left(\alpha,-\frac{L}{2}\sin\alpha\right) & G\left(\alpha,\frac{L}{2}\cos\alpha\right)\end{matrix}\right.\right)\bar{H}\left(\frac{1}{m_p\sin\alpha+\cos\alpha}-\cos\alpha\right)\bar{H}\left(\frac{m_p}{m_p\sin\alpha+\cos\alpha}-\sin\alpha\right)$
$\Gamma_2^a(\alpha)=\Theta\left(\alpha\left\vert\begin{matrix}F\left(\alpha,-\frac{L}{2}\sin\alpha\right) & G\left(\alpha,-\frac{L}{2}\cos\alpha\right)\\ F\left(\alpha,\frac{L}{2}\sin\alpha\right) & G\left(\alpha,\frac{L}{2}\cos\alpha\right)\end{matrix}\right.\right)H\left(\frac{1}{m_p\sin\alpha+\cos\alpha}-\cos\alpha\right)H\left(\frac{m_p}{m_p\sin\alpha+\cos\alpha}-\sin\alpha\right)$
$\Gamma_2^b(\alpha)=\Theta\left(\alpha\left\vert\begin{matrix}F\left(\alpha,-\frac{L}{2}\sin\alpha\right) & G\left(\alpha,\frac{L}{2}\cos\alpha\right)\\ F\left(\alpha,\frac{L}{2}\sin\alpha\right) & G\left(\alpha,-\frac{L}{2}\cos\alpha\right)\end{matrix}\right.\right)H\left(\frac{1}{m_p\sin\alpha+\cos\alpha}-\cos\alpha\right)\bar{H}\left(\frac{m_p}{m_p\sin\alpha+\cos\alpha}-\sin\alpha\right)$
$\Gamma_2^c(\alpha)=\Theta\left(\alpha\left\vert\begin{matrix}F\left(\alpha,\frac{L}{2}\sin\alpha\right) & G\left(\alpha,-\frac{L}{2}\cos\alpha\right)\\ F\left(\alpha,-\frac{L}{2}\sin\alpha\right) & G\left(\alpha,\frac{L}{2}\cos\alpha\right)\end{matrix}\right.\right)\bar{H}\left(\frac{1}{m_p\sin\alpha+\cos\alpha}-\cos\alpha\right)H\left(\frac{m_p}{m_p\sin\alpha+\cos\alpha}-\sin\alpha\right)$
$\Gamma_2^d(\alpha)=\Theta\left(\alpha\left\vert\begin{matrix}F\left(\alpha,\frac{L}{2}\sin\alpha\right) & G\left(\alpha,\frac{L}{2}\cos\alpha\right)\\ F\left(\alpha,-\frac{L}{2}\sin\alpha\right) & G\left(\alpha,-\frac{L}{2}\cos\alpha\right)\end{matrix}\right.\right)\bar{H}\left(\frac{1}{m_p\sin\alpha+\cos\alpha}-\cos\alpha\right)\bar{H}\left(\frac{m_p}{m_p\sin\alpha+\cos\alpha}-\sin\alpha\right)$
$\Gamma_3^a(\alpha)=\Theta\left(\alpha\left\vert\begin{matrix}F\left(\alpha,\frac{L}{2}\sin\alpha\right) & G\left(\alpha,\frac{L}{2}\cos\alpha\right)\\ F\left(\alpha,-\frac{L}{2}\sin\alpha\right) & G\left(\alpha,-\frac{L}{2}\cos\alpha\right)\end{matrix}\right.\right)H\left(\frac{1}{m_p\sin\alpha+\cos\alpha}-\cos\alpha\right)H\left(\frac{m_p}{m_p\sin\alpha+\cos\alpha}-\sin\alpha\right)$
$\Gamma_3^b(\alpha)=\Theta\left(\alpha\left\vert\begin{matrix}F\left(\alpha,\frac{L}{2}\sin\alpha\right) & G\left(\alpha,-\frac{L}{2}\cos\alpha\right)\\ F\left(\alpha,-\frac{L}{2}\sin\alpha\right) & G\left(\alpha,\frac{L}{2}\cos\alpha\right)\end{matrix}\right.\right)H\left(\frac{1}{m_p\sin\alpha+\cos\alpha}-\cos\alpha\right)\bar{H}\left(\frac{m_p}{m_p\sin\alpha+\cos\alpha}-\sin\alpha\right)$
$\Gamma_3^c(\alpha)=\Theta\left(\alpha\left\vert\begin{matrix}F\left(\alpha,-\frac{L}{2}\sin\alpha\right) & G\left(\alpha,\frac{L}{2}\cos\alpha\right)\\ F\left(\alpha,\frac{L}{2}\sin\alpha\right) & G\left(\alpha,-\frac{L}{2}\cos\alpha\right)\end{matrix}\right.\right)\bar{H}\left(\frac{1}{m_p\sin\alpha+\cos\alpha}-\cos\alpha\right)H\left(\frac{m_p}{m_p\sin\alpha+\cos\alpha}-\sin\alpha\right)$
$\Gamma_3^d(\alpha)=\Theta\left(\alpha\left\vert\begin{matrix}F\left(\alpha,-\frac{L}{2}\sin\alpha\right) & G\left(\alpha,-\frac{L}{2}\cos\alpha\right)\\ F\left(\alpha,\frac{L}{2}\sin\alpha\right) & G\left(\alpha,\frac{L}{2}\cos\alpha\right)\end{matrix}\right.\right)\bar{H}\left(\frac{1}{m_p\sin\alpha+\cos\alpha}-\cos\alpha\right)\bar{H}\left(\frac{m_p}{m_p\sin\alpha+\cos\alpha}-\sin\alpha\right)$
$\Gamma_4^a(\alpha)=\Theta\left(\alpha\left\vert\begin{matrix}F\left(\alpha,\frac{L}{2}\sin\alpha\right) & G\left(\alpha,-\frac{L}{2}\cos\alpha\right)\\ F\left(\alpha,-\frac{L}{2}\sin\alpha\right) & G\left(\alpha,\frac{L}{2}\cos\alpha\right)\end{matrix}\right.\right)H\left(\frac{1}{m_p\sin\alpha+\cos\alpha}-\cos\alpha\right)H\left(\frac{m_p}{m_p\sin\alpha+\cos\alpha}-\sin\alpha\right)$
$\Gamma_4^b(\alpha)=\Theta\left(\alpha\left\vert\begin{matrix}F\left(\alpha,\frac{L}{2}\sin\alpha\right) & G\left(\alpha,\frac{L}{2}\cos\alpha\right)\\ F\left(\alpha,-\frac{L}{2}\sin\alpha\right) & G\left(\alpha,-\frac{L}{2}\cos\alpha\right)\end{matrix}\right.\right)H\left(\frac{1}{m_p\sin\alpha+\cos\alpha}-\cos\alpha\right)\bar{H}\left(\frac{m_p}{m_p\sin\alpha+\cos\alpha}-\sin\alpha\right)$
$\Gamma_4^c(\alpha)=\Theta\left(\alpha\left\vert\begin{matrix}F\left(\alpha,-\frac{L}{2}\sin\alpha\right) & G\left(\alpha,-\frac{L}{2}\cos\alpha\right)\\ F\left(\alpha,\frac{L}{2}\sin\alpha\right) & G\left(\alpha,\frac{L}{2}\cos\alpha\right)\end{matrix}\right.\right)\bar{H}\left(\frac{1}{m_p\sin\alpha+\cos\alpha}-\cos\alpha\right)H\left(\frac{m_p}{m_p\sin\alpha+\cos\alpha}-\sin\alpha\right)$
$\Gamma_4^d(\alpha)=\Theta\left(\alpha\left\vert\begin{matrix}F\left(\alpha,-\frac{L}{2}\sin\alpha\right) & G\left(\alpha,\frac{L}{2}\cos\alpha\right)\\ F\left(\alpha,\frac{L}{2}\sin\alpha\right) & G\left(\alpha,-\frac{L}{2}\cos\alpha\right)\end{matrix}\right.\right)\bar{H}\left(\frac{1}{m_p\sin\alpha+\cos\alpha}-\cos\alpha\right)\bar{H}\left(\frac{m_p}{m_p\sin\alpha+\cos\alpha}-\sin\alpha\right)$

Proof*: See Appendix E.* □

Theorem 4.3 *The probability of occurrence of Event 4.3 is upper-bounded as follows:*

$$\Pr\{\text{Event3}\} \leqslant \min\{\Pr\{\text{Event1}\}, \Pr\{\text{Event2}\}\} \tag{4.17}$$

where $\Pr\{\text{Event1}\}$ *is formulated in Theorem 4.1 or Theorem 4.2, and* $\Pr\{\text{Event2}\}$ *is given in Proposition 4.1.*

Proof: The proof follows by applying the Frechet inequality [82]. □

Remark 4.2 *By comparing Theorem 4.1 and Theorem 4.2 against Theorem 4.3, we observe that the probability of being a reflector highly depends on the length of the typical object if is it not coated with a reconfigurable metasurfaces, while it is independent of it if it is coated with a reconfigurable metasurface. This is a major benefit of using reconfigurable metasurfaces in wireless networks. This outcome is determined by the assumption that the metasurfaces can modify the angle of reflection regardless of their length. The analysis of the impact of the constraints imposed by the size of the metasurface on its capability of obtaining a given set of angles of reflection as a function of the angle of incidence is an open but very important research issue, which is left to future research.* □

4.5 Numerical Results and Discussion: Validation Against Monte Carlo Simulations

The aim of this section is to validate the analytical frameworks developed in the previous sections against Monte Carlo simulations, and to study the potential of using reconfigurable metasurfaces in wireless networks. The results are illustrated either as a function of the length, L, of the typical object or as a function of the locations of the transmitter and receiver. The simulation setup is detailed in the caption of each figure.

In Fig. 4.7, we validate the proposed mathematical frameworks, against Monte Carlo simulations, as a function of the length, L, of the typical object. The results depicted in Fig. 4.7 confirm the good accuracy of the proposed analytical approach. More importantly, we observe the large gain that the presence of metasurfaces bring about: Especially for objects of small length, the presence of reconfigurable metasurfaces increases the probability of the typical object to be a reflector significantly. This is expected to bring major gains in terms of signal strength of the received signal thanks to the reflection generated by the randomly distributed reflectors. The presence of multiple reflectors, however, may also increase the level of interference. Therefore, the optimization of wireless networks in the presence of reconfigurable metasurfaces is a challenging and open research issue. Figure 4.7, in addition, confirms the main finding in Remark 4.2.

In Figs. 4.8 and 4.9, we depict the probability that the typical object is as a reflector as a function of the location of the transmitter and for a fixed length of the typical object. Once again, the proposed analytical frameworks are accurate, and the upper-bound in Theorem 4.3 is sufficiently accurate for the considered setups. Especially for small-size objects, we observe the large gains that employing

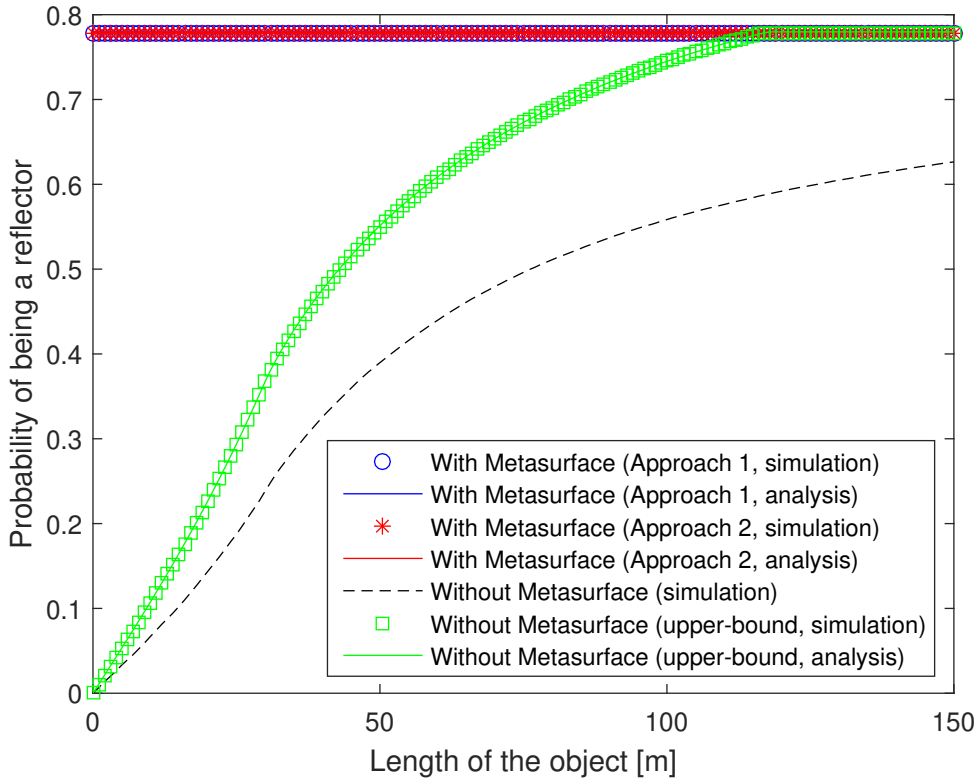

Figure 4.7 – Probability of being a reflector versus the length of the object. Setup: $R_{\text{net}} = 30$m, location of the transmitter $(0,3)$, location of the receiver $(20,20)$.

metasurfaces bring about. Even small-size objects can provide one with a relatively high probability of being a reflector, which is useful information in order to reduce the deployment cost of the metasurfaces over large-size environmental objects. Small-size metasurfaces, in fact, may be moved along large(size surfaces, and their location may be optimized in order to optimize the system performance.

4.6 Conclusion and Discussion

In this chapter, we have proposed the first analytical approach that provides one with the probability that a random object coated with reconfigurable metasurfaces acts as a reflector, and have compared it against the conventional setup in which the object is not coated with reconfigurable metasurfaces. This result has been obtained by modeling the environmental objects with a modified random line process with fixed length, and random orientations and locations. Our proposed analytical approach allows us to prove that the probability that an object is a reflector does not depend on the length of the object if it is coated with metasurfaces, while it strongly depends on it if the Snell's law of reflection needs to be applied. The reason of this major difference in system performance lies in the fact that the angles of incidence and reflection need to be the same according to the Snell's law of reflection.

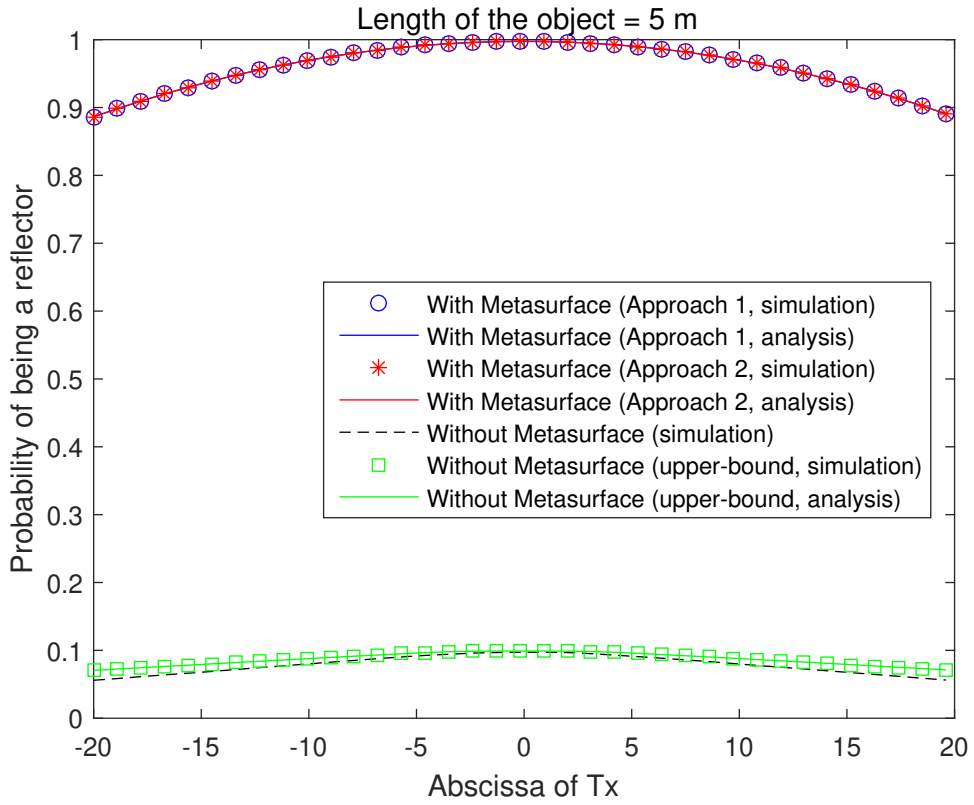

Figure 4.8 – Probability of being a reflector versus the horizontal location, x_{Tx}, of the transmitter. Setup: $R_{\text{net}} = 30$m, vertical location of the transmitter $y_{\text{Tx}} = 3$, location of the receiver $(0,0)$, length of the object $L = 5$m.

In spite of the novelty and contribution of the present chapter, it constitutes only a first attempt to quantify the potential of reconfigurable metasurfaces in large-scale wireless networks, and to develop a general analytical approach for understanding the ultimate performance limits, and to identify design guidelines for system optimization. For example, the performance trends are based on the assumption that, for any angle of incidence, an arbitrary angle of reflection can be synthetized. Due to practical constraints on implementing metasurfaces, only a finite subset of angles may be allowed, which needs to account for the concept of field-of-view of the metasurfaces. Also, the analytical models and the simulation results have been obtained by using ray tracing assumptions, and ignore, e.g., the radiation pattern of the metasurfaces, and near field effects. A major step is needed to obtain tractable analytical expressions of relevant performance metrics that are suitable to unveil scaling laws, amenable for optimization, and account for different functions applied by the metasurfaces (not just reflections).

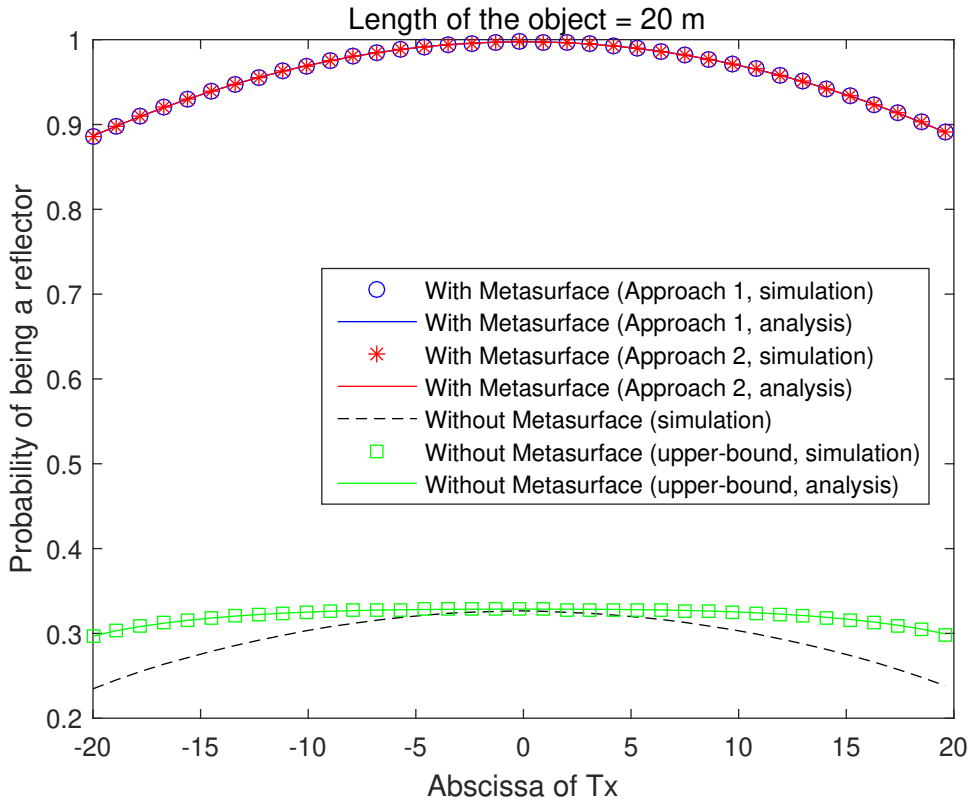

Figure 4.9 – Probability of being a reflector versus the horizontal location, x_{Tx}, of the transmitter. Setup: $R_{\text{net}} = 30$m, vertical location of the transmitter $y_{\text{Tx}} = 3$, location of the receiver $(0,0)$, length of the object $L = 20$m.

4.7 Appendices

4.7.1 Proof of Lemma 4.3

Let us rewrite (4.6) and (4.9) in the following standard forms:

$$\begin{aligned} mx - y + z &= 0 \\ m_p - y + z_p &= 0 \end{aligned} \tag{4.18}$$

where their slopes m and m_p are assumed to be non-zero.

With this formulation, the directional vector of each line is as follows:

$$\Delta = \begin{bmatrix} m \\ -1 \end{bmatrix}, \Delta_p = \begin{bmatrix} m_p \\ -1 \end{bmatrix} \tag{4.19}$$

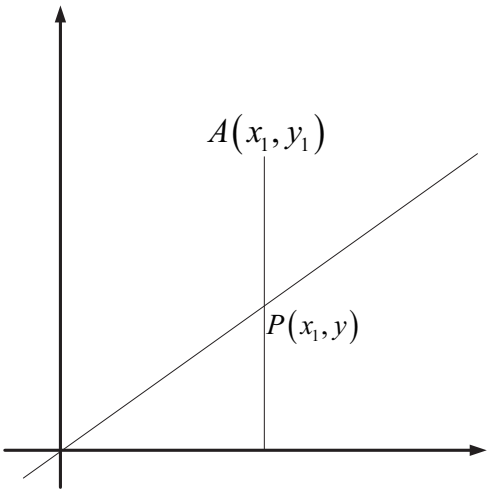

Figure 4.10 – Position of a point with respect to a line.

The two lines in (4.18) are perpendicular to each other if their directional vectors are orthogonal:

$$\Delta^T \Delta_p = 0 \tag{4.20}$$

which implies that the following identity need to be fulfilled $mm_p = -1$.

The coordinates of the mid-point of the line segment between the transmitter and receiver is $\left(\frac{x_{\text{Tx}}+x_{\text{Rx}}}{2}, \frac{y_{\text{Tx}}+y_{\text{Rx}}}{2}\right)$, which is located on the line $y = m_p x + z_p$. By substituting this latter point in (4.18), z_p turns out to be the following:

$$z_p = \frac{1}{2m}\left(x_{\text{Tx}} + x_{\text{Rx}}\right) + \frac{1}{2}\left(y_{\text{Tx}} + y_{\text{Rx}}\right) \tag{4.21}$$

This concludes the proof.

4.7.2 Proof of Lemma 4.5

Let $A(x_1, y_1)$ be a point that does not lie on a generic line $ax + by + c = 0$. Let us draw the perpendicular line from $A(x_1, y_1)$ to the horizontal axis. This perpendicular line intersect the generic line $ax + by + c = 0$ in the point $P(x_1, y)$, as shown in Fig. 4.10. By construction, x_1 is the abscissa of the point P. Let y denote the ordinate of P. Since P is on the line $ax + by + c = 0$, then the following equation needs to be satisfied:

$$ax_1 + by + c = 0 \Rightarrow y = -\frac{ax_1 + c}{b} \tag{4.22}$$

Let us compute the difference $y_1 - y$, as follows:

$$y_1 - y = y_1 - \left(-\frac{ax_1 + c}{b}\right) \Rightarrow y_1 - y = \frac{ax_1 + by_1 + c}{b} \tag{4.23}$$

Therefore, the following conclusions can be drawn:

- If the point A is above the line, then it must be $y_1 - y > 0$. From equation (4.23), we evince that $y_1 - y > 0$ if $\frac{ax_1+by_1+c}{b} > 0$. This, in turn, corresponds to the following: i) either $ax_1 + by_1 + c > 0$ and $b > 0$ or ii) $ax_1 + by_1 + c < 0$ and $b < 0$.

- If the point A is below the line, then it must be $y_1 - y < 0$. From equation (4.23), we evince that $y_1 - y < 0$ if $\frac{ax_1+by_1+c}{b} < 0$. This, in turn, corresponds to the following: i) either $ax_1 + by_1 + c < 0$ and $b > 0$ or ii) $ax_1 + by_1 + c > 0$ and $b < 0$.

This concludes the proof.

4.7.3 Proof of Theorem 4.1

From (4.10) in Lemma 4.4, the probability of the complement of Event 4.1 can be formulated as follows:

$$\begin{aligned}
&\Pr\left\{\overline{\text{Event 1}}\right\} = \Pr\begin{Bmatrix} \min\left(x_{\text{Tx}}, x_{\text{Rx}}\right) \leqslant x^* \leqslant \max\left(x_{\text{Tx}}, x_{\text{Rx}}\right) \\ \cap \min\left(y_{\text{Tx}}, y_{\text{Rx}}\right) \leqslant y^* \leqslant \max\left(y_{\text{Tx}}, y_{\text{Rx}}\right) \end{Bmatrix} \\
&= \Pr\begin{Bmatrix} \min\left(x_{\text{Tx}}, x_{\text{Rx}}\right) \leqslant \dfrac{p - z\sin\alpha}{m\sin\alpha + \cos\alpha} \leqslant \max\left(x_{\text{Tx}}, x_{\text{Rx}}\right) \\ \cap \min\left(y_{\text{Tx}}, y_{\text{Rx}}\right) \leqslant m\dfrac{p - z\sin\alpha}{m\sin\alpha + \cos\alpha} + z \leqslant \max\left(y_{\text{Tx}}, y_{\text{Rx}}\right) \end{Bmatrix}
\end{aligned} \tag{4.24}$$

Based on the sign of $m\sin\alpha + \cos\alpha$ and m, four cases can be identified.

Case 1

If $m\sin\alpha+\cos\alpha \geqslant 0$ and $m \geqslant 0$, we obtain the following:

$$\begin{aligned}
\Pr\left\{\overline{\text{Event 1}}\right\} &= \Pr\left\{\begin{matrix}\min\left(x_{\text{Tx}},x_{\text{Rx}}\right) \leqslant \dfrac{p-z\sin\alpha}{m\sin\alpha+\cos\alpha} \leqslant \max\left(x_{\text{Tx}},x_{\text{Rx}}\right) \\ \cap \min\left(y_{\text{Tx}},y_{\text{Rx}}\right) \leqslant m\dfrac{p-z\sin\alpha}{m\sin\alpha+\cos\alpha}+z \leqslant \max\left(y_{\text{Tx}},y_{\text{Rx}}\right)\end{matrix}\right\} \\
&\overset{(a)}{=} \Pr\left\{\begin{matrix}f\left(\alpha,\min\left(x_{\text{Tx}},x_{\text{Rx}}\right)\right) \leqslant v \leqslant f\left(\alpha,\max\left(x_{\text{Tx}},x_{\text{Rx}}\right)\right) \\ \cap g\left(\alpha,\min\left(y_{\text{Tx}},y_{\text{Rx}}\right)\right) \leqslant v \leqslant g\left(\alpha,\max\left(y_{\text{Tx}},y_{\text{Rx}}\right)\right)\end{matrix}\right\} \\
&= \frac{1}{2\pi}\int_0^{2\pi}\int_{\max\left\{f\left(\alpha,\min\left(x_{\text{Tx}},x_{\text{Rx}}\right)\right),g\left(\alpha,\min\left(y_{\text{Tx}},y_{\text{Rx}}\right)\right),0\right\}}^{\min\left\{f\left(\alpha,\max\left(x_{\text{Tx}},x_{\text{Rx}}\right)\right),g\left(\alpha,\max\left(y_{\text{Tx}},y_{\text{Rx}}\right)\right),1\right\}} f_v(v)\times H\left(m\sin\alpha+\cos\alpha\right)H(m) \\
&\times H\begin{pmatrix}\min\left\{f\left(\alpha,\max\left(x_{\text{Tx}},x_{\text{Rx}}\right)\right),g\left(\alpha,\max\left(y_{\text{Tx}},y_{\text{Rx}}\right)\right),1\right\} \\ -\max\left\{f\left(\alpha,\min\left(x_{\text{Tx}},x_{\text{Rx}}\right)\right),g\left(\alpha,\min\left(y_{\text{Tx}},y_{\text{Rx}}\right)\right),0\right\}\end{pmatrix} dv\,d\alpha \\
&\overset{(b)}{=} \frac{1}{2\pi}\left\{\int_0^{\delta_1}\theta_1\left(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}}\right)d\alpha+\int_{\delta_2}^{2\pi}\theta_1\left(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}}\right)d\alpha\right\}
\end{aligned} \tag{4.25}$$

where (a) follows from $p = R_{\text{net}}\sqrt{u} = R_{\text{net}}v$, and (b) follows by computing the integral with respect to v, whose probability density function is $f_v(v) = 2v$, since u is uniformly distributed in $[0,1]$.

The integration limits in (b) are determined from the conditions $m\sin\alpha+\cos\alpha \geqslant 0$, which implies $0 \leqslant \alpha \leqslant \delta_1$ and $\delta_2 \leqslant \alpha \leqslant 2\pi$, where $\delta_1 = 2\tan^{-1}\left(m+\sqrt{1+m^2}\right)$ and $\delta_2 = 2\pi+2\tan^{-1}\left(m-\sqrt{1+m^2}\right)$.

The other three case studies can be obtained by using the same approach as for Case 1. Thus, the details are omitted and only the final result is reported.

Case 2

If $m\sin\alpha+\cos\alpha \geqslant 0$ and $m < 0$, we have the following:

$$\begin{aligned}
\Pr\left\{\overline{\text{Event 1}}\right\} &= \Pr\left\{\begin{matrix}\min\left(x_{\text{Tx}},x_{\text{Rx}}\right) \leqslant \dfrac{p-z\sin\alpha}{m\sin\alpha+\cos\alpha} \leqslant \max\left(x_{\text{Tx}},x_{\text{Rx}}\right) \\ \cap \min\left(y_{\text{Tx}},y_{\text{Rx}}\right) \leqslant m\dfrac{p-z\sin\alpha}{m\sin\alpha+\cos\alpha}+z \leqslant \max\left(y_{\text{Tx}},y_{\text{Rx}}\right)\end{matrix}\right\} \\
&= \frac{1}{2\pi}\left\{\int_0^{\delta_1}\theta_2\left(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}}\right)d\alpha+\int_{\delta_2}^{2\pi}\theta_2\left(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}}\right)d\alpha\right\}
\end{aligned} \tag{4.26}$$

Case 3

If $m\sin\alpha+\cos\alpha<0$ and $m\geqslant 0$, we have the following:

$$\begin{aligned}\Pr\left\{\overline{\text{Event 1}}\right\}&=\Pr\left\{\begin{array}{l}\min\left(x_{\text{Tx}},x_{\text{Rx}}\right)\leqslant\dfrac{p-z\sin\alpha}{m\sin\alpha+\cos\alpha}\leqslant\max\left(x_{\text{Tx}},x_{\text{Rx}}\right)\\ \cap\min\left(y_{\text{Tx}},y_{\text{Rx}}\right)\leqslant m\dfrac{p-z\sin\alpha}{m\sin\alpha+\cos\alpha}+z\leqslant\max\left(y_{\text{Tx}},y_{\text{Rx}}\right)\end{array}\right\}\\&=\frac{1}{2\pi}\left\{\int_{\delta_1}^{\pi}\theta_3\left(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}}\right)d\alpha+\int_{\pi}^{\delta_2}\theta_3\left(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}}\right)d\alpha\right\}\\&=\frac{1}{2\pi}\int_{\delta_1}^{\delta_2}\theta_3\left(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}}\right)d\alpha\end{aligned}\tag{4.27}$$

Case 4

If $m\sin\alpha+\cos\alpha<0$ and $m<0$, we have the following:

$$\begin{aligned}\Pr\left\{\overline{\text{Event 1}}\right\}&=\Pr\left\{\begin{array}{l}\min\left(x_{\text{Tx}},x_{\text{Rx}}\right)\leqslant\dfrac{p-z\sin\alpha}{m\sin\alpha+\cos\alpha}\leqslant\max\left(x_{\text{Tx}},x_{\text{Rx}}\right)\\ \cap\min\left(y_{\text{Tx}},y_{\text{Rx}}\right)\leqslant m\dfrac{p-z\sin\alpha}{m\sin\alpha+\cos\alpha}+z\leqslant\max\left(y_{\text{Tx}},y_{\text{Rx}}\right)\end{array}\right\}\\&=\frac{1}{2\pi}\left\{\int_{\delta_1}^{\pi}\theta_4\left(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}}\right)d\alpha+\int_{\pi}^{\delta_2}\theta_4\left(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}}\right)d\alpha\right\}\\&=\frac{1}{2\pi}\int_{\delta_1}^{\delta_2}\theta_4\left(\alpha,x_{\text{Tx}},x_{\text{Rx}},y_{\text{Tx}},y_{\text{Rx}}\right)d\alpha\end{aligned}\tag{4.28}$$

This concludes the proof.

4.7.4 Proof of Theorem 4.2

The proof is based on Lemma 4.5. In particular, the following cases need to be examined.

- Case 1: The location of the transmitter $\left(x_{\text{Tx}},y_{\text{Tx}}\right)$ is above the line $x\cos\alpha+y\sin\alpha-p=0$ given $\sin\alpha>0$.
- Case 2: The location of the transmitter $\left(x_{\text{Tx}},y_{\text{Tx}}\right)$ is above the line $x\cos\alpha+y\sin\alpha-p=0$ given $\sin\alpha<0$.
- Case 3: The location of the receiver $\left(x_{\text{Rx}},y_{\text{Rx}}\right)$ is above the line $x\cos\alpha+y\sin\alpha-p=0$ given $\sin\alpha>0$.
- Case 4: The location of the receiver $\left(x_{\text{Rx}},y_{\text{Rx}}\right)$ is above the line $x\cos\alpha+y\sin\alpha-p=0$ given $\sin\alpha<0$.
- Case 5: The location of the transmitter $\left(x_{\text{Tx}},y_{\text{Tx}}\right)$ is below the line $x\cos\alpha+y\sin\alpha-p=0$ given $\sin\alpha>0$.
- Case 6: The location of the transmitter $\left(x_{\text{Tx}},y_{\text{Tx}}\right)$ is below the line $x\cos\alpha+y\sin\alpha-p=0$ given $\sin\alpha<0$.

- Case 7: The location of the receiver $(x_{\mathrm{Rx}}, y_{\mathrm{Rx}})$ is below the line $x\cos\alpha + y\sin\alpha - p = 0$ given $\sin\alpha > 0$.
- Case 8: The location of the receiver $(x_{\mathrm{Rx}}, y_{\mathrm{Rx}})$ is below the line $x\cos\alpha + y\sin\alpha - p = 0$ given $\sin\alpha < 0$.

From (4.3), the probability of Event 4.1 can be formulated as follows:

$$\begin{aligned}
&\Pr\{\text{Event 1}\} = \Pr\left\{\begin{matrix}\{[\text{Tx is above the line}] \cap [\text{Rx is above the line}]\} \\ \cup \{[\text{Tx is below the line}] \cap [\text{Rx is below the line}]\}\end{matrix}\right\} \\
&= \Pr\left\{\begin{matrix}[\text{Case 1} \cup \text{Case 2}] \cap [\text{Case 3} \cup \text{Case 4}] \\ \cup [\text{Case 5} \cup \text{Case 6}] \cap [\text{Case 7} \cup \text{Case 8}]\end{matrix}\right\} \\
&= \Pr\{[\text{Case 1} \cup \text{Case 2}] \cap [\text{Case 3} \cup \text{Case 4}]\} \\
&+ \Pr\{[\text{Case 5} \cup \text{Case 6}] \cap [\text{Case 7} \cup \text{Case 8}]\} \\
&= \Pr\{\text{Case 1} \cap \text{Case 3}\} + \Pr\{\text{Case 2} \cap \text{Case 4}\} \\
&+ \Pr\{\text{Case 5} \cap \text{Case 7}\} + \Pr\{\text{Case 6} \cap \text{Case 8}\}
\end{aligned} \tag{4.29}$$

Therefore, four probabilities need to be computed. Let us start with the first one:

$$\begin{aligned}
&\Pr\{\text{Case 1} \cap \text{Case3}\} \\
&= \Pr\left\{\left(x_{\mathrm{Tx}}\cos\alpha + y_{\mathrm{Tx}}\sin\alpha - p > 0 \cap \sin\alpha > 0\right) \cap \left(x_{\mathrm{Rx}}\cos\alpha + y_{\mathrm{Rx}}\sin\alpha - p > 0 \cap \sin\alpha > 0\right)\right\} \\
&= \Pr\left\{\left[\left(x_{\mathrm{Tx}}\cos\alpha + y_{\mathrm{Tx}}\sin\alpha - p > 0\right) \cap \left(x_{\mathrm{Rx}}\cos\alpha + y_{\mathrm{Rx}}\sin\alpha - p > 0\right)\right] \cap \sin\alpha > 0\right\} \\
&\overset{(a)}{=} \Pr\left\{\left[\left(v < \frac{x_{\mathrm{Tx}}\cos\alpha + y_{\mathrm{Tx}}\sin\alpha}{R_{\mathrm{net}}}\right) \cap \left(v < \frac{x_{\mathrm{Rx}}\cos\alpha + y_{\mathrm{Rx}}\sin\alpha}{R_{\mathrm{net}}}\right)\right] \cap \sin\alpha > 0\right\} \\
&= \frac{1}{2\pi}\int_0^{\pi}\int_0^{\min\left\{\frac{x_{\mathrm{Tx}}\cos\alpha + y_{\mathrm{Tx}}\sin\alpha}{R_{\mathrm{net}}}, \frac{x_{\mathrm{Rx}}\cos\alpha + y_{\mathrm{Rx}}\sin\alpha}{R_{\mathrm{net}}}, 1\right\}} f_v(v) \\
&\times H\left(\min\left\{\frac{x_{\mathrm{Tx}}\cos\alpha + y_{\mathrm{Tx}}\sin\alpha}{R_{\mathrm{net}}}, \frac{x_{\mathrm{Rx}}\cos\alpha + y_{\mathrm{Rx}}\sin\alpha}{R_{\mathrm{net}}}, 1\right\}\right) dv d\alpha \\
&\overset{(b)}{=} \frac{1}{2\pi}\int_0^{\pi} \rho_1\left(\alpha, x_{\mathrm{Tx}}, y_{\mathrm{Tx}}, x_{\mathrm{Rx}}, y_{\mathrm{Rx}}\right) d\alpha
\end{aligned} \tag{4.30}$$

where (a) follows from $p = R_{\mathrm{net}}\sqrt{u} = R_{\mathrm{net}}v$, and (b) follows by solving the integral with respect to v whose probability density function is $f_v(v) = 2v$, since u is uniformly distributed in $[0,1]$.

By using a similar approach, we we obtain the following results:

$$\begin{aligned}
\Pr\{\text{Case 2} \cap \text{Case 4}\} &= \frac{1}{2\pi}\int_{\pi}^{2\pi} \rho_2\left(\alpha, x_{\mathrm{Tx}}, y_{\mathrm{Tx}}, x_{\mathrm{Rx}}, y_{\mathrm{Rx}}\right) d\alpha \\
\Pr\{\text{Case 5} \cap \text{Case 7}\} &= \frac{1}{2\pi}\int_{0}^{\pi} \rho_2\left(\alpha, x_{\mathrm{Tx}}, y_{\mathrm{Tx}}, x_{\mathrm{Rx}}, y_{\mathrm{Rx}}\right) d\alpha \\
\Pr\{\text{Case 6} \cap \text{Case 8}\} &= \frac{1}{2\pi}\int_{\pi}^{2\pi} \rho_1\left(\alpha, x_{\mathrm{Tx}}, y_{\mathrm{Tx}}, x_{\mathrm{Rx}}, y_{\mathrm{Rx}}\right) d\alpha
\end{aligned} \tag{4.31}$$

This concludes the proof.

4.7.5 Proof of Proposition 4.1

From (4.8) in Lemma 4.2 and (4.11) in Lemma 4.4, the probability of Event 4.2 can be formulated as follows:

$$\begin{aligned}\Pr\{\text{Event 2}\} &= \Pr\left\{\begin{matrix}\min(x_{\text{end1}}, x_{\text{end2}}) \leqslant x_* \leqslant \max(x_{\text{end1}}, x_{\text{end2}}) \\ \cap \min\left(y_{\text{end1}}, y_{\text{end2}}\right) \leqslant y_* \leqslant \max\left(y_{\text{end1}}, y_{\text{end2}}\right)\end{matrix}\right\} \\ &= \Pr\left\{\begin{matrix}\min(x_{\text{end1}}, x_{\text{end2}}) \leqslant \dfrac{p - z_p \sin\alpha}{m_p \sin\alpha + \cos\alpha} \leqslant \max(x_{\text{end1}}, x_{\text{end2}}) \\ \cap \min\left(y_{\text{end1}}, y_{\text{end2}}\right) \leqslant m_p \dfrac{p - z_p \sin\alpha}{m_p \sin\alpha + \cos\alpha} + z_p \leqslant \max\left(y_{\text{end1}}, y_{\text{end2}}\right)\end{matrix}\right\}\end{aligned} \tag{4.32}$$

In order to compute this probability, we need to examine four cases depending on the relationship between x_{end1} and x_{end2}, as well as y_{end1} and y_{end2}.

Case 1

If $x_{\text{end1}} > x_{\text{end2}}$ and $y_{\text{end1}} > y_{\text{end2}}$, which implies $\sin\alpha < 0$ and $\cos\alpha > 0$, we obtain the following:

$$\begin{aligned}\Pr\{\text{Event 2}\} &= \Pr\left\{x_{\text{end2}} \leqslant \frac{p - z_p \sin\alpha}{m_p \sin\alpha + \cos\alpha} \leqslant x_{\text{end1}} \cap y_{\text{end2}} \leqslant m_p \frac{p - z_p \sin\alpha}{m_p \sin\alpha + \cos\alpha} + z_p \leqslant y_{\text{end1}}\right\} \\ &= \Pr\left\{\begin{matrix}\dfrac{L}{2}\sin\alpha + \dfrac{z_p \sin\alpha}{m_p \sin\alpha + \cos\alpha} \leqslant \left(\dfrac{1}{m_p \sin\alpha + \cos\alpha} - \cos\alpha\right) p \leqslant -\dfrac{L}{2}\sin\alpha + \dfrac{z_p \sin\alpha}{m_p \sin\alpha + \cos\alpha} \cap \\ -\dfrac{L}{2}\cos\alpha + \dfrac{m_p z_p \sin\alpha}{m_p \sin\alpha + \cos\alpha} - z_p \leqslant \left(\dfrac{m_p}{m_p \sin\alpha + \cos\alpha} - \sin\alpha\right) p \leqslant \dfrac{L}{2}\cos\alpha + \dfrac{m_p z_p \sin\alpha}{m_p \sin\alpha + \cos\alpha} - z_p\end{matrix}\right\}\end{aligned} \tag{4.33}$$

Depending on the sign of $\left(\frac{1}{m_p \sin\alpha + \cos\alpha} - \cos\alpha\right)$ and $\left(\frac{m_p}{m_p \sin\alpha + \cos\alpha} - \sin\alpha\right)$, four sub-cases need to be studied.

Case 1-a

If $\left(\frac{1}{m_p \sin\alpha+\cos\alpha} - \cos\alpha\right) > 0$ and $\left(\frac{m_p}{m_p \sin\alpha+\cos\alpha} - \sin\alpha\right) > 0$, we have the following:

$$\begin{aligned}
&\Pr\{\text{Event 2}\} = \Pr\left\{x_{\text{end2}} \leqslant \frac{p - z_p \sin\alpha}{m_p \sin\alpha + \cos\alpha} \leqslant x_{\text{end1}} \cap y_{\text{end2}} \leqslant m_p \frac{p - z_p \sin\alpha}{m_p \sin\alpha + \cos\alpha} + z_p \leqslant y_{\text{end1}}\right\} \\
&\overset{(a)}{=} \Pr\left\{F\left(\alpha, \frac{L}{2}\sin\alpha\right) \leqslant v \leqslant F\left(\alpha, -\frac{L}{2}\sin\alpha\right) \cap G\left(\alpha, -\frac{L}{2}\cos\alpha\right) \leqslant v \leqslant G\left(\alpha, \frac{L}{2}\cos\alpha\right)\right\} \\
&= \frac{1}{2\pi}\int_0^{2\pi} \int_{\max\{F(\alpha,\frac{L}{2}\sin\alpha), G(\alpha,-\frac{L}{2}\cos\alpha), 0\}}^{\min\{F(\alpha,-\frac{L}{2}\sin\alpha), F(\alpha,-\frac{L}{2}\sin\alpha), 1\}} f_v(v) H\left(\frac{1}{m_p \sin\alpha + \cos\alpha} - \cos\alpha\right) H\left(\frac{m_p}{m_p \sin\alpha + \cos\alpha} - \sin\alpha\right) \\
&\times H\begin{pmatrix} \min\left\{F\left(\alpha, -\frac{L}{2}\sin\alpha\right), F\left(\alpha, -\frac{L}{2}\sin\alpha\right), 1\right\} \\ -\max\left\{F\left(\alpha, \frac{L}{2}\sin\alpha\right), G\left(\alpha, -\frac{L}{2}\cos\alpha\right), 0\right\} \end{pmatrix} \bar{H}(\sin\alpha)\, H(\cos\alpha)\, dv d\alpha \\
&\overset{(b)}{=} \frac{1}{2\pi}\int_0^{2\pi} \Gamma_1^a(\alpha) \bar{H}(\sin\alpha)\, H(\cos\alpha)\, d\alpha = \frac{1}{2\pi}\int_{\frac{3\pi}{2}}^{2\pi} \Gamma_1^a(\alpha) d\alpha
\end{aligned} \tag{4.34}$$

where (a) follows from $p = R_{\text{net}}\sqrt{u} = R_{\text{net}} v$, and (b) follows by solving the integral with respect to v whose probability density function is $f_v(v) = 2v$, since u is uniformly distributed in $[0, 1]$.

By using a similar approach, we can study the remaining three sub-cases.

Case 1-b

If $\left(\frac{1}{m_p \sin\alpha+\cos\alpha} - \cos\alpha\right) > 0$ and $\left(\frac{m_p}{m_p \sin\alpha+\cos\alpha} - \sin\alpha\right) < 0$, we have the following:

$$\Pr\{\text{Event 2}\} = \frac{1}{2\pi}\int_0^{2\pi} \Gamma_1^b(\alpha) \bar{H}(\sin\alpha)\, H(\cos\alpha)\, d\alpha = \frac{1}{2\pi}\int_{\frac{3\pi}{2}}^{2\pi} \Gamma_1^b(\alpha) d\alpha \tag{4.35}$$

Case 1-c

If $\left(\frac{1}{m_p \sin\alpha+\cos\alpha} - \cos\alpha\right) < 0$ and $\left(\frac{m_p}{m_p \sin\alpha+\cos\alpha} - \sin\alpha\right) > 0$, we have the following:

$$\Pr\{\text{Event2}\} = \frac{1}{2\pi}\int_0^{2\pi} \Gamma_1^c(\alpha) \bar{H}(\sin\alpha)\, H(\cos\alpha)\, d\alpha = \frac{1}{2\pi}\int_{\frac{3\pi}{2}}^{2\pi} \Gamma_1^c(\alpha) d\alpha \tag{4.36}$$

Case 1-d

If $\left(\frac{1}{m_p \sin\alpha+\cos\alpha} - \cos\alpha\right) < 0$ and $\left(\frac{m_p}{m_p \sin\alpha+\cos\alpha} - \sin\alpha\right) < 0$, we have the following:

$$\Pr\{\text{Event 2}\} = \frac{1}{2\pi}\int_0^{2\pi} \Gamma_1^d(\alpha) \bar{H}(\sin\alpha)\, H(\cos\alpha)\, d\alpha = \frac{1}{2\pi}\int_{\frac{3\pi}{2}}^{2\pi} \Gamma_1^d(\alpha) d\alpha \tag{4.37}$$

Therefore, eventually, Pr{Event2} can be formulated as follows:

$$\Pr\{\text{Event 2}\} = \frac{1}{2\pi}\int_{\frac{3\pi}{2}}^{2\pi} \Gamma_1(\alpha)d\alpha = \frac{1}{2\pi}\int_{\frac{3\pi}{2}}^{2\pi} \left[\Gamma_1^a(\alpha)+\Gamma_1^b(\alpha)+\Gamma_1^c(\alpha)+\Gamma_1^d(\alpha)\right]d\alpha \tag{4.38}$$

By using a similar line of thought, the remaining three cases can be studied. The final result is reported in the following sections.

Case 2

If $x_{\text{end1}} > x_{\text{end2}}$, and $y_{\text{end1}} < y_{\text{end2}}$, which implies $\sin\alpha < 0$ and $\cos\alpha < 0$, we obtain the following:

$$\Pr\{\text{Event 2}\} = \frac{1}{2\pi}\int_{\pi}^{\frac{3\pi}{2}} \Gamma_2(\alpha)d\alpha = \frac{1}{2\pi}\int_{\pi}^{\frac{3\pi}{2}} \left[\Gamma_2^a(\alpha)+\Gamma_2^b(\alpha)+\Gamma_2^c(\alpha)+\Gamma_2^d(\alpha)\right]d\alpha \tag{4.39}$$

Case 3

If $x_{\text{end1}} < x_{\text{end2}}$, and $y_{\text{end1}} > y_{\text{end2}}$, which implies $\sin\alpha > 0$ and $\cos\alpha > 0$, we have the following:

$$\Pr\{\text{Event 2}\} = \frac{1}{2\pi}\int_{0}^{\frac{\pi}{2}} \Gamma_3(\alpha)d\alpha = \frac{1}{2\pi}\int_{0}^{\frac{\pi}{2}} \left[\Gamma_3^a(\alpha)+\Gamma_3^b(\alpha)+\Gamma_3^c(\alpha)+\Gamma_3^d(\alpha)\right]d\alpha \tag{4.40}$$

Case 4

If $x_{\text{end1}} < x_{\text{end2}}$, and $y_{\text{end1}} < y_{\text{end2}}$, which implies $\sin\alpha > 0$ and $\cos\alpha < 0$, we have the following:

$$\Pr\{\text{Event 2}\} = \frac{1}{2\pi}\int_{0}^{\frac{\pi}{2}} \Gamma_4(\alpha)d\alpha = \frac{1}{2\pi}\int_{\frac{\pi}{2}}^{\pi} \left[\Gamma_4^a(\alpha)+\Gamma_4^b(\alpha)+\Gamma_4^c(\alpha)+\Gamma_4^d(\alpha)\right]d\alpha \tag{4.41}$$

This concludes the proof.

5 Reconfigurable Intelligent Surfaces vs. Multi-Antenna Relaying: Differences, Similarities, and Performance Comparison

Chapter 5. Reconfigurable Intelligent Surfaces vs. Multi-Antenna Relaying: Differences, Similarities, and Performance Comparison

Reconfigurable intelligent surfaces (RISs) have the potential of realizing the emerging concept of smart radio environments by leveraging the unique properties of meta-surfaces. In this chapter, we discuss the need and potential applications of RISs in wireless networks that operate at high-frequency bands, e.g., millimeter wave (30-100 GHz) and sub-millimeter wave (greater than 100 GHz) frequencies; we overview the technology enabler of reconfigurable meta-surfaces; we elaborate on the working principle of RISs when used as low-complexity transmitters and anomalous mirrors; and we describe promising use cases that encompass signal, interference, security, and scattering engineering. When used as anomalous mirrors, the RISs resemble multiple-antenna relay stations. Thus, we develop the analytical frameworks of the end-to-end signal-to-noise ratio and the rate of half-duplex, full-duplex, amplify-and-forward, and decode-and-forward relay-aided communications, and the signal-to-noise ratio and the rate of the RISs. Towards the end, we elaborate on the key differences and similarities between RISs and relaying, and, notably, illustrate numerical results that highlight the spectral efficiency gains of RISs whose transverse size is large as compared with the wavelength of the radio wave. Finally, we discuss the key open issues that need to be addressed in order to unlock the potential of RISs, and, in particular, to unveil their ultimate performance limits and to optimize their deployment in future wireless networks.

5.1 Introduction

Over the last years, there has been a continuous increase of the demand for higher data rates driven by the proliferation of smartphones and the introduction of broadband services and applications, such as virtual and augmented reality, the Internet of Things, and mobile live streaming. By 2022, it is expected that the global mobile data traffic will reach a monthly run of 77 exabytes, which corresponds to a 7-fold growth compared with the monthly run of 2017. Such demands cannot be accommodated by current cellular standards that utilize sub-6 GHz frequency bands to convey information, e.g., the Long-Term Evolution Advanced (LTE-A) standard that can offer peak data rates up to 1 Gbps. Motivated by these overwhelming demands for ever-increasing data rates, a key feature of future wireless networks is the migration to higher frequencies, e.g., the millimeter (30-100 GHz) and sub-millimeter (above 100 GHz) wave bands [83].

Extensive measurements have been conducted at the millimeter wave band and, more recently, the sub-millimeter wave band. These have demonstrated that the use of highly directional steerable antennas can enable mobile communications at such high frequencies [83]. In particular, early measurement campaigns executed at 28 GHz and conducted in New York City in 2012 revealed that, if directional antennas are used at the transmit and receive end-points, distances up to 200 meters can be reached. However, millimeter and sub-millimeter wave frequency bands are highly susceptible to blockages from large-size structures, e.g., buildings, on the radio path [83, Tables 4, 5]. Also, millimeter- and sub-millimeter wave signals are severely attenuated by the presence of small-size objects, e.g., human bodies and foliage. For example, due to the human body, penetration losses of around 20-40 dB have been measured at the millimeter wave frequency band. In spite of their high penetration losses, such objects may act as good reflectors and scatterers for the impinging radio waves. Thanks to the resulting secondary paths, the receiver may be able to capture a sufficient signal power. To capitalize on such paths, however, the receiver generally needs to rely upon antennas with a wide beamwidth, which may result in low antenna gains. This inevitably reduces the maximum range of communication under non-line-of-sight (NLOS) conditions.

The problems highlighted above become more pronounced in dense urban environments because of the highly dynamic nature of the radio environment. Many factors, such as moving cars and humans, render, in fact, the establishment of reliable communication links challenging due to the rapidly varying propagation conditions at high frequencies. A major question is, therefore, whether such limitations can be curtailed, enabling high-rate and ultra-reliable wireless access networks at high-frequency bands and in the presence of mobility.

A possible approach for circumventing the unreliability of high-frequency channels is to sense the environment and to identify, on a real-time basis, alternative propagation routes through which the same information-bearing signal can be received. This is a realistic proposition, especially at high frequencies, since the potential interference caused by the presence of multiple propagation paths can be kept under control by the large path-loss of longer propagation routes. An established method to create additional routes for receiving the same information-bearing signal is the deployment of relay stations that capitalize on the concept of (distributed) cooperative diversity [84]. Relay stations can effectively turn a single NLOS link into multiple line-of-sight (LOS) links. This approach necessitates each relay station to be equipped with a dedicated power source and with the necessary front-end circuitry for reception, processing, and re-transmission. Thus, the use of relay stations usually reduces the network energy efficiency and increases the network complexity at run time, while requiring a larger capital expenditure for deployment.

Apart from the cost-related drawbacks highlighted above, the network spectral efficiency of relay-aided systems is reduced if a half-duplex (HD) relaying protocol is employed, since transmitters and relay stations are not allowed to transmit concurrently on the same physical resource. This issue can be overcome by employing a full-duplex (FD) relaying protocol, but at the cost of: (i) introducing high *loop-back self-interference* at the relay station because of the concurrent transmission and reception of signals; (ii) generating *co-channel interference* at the destination, since relay stations and transmitters emit different information data on the same physical resource; and (iii) increasing the *signal processing complexity* and the *power consumption* of the relay stations.

When the LOS path is of insufficient quality, another approach for establishing alternative routes is through *passive non-reconfigurable specular reflectors*, e.g., dielectric mirrors. This method for coverage enhancement has the potential of being more cost efficient as compared with relaying, especially at high-frequency bands. But a main limitation of non-reconfigurable reflectors is that they cannot enable the dynamic shaping of the impinging waves, since their operation cannot be modified after fabrication, i.e., at the time of deployment and operation. Due to the highly dynamic nature of the wireless environment at high frequencies, it is essential that such reflectors be capable of adaptively shaping the impinging radio waves based on actual blockage and environmental conditions. This would ideally enable the network to maximize the network coverage and throughput on a real-time basis.

Propitiously, electromagnetic-based reconfigurable structures that are capable of applying specified transformations to the impinging radio waves do exist and can operate at different frequency bands [85], [86]. They are referred to as **reconfigurable intelligent surfaces (RISs)**, and, when deployed in wireless networks, have the potential of turning the wireless environment, which is highly probabilistic in nature, into a programmable and partially deterministic space, which is referred to as **smart radio environment** [16]. The aim of this paper is to provide an introduction to this topic, with focus on the differences with relay-aided systems.

What is a Reconfigurable Intelligent Surface?

An RIS is an artificial surface, made of electromagnetic material, that is capable of altering the propagation of the radio waves impinging upon it. An example consists of reflecting surfaces whose constituent (often discrete) elements can be programmed in order to apply arbitrary phase shifts to the signals [87], [88]. These solutions are referred to as **intelligent reflecting surfaces (IRSs)**. Large-scale testbeds for IRSs have recently been built, and have empirically proved that IRSs can enhance the signal strength by a factor of 10 and can increase the median data rate by a factor of 2 [89], [90].

A conceptually and technologically different solution consists of realizing RISs based on the two-dimensional equivalent of meta-materials, which are referred to as **meta-surfaces**. RISs based on meta-surfaces are very thin – their thickness is much smaller than the wavelength – sheets of electromagnetic material that are engineered to possess peculiar properties that cannot be found in naturally occurring materials [16], [85]-[86]. In contrast to IRSs, RISs based on meta-surfaces are not only capable of modifying the phase of the impinging radio waves, but they can shape their wavefront in more general ways. This can be realized by reflecting and refracting the impinging signals towards anomalous directions not predicted by the Snell's law (see Fig. 5.1), or by changing their polarization and waveform [16]. RISs based on meta-surfaces are better viewed as continuous surfaces of electromagnetic material that modify the wavefront of the radio waves by introducing an appropriately designed phase gradient that depends on the specific function to be realized. As far as the propagation of the radio waves is

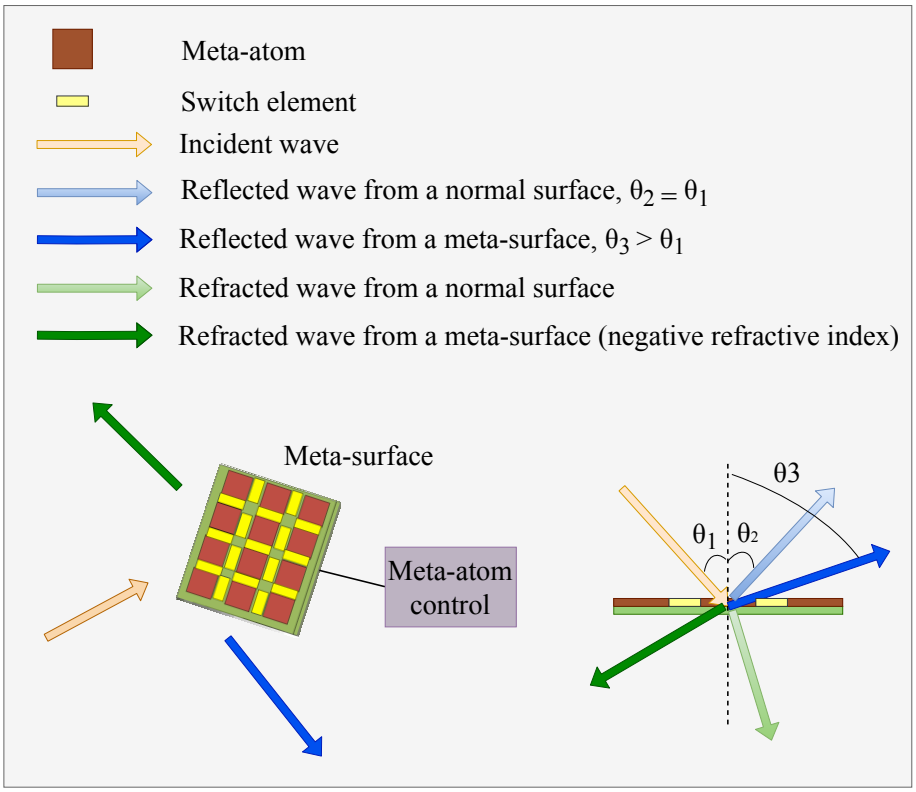

Figure 5.1 – Working principle of a reconfigurable meta-surface.

concerned, an RIS based on meta-surfaces acts as an abrupt electromagnetic discontinuity that alters the scattered field. Prototypes of meta-surfaces that realize various reconfigurable functions exist [91], and their potential and achievable gains in complex reverberating media, e.g., indoor environments, have been proved since 2014 [92].

As discussed, the key element to realize multi-function RISs is a meta-surface. A meta-surface is a sub-wavelength array formed by sub-wavelength metallic or dielectric scattering particles that are referred to as meta-atoms [85]-[86]. It can be described as an electromagnetic discontinuity that is sub-wavelength in thickness, with typical values ranging from 1/10 to 1/5 of the wavelength, and is electrically large in transverse size. Its unique properties lie in its capability of shaping the electromagnetic waves, for reflection and transmission, in very general ways, which overcome the limitations of the phased-array principles of reflection and refraction (also known as the **generalized Snell's laws**) [93]. This is illustrated in Fig. 5.1, which depicts the expected reflection and refraction response of a normal surface and a meta-surface for a given angle of incidence of the electromagnetic wave.

It should be noted that meta-surfaces can be either reconfigurable or not. As mentioned, the highly dynamic nature of wireless environments requires reconfigurable meta-surfaces. In non-reconfigurable meta-surfaces, the meta-atoms have fixed structural and geometrical arrangements, which result in static interactions with the impinging radio waves that cannot be modified once they are manufactured. In **reconfigurable meta-surfaces**, the arrangements of the meta-atoms can be modified and programmed based on external stimuli. The reconfigurability can be enabled by electronic phase-changing components, such as semiconductors or graphene, which are used as switches or tunable reactive and resistive elements. They can be either inserted between adjacent meta-atoms (see Fig. 5.1)

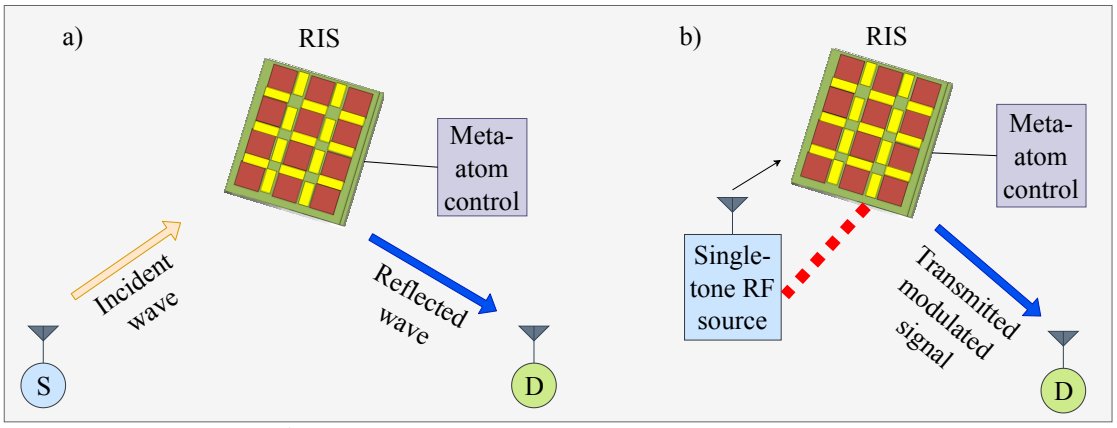

Figure 5.2 – An RIS operating as a multi-function "anomalous mirror" (a) and as a single-RF transmitter (b).

or can be used to adjust the properties of individual meta-atoms. This way, the wavefront of the radio waves scattered by a meta-surface is manipulated by controlling the status of the switches, in order for different functions to be realized. The functions can be optimized by a central controller through software-defined networking technologies, as recently demonstrated in [85]. A major difference between static and reconfigurable meta-surfaces lies in their associated power consumption. Static meta-surfaces can be fully passive, since no active electronic circuits are needed. Reconfigurable meta-surfaces can only be **nearly passive**, since some energy is needed to control the switches. No dedicated power supply is, in general, needed for signal transmission.

In summary, based on the mentioned distinctive properties of the meta-surfaces, an RIS can be thought of as a **multi-function "anomalous mirror"**. In this paper, we are primarily interested in this use of the RISs, which is illustrated in Fig. 5.2(a). Also, we restrict the discussion to planar RISs, called *meta-mirrors*, but we emphasize that the meta-surfaces need not to be planar but can be made conformal to curved surfaces.

We specifically model an RIS as a collection of meta-surfaces that are capable of applying distinct functions to the impinging radio waves, such as reflecting the signals towards different directions, modifying the polarization, and applying different phase shifts. In practice, this is realized via appropriate arrangements of the meta-atoms on the same RIS.

Another promising application of the RISs in wireless systems, not covered in much detail in this paper, is the realization of **single-RF transmitters** [94], which is illustrated in Fig. 5.2(b). In typical multiple-antenna transmitters, the number of independent information streams that can be transmitted is equal to the number of radio frequency (RF) chains. Each RF chain usually consists of an independent power amplifier, a filter, a digital-to-analog converter (DAC), and a mixer. As a result, the complexity, power consumption, and cost of a multiple-antenna transmitter is larger than for its single-antenna counterpart [86]. Recently, it has been experimentally demonstrated that a multiple-stream transmitter can be realized by using a single-RF feeder and an RIS [94]. As illustrated in Fig. 5.2(b), a signal generator produces an un-modulated carrier that is radiated through an antenna towards an RIS. The un-modulated signal emitted by the antenna impinges upon the RIS, where it is phase-modulated by appropriately controlling the meta-atoms. For a given un-modulated signal impinging upon the RIS, multiple phase-modulated signals can be reflected off it by appropriately grouping and controlling the meta-atoms. If N is the number of independently controllable phases, then an N-stream transmitter

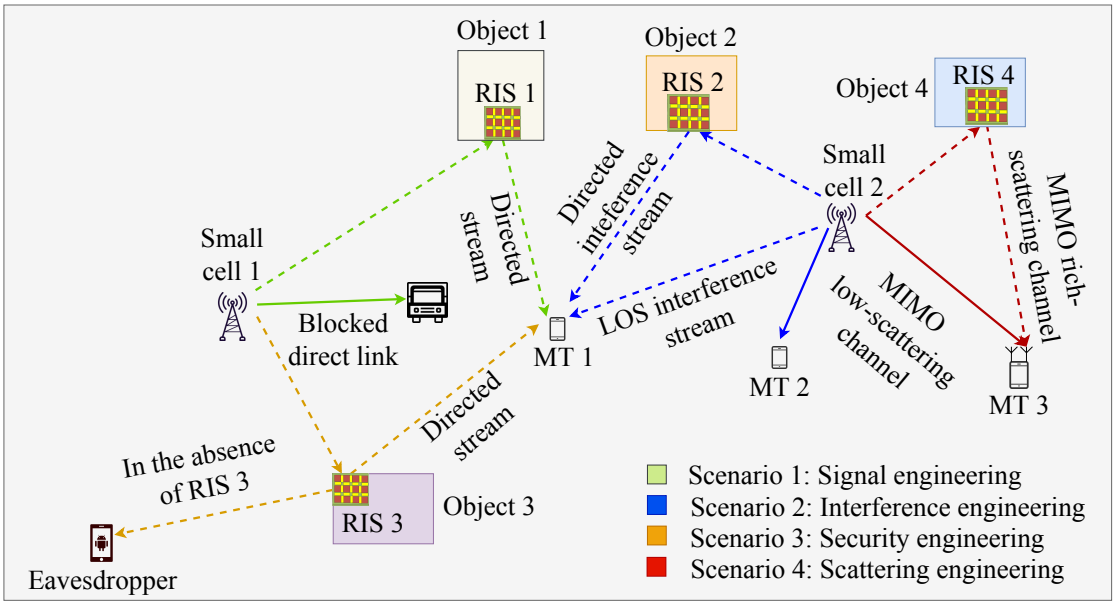

Figure 5.3 – Example of smart radio environment.

can be realized with a single-RF chain. As elaborated in [86], this concept is also known as *media-based modulation*. In this implementation, the single-RF feeder and the RIS are viewed as an integral part of the transmitter, and they may be even directly connected with each other, as represented in Fig. 5.2(b) by the dashed line.

5.2 Wireless 2.0: Smart Radio Environments

From the viewpoint of the communication engineer, the wireless environment is conventionally modeled as an exogenous entity that cannot be controlled, but only adapted to. To this end, communication engineers can only design the transmitters, the receivers, and the transmission protocols in order to achieve the desired performance. Common approaches to capitalize on the properties of the wireless environment and to mitigate its impairments include using multiple antennas, employing complex encoding and decoding algorithms at the end-points of the communication link, and adding additional network infrastructure, e.g., relay stations [84], in an attempt to make the transmission of signals more reliable. These solutions, however, increase the network complexity, the power consumption, and the deployment cost. In deployments such as the Internet of Things, these approaches may not be adopted due to the design constraints at the devices [90].

RISs provide wireless researchers and engineers with a radically **new view of the wireless environment**. Since the RISs are capable of shaping the wavefront of the radio waves throughout the network, the **wireless environment can be in principle customized** to suit the system requirements. The wireless environment is not to be treated as a random uncontrollable entity, but rather as part of the network design parameters that are subject to optimization in order to support diverse performance metrics, such as rate, latency, reliability, energy efficiency, privacy, and massive connectivity.

The overarching vision is that environmental objects and devices are coated with software-controlled RISs, and, through environmental sensing and software-defined networking protocols, they can be programmed for shaping the radio propagation environment and meeting the desired system require-

ments [16], [85]. An illustrative example of this emerging wireless future is sketched in Fig. 5.3. Four application scenarios are identified: (i) **signal engineering**; (ii) **interference engineering**; (iii) **security engineering**; and (iv) **scattering engineering**.

Signal engineering: Assume that small cell 1 wishes to communicate with mobile terminal (MT) 1, but the LOS link is blocked by an object. In this case, small cell 1 redirects the transmitted beam towards RIS 1 that coats object 1, and assists the communication by shaping the incident wave towards MT 1 so that the received signal strength is maximized.

Interference engineering: While small cell 1 communicates with MT 1, small cell 2 communicates with MT 2. Therefore, an interfering signal reaches MT 1 from small cell 2. To suppress it at MT 1, RIS 2 is programmed to shape the impinging radio wave from small cell 2 towards MT 1 in a way that the two signals are destructively combined at MT 1.

Security engineering: In the absence of RIS 3, the signal emitted by small cell 1 and intended to MT 1 is reflected from object 3 towards a malicious user that overhears it. To avoid this, RIS 3 is programmed to shape the reflection towards MT 1 so that it is steered away from the malicious user while being decoded more reliably, via diversity combining, at MT 1.

Scattering engineering: The multiple-antenna small cell 2 wishes to convey information to the multiple-antenna MT 3 with the aid of multiple-input multiple-output transmission. The channel between small cell 2 and MT 3 has, however, a low rank (low scattering environment), which negatively affects the attainable data rate. To avoid this issue, small cell 2 directs the signal intended to MT 3 towards RIS 4, which appropriately shapes it so as to create a rich-scattering environment (high rank channel) for high data rate transmission.

From the analysis of the these four scenarios, it is apparent that, with the aid of RISs, the propagation of radio waves in wireless networks can be engineered and optimized, at a low complexity, in a way that benefits the network.

Let us further analyze the application for **interference engineering**. A widespread approach to deal with the interference from a small cell is to equip it with an adequate number of multiple antennas and corresponding RF chains, so that it can beamform the signal towards the desired direction, and, at the same time, it can suppress the interference towards the unwanted directions. This approach poses size and power consumption issues, since small cells need to be deployed on a wide scale in places that are size-limited, e.g., lampposts. By leveraging RISs, the cost of deploying small cells can be substantially reduced. The RISs, in fact, are nearly passive and can be readily deployed on the facades of buildings and walls.

Reconfigurable Intelligent Surfaces vs. Relaying

In this section, we elaborate on differences and similarities between RISs that are employed as anomalous mirrors and relays. The comparison is made here on a qualitative basis, and is complemented, in the next section, with results that showcase scenarios in which the RISs constitute a promising alternative to relaying. For ease of description, we focus our attention on two-hop repetition-based relaying, decode-and-forward (DF) and amplify-and-forward (AF) protocols, and HD and FD duplexing methods. For illustrative purposes, the system scenario in Fig. 5.4 is considered as a reference.

5.2.1 Hardware Complexity

Relay stations are active devices that need a dedicated power source for operation. They are equipped with active electronic components, such as DACs and analog-to-digital converters (ADCs), mixers, power amplifiers for transmission, and low-noise amplifiers for reception. The electronic components that are typically needed for implementing DF and AF relaying are discussed and illustrated in [84, Ch. 5]. The deployment space of relays can, thus, be costly and power-consuming, especially for realizing multiple-antenna designs at millimeter and sub-millimeter wave frequency bands [83]. If, in addition, FD relays are used, the complexity is further increased due to the need of canceling the loop-back self-interference by using tailored antennas, and analog/digital signal processing methods.

In contrast, RISs are meant to be realized with minimal hardware complexity not requiring dedicated power amplifiers, mixers, and DACs/ADCs. As discussed, the operating principle of RISs is based on appropriately designing the distribution of the currents induced by the incident waves on their surface, in order to implement the desired function [91]-[93]. RISs are composite material layers that are made of metallic or dielectric patches printed on a grounded dielectric substrate. Their reconfigurability is ensured through low-power and low-complexity electronic circuits (switches or varactors). As for the hardware complexity, the RISs are aimed to be of much lower complexity than relay stations, especially at mass production and if realized by using inexpensive large-area electronics. An example of large-size RIS made of inexpensive antennas can be found in [90]. In contrast to other implementations, e.g., tunable reflect-arrays, the use of sub-wavelength reconfigurable meta-surfaces ensures a finer-grained control and shaping of the radio waves in both the near and far fields.

5.2.2 Noise

The active electronic components used in relay stations are responsible for the presence of additive noise that negatively affects the performance of relaying protocols. In AF relaying, the noise is amplified at the relay stations. The impact of additive noise can be mitigated by employing DF relaying, at the expense of decoding and re-encoding (regeneration) the signal at the relay stations and increasing the signal processing complexity and power consumption. In FD relaying, the impact of residual loop-back self-interference further deteriorates the system performance. The RISs behave, on the other hand, as anomalous mirrors, and, therefore, they are not affected by additive noise. The RISs may, however, be affected by phase noises and they cannot amplify or regenerate the signals [86].

5.2.3 Spectral Efficiency

The spectral efficiency of relaying systems depends on the adopted duplexing protocol. Under HD relaying, the achievable rate is generally scaled down by a factor of two since different physical resources are used for the data emitted by the transmitter and by the relay station. The end-to-end signal-to-noise ratio can be increased by optimally combining the direct and relayed signals. Under FD relaying, the achievable rate is not scaled down by a factor of two, but the relay station is affected by the residual loop-back self-interference, and the receiver is impaired by the interference generated by the concurrent transmission of the transmitter. RISs behaving as anomalous mirrors are not subject to the half-duplex constraint and the loop-back self-interference. Also, the phase gradient of the meta-surfaces can be designed for optimally combining the signals received from the transmitter and the RIS.

5.2.4 Power Budget

Relay stations need an independent power source for operation, which is used for transmitting the signals (RF power) and for powering up their electronic components. In contrast, the RISs are suitable for nearly passive implementations, since non-reconfigurable meta-surfaces can be realized with fully passive components, and low-power active components (switches or varactors) are needed only for ensuring their reconfigurability. The low-power nature of switches and varactors makes the use of energy harvesting a suitable candidate for realizing close-to-passive implementations. For relaying, it is usually assumed that the total RF power is allocated between the transmitter and the relay, so as to ensure a total power constraint. In RISs, the transmitter uses the total RF power. Also, the power reflected or scattered by the RIS depends on its transmittance, which can be made close to one if the meta-surface is appropriately designed and optimized [95]. In the ideal case, the total power reflected or scattered by the RIS is the same as the total RF power received from the transmitter.

5.2.5 Average Signal-to-Noise Ratio vs. Number of Elements

Let us consider a multiple-antenna relay station, as illustrated in Fig. 5.4, and that it employs maximum ratio weighting for reception and transmission. If N antennas are used at the relay, the average end-to-end signal-to-noise ratio increases *linearly* with N [84]. Recently, it has been analytically proved in [86] and empirically substantiated, for discrete antennas, in [90] that the average end-to-end signal-to-noise ratio of an RIS constituted by N individually reconfigurable elements increases *quadratically* with N, while still being subject to the energy conservation principle. This is due to the fact that, in relay stations, the RF power is allocated among the N antennas so that the total RF power is kept constant. In RISs, in contrast, each constituent meta-surface acts as a separable anomalous mirror, which reflects, after scaling by the transmittance and with no noise addition, the same amount of RF power received from the transmitter. It is worth mentioning, however, that the more favorable scaling law as a function on N does not necessarily imply that the RISs outperform relaying. For a fixed total power constraint, in fact, the path-loss as a function of the transmission distance cannot be overlooked. This is discussed next by considering, for ease of exposition, a free-space propagation model.

5.2.6 Average Signal-to-Noise Ratio vs. Transmission Distance

In relay stations with N antenna elements, the distance between adjacent antennas is, usually, of the order of $\lambda/2$, where λ is the wavelength of the radio wave. By optimistically excluding the space needed for accommodating multiple RF chains (amplifiers, ADCs/DACs, mixers), the geometric area occupied by a planar deployment would be of the order of $N\lambda^2/4$. In this case, the end-to-end power received from an AF relay scales with the product of the transmitter-to-relay distance and the relay-to-receiver distance [84], i.e., as $\left(d_{\mathrm{SR}}^2 d_{\mathrm{RD}}^2\right)^{-1}$ by using the notation in Fig. 5.4. When considering the effect of noise, the end-to-end signal-to-noise ratio of both DF and AF relaying scales with the distance of the weakest of the two paths, i.e., as $\min\left\{d_{\mathrm{SR}}^{-2}, d_{\mathrm{RD}}^{-2}\right\}$. By coherently combining the signals from the N antennas at the relay, the scaling law is, therefore, $N\min\left\{d_{\mathrm{SR}}^{-2}, d_{\mathrm{RD}}^{-2}\right\}$ [84].

The total power scattered by an RIS, and, therefore, the scaling law of the received power as a function of the distance, depends on the relation between the geometric size of each meta-surface of the RIS and the wavelength of the radio wave. In this paper, two regimes need to be distinguished: (i) **anomalous reflection** and (ii) **diffuse scattering**. In the sequel, these two regimes are also referred to as "**electrically large RISs**" and "**electrically small RISs**", respectively.

Anomalous reflection: If the geometric size of each constituent meta-surface of the RIS is large enough

as compared with the wavelength of the radio wave, e.g., it is ten times (one order of magnitude) larger than λ [85], [86], then each meta-surface behaves, approximately, as an anomalous reflector (mirror). In this case, the power received from each meta-surface and the end-to-end average signal-to-noise ratio scale, as function of the distance, as $(d_{\mathrm{SR}} + d_{\mathrm{RD}})^{-2}$ [95]. This scaling law is employed in state-of-the-art software for ray tracing [83]. Anomalous reflection is, in fact, often referred to as *geometric scattering*, and the laws of geometric optics are sufficient to analyze it. If the meta-surfaces are uniform, i.e., they do not introduce any phase gradients, they behave, approximately, as specular reflectors (i.e., their response adhere to Snell's law). By combining the signals from the N meta-surfaces of the RIS, the scaling law is $N^2(d_{\mathrm{SR}} + d_{\mathrm{RD}})^{-2}$.

Diffuse scattering: If the size of each meta-surface of the RIS is of the order of the wavelength of the radio wave, then each of them behaves, approximately, as a dipole scatterer, i.e., a diffuser [96]. If the signals from the N meta-surfaces of the RIS are combined together, the received power and the end-to-end average signal-to-noise ratio scale as $N^2\left(d_{\mathrm{SR}}^2 d_{\mathrm{RD}}^2\right)^{-1}$ [83], [96]. As a function of the distance, this is the same scaling law as for the received power of AF relaying.

It is worth noting that the behavior of an RIS as an anomalous mirror and as a diffuser has some resemblance, even though the underlying physical phenomena are slightly different, with the propagation mechanisms of specular reflections and diffuse scattering that occur when the radio waves impinge on the surfaces of objects [83, Fig. 2], [84, Fig. 2.2].

Main Takes: For simplicity, let us ignore the direct path in Fig. 5.4, and assume $d_{\mathrm{SR}} = d_{\mathrm{RD}} = d/2$, i.e., the RIS/relay is located equidistantly from the transmitter and receiver. Then, the average end-to-end signal-to-noise ratio scales as: (i) $4\mathrm{N}d^{-2}$ for a relay station; (ii) $4N^2d^{-2}$ for an electrically large RIS; and (iii) $16N^2d^{-4}$ for an electrically small RIS.

Take 1: From the discussed scaling behavior, it is apparent that an electrically large RIS has the potential of outperforming relaying because of the better scaling law of the received power as a function of N. Even if $N = 1$, an electrically large RIS can potentially outperform a relay due to the absence of noise amplification, loop-back self-interference, and half-duplex constraint. To appreciate the practical value of electrically large RISs, it is sensible to investigate their expected geometric size as compared with relays. As mentioned, the typical size of a relay with N antennas is of the order of $N\lambda^2/4$, and the typical size of an RIS with N reconfigurable meta-surfaces that act as anomalous mirrors is of the order of $N(10\lambda)^2$. Even when the space necessary for the N RF chains, power amplifiers, and associated circuitry of the relays is not taken into account, it is reasonable to argue that the number of meta-surfaces that act as independent anomalous mirrors cannot be too large for practically viable geometric sizes of the RISs. For example, an RIS of 1 m^2 in size can accommodate only one reconfigurable meta-surface if it operates at 3 GHz and 100 reconfigurable meta-surfaces if it operates at 30 GHz.

To sum up, electrically large RISs are a suitable technology for implementing low-complexity anomalous mirrors at millimeter wave and sub-millimeter wave frequency bands.

Take 2: In spite of the quadratic scaling law of the average signal-to-noise ratio as a function of N, electrically small RISs are expected to provide worse performance than relays because of the unfavorable scaling law as a function of the distance. Electrically small RISs are more suitable either for implementing single-RF multi-stream transmitters (see Fig. 5.2(b)), since the distance from the feeder to the RIS can be ignored, or when they are deployed sufficiently close to the receiver for local coverage enhancement [87], [88]. In the former case, the average signal-to-noise ratio per information stream scales as $d_{\mathrm{RD}}^{-2} = 4d^{-2}$. In conventional (equipped with multi-RF chains) transmitters, by contrast, the

average signal-to-noise ratio per information stream scales as $d_{\mathrm{RD}}^{-2}/N = 4d^{-2}/N$. Electrically small RISs thus have the potential of reducing the implementation complexity and the power consumption, as well as outperforming transmitters with multiple RF chains.

To sum up, electrically small RISs are a suitable technology for implementing low-complexity transmitters at sub-6 GHz, millimeter, and sub-millimeter wave frequency bands.

5.3 Performance Comparison – System Model

Let us assume that a single-antenna source wishes to communicate with a single-antenna destination that is located at a distance d_{SD}. Due to the possibility of the source-to-destination link of being of poor quality due to blockages, such as fixed and moving objects, which can frequently be the case for communication at high-frequency bands, such as millimeter wave (30-100 GHz) and sub-millimeter wave (greater than 100 GHz) bands, a multiple-antenna relay [84] or a multiple-element reconfigurable intelligent surface (RIS) [16], [86] can assist the communication between the source and the destination, as depicted in Fig. 5.4. As far as the relay is concerned, we consider both half-duplex (HD) and full-duplex (FD) duplexing schemes, and decode-and-forward (DF) and amplify-and-forward (FD) relaying protocols. The distance of the source-to-relay/RIS link is denoted by d_{SR} and the distance of the relay/RIS-to-destination link is denoted by d_{RD}. In scenarios where the source-to-destination link can frequently be under non-line-of-sight (NLOS) propagation conditions, the relay/RIS should be placed at positions where a line-of-sight (LOS) link is secured for both the source-to-relay/RIS and relay/RIS-to-destination links.

Furthermore, we denote the number of antennas and elements of the relay and the RIS by N_R and N_{RIS}, respectively. In addition, we denote the path-loss at a distance d by $P_L(d)$, and the transmitted symbol from the source by s. Finally, we consider additive white Gaussian noise affecting the relay and destination, whose variance is equal to $N_0 = -174 + 10log_{10}(BW) + NF\ (dBm)$, where BW and NF are the signal bandwidth and noise figure, respectively.

5.3.1 Relay-Assisted Communication

In this section, we introduce the mathematical formulation of the end-to-end signal-to-noise ratio (SNR) expressions for relay-assisted communication. By $h_{SD} \in \mathbb{C}$, $\mathbf{h}_{SR} \in \mathbb{C}^{N_R \times 1}$ and $\mathbf{h}_{RD} \in \mathbb{C}^{1 \times N_R}$ we denote the fast-fading complex envelopes of the source-to-destination, source-to-relay, and relay-to-destination links, respectively. We further assume that the relay and the destination have complete knowledge of the instantaneous channel-state information (CSI). This serves as an upper bound for the resulting performance. Finally, with p_S and p_R we denoted the power levels used for the transmission from the source and the relay, respectively. A total power budget equal to p_{tot} is considered, which implies $p_{tot} = p_S + p_R$.

HD Relaying

Under HD relaying, the communication between the source and the destination is realized in two time slots. In the first time slot, only the source transmits and its symbol is received by the relay, which performs maximal ratio combining (MRC) [97], and by the destination. In the second time slot, only the relay transmits by using the maximal ratio transmission (MRT) [97] principle and the signal received

at the destination is combined, by leveraging the maximal ratio combining principle, with the signal received in the first time slot.

DF Relaying

First time slot – In the first time slot, the signal received at the destination, which we denote by $y_{D_{DF}}^{(HD)(1)}$, is given by

$$y_{D_{DF}}^{(HD)(1)} = \sqrt{p_S P_L(d_{SD})} h_{SD} s + n_D^{(1)}. \tag{5.1}$$

Hence, the SNR at the destination in the first time slot, which we denote by $\gamma_{D_{DF}}^{(HD)(1)}$, is given by

$$\gamma_{D_{DF}}^{(HD)(1)} = \frac{p_S P_L(d_{SD}) |h_{SD}|^2}{N_0}. \tag{5.2}$$

The signal received at the relay, which we denote by $y_{R_{DF}}^{(HD)}$, is given by

$$y_{R_{DF}}^{(HD)} = \sqrt{p_S P_L(d_{SR})} \mathbf{w}_{R_{MRC}}^{(HD)} \mathbf{h}_{SR} s + \mathbf{w}_{R_{MRC}}^{(HD)} \mathbf{n}_R, \tag{5.3}$$

where

$$\mathbf{w}_{R_{MRC}}^{(HD)} = \frac{\mathbf{h}_{SR}^H}{\|\mathbf{h}_{SR}\|} \tag{5.4}$$

is the MRC combining vector at the relay. Consequently, the SNR at the relay, which we denote by $\gamma_{R_{DF}}^{(HD)}$, is given by

$$\gamma_{R_{DF}}^{(HD)} = \frac{p_S P_L(d_{SR}) \|\mathbf{h}_{SR}\|^2}{N_0}. \tag{5.5}$$

Second time slot – Let us denote the remodulated symbol at the relay after decoding by $\bar{s}$. According to the MRT principle, the signal received at the destination in the second time slot, which is denoted by $y_{D_{DF}}^{(HD)(2)}$, is given by

$$y_{D_{DF}}^{(HD)(2)} = \sqrt{p_R P_L(d_{RD})} \mathbf{h}_{RD} \mathbf{w}_{R_{MRT}}^{(HD)} \bar{s} + n_D^{(2)}, \tag{5.6}$$

where

$$\mathbf{w}_{R_{MRT}}^{(HD)} = \frac{\mathbf{h}_{RD}^H}{\|\mathbf{h}_{RD}\|} \tag{5.7}$$

is the MRT combining vector at the relay. Hence, the SNR at the destination in the second time slot, which is denoted by $\gamma_{D_{DF}}^{(HD)(2)}$, is given by

$$\gamma_{D_{DF}}^{(HD)(2)} = \frac{p_R P_L(d_{RD}) \|\mathbf{h}_{RD}\|^2}{N_0}. \tag{5.8}$$

By combining $y_{D_{DF}}^{(HD)(1)}$ and $y_{D_{DF}}^{(HD)(2)}$ at the receiver, the resulting combined signal, which we denote by

$y_{D_{DF}}^{(HD)}$, is given by

$$y_{D_{DF}}^{(HD)} = \begin{bmatrix} y_{D_{DF}}^{(HD)(1)} & y_{D_{DF}}^{(HD)(2)} \end{bmatrix} \mathbf{w}_{D_{MRC}}, \tag{5.9}$$

where

$$\mathbf{w}_{D_{MRC}} = \frac{\begin{bmatrix} h_{SD} & \|\mathbf{h}_{RD}\| \end{bmatrix}^H}{\left\| \begin{bmatrix} h_{SD} & \|\mathbf{h}_{RD}\| \end{bmatrix} \right\|} \tag{5.10}$$

is the MRC combining vector at the destination. Consequently, the resulting SNR at the destination, which we denote by $\gamma_{D_{DF}}^{(HD)}$, is given by

$$\gamma_{D_{DF}}^{(HD)} = \frac{p_S P_L(d_{SD}) |h_{SD}|^2}{N_0} + \frac{p_R P_L(d_{RD}) \|\mathbf{h}_{RD}\|^2}{N_0} = \gamma_{D_{DF}}^{(HD)(1)} + \gamma_{D_{DF}}^{(HD)(2)}. \tag{5.11}$$

AF Relaying

First time slot – In the first time slot, the signals received at the destination and the relay, which we denote by $y_{D_{AF}}^{(HD)(1)}$ and $y_{R_{AF}}^{(HD)}$, are given by (5.1) and (5.3), respectively. Consequently, the corresponding SNR, which we denote by $\gamma_{D_{AF}}^{(HD)(1)}$ and $\gamma_{R_{AF}}^{(HD)}$, are given by (5.2) and (5.5), respectively.

Second time slot – By assuming a variable-gain relay, in the second time slot the signal received at the destination, which we denote by $y_{D_{AF}}^{(HD)(2)}$, is given by

$$y_{D_{AF}}^{(HD)(2)} = G_R^{(HD)} y_{R_{AF}}^{(HD)} \sqrt{p_R P_L(d_{RD})} \mathbf{h}_{RD} \mathbf{w}_{R_{MRT}}^{(HD)} + n_D^{(2)} = G_R^{(HD)} y_{R_{AF}}^{(HD)} \sqrt{p_R P_L(d_{RD})} \|\mathbf{h}_{RD}\| + n_D^{(2)}, \tag{5.12}$$

where

$$G_R^{(HD)} = \sqrt{\frac{1}{p_S P_L(d_{SR}) \|\mathbf{h}_{SR}\|^2 + N_0}} \tag{5.13}$$

is the gain of the relay. As a result, the resulting SNR at the destination, which we denote by $\gamma_{D_{AF}}^{(HD)(2)}$, is given by

$$\gamma_{D_{AF}}^{(HD)(2)} = \frac{\gamma_{R_{AF}}^{(HD)} \gamma_{D_{AF}}^{'(HD)(2)}}{\gamma_{R_{AF}}^{(HD)} + \gamma_{D_{AF}}^{'(HD)(2)} + 1}, \tag{5.14}$$

where $\gamma_{D_{AF}}^{'(HD)(2)} = \gamma_{D_{DF}}^{(HD)(2)}$. After MRC combining at the destination of $y_{D_{AF}}^{(HD)(1)}$ and $y_{D_{AF}}^{(HD)(2)}$, the resulting SNR, which we denote by $\gamma_{D_{AF}}^{(HD)}$, is given by

$$\gamma_{D_{AF}}^{(HD)} = \gamma_{D_{AF}}^{(HD)(1)} + \gamma_{D_{AF}}^{(HD)(2)}. \tag{5.15}$$

FD Relaying

As for HD relaying, under FD relaying the signal transmission is realized in two time slots. However, in contrast to the HD case, both the source and the relay are allowed to transmit concurrently on th same physical resource, i.e., a time slot [98], [99]. Due to this protocol, by assuming that in time slot k the source transmits the symbol s_k, the signal received at the relay is affected by the loop-back self-interference due to the concurrent transmission from the relay of the signal that corresponds to s_{k-1}, which is the symbol transmitted by the source in time slot $k-1$. Furthermore, the signal received by the destination from the relay in time slot $k+1$, which corresponds to the symbol s_k, is affected by the interference resulting from the concurrent transmission of the symbol s_{k+1} by the source. Hence, after M signaling transmissions the resulting symbol rate under the FD protocol is equal to $\frac{M}{M+1} \rightarrow 1$, as $M \rightarrow \infty$, which is the same as the symbol rate of a single-input single output (SISO) transmission. Such a rate advantage of FD relaying compared to its HD counterpart comes at the cost of an increased interference, as explained.

If $N_R \geq 2$ antennas are available at the relay, a way to perform FD operation is to allocate half of the antennas for reception and the rest half for transmission. Another way is to simultaneously use the N_R antennas for both transmission and reception by using circuits that are called circulators. In this work, we consider the former approach. Accordingly, by $\tilde{\mathbf{h}}_{SR} \in \mathbb{C}^{\frac{N_R}{2} \times 1}$ and $\tilde{\mathbf{h}}_{RD} \in \mathbb{C}^{1 \times \frac{N_R}{2}}$ we denote the fast-fading complex envelopes of the source-to-relay and relay-to-destination links, respectively.

The loop-back self-interference consists of two parts: (i) the direct-path loop-back self-interference propagating directly from the transmit to the receive chain. It can either comprise the LOS signal propagating directly from the transmit to the receive antennas in a separate antenna structure or the signal reaching the receive chain due to the circulators' leakage (due to antenna mismatching, for instance) in a shared antenna structure. The power impact of this component can be much larger than the power of the received signal, which renders the realization of FD-relaying challenging. Hence, it is essential that such a component is substantially suppressed so that it is comparable or smaller than the noise level. This can be realized by the use of highly directional antennas or significant isolation between the transmit and receive antennas, as well as analog and digital cancellation techniques [99]; and (ii) a multipath part due to scattering and reflections from objects in the radio path. We assume that this fast-fading coefficient for each transmit-receive antenna pair of the relay is described by a zero-mean complex Gaussian random variable with variance that depends on the distance-based path-loss and the shadowing effects that the scattering components experience before reaching the receive antennas.

Based on the above, a widespread approach for modeling the loop-back self-interference consists of assuming that the residual loop-back self-interference, which is the remaining interference after all cancellation stages are implemented, is as a complex Ricean random variable [100]. The strong LOS component of such a process represents the residual level of the direct-path loop-back self-interference. By $\tilde{\mathbf{h}}_{LI} \in \mathbb{C}^{\frac{N_R}{2} \times 1}$, we denote the complex envelope representing such a process. Furthermore, we assume that the relay perfectly knows $\tilde{\mathbf{h}}_{LI}$ through, for example, an estimation phase prior to data transmission.

DF Relaying

First time slot – In the first time slot, the signal received at the relay, which we denote by $y_{R_{DF}}^{(FD)}$, is given

by

$$y_{R_{DF}}^{(FD)} = \sqrt{p_S P_L(d_{SR})}\mathbf{w}_{R_{MRC}}^{(FD)}\tilde{\mathbf{h}}_{SR}s + \sqrt{p_R}\mathbf{w}_{R_{MRC}}^{(FD)}\tilde{\mathbf{h}}_{LI}\tilde{s}' + \mathbf{w}_{R_{MRC}}^{(FD)}\mathbf{n}_R, \tag{5.16}$$

where $\tilde{s}'$ is the remodulated symbol, which was conveyed by the source in the previous time slot, after decoding and

$$\mathbf{w}_{R_{MRC}}^{(FD)} = \frac{\tilde{\mathbf{h}}_{SR}^H}{\left\|\tilde{\mathbf{h}}_{SR}\right\|}. \tag{5.17}$$

is the MRC combining vector at the relay. Consequently, the signal-to-interference-plus-noise ratio (SINR) at the relay, which we denote by $\gamma_{R_{DF}}^{(FD)}$, is given by

$$\gamma_{R_{DF}}^{(FD)} = \frac{p_S P_L(d_{SR})\left\|\tilde{\mathbf{h}}_{SR}\right\|^2}{p_R\left|\mathbf{w}_{R_{MRC}}^{(FD)}\tilde{\mathbf{h}}_{LI}\right|^2 + N_0}. \tag{5.18}$$

Second time slot – In the second time slot, the signal received at the destination and that corresponds to the decoded and remodulated symbol $\tilde{s}$, which is denoted by $y_{D_{DF}}^{(FD)}$, is given by

$$y_{D_{DF}}^{(FD)} = \sqrt{p_R P_L(d_{RD})}\tilde{\mathbf{h}}_{RD}\mathbf{w}_{R_{MRT}}^{(FD)}\tilde{s} + \sqrt{p_S P_L(d_{SD})}h_{SD}s'' + n_D, \tag{5.19}$$

where s'' is the symbols conveyed by the source in the current time slot and

$$\mathbf{w}_{R_{MRT}}^{(FD)} = \frac{\tilde{\mathbf{h}}_{RD}^H}{\left\|\tilde{\mathbf{h}}_{RD}\right\|} \tag{5.20}$$

is the MRT combining vector at the relay. Hence, the resulting SINR at the destination, which we denote by $\gamma_{D_{DF}}^{(FD)}$, is given by

$$\gamma_{D_{DF}}^{(FD)} = \frac{p_R P_L(d_{RD})\left\|\tilde{\mathbf{h}}_{RD}\right\|^2}{p_S P_L(d_{SD})|h_{SD}|^2 + N_0}. \tag{5.21}$$

AF Relaying

First time slot – In the first time slot, the signal received at the relay, which we denote by $y_{R_{AF}}^{(FD)}$, is given by

$$y_{R_{AF}}^{(FD)} = \sqrt{p_S P_L(d_{SR})}\mathbf{w}_{R_{MRC}}^{(FD)}\tilde{\mathbf{h}}_{SR}s + \sqrt{p_R}\mathbf{w}_{R_{MRC}}^{(FD)}\tilde{\mathbf{h}}_{LI}G_R^{(FD)}y_{R_{AF}}^{\prime(FD)} + \mathbf{w}_{R_{MRC}}^{(FD)}\mathbf{n}_R, \tag{5.22}$$

where $y_{R_{AF}}^{\prime(FD)}$ is the signal received at the relay in the previous time slot and

$$G_R^{(FD)} = \sqrt{\frac{1}{p_S P_L(d_{SR})\left\|\tilde{\mathbf{h}}_{SR}\right\|^2 + p_R\left|\mathbf{w}_{R_{MRC}}^{(FD)}\tilde{\mathbf{h}}_{LI}\right|^2 + N_0}} \tag{5.23}$$

is the gain of the relay. Consequently, the resulting SINR at the relay, which we denote by $\gamma_{R_{AF}}^{(FD)}$, is given by

$$\gamma_{R_{AF}}^{(FD)} = \frac{p_S P_L(d_{SR}) \left\| \tilde{\mathbf{h}}_{SR} \right\|^2}{p_R \left| \mathbf{w}_{R_{MRC}}^{(FD)} \tilde{\mathbf{h}}_{LI} \right|^2 + N_0}. \tag{5.24}$$

Second time slot – In the second time slot, the signal received at the destination, which we denote by $y_{D_{AF}}^{(FD)}$, is given by

$$y_{D_{AF}}^{(FD)} = G_R y_{R_{AF}}^{(FD)} \sqrt{p_R P_L(d_{RD})} \tilde{\mathbf{h}}_{RD} \mathbf{w}_{R_{MRT}}^{(FD)} + \sqrt{p_S P_L(d_{SD})} h_{SD} s'' + n_D \tag{5.25}$$

After some algebraic manipulations, the SINR at the destination, which we denote by $\gamma_{D_{AF}}^{(FD)}$, is given by

$$\gamma_{D_{AF}}^{(FD)} = \frac{\gamma_{R_{AF}}^{(FD)} \gamma_{D_{AF}}^{(FD)(2)}}{\gamma_{R_{AF}}^{(FD)} + \gamma_{D_{AF}}^{(FD)(2)} + 1}, \tag{5.26}$$

where $\gamma_{D_{AF}}^{(FD)(2)} = \gamma_{D_{DF}}^{(FD)}$.

5.3.2 RIS-Assisted Communication

Let us denote by $\mathbf{h}_{S,RIS} \in \mathbb{C}^{N_{RIS} \times 1}$ and $\mathbf{h}_{RIS,D} \in \mathbb{C}^{1 \times N_{RIS}}$ the fast-fading complex envelopes of the source-to-RIS and RIS-to-destination links, respectively. In addition, $\Phi = \text{diag}\left(e^{j\varphi_1}, \ldots, e^{j\varphi_{N_{RIS}}}\right)$ denotes the diagonal matrix representing the reflective and phase-shifting values of the RIS elements. It holds that $R_A \in (0,1]$, with $R_A = 1$ corresponding to a lossless reflection. Based on the above, the received signal at the destination consists of the sum of the direct link signal and of the signal that is reflected by the RIS. By denoting it by y_D^{RIS}, it holds that

$$y_D^{RIS} = \sqrt{p_{tot} P_L(d_{SD})} h_{SD} s + \sqrt{p_{tot} P_L^{RIS}} \mathbf{h}_{S,RIS}^T \Phi \mathbf{h}_{RIS,D}^T s + n_D, \tag{5.27}$$

where P_L^{RIS} is the path-loss in the RIS case. Hence, the resulting SNR, which we denote by γ_D^{RIS}, is given by

$$\gamma_D^{RIS} = \frac{\left| \sqrt{p_{tot} P_L(d_{SD})} h_{SD} + \sqrt{p_{tot} P_L^{RIS}} \mathbf{h}_{S,RIS}^T \Phi \mathbf{h}_{RIS,D}^T \right|^2}{N_0} \tag{5.28}$$

The maximum γ_D^{RIS}, which we denote by $\gamma_{D_{max}}^{RIS}$, is achieved by adjusting the phase shifts $\varphi_1, \ldots, \varphi_{N_{RIS}}$ in a way that the direct path signal and the reflected signals from the RIS are co-phased. Based on this, the optimal phase shifts employed by the RIS, which we denote by $\varphi_1^{opt}, \ldots, \varphi_{N_{RIS}}^{opt}$, are given by

$$\begin{pmatrix} \varphi_1^{opt} & \cdots & \varphi_{N_{RIS}}^{opt} \end{pmatrix} = \begin{pmatrix} \arg[h_{SD}] - \arg\left[\left[h_{S,RIS}\right]_1 \left[h_{RIS,D}\right]_1\right] & \cdots & \arg[h_{SD}] - \arg\left[\left[h_{S,RIS}\right]_{N_{RIS}} \left[h_{RIS,D}\right]_{N_{RIS}}\right] \end{pmatrix}. \tag{5.29}$$

As a result, it holds that

$$\gamma_{D_{\max}}^{RIS} = \frac{p_{tot}\left(\sqrt{P_L(d_{SD})}|h_{SD}| + \sqrt{P_L^{RIS}}R_A \sum_{n=1}^{N_{RIS}} [h_{S,RIS}]_n [h_{RIS,D}]_n\right)^2}{N_0}. \tag{5.30}$$

As far as P_L^{RIS} and the resulting instantaneous maximum SNR are concerned, we consider the two extreme cases where the elements of the RIS act as anomalous reflectors and diffuse scatterers.

Anomalous Reflection

This case occurs when the size of the RIS elements is sufficiently large so that they act as anomalous reflectors, and the theory of geometric optics can be applied to model the interactions between the RIS and the signals [86]. In such a case, according to the theory of geometric optics, the received power under free-space propagation is expected to scale with $\left[(d_{SR}+d_{RD})^2\right]^{-1}$, which is the sum-distance law. Hence, it holds that $P_L^{RIS} \approx P_L(d_{SR}+d_{RD})$. As a result, the instantaneous maximum SNR, which we denote by $\gamma_{D_{\max}}^{RIS,\,anomalous\,reflection}$, is approximated as

$$\gamma_{D_{\max}}^{RIS,\,anomalous\,reflection} \approx \frac{p_{tot}\left(\sqrt{P_L(d_{SD})}|h_{SD}| + \sqrt{P_L(d_{SR}+d_{RD})}R_A \sum_{n=1}^{N_{RIS}} [h_{S,RIS}]_n [h_{RIS,D}]_n\right)^2}{N_0}. \tag{5.31}$$

Diffuse Scattering

This case occurs when the size of the elements of the RIS is comparable to the wavelength. In such a case, the elements of the RIS are expected to act as diffusers (dipole scatterers). Due to the incoherent summation at the receiver of the diffuse components, the received power under free-space propagation is expected to scale with $\left(d_{SR}^2 d_{RD}^2\right)^{-1}$, which is the same product law of the distance as in backscattering communication, for instance, which is characterized by the incoherent sum of diffuse components. Hence, it holds that $P_L^{RIS} \approx P_L(d_{SR})\,P_L(d_{RD})$. As a result, the instantaneous maximum SNR, which we denote by $\gamma_{D_{\max}}^{RIS,\,diffuse\,scattering}$, is approximated as

$$\gamma_{D_{\max}}^{RIS,\,diffuse\,scattering} \approx \frac{p_{tot}\left(\sqrt{P_L(d_{SD})}|h_{SD}| + \sqrt{P_L(d_{SR})\,P_L(d_{RD})}R_A \sum_{n=1}^{N_{RIS}} [h_{S,RIS}]_n [h_{RIS,D}]_n\right)^2}{N_0}. \tag{5.32}$$

The end-to-end SNR for relay-assisted and RIS-assisted communications is summarized in Table 5.1.

Table 5.1 – End-to-end SNR of relay-assisted and RIS-assisted communications

HD relaying under the DF protocol	$\gamma_{R_{DF}}^{(HD)} = \frac{p_S P_L(d_{SR})\|\mathbf{h}_{SR}\|^2}{N_0}$, $\gamma_{D_{DF}}^{(HD)} = \frac{p_S P_L(d_{SD})\|h_{SD}\|^2}{N_0} + \frac{p_R P_L(d_{RD})\|\mathbf{h}_{RD}\|^2}{N_0}$
HD relaying under the AF protocol	$\gamma_{D_{AF}}^{(HD)} = \frac{p_S P_L(d_{SD})\|h_{SD}\|^2}{N_0} + \frac{\frac{p_S P_L(d_{SR})\|\mathbf{h}_{SR}\|^2}{N_0}\frac{p_R P_L(d_{RD})\|\mathbf{h}_{RD}\|^2}{N_0}}{\frac{p_S P_L(d_{SR})\|\mathbf{h}_{SR}\|^2}{N_0} + \frac{p_R P_L(d_{RD})\|\mathbf{h}_{RD}\|^2}{N_0} + 1}$
FD relaying under the DF protocol	$\gamma_{R_{DF}}^{(FD)} = \frac{\frac{p_S P_L(d_{SR})\|\tilde{\mathbf{h}}_{SR}\|^2}{N_0}}{\frac{p_R\left\|\mathbf{w}_{R_{MRC}}^{(FD)}\tilde{\mathbf{h}}_{LI}\right\|^2}{N_0}+1}$, $\gamma_{D_{DF}}^{(FD)} = \frac{\frac{p_R P_L(d_{RD})\|\tilde{\mathbf{h}}_{RD}\|^2}{N_0}}{\frac{p_S P_L(d_{SD})\|h_{SD}\|^2}{N_0}+1}$
FD relaying under the AF protocol	$\gamma_{D_{AF}}^{(FD)} = \frac{\frac{\frac{p_S P_L(d_{SR})\|\tilde{\mathbf{h}}_{SR}\|^2}{N_0}}{\frac{p_R\left\|\mathbf{w}_{R_{MRC}}^{(FD)}\tilde{\mathbf{h}}_{LI}\right\|^2}{N_0}+1}\frac{\frac{p_R P_L(d_{RD})\|\tilde{\mathbf{h}}_{RD}\|^2}{N_0}}{\frac{p_S P_L(d_{SD})\|h_{SD}\|^2}{N_0}+1}}{\frac{\frac{p_S P_L(d_{SR})\|\tilde{\mathbf{h}}_{SR}\|^2}{N_0}}{\frac{p_R\left\|\mathbf{w}_{R_{MRC}}^{(FD)}\tilde{\mathbf{h}}_{LI}\right\|^2}{N_0}+1} + \frac{\frac{p_R P_L(d_{RD})\|\tilde{\mathbf{h}}_{RD}\|^2}{N_0}}{\frac{p_S P_L(d_{SD})\|h_{SD}\|^2}{N_0}+1}+1}$
RIS as an anomalous reflector	$\gamma_{D_{\max}}^{RIS,\,anomalous\,reflection} \approx \frac{p_{tot}\left(\sqrt{P_L(d_{SD})}\|h_{SD}\|+\sqrt{P_L(d_{SR}+d_{RD})}R_A\sum_{n=1}^{N_{RIS}}[h_{S,RIS}]_n[h_{RIS,D}]_n\right)^2}{N_0}$
RIS as a dipole scatterer	$\gamma_{D_{\max}}^{RIS,\,diffuse\,scattering} \approx \frac{p_{tot}\left(\sqrt{P_L(d_{SD})}\|h_{SD}\|+\sqrt{P_L(d_{SR})P_L(d_{RD})}R_A\sum_{n=1}^{N_{RIS}}[h_{S,RIS}]_n[h_{RIS,D}]_n\right)^2}{N_0}$

5.4 Performance Comparison – Achievable Rate

In this section, we provide the mathematical formulation for the achievable rate for both relay-assisted and RIS-assisted communications. For the relaying case, we derive the optimal allocation of p_S and p_R so that the achievable rate is maximized.

5.4.1 Relay-Assisted Communication

Let us first define the following parameters that are used in the computation of the optimal power allocation for each of the following case studies.

$$\begin{aligned}
A &= \frac{P_L(d_{SD})\,|h_{SD}|^2}{N_0} \\
B^{(HD)} &= \frac{P_L(d_{SR})\,\|\mathbf{h}_{SR}\|^2}{N_0} \\
C^{(HD)} &= \frac{P_L(d_{RD})\,\|\mathbf{h}_{RD}\|^2}{N_0} \\
B^{(FD)} &= \frac{P_L(d_{SR})\,\left\|\tilde{\mathbf{h}}_{SR}\right\|^2}{N_0} \\
C^{(FD)} &= \frac{P_L(d_{RD})\,\left\|\tilde{\mathbf{h}}_{RD}\right\|^2}{N_0} \\
D &= \frac{\left|\mathbf{w}_{R_{MRC}}^{(FD)}\tilde{\mathbf{h}}_{LI}\right|^2}{N_0}
\end{aligned} \tag{5.33}$$

HD Relaying

DF Operation The achievable rate under HD-DF relaying, which we denote by $R_{DF}^{(HD)}$, is given by

$$R_{DF}^{(HD)} = \frac{1}{2}\log_2\left(1+\min\left(\gamma_{R_{DF}}^{(HD)}, \gamma_{D_{DF}}^{(HD)(1)}+\gamma_{D_{DF}}^{(HD)(2)}\right)\right). \tag{5.34}$$

The maximum rate, which we denote by $R_{DF_{max}}^{(HD)}$, is given by

$$R_{DF_{\max}}^{(HD)} = \max_{p_S,p_R}\frac{1}{2}\log_2\left(1+\min\left(\gamma_{R_{DF}}^{(HD)}, \gamma_{D_{DF}}^{(HD)(1)}+\gamma_{D_{DF}}^{(HD)(2)}\right)\right) = \frac{1}{2}\log_2\left(1+\max_{p_S,p_R}\min\left(\gamma_{R_{DF}}^{(HD)}, \gamma_{D_{DF}}^{(HD)(1)}+\gamma_{D_{DF}}^{(HD)(2)}\right)\right), \tag{5.35}$$

where the last equality holds due to the concavity of the log function. We distinguish two cases.

Case 1: $\gamma_{D_{DF}}^{(HD)(1)} > \gamma_{R_{DF}}^{(HD)}$. In this case, relaying is suboptimal since the direct source-to-destination link achieves a higher rate. Consequently, $p_S = p_{tot}$ and the achievable rate R_{SISO} is given by

$$R_{SISO} = \log_2\left(1+\gamma_{D_{DF}}^{(HD)(1)}\right). \tag{5.36}$$

Case 2: $\gamma_{D_{DF}}^{(HD)(1)} \leq \gamma_{R_{DF}}^{(HD)}$. In this case, the term $\min\left(\gamma_{R_{DF}}^{(HD)}, \gamma_{D_{DF}}^{(HD)(1)}+\gamma_{D_{DF}}^{(HD)(2)}\right)$ in (5.35) is maximized for $\gamma_{R_{DF}}^{(HD)} = \gamma_{D_{DF}}^{(HD)(1)}+\gamma_{D_{DF}}^{(HD)(2)}$. The obtained optimal value of p_S, which we denote by $p_{S_{DF}}^{opt(HD)}$, is given by

$$p_{S_{DF}}^{opt(HD)} = \frac{p_{tot}C^{(HD)}}{B^{(HD)}-A+C^{(HD)}}. \tag{5.37}$$

The optimal value of p_R, which we denote by $p_{R_{DF}}^{opt(HD)}$, is equal to $p_{R_{DF}}^{opt(HD)} = p_{tot} - p_{S_{DF}}^{opt(HD)}$.

AF Operation The achievable rate under HD-AF relaying, which we denote by $R_{AF}^{(HD)}$, is given by

$$R_{AF}^{(HD)} = \frac{1}{2}\log_2\left(1+\gamma_{D_{AF}}^{(HD)}\right). \tag{5.38}$$

The maximum rate, which we denote by $R_{AF_{max}}^{(HD)}$, is given by

$$R_{AF_{\max}}^{(HD)} = \max_{p_S,p_R}\frac{1}{2}\log_2\left(1+\gamma_{D_{AF}}^{(HD)}\right) = \frac{1}{2}\log_2\left(1+\max_{p_S,p_R}\gamma_{D_{AF}}^{(HD)}\right). \tag{5.39}$$

After replacing p_R at $\gamma_{D_{AF}}^{(HD)}$ with $p_{tot}-p_S$ and taking the first-order derivative with respect to p_S and setting it to zero, the optimum value of p_S, which we denote by $p_{S_{AF}}^{opt(HD)}$, is given by

$$p_{S_{AF}}^{opt(HD)} = \frac{\left[AB^{(HD)}-\left(A+B^{(HD)}\right)C^{(HD)}\right]\left(1+C^{(HD)}p_{tot}\right) \pm \sqrt{B^{(HD)}C^{(HD)}\left[\left(A+B^{(HD)}\right)C^{(HD)}-AB^{(HD)}\right]\left(1+B^{(HD)}p_{tot}\right)\left(1+C^{(HD)}p_{tot}\right)}}{\left(B^{(HD)}-C^{(HD)}\right)\left[\left(A+B^{(HD)}\right)C^{(HD)}-AB^{(HD)}\right]}. \tag{5.40}$$

From the two values of $p_{S_{AF}}^{opt(HD)}$, we keep the one for which it holds that $0 < p_{S_{AF}}^{opt(HD)} \leq p_{tot}$. The optimal value of p_R, which we denote by $p_{R_{AF}}^{opt(HD)}$, is equal to $p_{R_{AF}}^{opt(HD)} = p_{tot} - p_{S_{AF}}^{opt(HD)}$.

FD Relaying

DF Operation The achievable rate under FD-DF relaying, which we denote by $R_{DF}^{(FD)}$, is given by

$$R_{DF}^{(FD)} = \log_2\left(1+\min\left(\gamma_{R_{DF}}^{(FD)},\gamma_{D_{DF}}^{(FD)}\right)\right). \tag{5.41}$$

The maximum achievable rate, which we denote by $R_{DF_{max}}^{(FD)}$, is given by

$$R_{DF_{max}}^{(FD)} = \max_{p_S,p_R} \log_2\left(1+\min\left(\gamma_{R_{DF}}^{(FD)},\gamma_{D_{DF}}^{(FD)}\right)\right) = \log_2\left(1+\max_{p_S,p_R}\min\left(\gamma_{R_{DF}}^{(FD)},\gamma_{D_{DF}}^{(FD)}\right)\right). \tag{5.42}$$

Hence, $R_{DF}^{(FD)}$ is maximized if p_S and p_R take the values for which $\gamma_{R_{DF}}^{(FD)} = \gamma_{D_{DF}}^{(FD)}$ holds. Accordingly, the optimum value of p_S, which we denote by $p_{S_{DF}}^{opt(FD)}$, is given by

$$p_{S_{DF}}^{opt(FD)} = \frac{-b_{DF}^{(FD)} \pm \sqrt{\left(b_{DF}^{(FD)}\right)^2 - 4a_{DF}^{(FD)}c_{DF}^{(FD)}}}{2a_{DF}^{(FD)}}, \tag{5.43}$$

where

$$\begin{aligned} a_{DF}^{(FD)} &= AB^{(FD)} - DC^{(FD)} \\ b_{DF}^{(FD)} &= B^{(FD)} + C^{(FD)} + 2p_{tot}DC^{(FD)}. \\ c_{DF}^{(FD)} &= -p_{tot}^2 DC^{(FD)} - p_{tot}C^{(FD)} \end{aligned} \tag{5.44}$$

From the two values of $p_{S_{DF}}^{opt(FD)}$, we keep the one for which it holds that $0 < p_{S_{DF}}^{opt(FD)} \leq p_{tot}$. The optimal value of p_R, which we denote by $p_{R_{DF}}^{opt(FD)}$, is equal to $p_{R_{DF}}^{opt(FD)} = p_{tot} - p_{S_{DF}}^{opt(FD)}$.

AF Operation The achievable rate under FD-AF relaying, which we denote by $R_{AF}^{(FD)}$, is given by

$$R_{AF}^{(FD)} = \log_2\left(1+\gamma_{D_{AF}}^{(FD)}\right). \tag{5.45}$$

The maximum achievable rate, which we denote by $R_{DF_{max}}^{(FD)}$, is given by

$$R_{AF_{max}}^{(FD)} = \max_{p_S,p_R} \log_2\left(1+\gamma_{D_{AF}}^{(FD)}\right) = \log_2\left(1+\max_{p_S,p_R}\gamma_{D_{AF}}^{(FD)}\right). \tag{5.46}$$

Consequently, to maximize $R_{AF_{max}}^{(FD)}$ we replace p_R at $\gamma_{D_{AF}}^{(FD)}$ with $p_{tot} - p_S$ and then we take the first-order derivative with respect to p_S and set it to zero. According to the solution, the optimum value of p_S, which we denote by $p_{S_{AF}}^{opt(FD)}$, is given by

$$p_{S_{AF}}^{opt(FD)} = \frac{-b_{AF}^{(FD)} \pm \sqrt{\left(b_{AF}^{(FD)}\right)^2 - 4a_{AF}^{(FD)}c_{AF}^{(FD)}}}{2a_{AF}^{(FD)}}, \tag{5.47}$$

where

$$\begin{aligned} a_{AF}^{(FD)} &= \left(AB^{(FD)} - C^{(FD)}D\right)p_{tot} + A + B^{(FD)} - C^{(FD)} - D \\ b_{AF}^{(FD)} &= 2C^{(FD)}Dp_{tot}^2 + 2\left(C^{(FD)} + D\right)p_{tot} + 2. \\ c_{AF}^{(FD)} &= -C^{(FD)}Dp_{tot}^3 - \left(C^{(FD)} + D\right)p_{tot}^2 - p_{tot} \end{aligned} \tag{5.48}$$

From the two values of $p_{S_{AF}}^{opt(FD)}$, we keep the one for which it holds that $0 < p_{S_{AF}}^{opt(FD)} \leq p_{tot}$. The optimal value of p_R, which we denote by $p_{R_{AF}}^{opt(FD)}$, is equal to $p_{R_{AF}}^{opt(FD)} = p_{tot} - p_{S_{AF}}^{opt(FD)}$.

5.4.2 RIS-Assisted Communication

As mentioned, RISs are an emerging field of research in wireless communications. The interested readers can refer to [16] and [86] for further information. Depending on the ratio between the geometric size and the wavelength of the radio wave, two limiting cases can be considered.

Anomalous Reflection

If the geometric size of the RIS is sufficiently large as compared with the wavelength of the radio wave, then each element of the RIS behaves like a mirror that can reflect the signals towards directions that are different from the direction of arrival. This case study is referred to as anomalous reflection.

The maximum achievable rate in this case, which we denote by $R_{\max}^{RIS,\,anomalous\,reflection}$, is given by

$$\begin{aligned} R_{\max}^{RIS,\,anomalous\,reflection} &= \log_2\left(1 + \gamma_{D_{\max}}^{RIS,\,anomalous\,reflection}\right) \\ &\approx \log_2\left(1 + \frac{p_{tot}\left(\sqrt{P_L(d_{SD})}|h_{SD}| + \sqrt{P_L(d_{SR}+d_{RD})}R_A\sum_{n=1}^{N_{RIS}}\left[h_{S,RIS}\right]_n\left[h_{RIS,D}\right]_n\right)^2}{N_0}\right). \end{aligned} \tag{5.49}$$

Diffuse Scattering

If the geometric size of the RIS is of the same order as or smaller than the wavelength of the radio wave, then the RIS behaves as a dipole scatterer that diffuses the signals towards all possible directions. This case study is referred to as diffuse scattering.

The maximum achievable rate in this case, which we denote by $R_{\max}^{RIS,\,diffuse\,scattering}$, is given by

$$\begin{aligned} R_{\max}^{RIS,\,diffuse\,scattering} &= \log_2\left(1 + \gamma_{D_{\max}}^{RIS,\,diffuse\,scattering}\right) \\ &\approx \log_2\left(1 + \frac{p_{tot}\left(\sqrt{P_L(d_{SD})}|h_{SD}| + \sqrt{P_L(d_{SR})P_L(d_{RD})}R_A\sum_{n=1}^{N_{RIS}}\left[h_{S,RIS}\right]_n\left[h_{RIS,D}\right]_n\right)^2}{N_0}\right). \end{aligned} \tag{5.50}$$

Table 5.2 – Achievable rate of relay-assisted and RIS-assisted communications

HD relaying under the DF protocol	$R_{DF_{\max}}^{(HD)} = \frac{1}{2}\log_2\left(1+\max_{p_S,p_R}\min\left(\gamma_{R_{DF}}^{(HD)},\gamma_{D_{DF}}^{(HD)(1)}+\gamma_{D_{DF}}^{(HD)(2)}\right)\right)$, if $\gamma_{D_{DF}}^{(HD)(1)} \leq \gamma_{R_{DF}}^{(HD)}$
HD relaying under the AF protocol	$R_{AF_{\max}}^{(HD)} = \frac{1}{2}\log_2\left(1+\max_{p_S,p_R}\gamma_{D_{AF}}^{(HD)}\right)$
FD relaying under the DF protocol	$R_{DF_{max}}^{(FD)} = \log_2\left(1+\max_{p_S,p_R}\min\left(\gamma_{R_{DF}}^{(FD)},\gamma_{D_{DF}}^{(FD)}\right)\right)$
FD relaying under the AF protocol	$R_{AF_{max}}^{(FD)} = \log_2\left(1+\max_{p_S,p_R}\gamma_{D_{AF}}^{(FD)}\right)$
RIS as an anomalous reflector	$R_{\max}^{RIS,\,anomalous\,reflection} \approx \log_2\left(1+\frac{p_{tot}\left(\sqrt{P_L(d_{SD})}\lvert h_{SD}\rvert+\sqrt{P_L(d_{SR}+d_{RD})}R_A\sum_{n=1}^{N_{RIS}}[h_{S,RIS}]_n[h_{RIS,D}]_n\right)^2}{N_0}\right)$
RIS as a dipole scatterer	$R_{\max}^{RIS,\,diffuse\,scattering} \approx \log_2\left(1+\frac{p_{tot}\left(\sqrt{P_L(d_{SD})}\lvert h_{SD}\rvert+\sqrt{P_L(d_{SR})P_L(d_{RD})}R_A\sum_{n=1}^{N_{RIS}}[h_{S,RIS}]_n[h_{RIS,D}]_n\right)^2}{N_0}\right)$

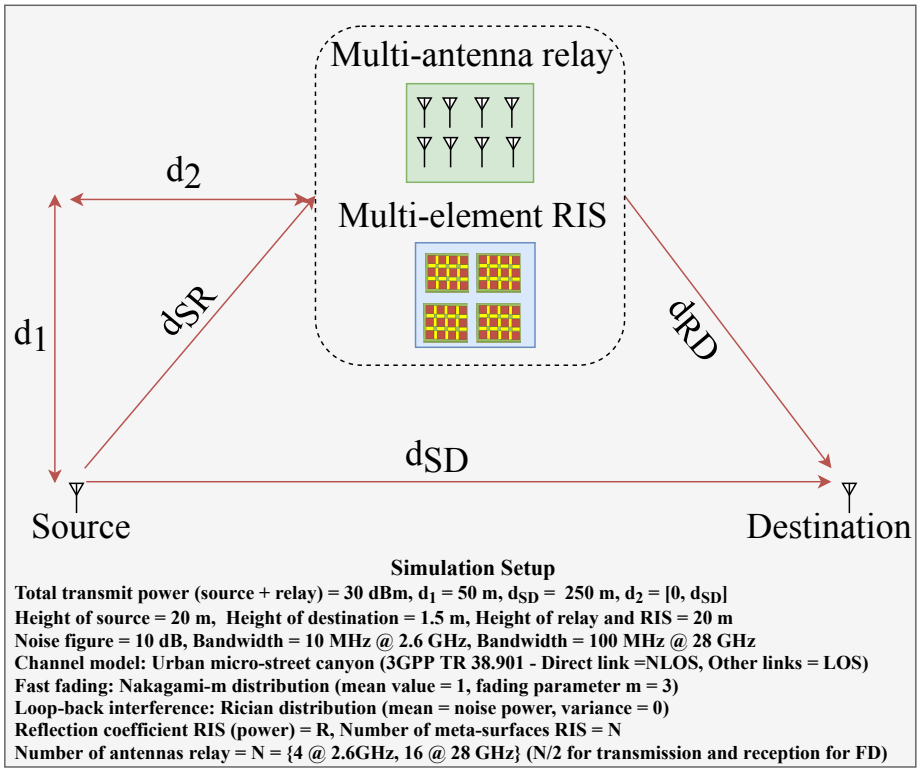

Figure 5.4 – Simulation scenario and setup.

The formulas of the achievable rate for relay-assisted and RIS-assisted communications are summarized in Table 5.2.

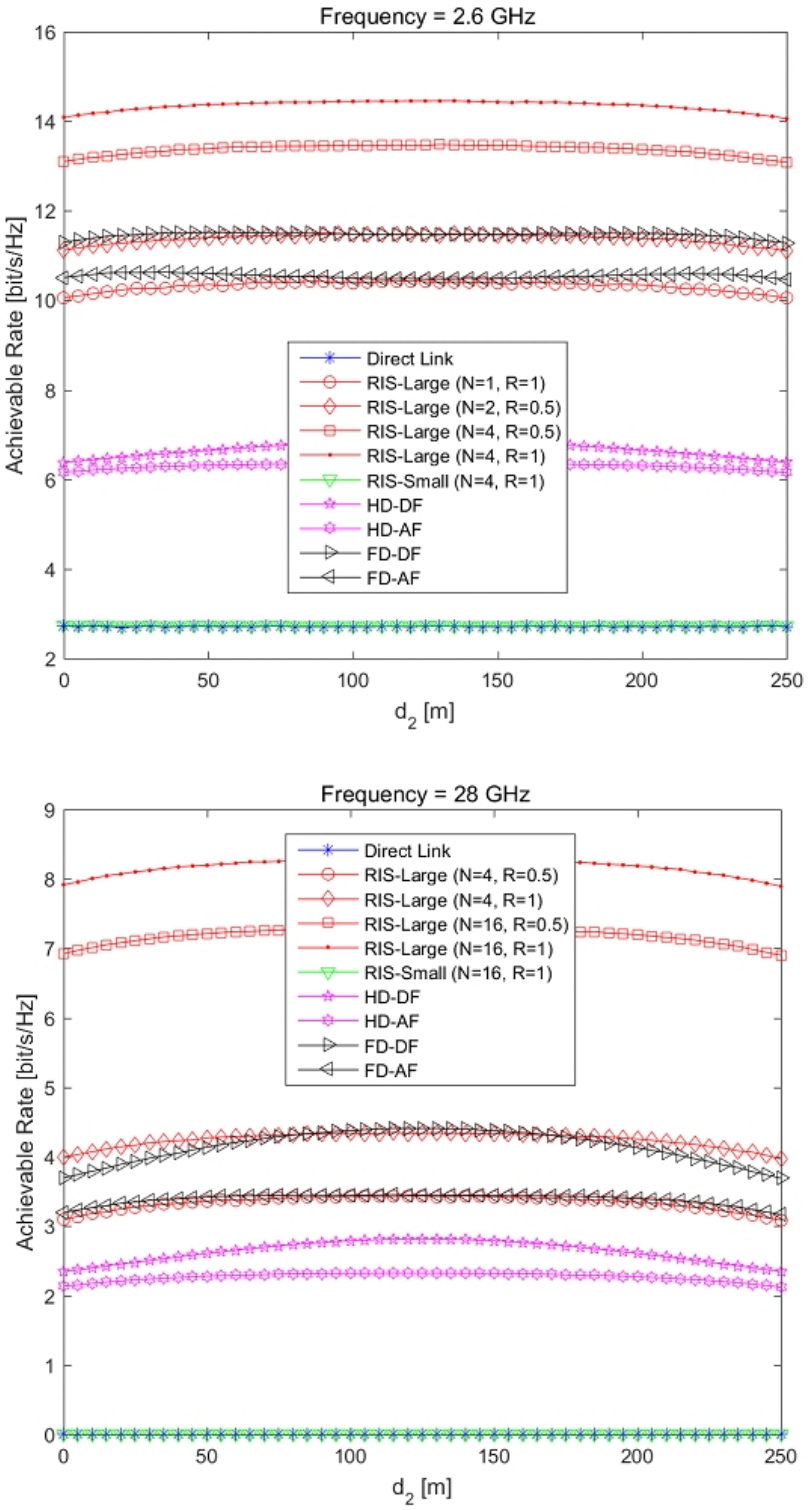

Figure 5.5 – Achievable data rate (bit/s/Hz) of RISs and relaying versus the distance d_2 for the setup in Fig. 5.4.

5.5 Numerical Results

In Fig. 5.5, we compare the performance of RISs and relays based on the simulation setup illustrated in Fig. 5.4. The results are obtained by employing the analytical formulas obtained in the previous sections. The chosen setup ensures a fair comparison between RISs and relays, since different reflection coefficients for the RISs are analyzed; an optimistic (equal to the noise power) residual loop-back self-interference model is assumed for FD relaying; and, at low frequencies, RISs with a few (even a single) reconfigurable meta-surfaces are considered. For relaying, the transmit powers at the transmitter and relay are optimized to maximize the achievable rate while enforcing a total power constraint. In all cases, perfect channel state information is assumed.

The results unveil the promising performance offered by *electrically large RISs*, even if the reflection coefficient R is small (i.e., $R = 0.5$ so that only 50% of the total impinging power is reflected by the RIS), and $N = 1$ or $N = 2$ reconfigurable meta-surfaces can be accommodated at sub-6 GHz. We highlight three main conclusions from the analysis of Fig. 5.5: (i) at 2.6 GHz, RISs with $N = 1$, $R = 1$ and $N = 2$, $R = 0.5$ offer rates similar to FD-AF and FD-DF relaying with 4 antennas, optimal power allocation, and optimal transmission/reception vectors, respectively; (ii) at 28 GHz, RISs with $N = 4$, $R = 0.5$ and $N = 4$, $R = 1$ offer rates similar to 16-antenna FD-AF and FD-DF relaying with the same optimal setup as for 2.6 GHz; and (iii) RISs with $N = 4$ at 2.6 GHz and $N = 16$ at 28 GHz, respectively, outperform relaying and double the data rate in the latter case study. It is worth noting that the area occupied by the considered RISs is of the order of 1.33 m^2 for $N = 1$; 2.66 m^2 for $N = 2$; and 5.32 m^2 for $N = 4$ at 2.6 GHz; and only 0.046 m^2 for $N = 4$; and 0.18 m^2 for $N = 16$ at 28 GHz. RISs of 6 m^2 in size have recently been built for operation at 2.45 GHz [90]. These results are encouraging, and highlight the potential benefits of employing RISs in wireless networks.

5.6 The Road Ahead

Theoretical and experimental research on RISs is still at its infancy. Fundamental issues need to be understood to unlock the potential of this emerging technology, to identify the ultimate performance limits, and to optimally design RIS-aided wireless networks. Three fundamental and open research issues deserve, in our opinion, more attention than others.

Physics-Based Modeling. Current research on RISs relies on simplified assumptions and models about the interactions of the radio waves with the meta-surfaces. No equivalent models for the meta-surfaces are available that are suitable for application to wireless networks. Hence, there is a compelling need for developing sufficiently accurate but analytically tractable models for the meta-surfaces whose foundation is to be built on the fundamental laws of electromagnetism and physics.

Experimental Validation. To be accepted by the wireless community, these equivalent models need to be validated through hardware testbeds and empirical measurements. Our analysis reveals that the potential gains and applications of RISs in wireless networks depend on the scaling law of the received power as a function of the distance. There exist, however, no experimental results that allow us to infer, as a function of the wavelength of the radio waves and the geometric size of the meta-surfaces, the most appropriate model for an RIS, i.e., electrically large or electrically small.

Constrained System Design. The potential gains and applications of RISs in wireless networks depend on their nearly passive implementation. This imposes stringent constraints on the development of efficient signal processing algorithms and communication protocols. The absence of power amplifiers

and channel estimation units on the RISs implies, for example, that no channel estimation can be performed at the RISs, and new and efficient (low overhead) protocols need to be developed for acquiring the necessary environmental information for optimally programming their operation. The impact of these constrains is exacerbated in high-mobility environments, e.g., in outdoors (see Fig. 5.3) in contrast to low-mobility environments, e.g., in indoors [92].

5.7 Conclusion

RISs are an emerging and little understood technology with a strong potential for application to wireless networks. The main feature of RISs lies in turning the wireless environment into a customizable space by coating objects and devices with reconfigurable meta-surfaces that are capable of shaping the radio waves impinging upon them. In this article, we have discussed differences and similarities between RISs and relaying, and have provided qualitative and quantitative arguments showing that electrically large RISs can outperform relay-aided communications in terms of data rate, while at the same time reducing the implementation complexity. Electrically large RISs that act as anomalous mirrors can be a suitable technology for wireless communications at high radio frequencies, namely at millimeter and sub-millimeter wave bands. We have briefly discussed the potential applications of electrically small RISs to realize single-RF multi-stream transmitters with low complexity, and have elaborated on their potential for application to sub-6 GHz and higher frequencies. The generalization of our analysis to account for the spatial distribution of nodes with the aid of stochastic geometry is a promising research venue.

6 Conclusion

This chapter closes the thesis with general conclusions in *Section 6.1* and summaries of future work in *Section 6.2*.

6.1 Conclusion

In this thesis, new contributions on modeling, evaluating and optimizing the next generation (5G) cellular networks by using stochastic geometry have been reported. In addition, the emerging technology of Reconfigurable Intelligent Surfaces (RISs), as a promising enabler for wireless communications beyond 5G, has been introduced and inspected. More specifically, the contribution of this thesis can be summarized as follows.

- In chapter 2, a new tractable mathematical framework of coverage probability to evaluate the performance of 3D cellular networks with antenna height has been proposed. To gain insights, a closed-form upper bounds of the coverage probability have been developed, which yield explicit expressions of the optimal BSs' height and density to optimize the coverage probability under some conditions. The impact of key system parameters on the obtained expressions of the optimal height and density has been examined, which is shown to be consistent with practical setups in engineering applications. The performance trends and optimization findings have been substantiated with aid of numerical simulations.

- In chapter 3, optimal resource management mechanisms motivated by the emerging principle of network slicing in RANs has been designed. We have proposed a closed-from utility function of the spectral efficiency, which was proved to be jointly concave as a function of the transmit power and bandwidth. The optimization problem has been formulated as a MINLP and solved by using a Generalized Benders Decomposition approach where the global optimality has been retained compared with the brute-force search approach. The numerical results have further verified the accuracy of the adopted algorithm, and the visible performance gain of sum PSE has been achieved by comparing with the Greedy algorithm.

- In chapter 4, we have proposed the first analytical approach that provides one with the probability that a random object coated with reconfigurable metasurfaces acts as a reflector, and have compared it against the conventional setup in which the object is not coated with reconfigurable metasurfaces. This result has been obtained by modeling the environmental objects with a modified random line process with fixed length, and random orientations and locations. Our proposed analytical approach allows us to prove that the probability that an object is a reflector does not depend on the length of the object if it is coated with metasurfaces, while it strongly depends on it if the Snell's law of reflection needs to be applied. The reason of this major difference in system performance lies in the fact that the angles of incidence and reflection need to be the same according to the Snell's law of reflection.

- In chapter 5, the end-to-end signal-to-noise ratio and achievable rate of half-duplex, full-duplex, amplify-and-forward, and decode-and-forward relay-assisted communication, as well as the signal-to-noise ratio and achievable rate formulas of the emerging technology known as reconfigurable intelligent surfaces have been developed. Furthermore, we have discussed differences and similarities between RISs and relaying, and have provided qualitative and quantitative arguments showing that electrically large RISs can outperform relay-aided communications in terms of data rate, while at the same time reducing the implementation complexity.

6.2 Future Work

Many topics of interest in the field of stochastic geometry analysis of wireless communications and cellular networks remain open. Relying on the findings of this thesis, possible directions for research can be the following.

6.2.1 UAVs-Aided Wireless Networks

Beyond terrestrial access points - Current communication networks are designed and optimized based on the availability of terrestrial base stations. This has been the status quo so far but it is not sufficient anymore. At present, we deploy the access points based on some priori information on the network traffic. Once the base stations are installed, they are usually kept there forever. This is because of the cost of renting sites where to deploy them is significant. This strategy has been successful since data traffic usually changes very slowly and usually more access points are needed in densely deployed areas. In the future, this status quo will change for several reasons, since new and emerging applications will require connectivity on an opportunistic and capillary manner rather than conventional communication networks. The most typical application scenario, but it is not the only one, is when disasters of various nature occur in both densely urban or more rural scenarios. It is known that communication networks are usually unreliable if such events occur.

The deployment of aerial access points, often known as drones or unmanned aerial vehicles, offers a suitable solution for providing ad hoc connectivity. Other scenarios are rural areas or events that occur in different places but where a large amount of people aggregate and necessitate a reliable communication infrastructure. In all these scenarios, it is not cost-efficient to deploy terrestrial infrastructure due to the associated cost and the fact that it will become obsolete at the end of the event of interest. One may even envision a future where only aerial access points will be available in order to totally avoid the cost of deploying cellular infrastructure and to provide connectivity when and if it is needed. In order to enable this vision, the drones will have to be energy-neural since it may not be possible to re-charge them or they may not have access to reliable power sources.

The design and optimization of a communication network solely based on drones or relying on both terrestrial and aerial access points is a challenging and open research issue. This is because the service depends on several factors, such as the density, the altitude, the velocity, etc. of the drones, which change from scenario to scenario. In addition, the propagation channels of terrestrial and aerial base stations are complexity different, which affect performance and optimization. In addition, the drones may have local storage capabilities and may, on the other hand, have very strict power requirements and transmission range constraints. Where and how to deploy the drones in urban and rural areas is unknown to date. The new methodologies for network modeling and design will have to account for this new generation of ad hoc access points and their potential application in the IoT market. At present, no clear approach for modeling this scenario is available.

6.2.2 Traffic Load Modeling and Caching

The current methodologies for studying spatial networks suffer from several simplifying assumptions, for the sake of mathematical tractability, which are challenged by researchers working in academia and, especially, industry. The fundamental assumptions is that, not only access points are often distributed totally random, but the mobile users and, more recently, the "trillion of things" of the IoT are randomly

distributed as well. This is not the case in practice, since the traffic demands are spatially correlated, as well as the spatial distributions of access points and mobile users and devices are mutually coupled with each other. This implies that not only spatial attractions and repulsions of the access points need to be accounted for, as mentioned before, but the joint spatial characteristics of users and access points need to be accounted for, in order to design future network, where the new analytical approach needs to be provided.

In current and future networks the number of operators and the role of mobile operators will change dramatically, to an extent that network resources will be shared among them according to the characteristics of the data traffic, which, in turn, will change in the future due to the emergence of IoT. At present, there is no mathematical approach that is capable of handling this emerging scenario, which goes far beyond modeling the spatial characteristics of the access points (mentioned just above as an open problem itself). On top of that, current networks assume that access points and devices have no local storage capabilities. This implies that every single time some data is requested by the users, even though they are closely located, it is accessed in the core networks. The devices used by the mobile users, however, are becoming very powerful and have storage capabilities that were unforseen in the past. This opens the possibility of performing local caching of the data and of exchanging it among local neighbors without putting pressure to the core network. The potential of caching highly depends on the characteristics of the data traffic and of the spatial distribution of nodes, which cannot be assumed to be distributed at random, otherwise even the concept of caching will become meaningless.

To model this scenario where the network deployments are likely to exhibit some degree of interactions among the locations of the base stations, which include spatial inhibition, i.e., repulsion, and spatial aggregation, i.e., clustering, it will mainly relies on the application of analytical tools and theory of stochastic geometry beyond PPP. This is because the PPP-based abstraction model assumes the mutual independence among all points in the point process, where two points could be arbitrarily close to each other, which fails to capture the characteristics that the locations of points are spatially correlated.

6.2.3 Wireless Communications with Deep Learning

All the previous and current generations of wireless networks are based on the analytical models, such as stochastic geometry, random matrix, game theory, etc., or from field measurement campaigns. Analytical modelings behave as a fundamental role in all phases of network design, presenting in quantitative terms which allow us to evaluate the system performance under the impact of different parameters. The initial network planning and deployment for network resource management, network maintenance and control mainly relies on analytical models, where the infrastructure nodes are statistically deployed to cover and manage fixed geographical areas, and the centralized resource allocation schemes are utilized in traditional network optimizations. This traditional approach, on the other hand, has at least the following two drawbacks: i) The available tractable analytical models, formulated in an accepted complexity, may not be accurate enough to specify the realistic problem. In addition, even if such models are available, a trade-off exists between the accuracy and its complexity. ii) The heterogeneous service requirements and random evolving environments with unpredictable on-demand connectivity requests exist where the current static network infrastructure may not to flexible enough to satisfy.

To overcome this, as a particular machine learning technique that implementing learning process elaborating the data via Artificial Neural Networks (ANNs), deep learning and data-oriented approaches have gained attentions in their application to communication system and wireless networks. Moreover,

there are also some recent technological advancements that enable deep learning a reliable technology for future wireless communication networks:

- The exponential increase of wireless devices in current and future wireless networks results in a corresponding growth of data that can be used. The availability of large dataset, therefore, necessitate the efficient algorithms of deep learning.
- The larger and more complex algorithms can be executed much faster due to the modern advancements in computing capacity. In particular, the application of Graphics Processing Unites (GPUs) makes deep learning algorithms execute at speeds many times faster than traditional processor chips.
- The concept of smart radio environment is a fundamental paradigm shift with respect to the design of current wireless networks. In this scenario, the operation of each environment object may be optimized besides the operation of transmitters, receivers and implementations of communication protocols. Accurately modeling this network environment and optimizing it in real time at a lower mathematical complexity is an open issue.

Motivated by these considerations, we argue that the design and optimization of next generation wireless networks, namely 6G, need to be tackled by taking the benefits of data-driven and AI into the wireless network modeling.

Bibliography

[1] B. Clerckx, A. Lozano, S. Sesia, C. van Rensburg, and C. B. Papadi, "3GPP LTE and LTE-Advanced", *EURASIP Journal on Wireless Communications and Networking*,vol. 2009, no. 1, pp. 472124, September 2009.

[2] E. Dahlman, S. Parkvall, and J. Skold, "5G NR: The next generation wireless access technology", *Academic Press*, 2018.

[3] J. G. Andrews et al., "What Will 5G Be?", *IEEE Journal On Selected Areas In Communication*, vol. 32, no. 6, pp. 1065-1082, June 2014.

[4] 3GPP TR 38.913, "Study on Scenarios and Requirements for Next Generation Access Technologies, V14.1.0", 2017.

[5] 3GPP TR 38.802, "Study on New Radio (NR) Access Technology - Physical Layer Aspects - Release 14", 2017.

[6] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-Dense Networks: A Survey", *IEEE Communications Surveys & Tutorials*, vol. 18, no. 4, pp. 2522-2545, 2016.

[7] Jialing Liu, Weimin Xiao, Chih-Lin I, Chenyang Yang, Anthony Soong, "Ultra-Dense Networks (UDNs) for 5G", *IEEE 5G Tech Focus*, vol. 1, no. 1, March 2017. Online: https://5g.ieee.org/techfocus/march-2017/ultra-dense-networks-udns-for-5g

[8] M. Ding, D. López-Pérez, G. Mao, P. Wang, and Z. Lin, "Will the area spectral efficiency monotonically grow as small cells go dense?", *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, San Diego, CA, USA, pp. 1-7, 2015.

[9] J. Song, L. T. Tu, and M. D. Renzo, "On the feasibility of interference alignment in ultra-dense millimeter-wave cellular networks", *Proc. Asilomar Conf. Signals Syst. Comput.*, Pacific Grove, CA, USA, pp. 1176-1180, November 2016.

[10] T. S. Rappaport et al., "Millimeter wave mobile communications for 5G cellular: It will work!", *IEEE Access*, vol. 1, no. 1, pp. 335-349, August 2013.

[11] S. Rangan, T. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges", *Proc. IEEE*, vol. 102, no. 3, pp. 366-385, March 2014.

[12] J. G. Andrews, T. Bai, M. N. Kulkarni, A. Alkhateeb, A. K. Gupta, and R. W. Heath, Jr., "Modeling and analyzing millimeter wave cellular systems", *IEEE Transactions on Communications*, vol. 65, no. 1, pp. 403-430, January 2017.

[13] P. Popovski, K. F. Trillingsgaard, O. Simeone, and G. Durisi, "5G Wireless Network Slicing for eMBB, URLLC, and mMTC: A Communication-Theoretic View", *IEEE Access*, vol. 6, pp. 55765-55779, September 2018.

[14] H. Zhang, N. Liu, X. Chu, K. Long, A.-H. Aghvami, and V. C. M. Leung, "Network slicing based 5G and future mobile networks: Mobility, resource management, and challenges", *IEEE Communications Magazine*, vol. 55, no. 8, pp. 138-145, August 2017.

[15] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems", *arXiv preprint arXiv:1902.1*, 2019.

[16] M. Di Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin et al., "Smart radio environments empowered by reconfigurable AI meta-surfaces: an idea whose time has come", *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 129, 2019.

[17] 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.

[18] M. D. Renzo, S. Wang, and X. Xi, "Inhomogeneous double thinning – modeling and analysis of cellular networks by using inhomogeneous poisson point processes", *IEEE Trans. Wireless Commun.*, 2018.

[19] 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, Mar. 2004.

[20] S. M. Yu and S.-L. Kim, "Downlink capacity and base station density in cellular networks", *IEEE Workshop on Spatial Stochastic Models for Wireless Networks*, pp. 1-7, May 2013. [Online]. Available: http://arxiv.org/pdf/1109.2992.pdf.

[21] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5g multi-tier cellular wireless networks: An interference management perspective", *Wireless Communications, IEEE*, vol. 21, no. 3, pp. 118-127, June 2014.

[22] 3GPP TR 36.814, "Further advancements for E-UTRA physical layer aspects (Release 9),", http://www.3gpp.org/ftp/Specs/archive/36series/36.814/36814-900.zip, Mar. 2010.

[23] T. Bai, R. Vaze, and R. W. Heath Jr., "Analysis of blockage effects on urban cellular networks", *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 5070-5083, Sep. 2014.

[24] T. Bai, R. Vaze, and R. W. Heath Jr., "Coverage and rate analysis for millimeter wave cellular networks", *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 1100-1114, Oct. 2014.

[25] M. Di Renzo, W. Lu, and P. Guan, "The intensity matching approach: A tractable stochastic geometry approximation to system-level analysis of cellular networks,", *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 5963-5983, Sep. 2016.

[26] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP altitude for maximum coverage", *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 569-572, Dec. 2014.

[27] M. Ding and D. L. Perez, "Please lower small cell antenna heights in 5G", *IEEE GLOBECOM*, pp. 1-6, Dec 2016.

[28] I. Atzeni, J. Arnau, and M. Kountouris, "Downlink cellular network analysis with LOS/NLOS propagation and elevated base stations,", *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 142-156, Jan. 2018.

[29] H. Cho, C. Liu, J. Lee, T. Noh, and T. Q. Quek, "Impact of elevated base stations on the ultra-dense networks", *IEEE Commun. Lett.*, vol. 22, no. 6, pp. 1268-1271, Jun. 2018.

[30] J. Yang, M. Ding, G. Mao, Z. Lin, D. Zhang and T. H. Luan, "Optimal Base Station Antenna Downtilt in Downlink Cellular Networks", *IEEE Trans. Wireless Commun.*, Submitted on 21 Feb 2018.

[31] M. Ding, D. López-Pérez, and G. Mao, "Ultra-dense networks: Is there a limit to spatial spectrum reuse?", in *Proc. IEEE ICC, May 2018*, pp. 1–7. [Online]. Available: http://arxiv.org/abs/1704.00399.

[32] M. Filo, C. H. Foh, S. Vahid, and R. Tafazolli, "Stochastic geometry analysis of ultra-dense networks: Impact of antenna height and performance limits,", arXiv preprint arXiv:1712.02235, 2017.

[33] F. Baccelli and B. Blaszczyszyn, *Stochastic Geometry and Wireless Networks, Part I: Theory*, Now Publishers, Sep. 2009.

[34] M. D. Renzo, A. Zappone, T. T. Lam, and M. Debbah, "Systemlevel modeling and optimization of the energy efficiency in cellular networks - a stochastic geometry framework", in *IEEE Transactions on Wireless Communications*, 2018.

[35] N. Alliance, "Description of network slicing concept", in *http://www.ngmn.org/publications/technical.html*, 2015.

[36] B. Han et al., "A Utility-driven Multi-Queue Admission Control Solution for Network Slicing", in *Proceedings of IEEE INFOCOM*, April 2019.

[37] R. Pries et al., "Network as a Service - A Demo on 5G Network Slicing", in *International Teletraffic Congress*, 2016.

[38] X. Foukas, G. Patounas, A. Elmokashfi, and M. K. Marina, "Network Slicing in 5G: Survey and Challenges", in *IEEE Communications Magazine*, vol. 55, pp. 80-87, May 2017.

[39] V. Sciancalepore et al., "Mobile Traffic Forecasting for Maximizing 5G Network Slicing Resource Utilization", in *Proceedings of IEEE INFOCOM*, pp. 4883-4888, May 2017.

[40] J. Salvat, L. Zanzi, A. Garcia-Saavedra, V. Sciancalepore, and X. Costa Perez, "Overbooking network slices through yield-driven end-to-end orchestration", in *ACM CONEXT 2018*, pp. 1-12, December 2018.

[41] K. Samdanis, X. Costa-Perez, and V. Sciancalepore, "From network sharing to multi-tenancy: The 5g network slice broker", in *IEEE Communications Magazine*, vol. 54, no. 7, pp. 32-39, July 2016.

[42] R. Wen, G. Feng, J. Tang, T. Q. S. Quek, G. Wang, W. Tan, and S. Qin, "On robustness of network slicing for next generation mobile networks", in *IEEE Transactions on Communications*, 2018.

[43] H. ElSawy, A. Sultan-Salem, M. S. Alouini, and M. Z. Win, "Modeling and analysis of cellular networks using stochastic geometry: A tutorial", in *IEEE Communications & Surveys Tutorials*, vol. 19, pp. 167-203, 2017.

[44] V. Sciancalepore, M. Di Renzo, and X. Costa-Perez, "STORNS: Stochastic Radio Access Network Slicing", in *IEEE International Conference on Communications (ICC)*, Submitted, 2019. Available: https://arxiv.org/abs/1901.05336.

[45] S. Boyd and L. Vandenberghe, "Convex Optimization", in *Cambridge University Press*, 2004.

[46] C. A. Floudas, "Nonlinear and mixed-integer optimization: fundamentals and applications", *Oxford University Press*, 1995.

[47] A. Geoffrion, "Generalized Benders decomposit", *Journal of Optimization Theory and Applications*, vol. 10, no. 4, 1972.

[48] J.K. Peng, A.M. Manthanwar, D.J. Chmielewski, "On the Tuning of Predictive Controllers: Application of Generalized Benders Decomposition to the ELOC Problem", *Computers & Chemical Engineering*, vol. 82, pp. 105-114, July 2015.

[49] Jin Zhang and Donald J. Chmielewski, "Profit-Based Sensor Network Design Using the Generalized Benders Decomposition", *IEEE 2017 American Control Conference (ACC)*, pp. 3894-3899, 2017.

[50] OPTI Toolbox, "https://inverseproblem.co.nz/OPTI/".

[51] T. Achterberg, "SCIP: Solving constraint integer programs", *Mathematical Programming Computation*, vol. 1, no. 1, pp. 1-41, 2009.

[52] A. Wächter and L. T. Biegler, "On the Implementation of a Primal-Dual Interior Point Filter Line Search Algorithm for Large-Scale Nonlinear Programming", *Mathematical Programming*, vol. 106, no. 1, pp. 25-57, 2006.

[53] https://www.comsoc.org/ctn/what-will-6g-be.

[54] P. Hu, P. Zhang, M. Rostami, and D. Ganesan, "Braidio: An Integrated Active-Passive Radio for Mobile Devices with Asymmetric Energy Budgets", ACM SIGCOMM, Florianopolis, Brazil, Aug. 2016.

[55] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. F. Akyildiz, "Realizing Wireless Communication Through Software-Defined HyperSurface Environments", IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks, Crete, Greece, Jun. 2018.

[56] 5GPPP Vision on Software Networks and 5G SN WG, Jan. 2017.

[57] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. F. Akyildiz, "A New Wireless Communication Paradigm Through Software-Controlled Metasurfaces", IEEE Communications Magazine, Vol. 56, No. 9, pp. 162-169, Sep. 2018.

[58] L. Subrt and P. Pechac, "Controlling Propagation Environments Using Intelligent Walls", European Conference on Antennas and Propagation, Prague, Czech Republic, Mar. 2012.

[59] L. Subrt and P. Pechac, "Intelligent Walls as Autonomous Parts of Smart Indoor Environments", IET Communications, Vol. 6, No. 8, pp. 1004-1010, May 2012.

[60] X. Tan, Z. Sun, J. M. Jornet, and D. Pados, "Increasing Indoor Spectrum Sharing Capacity using Smart Reflect-Array", IEEE International Conference on Communications, Kuala Lumpur, Malaysia, May 2016.

[61] O. Abari, D. Bharadia, A. Duffield, and D. Katabi, "Enabling High-Quality Untethered Virtual Reality", USENIX Symposium on Networked Systems Design and Implementation, Boston, USA, Mar. 2017.

[62] A. Welkie, L. Shangguan, J. Gummeson, W. Hu, and K. Jamieson, "Programmable Radio Environments for Smart Spaces", ACM Workshop on Hot Topics in Networks, Palo Alto, USA, Nov. 30 - Dec. 1, 2017.

[63] R. Chandra and K. Winstein, "Programmable Radio Environments for Smart Spaces - HotNets-XVI Dialogue", ACM Workshop on Hot Topics in Networks, Palo Alto, USA, Nov. 30 - Dec. 1, 2017.

[64] S. Hu, F. Rusek, and O. Edfors, "Beyond Massive MIMO: The Potential of Data Transmission With Large Intelligent Surfaces", IEEE Transactions on Signal Processing, Vol. 66, No. 10, pp. 2746-2758, May 2018.

[65] X. Tan, Z. Sun, D. Koutsonikolas, and J. M. Jornet, "Enabling Indoor Mobile Millimeter-Wave Networks Based on Smart Reflect-Arrays", IEEE Conference on Computer Communications, Honolulu, USA, Apr. 2018.

[66] C. Liaskos, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. F. Akyildiz, "Using any Surface to Realize a New Paradigm for Wireless Communications", Communications of the ACM, Vol. 61 No. 11, pp. 30-33, Nov. 2018.

[67] A. Tsioliaridou, C. Liaskos, and S. Ioannidis, "Towards a Circular Economy via Intelligent Metamaterials", IEEE International Conference on Computer-Aided Modeling Analysis and Design of Communication Links and Networks, Barcelona, Spain, Sep. 2018.

[68] N. Yu, P. Genevet, M. A. Kats, F. Aieta, J.-P. Tetienne, F. Capasso, and Z. Gaburro, "Light Propagation with Phase Discontinuities: Generalized Laws of Reflection and Refraction", Science, Vol. 334, No. 6054, pp. 333-337, Oct. 2011.

[69] C. L. Holloway, E. F. Kuester, J. A. Gordon, J. O'Hara, J. Booth, and D. R. Smith, "An Overview of the Theory and Applications of Metasurfaces: The Two-Dimensional Equivalents of Metamaterials", IEEE Antennas and Propagation Magazine, Vol. 54, No. 2, pp. 10-35, Apr. 2012.

[70] L. Spada, "Metamaterials for Advanced Sensing Platforms", Research Journal on Optical Photonics, Vol. 1, No. 1, Oct. 2017.

[71] T. Nakanishi, T. Otani, Y. Tamayama, and M. Kitano, "Storage of Electromagnetic Waves in a Metamaterial That Mimics Electromagnetically Induced Transparency", Physical Review B, Vol. 87, No. 161110, Apr. 2013.

[72] A. Silva, F. Monticone, G. Castaldi, V. Galdi, A. Alu, and N. Engheta, "Performing Mathematical Operations with Metamaterials", Vol. 343, No. 6167, pp. 160-163, Jan. 2014.

[73] H2020 VISORSURF project, "A Hardware Platform for Software-Driven Functional Metasurfaces", http://www.visorsurf.eu/.

[74] J. Lee and F. Baccelli, "On the Effect of Shadowing Correlation on Wireless Network Performance", IEEE International Conference on Computer Communications, Honolulu, Hawaii, Apr. 2018.

[75] M. Di Renzo, A. Guidotti, G. E. Corazza, "Average Rate of Downlink Heterogeneous Cellular Networks over Generalized Fading Channels – A Stochastic Geometry Approach", IEEE Transactions on Communications, Vol. 61, No. 7, pp. 3050–3071, Jul. 2013.

[76] M. Di Renzo, "Stochastic Geometry Modeling and Analysis of Multi-Tier Millimeter Wave Cellular Networks", IEEE Transactions on Wireless Communications, Vol. 14, No. 9, pp. 5038-5057, Sep. 2015.

[77] W. Lu and M. Di Renzo, "Stochastic Geometry Modeling of Cellular Networks: Analysis, Simulation and Experimental Validation", ACM International Conference on Modeling Analysis and Simulation of Wireless and Mobile Systems, Nov. 2015.

[78] M. Di Renzo, S. Wang, and X. Xi, "Modeling and Analysis of Cellular Networks by Using Inhomogeneous Poisson Point Processes", IEEE Transactions on Wireless Communications, Vol. 17, No. 8, pp. 5162-5182, Aug. 2018.

[79] M. Di Renzo, T. T. Lam, A. Zappone, and M. Debbah, "A Tractable Closed-Form Expression of the Coverage Probability in Poisson Cellular Networks", IEEE Wireless Communications Letters, to appear, 2018.

[80] M. Di Renzo, A. Zappone, T. T. Lam, and M. Debbah, "Spectral-Energy Efficiency Pareto Front in Cellular Networks: A Stochastic Geometry Framework", IEEE Wireless Communications Letters, to appear, 2018.

[81] A. Narayanan, S. T. Veetil, and R. K. Ganti, "Coverage Analysis in Millimeter Wave Cellular Networks with Reflections", IEEE Global Communications Conference, Singapore, Singapore, Dec. 2017.

[82] M. Di Renzo and W. Lu, "System-level Analysis and Optimization of Cellular Networks With Simultaneous Wireless Information and Power Transfer: Stochastic Geometry Modeling", IEEE Transactions on Vehicular Technology, Vol. 66, No. 3, pp. 2251-2275, TBA.

[83] T. S. Rappaport *et al.*, "Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond", *IEEE Access*, vol. 7, pp. 78729-78757, 2019.

[84] M. Dohler and Y. Li, *Cooperative communications*, Wiley, Feb. 2010.

[85] C. Liaskos *et al.*, "A new wireless commun. paradigm through software-controlled metasurfaces", *IEEE Commun. Mag.*, vol. 56, Sep. 2018.

[86] E. Basar *et al.*, "Wireless communications through reconfigurable intelligent surfaces", *IEEE Access*, to appear. Available: arXiv:1906.09490.

[87] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design", *IEEE Trans. Wireless Commun.*, to appear. Available: arXiv:1905.00152.

[88] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network", *IEEE Commun. Mag.*, to appear. Available: arXiv:1809.01423.

[89] Z. Li *et al.*, "Towards programming the radio environment with large arrays of inexpensive antennas", *USENIX NSDI*, Feb. 2019.

[90] V. Arun and H. Balakrishnan, "RFocus: Practical beamforming for small devices", *ArXiv*, submitted. Available: arXiv:1905.05130.

[91] F. Liu *et al.*, "Intelligent metasurfaces with continuously tunable local surface impedance for multiple reconfigurable functions", *Physical Review Applied*, vol. 11, no. 4, Apr. 2019.

[92] N. Kaina *et al.*, "Shaping complex microwave fields in reverberating media with binary tunable metasurfaces", *Scien. Rep.*, vol. 4, Oct. 2014.

[93] S. A. Tretyakov, V. Asadchy, and A. Diaz-Rubio, "Metasurfaces for general control of reflection and transmission", *World Scientific Handbook of Metamaterials and Plasmonics*, pp. 249-293, 2018.

[94] W. Tang *et al.*, "Wireless communications with programmable metasurface: New paradigms, opportunities, and challenges on transceiver design", *IEEE Wireless Commun.*, Available: arXiv:1907.01956.

[95] R. Mehrotra *at al.*, "3D channel modeling and characterization for Hypersurface empowered indoor environment at 60 GHz millimeter-wave band", *ArXiv*, submitted. Available: arXiv:1907.00037.

[96] J. D. Griffin and G. D. Durgin, "Complete link budgets for backscatter-radio and RFID systems", *IEEE Ant. Propag. Mag.*, vol. 51, Apr. 2009.

[97] A. Goldsmith, *Wireless communications*, Cambridge New York: Cambridge University Press, 2005.

[98] T. Riihonen, S. Werner, and R. Wichman, "Hybrid full-duplex/half-duplex relaying with transmit power adaptation", *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3074-3085, Sep. 2011.

[99] G. Liu, F. R. Yu, H. Ji, V. C. M. Leung, and X. Li, "In-band full-duplex relaying: A survey, research issues and challenges", *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 500-524, 2nd Quart. 2015.

[100] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems", *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296-4307, Dec. 2012.

[101] M. Di Renzo and J. Song, "Reflection probability in wireless networks with metasurface-coated environmental objects: an approach based on random spatial processes", *EURASIP Journal on Wireless Communications and Networking*, April, 2019.

[102] K. Ntontin, M. Di Renzo, J. Song, F. Lazarakis, J. de Rosny, D.-T. Phan-Huy, O. Simeone, R. Zhang, M. Debbah, G. Lerosey, M. Fink, S. Tretyakov and S. Shamai, "Reconfigurable Intelligent Surfaces vs. Relaying: Differences, Similarities, and Performance Comparison", *IEEE Communications Magazines*, Submitted, August, 2019.

[103] J. Song, M. Di Renzo and D. Lopez-Perez, "System-Level Modeling and Optimization of the Coverage in 3D Poisson Cellular Networks - A Stochastic Geometry Approach", *IEEE Transactions on Wireless Communications*, in Submission, 2019.

[104] J. Song and M. Di Renzo, "Optimization of Resource Allocation in Multi-Tenants Radio Access Network Slicing: A Generalized Benders Decomposition Algorithm", *IEEE Wireless Communications Letters*, in Submission, 2019.

[105] J. Song and M. Di Renzo, "On the Stochastic Geometry Modeling and Optimization of Multi-Tenant Network Slicing", *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Invited, Sep. 2018.

French Summary

Vue d'ensemble - Réseaux cellulaires de prochaine génération

Le projet de partenariat de troisième génération (3GPP) a proposé la série de spécifications, Longue Term Evolution (LTE), qui a été développé pour répondre aux exigences de l'actuel système cellulaire 4G réseaux en 2009. Certaines des exigences de la prochaine génération (5G) peuvent également être satisfaites dans des scénarios de déploiement spécifiques à l'aide d'une telle évolution. Cependant, ils ne sont pas capable de répondre à toutes les demandes de trafic 5G et aux cas d'utilisation caractérisés par des débits de données, latence ultra-faible et prise en charge d'un grand nombre de connexions. Être plus précis, le réseau LTE fonctionne dans une bande maximale de 20 MHz par porteuse, ce qui limite les débit de données réalisable est fourni. En même temps, sa structure de cadre rigide rend difficile la réduire la latence aller-retour en dessous de 1 ms, et la désignation derrière LTE ne tient pas compte pour l'efficacité énergétique car les signaux du pilote sont toujours activés. De plus, la demande et l'internet des objets à bande étroite (NB-IoT), qui est appliqué pour servir une large gamme d'appareils massifs, est difficile à réaliser avec le fonctionnement actuel du LTE. Afin de surmonter les limites des réseaux cellulaires 4G actuellement déployés, le 3GPP définit une nouvelle technologie d'accès radio (RAT), 3GPP New Radio (NR), afin de proposer le nouveau technologies et approches par rapport aux exigences de la 5G et aux cas d'utilisation. En particulier, le déploiement de réseaux ultra-denses (UDN) à base de petites cellules offrira une couverture étendue, débit amélioré et allègement de la charge de trafic dans les macro-cellules. De plus, NR exploite de nouveaux spectre, bande de fréquence à onde millimétrique (mmWave), pour augmenter le débit connu de l'utilisateur, l'efficacité spectrale et de prendre en charge les nouvelles technologies telles que l'énorme entrée multiple et sortie multiple (MIMO). Du point de vue du réseau central, les nouveaux schémas ont introduit pour offrir le découpage en réseau et la virtualisation afin de répondre aux besoins multiples Level Agreements (SLAs) entre les différents locataires d'utilisateurs, notamment en ce qui concerne conception du mécanisme de gestion des ressources radio dans le RAN. Les technologies susmentionnées de la 5G seront discutées comme suit, associées à sa contexte général sur les motivations et les défis. De plus, la brève

introduction sur les questions 5G et un mécanisme prometteur au-delà de la 5G (6G) seront également expliqués. Autre candidats potentiels tels que MIMO massif sont hors de portée de cette thèse.

Réseaux ultra-denses (UDN)

UDN est l'un des moyens les plus efficaces de transmettre le débit binaire sans cesse croissant et d'améliorer l'expérience utilisateur. Il s'est avéré être une solution de premier plan au cours des dernières années, à savoir l'ajout de stations de base et de points d'accès Les points d'accès) permettant ainsi une réutilisation plus spatiale du spectre et libérant les demandes de trafic élevées. Cette tendance à la densification des infrastructures devrait se poursuivre en 5G et au-delà. Les UDN peuvent être définis comme les réseaux dont la densité de BS déployées est beaucoup plus élevée que celle des terminaux mobiles actifs (MT). Une autre définition du réseau UDN peut être énoncée quantitativement lorsque la densité d'un réseau candidat est mesurée à $\geqslant 10^3 \text{cells/km}^2$.

Comparés aux réseaux cellulaires classiques, les réseaux UDN possèdent les caractéristiques fondamentales suivantes. Premièrement, mis à part les Macro BS actuellement déployées, les éléments de réseau ou les nœuds d'accès dans le scénario UDN sont exploités principalement avec une faible puissance et une faible couverture, c'est-à-dire petites cellules (Picocells et Femtocells). Ainsi, la distance entre sites dans ce scénario serait de l'ordre de quelques mètres et la topologie du réseau est plus probablement hétérogène. Dans de tels environnements de couverture, les MT seraient très proches des BS. Par conséquent, la deuxième caractéristique du réseau UDN est la probabilité de transmission élevée en mode ligne de vue (LOS). En outre, en raison de la densité élevée de petites cellules, de nombreuses stations de base seraient inactives, en particulier sous un trafic léger. charge. Cela motive la proposition du concept de mode veille. Une autre caractéristique importante de l'UDN est l'interférence drastique entre les cellules adjacentes, compte tenu de la proximité étroite de petites cellules les unes par rapport aux autres.

Sur la base des caractéristiques ci-dessus, plusieurs problèmes de recherche associés apparaissent. Puisque le réseau devient hétérogène, les modèles spatiaux habituels, modèle en réseau ou modèle hexagonal, utilisés pour la comptabilisation de la topologie spatiale du réseau, ne parviennent pas à saisir leurs caractéristiques. Même si les modèles spatiaux conventionnels peuvent être appliqués dans certaines zones limitées dont la topologie de réseau est régulière, leur manque de capacité d'analyse compromettra leur efficacité. Afin de remplir les conditions de transmission proches de la LOS dans des circonstances UDN, la nécessité de différents modèles de propagation devient une nécessité. En particulier, le modèle de répartition de la perte d'énergie

avec une pente unique, largement adopté, ne peut pas caractériser la transmission à courte portée des réseaux UDN. En outre, le modèle actuel de réseau entièrement chargé suppose que les stations de base transmettent toujours des signaux vers les terminaux mobiles, surestimant ainsi les interférences. En ce sens, un modèle de charge approprié prenant en compte la densité de BS et le déploiement dense de BS est nécessaire pour examiner l'impact de l'ultra-densification.

Onde millimétrique

Outre l'ultra-densification, exploiter le spectre d'ondes millimétriques (mmWave) jamais adressé au-delà de 6 GHz, compris entre 30 et 300 GHz, constitue un autre moyen essentiel de répondre aux besoins sans précédent de la 5G. C'est parce que l'application ces bandes de fréquences, en tant que porteuses, permettent une plus grande allocation de bande passante, qui pourrait être traduite directement en débit de données élevé En outre, en augmentant la largeur de bande du canal de la radio mobile, il sera possible de prendre en charge un accès beaucoup plus performant basé sur Internet et des applications avec un temps de latence minimal, car la capacité de transmission de données est considérablement accrue et le temps de latence du trafic numérique considérablement réduit. De plus, les fréquences mmWave comprises entre 30 et 300 GHz, ce qui correspond à 1 - 10 mm, peuvent prendre en charge de nouvelles techniques de traitement spatial, telles que la technologie MIMO massive, en raison de la longueur d'onde beaucoup plus petite. Dans le même temps, les systèmes mmWave fonctionnent avec une antenne hautement directionnelle, ce qui permet d'isoler les utilisateurs et de réduire ainsi les interférences. Jusqu'à présent, en revanche, les bandes de fréquence mmWave n'étaient pas jugées utiles pour les applications commerciales dynamiques dans les communications sans fil telles que système cellulaire, même s'il a été étudié pour des systèmes de communication terrestre et par satellite à grande portée. Ceci est principalement dû aux nouveaux défis introduits dans la réalisation des propriétés de propagation et de la dégradation de canal dans les bandes de fréquences élevées.

En ce qui concerne la propagation mmWave pour les réseaux cellulaires de prochaine génération, le premier défi provient de plusieurs composants importants de la modélisation de canal, en particulier l'atténuation à grande échelle dans le cheminement omnidirectionnel en espace libre, qui croît avec le carré de la fréquence porteuse selon la transmission Friis. loi. Une préoccupation plus importante en ce qui concerne le canal mmWave est que les signaux transmis sont extrêmement susceptibles d'être ombrés. Dans les villes urbaines denses, par exemple, les matériaux utilisés à l'extérieur, tels que les briques et le corps humain, peuvent atténuer considérablement les signaux en raison de la décoloration par ombrage. Statistiquement, ce phénomène est modélisé comme une variable aléatoire log-normale distribuée dans les canaux

mmWave, ce qui entraîne une complexité accrue dans la formulation analytique et le calcul. De plus, l'incorporation de la communication mmWave dans les UDN nécessiterait une très forte intermittence. La connectivité et la communication doivent être adaptables en raison des fluctuations rapides du canal causées par les hautes fréquences.

Outre les défis mentionnés ci-dessus concernant les canaux sans fil, en raison de la gamme limitée de signaux mmWave, la majorité des réseaux cellulaires candidats du système mmWave sont déployés en petites cellules et en ultra-densification, comme indiqué dans les sections précédentes. En conséquence, la couverture en milieu urbain dense rencontrera le NLOS beaucoup plus fréquemment sous l'impact de blocages à grande échelle. De tels effets de blocage, résultant de la réflexion, de la diffusion, de la diffraction, de la réfraction et de l'absorption, altèrent profondément la fiabilité de la liaison des communications sans fil. Cependant, ces caractéristiques sont soit négligées, soit trop simplifiées dans la modélisation des réseaux cellulaires actuels. Bien que la théorie de la forme aléatoire puisse être utilisée pour identifier des bâtiments situés au hasard dans la zone bidimensionnelle, ce qui nous permet d'évaluer les performances du système dans une approche analytique flexible, le modèle probabiliste fourni introduit une nouvelle complexité dans l'analyse de l'influence de la hauteur du bâtiment. De plus, à l'exception de la recherche sur le trajet de propagation direct des effets de blocage, la modélisation de certaines autres propriétés telles que les réflexions est plus difficile et presque aucun document n'a abordé cette question.

Tranchage du réseau

Les besoins en services avancés des industries verticales nécessitent une conception nouvelle des réseaux de téléphonie mobile 5G, afin d'augmenter les revenus des fournisseurs de télécommunications et des opérateurs de réseaux. Pour relever ce défi, tel que la programmabilité et la virtualisation du réseau, est nécessaire. La programmabilité de réseau apporte les avantages de l'automatisation et de la réactivité des modules logiciels, ce qui permet de (re) configurer les réseaux mobiles de manière dynamique en cours de fonctionnement. La virtualisation de réseau surmonte les limites des infrastructures de réseau monolithiques, en résumant le concept de "fonction de réseau" et en offrant une flexibilité dans la composition, le placement et la gestion de ces fonctions. Dans ce contexte, la nouvelle définition du découpage en réseau dans les réseaux mobiles, principalement basée sur le réseau défini par logiciel (SDN) et la virtualisation de la fonction réseau (NFV), englobe ces nouvelles exigences et constitue un catalyseur des avantages économiques potentiels. De nouvelles industries verticales, telles que l'automobile, les réseaux intelligents, la santé en ligne, la sécurité publique, les usines, etc., font leur entrée sur le marché

des télécommunications et perturbent les modèles commerciaux traditionnels des opérateurs de télécommunications. Ils obligent les fournisseurs d'infrastructure à ouvrir leurs réseaux aux locataires, solution qui incite à monétiser la disponibilité de tranches de réseau isolées et sécurisées (virtualisées).

Ce nouveau concept perturbateur a suscité un intérêt pour la recherche dans les milieux universitaires et industriels. Cependant, sa réalisation nécessite la résolution d'un certain nombre de problèmes techniques qui, pour le moment, ne sont pas complètement traités et résolus, que ce soit dans le monde universitaire ou dans l'industrie. À l'avenir, les fournisseurs de télécommunications envisagent une demande croissante de tranches de réseau de bout en bout, impliquant des SLA hétérogènes comprenant différents indicateurs de performance clés, tels que le débit, la latence et la fiabilité. La possibilité d'avoir différents accords de niveau de service par tranche nécessite des protocoles de contrôle d'admission et d'allocation de ressources automatisés appropriés pour la conception de systèmes de gestion de travail efficaces. En particulier, la gestion des ressources du réseau d'accès radio (RAN) est l'un des aspects les plus difficiles à traiter. De nombreux problèmes doivent être résolus pour découper efficacement le RAN. Les ressources disponibles de l'interface radio RAN peuvent être divisées en plusieurs niveaux: dans les domaines de la fréquence, de l'heure et de la puissance. Découper le réseau à ce niveau de granularité nécessite de prendre en compte la topologie du réseau cellulaire, les autres interférences de cellules et les conditions de canal radio rencontrées par les utilisateurs de chaque client hébergé. Cependant, cette question difficile n'est jamais abordée et résolue.

Méta-surfaces - Réseaux sans fil au-delà de 5G (6G)

Les deux dernières décennies ont été témoins de la révolution extrêmement rapide des réseaux de communications sans fil et cellulaires mobiles, du 1G au 4G. De nombreux défis et problèmes pourraient être surmontés une fois que les technologies 5G seront exploitées d'ici 2020. Néanmoins, certains problèmes en suspens associés aux réseaux 5G n'ont pas encore été identifiés et nécessitent de nouveaux efforts de recherche. Tout d'abord, avec l'avènement de la 5G, principal moyen d'augmenter la capacité du réseau, on assistera au processus de densification du site. Cependant, cette approche commune courante posera des problèmes économiques susceptibles de ralentir le déploiement de la 5G dans l'espace et dans le temps. En attendant, le spectre sous-utilisé n'est pas illimité, et l'exploration de nouvelles bandes de fréquences créera toujours plus d'obstacles dans la conception des antennes, le traitement du signal numérique et la consommation d'énergie.

En résumé, l'énorme augmentation de la demande de trafic pour la révolution cellulaire consiste à utiliser plus d'énergie et à émettre plus d'ondes radio du point de

vue des opérateurs de réseau sans fil. Normalement, cela se fait en exploitant de nouvelles bandes de fréquences pour la transmission et en densifiant le déploiement des réseaux, ce qui coïncide avec les efforts actuels pour développer les technologies sans fil de nouvelle génération où il existe toujours plus d'énergie et d'efficacité spectrale que le précédent. la consommation d'énergie et l'émission d'ondes radio augmentent toujours par rapport à la précédente. En conséquence, les contraintes découlent du concept selon lequel plus de données nécessitent plus de puissance et plus d'émissions d'ondes radio pourraient être émises par des solutions alternatives pour surmonter le goulet d'étranglement des réseaux sans fil au-delà de la 5G. L'approche potentielle est accompagnée de deux questions fondamentales: 1) Et si l'environnement sans fil pouvait être contrôlé pour utiliser les dispositifs sous contrainte d'énergie afin de détecter et de mesurer les données sans transmettre de nouvelles ondes radio? 2) Et si un contrôleur logiciel à distance doté de capacités prédictives peut-il être utilisé pour personnaliser la propagation des ondes radio afin d'améliorer la couverture sans augmenter la consommation électrique?

Pour répondre aux deux questions ci-dessus et mieux répondre aux défis et aux attentes des futurs réseaux sans fil en 6G qui permettent à plus que des personnes, des appareils mobiles et des objets de communiquer, la création d'un environnement de radio intelligente constituerait un catalyseur potentiel. Plus précisément, les futurs réseaux sans fil ont besoin d'un environnement qui puisse lui-même être transformé en un espace reconfigurable pour le transfert et le traitement de l'information. En conséquence, le concept de méta-surfaces, capable de modifier la propagation des ondes radioélectriques de manière totalement personnalisable, est proposé comme aspect fondamental des environnements radioélectriques intelligents pour équiper les dispositifs à faible énergie et effectuer les opérations à faible consommation. des algorithmes complexes. Certains problèmes mis en évidence dans les réseaux 5G pourraient être résolus à l'aide des méta-surfaces reconfigurables. Par exemple, les interférences élevées générées par d'autres cellules dans les réseaux UDN peuvent être contrôlées et atténuées de manière adaptative. De plus, les signaux mmWave à haute sensibilité au blocage transmis par NLOS peuvent être réfléchis par un objet revêtu de méta-surface reconfigurable, la force du signal reçu par les utilisateurs mobiles étant améliorée en conséquence.

Les applications clés mentionnées ci-dessus reposent sur le fait que de tels matériaux sont fixés aux objets environnementaux comme des réflecteurs qui sont répartis dans l'espace selon des schémas spatiaux très complexes. Cependant, la modélisation, l'évaluation et l'optimisation des performances des environnements radio intelligents nécessiteraient la connaissance: 1) de la répartition des emplacements des méta-surfaces dans les réseaux à grande échelle, 2) des manipulations d'onde appliquées par les méta-surfaces en fonction 3) les localisations spatiales et les manipulations d'ondes appliquées par d'autres méta-surfaces distribuées de manière

aléatoire. L'absence de modèles exploitables dans la théorie de la communication et les réseaux sans fil rend l'analyse analytique encore plus ardue.

Résumé de la thèse et contributions majeures

Dans la présente thèse, motivés par les défis de recherche énoncés ci-dessus pour la 5G et au-delà de la 5G, nous étudions l'évaluation de la performance au niveau du système et l'optimisation des futurs réseaux cellulaires candidats à l'aide d'outils de géométrie stochastique, et mettons en évidence le gain de performance en termes de taux réalisable dans les communications sans fil avec méta-surface.

Les principales contributions de cette thèse sont les suivantes: i) cette thèse étudie l'optimisation au niveau du système en fonction de la hauteur et de la densité d'antenne des stations de base dans les réseaux cellulaires 3D, sur la base d'un cadre analytique modifiable;

ii) cette thèse développe un algorithme efficace basé sur l'approche de décomposition généralisée de bender (GBD) pour résoudre les problèmes de programmation mixte non entiers (MINLP) pour une allocation optimale des ressources RAN inspirée du concept de découpage en réseau;

iii) cette thèse propose une nouvelle approche modélisable de la probabilité de réflexion dans les réseaux sans fil en présence d'objets environnementaux revêtus d'une métasurface;

iv) cette thèse compare les différences, les similitudes et les performances de débit de bout en bout des communications assistées par relais multi-antennes et des surfaces intelligentes reconfigurables.

Travail futur

De nombreux sujets d'intérêt dans le domaine de l'analyse de la géométrie stochastique des communications sans fil et des réseaux cellulaires restent ouverts. En s'appuyant sur les résultats de cette thèse, les directions possibles pour la recherche peuvent être les suivantes.

Réseaux sans fil assistés par drones

Au-delà des points d'accès terrestres - Les réseaux de communication actuels sont conçus et optimisés en fonction de la disponibilité des stations de base terrestres. Jusqu'à présent, c'est le statu quo, mais ce n'est plus suffisant. À l'heure actuelle, nous déployons les points d'accès sur la base de certaines informations a priori sur le trafic réseau. Une fois les stations de base installées, elles y sont généralement conservées

pour toujours. Ceci est dû au coût de location des sites où les déployer est important. Cette stratégie a réussi car le trafic de données change généralement très lentement et généralement plus de points d'accès sont nécessaires dans les zones densément déployées. À l'avenir, ce statu quo changera pour plusieurs raisons, car les applications nouvelles et émergentes nécessiteront une connectivité de manière opportuniste et capillaire plutôt que des réseaux de communication conventionnels. Le scénario d'application le plus typique, mais ce n'est pas le seul, est lorsque des catastrophes de nature diverse se produisent dans des scénarios à la fois densément urbains ou plus ruraux. Il est connu que les réseaux de communication ne sont généralement pas fiables si de tels événements se produisent.

Le déploiement de points d'accès aériens, souvent appelés drones ou véhicules aériens sans pilote, offre une solution appropriée pour fournir une connectivité ad hoc. D'autres scénarios sont des zones rurales ou des événements qui se produisent à différents endroits mais où une grande quantité de personnes se regroupent et nécessitent une infrastructure de communication fiable. Dans tous ces scénarios, il n'est pas rentable de déployer une infrastructure terrestre en raison du coût associé et du fait qu'elle deviendra obsolète à la fin de l'événement d'intérêt. On peut même envisager un avenir où seuls des points d'accès aériens seront disponibles afin d'éviter totalement le coût du déploiement de l'infrastructure cellulaire et d'assurer la connectivité quand et si cela est nécessaire. Afin de permettre cette vision, les drones devront être à énergie neuronale car il ne sera peut-être pas possible de les recharger ou ils n'auront pas accès à des sources d'alimentation fiables.

La conception et l'optimisation d'un réseau de communication basé uniquement sur des drones ou reposant à la fois sur des points d'accès terrestres et aériens est un enjeu de recherche difficile et ouvert. C'est parce que le service dépend de plusieurs facteurs, tels que la densité, l'altitude, la vitesse, etc. des drones, qui changent de scénario en scénario. De plus, les canaux de propagation de la base terrestre et aérienne les stations sont de complexité différente, ce qui affecte les performances et l'optimisation. De plus, les drones peuvent avoir des capacités de stockage local et peuvent, d'autre part, avoir des exigences de puissance et des contraintes de portée de transmission très strictes. Le lieu et la manière de déployer les drones dans les zones urbaines et rurales sont inconnus à ce jour. Les nouvelles méthodologies de modélisation et de conception des réseaux devront tenir compte de cette nouvelle génération de points d'accès ad hoc et de leur application potentielle sur le marché de l'IoT. À l'heure actuelle, aucune approche claire pour la modélisation de ce scénario n'est disponible.

Modélisation et mise en cache de la charge de trafic

Les méthodologies actuelles pour étudier les réseaux spatiaux souffrent de plusieurs hypothèses simplificatrices, pour des raisons de tractabilité mathématique, qui sont contestées par les chercheurs travaillant dans le milieu universitaire et, en particulier, l'industrie. Les hypothèses fondamentales sont que non seulement les points d'accès sont souvent distribués de manière totalement aléatoire, mais que les utilisateurs mobiles et, plus récemment, le "billion de choses" de l'IoT sont également distribués de manière aléatoire. Ce n'est pas le cas dans la pratique, car les demandes de trafic sont spatialement corrélées, ainsi que les distributions spatiales des points d'accès et des utilisateurs et appareils mobiles sont mutuellement couplées les unes aux autres. Cela implique que non seulement les attractions spatiales et les répulsions des points d'accès doivent être prises en compte, comme mentionné précédemment, mais les caractéristiques spatiales communes des utilisateurs et des points d'accès doivent être prises en compte, afin de concevoir le futur réseau, où le nouveau système d'analyse doit être fournie.

Dans les réseaux actuels et futurs, le nombre d'opérateurs et le rôle des opérateurs mobiles changeront considérablement, dans la mesure où les ressources du réseau seront partagées entre eux en fonction des caractéristiques du trafic de données, qui, à son tour, changera à l'avenir en raison à l'émergence de l'IoT. À l'heure actuelle, il n'existe aucune approche mathématique capable de gérer ce scénario émergent, qui va bien au-delà de la modélisation des caractéristiques spatiales des points d'accès (mentionné ci-dessus comme un problème ouvert lui-même). En plus de cela, les réseaux actuels supposent que les points d'accès et les périphériques n'ont pas de capacités de stockage local. Cela implique que chaque fois que certaines données sont demandées par les utilisateurs, même si elles sont proches, elles sont accessibles dans les réseaux centraux. Les appareils utilisés par les utilisateurs mobiles deviennent cependant très puissants et disposent de capacités de stockage imprévues dans le passé. Cela ouvre la possibilité d'effectuer une mise en cache locale des données et de les échanger entre voisins locaux sans exercer de pression sur le cœur de réseau. Le potentiel de la mise en cache dépend fortement des caractéristiques du trafic de données et de la distribution spatiale des nœuds, qui ne peuvent pas être supposés distribués de manière aléatoire, sinon même le concept de mise en cache perdra tout son sens.

Pour modéliser ce scénario où les déploiements de réseau sont susceptibles de présenter un certain degré d'interactions entre les emplacements des stations de base, qui incluent l'inhibition spatiale, c'est-à-dire la répulsion et l'agrégation spatiale, c'est-à-dire le regroupement, il s'appuiera principalement sur l'application de l'analyse outils et théorie de la géométrie stochastique au-delà du PPP. En effet, le modèle d'abstraction basé sur PPP suppose l'indépendance mutuelle entre tous les points du processus de points, où deux points pourraient être arbitrairement proches l'un de l'autre, ce qui ne permet pas de saisir les caractéristiques selon lesquelles les emplace-

ments des points sont spatialement corrélés.

Communications sans fil avec apprentissage en profondeur

Toutes les générations précédentes et actuelles de réseaux sans fil sont basées sur des modèles analytiques, tels que la géométrie stochastique, la matrice aléatoire, la théorie des jeux, etc., ou à partir de campagnes de mesure sur le terrain. Les modélisations analytiques se comportent comme un rôle fondamental dans toutes les phases de la conception du réseau, se présentant en termes quantitatifs qui nous permettent d'évaluer les performances du système sous l'impact de différents paramètres. La planification et le déploiement initiaux du réseau pour la gestion des ressources du réseau, la maintenance et le contrôle du réseau reposent principalement sur des modèles analytiques, où les nœuds d'infrastructure sont déployés statistiquement pour couvrir et gérer des zones géographiques fixes, et les schémas centralisés d'allocation des ressources sont utilisés dans les optimisations de réseau traditionnelles. Cette approche traditionnelle, d'autre part, présente au moins les deux inconvénients suivants: i) Les modèles analytiques traitables disponibles, formulés dans une complexité acceptée, peuvent ne pas être suffisamment précis pour préciser le problème réaliste. De plus, même si de tels modèles sont disponibles, un compromis existe entre la précision et sa complexité. ii) Les exigences de service hétérogènes et les environnements à évolution aléatoire avec des demandes de connectivité à la demande imprévisibles existent là où l'infrastructure de réseau statique actuelle peut ne pas être suffisamment flexible pour satisfaire.

Pour surmonter cela, en tant que technique d'apprentissage automatique particulière mettant en œuvre un processus d'apprentissage élaborant les données via des réseaux de neurones artificiels (ANN), l'apprentissage en profondeur et les approches orientées données ont attiré l'attention dans leur application au système de communication et aux réseaux sans fil. De plus, il y a aussi quelques avancées technologiques récentes qui permettent l'apprentissage en profondeur d'une technologie fiable pour les futurs réseaux de communication sans fil:

• L'augmentation exponentielle des appareils sans fil dans les réseaux sans fil actuels et futurs se traduit par une croissance correspondante des données qui peuvent être utilisées. La disponibilité d'un grand ensemble de données nécessite donc les algorithmes efficaces d'apprentissage en profondeur.

• Les algorithmes plus grands et plus complexes peuvent être exécutés beaucoup plus rapidement en raison des progrès modernes de la capacité de calcul. En particulier, l'application de Graphics Processing Unites (GPU) permet aux algorithmes d'apprentissage en profondeur de s'exécuter à des vitesses plusieurs fois plus rapides que les puces de processeur traditionnelles.

• Le concept d'environnement radio intelligent est un changement de paradigme fondamental en ce qui concerne la conception des réseaux sans fil actuels. Dans ce scénario, le fonctionnement de chaque objet d'environnement peut être optimisé en plus du fonctionnement des émetteurs, des récepteurs et des implémentations de protocoles de communication. La modélisation précise de cet environnement réseau et son optimisation en temps réel à une complexité mathématique plus faible est un problème ouvert.

Motivés par ces considérations, nous soutenons que la conception et l'optimisation des réseaux sans fil de prochaine génération, à savoir la 6G, doivent être abordées en intégrant les avantages de l'IA pilotée par les données et de l'IA à la modélisation des réseaux sans fil.

université PARIS-SACLAY
ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre : Analyse et Optimisation des Réseaux Cellulaires par la Géométrie Stochastique

Mots clés : Réseaux cellulaires, géométrie stochastique, processus de Poisson, optimisation, surfaces intelligentes reconfigurables.

Résumé : Cette thèse porte principalement sur la modélisation, l'évaluation des performances et l'optimisation au niveau système des réseaux cellulaires de nouvelle génération à l'aide de la géométrie stochastique. En plus, la technologie émergente des surfaces intelligentes reconfigurables (RISs) est étudiée pour l'application aux futurs réseaux sans fil. En particulier, reposant sur un modèle d'abstraction basé sur la loi de Poisson pour la distribution spatiale des nœuds et des points d'accès, cette thèse développe un ensemble de nouveaux cadres analytiques pour le calcul d'importantes métriques de performance, telles que la probabilité de couverture et l'efficacité spectrale potentielle, qui peuvent être utilisés pour l'analyse et l'optimisation au niveau système. Plus spécifiquement, une nouvelle méthodologie d'analyse pour l'analyse de réseaux cellulaires tridimensionnels est introduite et utilisée pour l'optimisation du système. Un nouveau problème d'allocation de ressources est formulé et résolu en combinant pour la première fois géométrie stochastique et programmation non linéaire mixte en nombres entiers. L'impact du déploiement de surfaces réfléchissantes intelligentes sur un réseau sans fil est quantifié à l'aide de processus ponctuels, et les avantages potentiels des RISs contre le relais sont étudiés à l'aide de simulations numériques.

Title : A Stochastic Geometry Approach to the Analysis and Optimization of Cellular Networks

Keywords : Cellular Networks, Stochastic Geometry, Poisson Point Process, Optimization, Reconfigurable Intelligent Surfaces.

Abstract : The main focus of this thesis is on modeling, performance evaluation and system-level optimization of next-generation cellular networks by using stochastic geometry. In addition, the emerging technology of Reconfigurable Intelligent Surfaces (RISs) is investigated for application to future wireless networks. In particular, relying on a Poisson-based abstraction model for the spatial distribution of nodes and access points, this thesis develops a set of new analytical frameworks for the computation of important performance metrics, such as the coverage probability and potential spectral efficiency, which can be used for system-level analysis and optimization. More specifically, a new analytical methodology for the analysis of three-dimensional cellular networks is introduced and employed for system optimization. A novel resource allocation problem is formulated and solved by jointly combining for the first time stochastic geometry and mixed-integer non-linear programming. The impact of deploying intelligent reflecting surfaces throughout a wireless network is quantified with the aid of line point processes, and the potential benefits of RISs against relaying are investigated with the aid of numerical simulations.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France