The Dissertation Committee for Xingqin Lin
certifies that this is the approved version of the following dissertation:

Integrated Cellular and Device-to-Device Networks

Committee:

Jeffrey G. Andrews, Supervisor

François Baccelli

Robert W. Heath, Jr.

Mihai Sirbu

Haris Vikalo

Integrated Cellular and Device-to-Device Networks

by

Xingqin Lin, B.E.; M.Philosophy

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

December 2014

Dedicated to my family.

Acknowledgments

The writing of this dissertation has been an amazing journey in which I owe my deep gratitude to many people for their help with this dissertation.

I would first like to thank my dissertation advisor Prof. Jeffrey G. Andrews for his guidance and support of my research. Jeff unhesitatingly admitted me into his research group in Fall 2011, and since then his standards for high-quality research have shaped my research goals and attitude. Jeff believes in my research abilities and gives me great freedom in choosing interesting research problems. I also learned very much from Jeff's great personality, enthusiasm, presentation and writing skills. He is a model for me. I sincerely appreciate his trust, caring and dedication during the journey.

I thank Prof. François Baccelli, Prof. Robert W. Heath Jr., Prof. Mihai Sirbu and Prof. Haris Vikalo for serving on my dissertation committee and providing me with invaluable feedback. I greatly benefited from Prof. François Baccelli's course on Stochastic Geometry, which has led to the multicast D2D work in this dissertation. A special thanks to Prof. Robert W. Heath Jr. for his two courses Wireless Communications Lab and Space-time Communications, which have led to our joint work on massive MIMO with D2D. I am grateful to Prof. Mihai Sirbu for teaching me Theory of Probability, which provided me with a rigorous background of mathematical probability. I thank Prof. Haris

Vikalo for allowing me to audit his course Estimation Theory, which helped me develop a systematic understanding of estimation problems.

I would like to acknowledge Nokia Siemens Networks (now Nokia) and National Science Foundation (under the Grant CIF-1016649) for supporting my doctorate. I also thank the Graduate School of UT Austin for awarding me the Microelectronics and Computer Development (MCD) fellowship.

I would like to take this opportunity to acknowledge my industry collaborators. A special thanks to Dr. Amitava Ghosh and his group at Nokia for their invaluable technical inputs to this dissertation. I also thank Dr. Amitava Ghosh for hosting me at Nokia Siemens Networks, Arlington Heights, IL, during summer 2013. Thanks to Dr. Harish Viswanathan for hosting me at Alcatel-Lucent Bell Labs, Murray Hill, NJ, during summer 2012, and to Dr. Shailesh Patil and Dr. Libin Jiang for hosting me at Qualcomm Flarion Technologies Inc., Bridgewater, NJ, during summer 2014. The internship with Qualcomm inspired my work on asynchronous multicarrier wireless networks in this dissertation.

I thank my Chinese friends at UT Austin with whom I had a lot of fun during the journey: Beiyu, Chao, Chun-Hung, Hongbo, Jing, Ping, Qiaoyang, Tianyang, Yan, Yingxi, Yingzhe, Yudong, Yuhuan, Xinchen and Zheng. I would especially like to thank Qiaoyang, who joined Jeff's group roughly at the same time as I did, for encouragement along our paths towards doctorate work. I thank Harpreet and Sarabjot for various technical discussions. A special thanks to Avhishek for being a special friend and teaching me a lot

about Indian culture. I owe a deep thank you to my friend Namyoon for sharing with me his enormous experiences. I would also like to acknowledge Melanie Gulick from ECE graduate office and Janet Preuss, Jennifer Graham, Karen Little, and Lauren Bringle from WNCG office, for taking care of all the paperwork.

Finally, I would like to thank my parents for their love and encouragement. A special thanks to my grandfather for his stress on quality education that has made me the person I am today, and to my grandmother for loving me unconditionally. The last word of acknowledgment goes to my dear wife Xi Yang, whose love and support allowed me to finish this journey. Xi, I owe you a honeymoon.

Integrated Cellular and Device-to-Device Networks

Xingqin Lin, Ph.D.
The University of Texas at Austin, 2014

Supervisor: Jeffrey G. Andrews

Device-to-device (D2D) networking enables direct discovery and communication between cellular subscribers that are in proximity, thus bypassing the base stations (BSs). In principle, exploiting direct communication between nearby mobile devices will improve spectrum utilization, overall throughput, and energy consumption, while enabling new peer-to-peer and location-based applications and services. D2D-enabled broadband communication technology is also required by public safety networks that must function when cellular networks are not available. Integrating D2D into cellular networks, however, poses many challenges and risks to the long-standing cellular architecture, which is centered around the BSs. This dissertation identifies outstanding technical challenges in D2D-enabled cellular networks and addresses them with novel models and fundamental analysis.

First, this dissertation develops a baseline hybrid network model consisting of both ad hoc nodes and cellular infrastructure. This model uses Poisson point processes to model the random and unpredictable locations of mobile

users. It also captures key features of multicast D2D including multicast receiver heterogeneity and retransmissions while being tractable for analytical purpose. Several important multicast D2D metrics including coverage probability, mean number of covered receivers per multicast session, and multicast throughput are analytically characterized under the proposed model.

Second, D2D mode selection which means that a potential D2D pair can switch between direct and cellular modes is incorporated into the hybrid network model. The extended model is applied to study spectrum sharing between cellular and D2D communications. Two spectrum sharing models, overlay and underlay, are investigated under a unified analytical framework. Analytical rate expressions are derived and applied to optimize the design of spectrum sharing. It is found that, from an overall mean-rate perspective, both overlay and underlay bring performance improvements (vs. pure cellular).

Third, the single-antenna hybrid network model is extended to multi-antenna transmission to study the interplay between massive MIMO (multi-input multiple-output) and underlaid D2D networking. The spectral efficiency of such multi-antenna hybrid networks is investigated under both perfect and imperfect channel state information (CSI) assumptions. Compared to the case without D2D, there is a loss in cellular spectral efficiency due to D2D underlay. With perfect CSI, the loss can be completely overcome if the number of canceled D2D interfering signals is scaled appropriately. With imperfect CSI, in addition to pilot contamination, a new asymptotic underlay contamination effect arises.

Finally, motivated by the fact that transmissions in D2D discovery are usually not or imperfectly synchronized, this dissertation studies the effect of asynchronous multicarrier transmission and proposes a tractable signal-to-interference-plus-noise ratio (SINR) model. The proposed model is used to analytically characterize system-level performance of asynchronous wireless networks. The loss from lack of synchronization is quantified, and several solutions are proposed and compared to mitigate the loss.

Table of Contents

Acknowledgments **v**

Abstract **viii**

List of Tables **xv**

List of Figures **xvi**

Chapter 1. Introduction **1**
1.1 Background 2
1.2 Motivation 6
1.3 Contributions and Organization 9

Chapter 2. Multicast D2D Transmissions **15**
2.1 Related Work 16
2.2 Contributions and Main Outcomes 18
2.3 System Model 20
2.3.1 Distributions of Network Nodes 20
2.3.2 Multicast Transmission 21
2.3.3 Channel Model 23
2.3.4 Performance Metrics 25
2.4 Multicast without Network Assistance 26
2.4.1 Coverage Probability 27
2.4.2 Spatial Correlation in Multicast D2D Transmissions . . 28
2.4.3 Mean Number of Covered Receivers 31
2.4.4 Multicast Throughput 36
2.4.5 Impact of Dynamics 38
2.5 Multicast with Network Assistance 42
2.6 Optimizing Multicast Transmissions 46

2.7 Summary 53
2.8 Appendix 53
2.8.1 Proof of Theorem 2.1 53
2.8.2 Proof of Proposition 2.1 56
2.8.3 Proof of Proposition 2.2 57
2.8.4 Proof of Proposition 2.3 58
2.8.5 Proof of Proposition 2.4 58
2.8.6 Proof of Lemma 2.1 59
2.8.7 Proof of Proposition 2.6 60
2.8.8 Proof of Proposition 2.7 61

Chapter 3. Spectrum Sharing between Cellular and D2D Communications 62

3.1 Related Work 63
3.2 Contributions and Main Outcomes 65
3.3 System Model 67
3.3.1 Network Model 67
3.3.2 Transmission Scheduling 71
3.3.3 Performance Metrics 72
3.4 Preliminary Analysis 73
3.4.1 A Unified Analytical Approach 73
3.4.2 Transmit Power Analysis 75
3.5 Analysis of Overlay In-Band D2D 78
3.5.1 Link Spectral Efficiency 78
3.5.2 Optimizing Spectrum Partition 84
3.6 Analysis of Underlay In-Band D2D 87
3.6.1 Link Spectral Efficiency 87
3.6.2 Optimizing Spectrum Access 90
3.7 Overlay vs. Underlay: A Case Study 93
3.8 Summary 96
3.9 Appendix 96
3.9.1 Proof of Lemma 3.1 96
3.9.2 Proof of Proposition 3.1 97

3.9.3 Proof of Proposition 3.2 98
3.9.4 Proof of Proposition 3.3 99
3.9.5 Proof of Corollary 3.2 101
3.9.6 Proof of Proposition 3.4 102
3.9.7 Proof of Proposition 3.5 103
3.9.8 Proof of Proposition 3.6 104

Chapter 4. Massive MIMO Systems with D2D Underlay 105
4.1 Related Work 106
4.2 Contributions and Main Outcomes 108
4.3 Mathematical Models 110
4.3.1 Network Model 110
4.3.2 Baseband Channel Models 111
4.3.3 Receive Filters 113
4.4 Spectral Efficiency with Perfect Channel State Information . . 115
4.4.1 Asymptotic Cellular Spectral Efficiency 115
4.4.2 Non-asymptotic Cellular Spectral Efficiency 120
4.4.3 D2D Spectral Efficiency 122
4.5 Spectral Efficiency with Imperfect Channel State Information . 124
4.5.1 Estimating UE-BS Channels 124
4.5.2 Asymptotic Cellular Spectral Efficiency 127
4.6 Simulation and Numerical Results 131
4.7 Summary 138
4.8 Appendix 138
4.8.1 Proof of Proposition 4.1 138
4.8.2 Proof of Proposition 4.2 142
4.8.3 Proof of Proposition 4.4 144
4.8.4 Proof of Lemma 4.1 146
4.8.5 Proof of Proposition 4.6 147

Chapter 5. Asynchronous Multicarrier Wireless Networks 149
5.1 Background and Related Work 150
5.2 Contributions and Main Outcomes 152
5.3 System Model 153
5.4 Tractable SINR Model for Asynchronous Networks 155
5.4.1 Link-Level Timing Misalignment Analysis 155
5.4.2 From Link-Level to System-Level Studies 159
5.5 On the Decodable Transmitters of a Typical Receiver 162
5.5.1 Mean Number of Decodable Transmitters 164
5.5.2 An Upper Bound on the Distribution of the Number of Decodable Transmitters 170
5.5.3 On Decoding the Nearest Transmitter 174
5.5.4 Optimizing System Throughput 175
5.6 Solutions for Mitigating the Loss of Asynchronous Transmissions 177
5.7 Summary 180
5.8 Appendix 181
5.8.1 Derivation of Equations (5.4) and (5.7) 181
5.8.2 Proof of Proposition 5.1 183
5.8.3 Proof of Proposition 5.2 184
5.8.4 Proof of Proposition 5.3 186

Chapter 6. Conclusions 188
6.1 Summary 188
6.2 Future Directions 190
6.2.1 Millimeter Wave D2D 190
6.2.2 D2D in Heterogeneous Networks 192

Bibliography 194

Vita 214

List of Tables

2.1 Simulation/Numerical Parameters for Multicast D2D 22

3.1 Simulation/Numerical Parameters for Spectrum Sharing between Cellular and D2D 78

4.1 Simulation/Numerical Parameters for Massive MIMO with D2D Underlay 132

5.1 Simulation/Numerical Parameters for Asynchronous OFDM Networks 169

List of Figures

1.1 Illustration of possible D2D use cases and potential benefits. . 5

1.2 A hybrid network consisting of both cellular and D2D links. Solid triangles denote BSs. Blue solid and red solid lines respectively denote cellular and D2D links. The cellular UE-BS association is formed based on maximum received power with pathloss and shadowing considered. 7

2.1 A sample realization of the network nodes: Blue solid triangles, red solid squares and green dots denote BSs, multicast D2D transmitters and receivers, respectively. 21

2.2 Coverage probability versus detection threshold without network assistance: The numbers, 50 m, 150 m, 250 m, indicate three different D2D Tx-Rx distances. 29

2.3 Normalized mean number of covered receivers versus multicast times without network assistance. 34

2.4 Multicast throughput versus detection threshold: $R = 150$ m. . 38

2.5 Tradeoff between efficiency – multicast throughput – and reliability – mean number of covered receivers. 39

2.6 Dynamics increase the mean number of covered receivers. . . . 41

2.7 Network assistance helps increase the mean number of covered receivers. 45

2.8 Impact of network assistance vs. dynamics: $R = 250$ m. 46

2.9 Network assistance statistics in the case of optimized multicast transmissions: $\eta = 95\%, B = 2$. 52

3.1 Different D2D spectrum sharing scenarios 63

3.2 A hybrid network consisting of both cellular and D2D links. Solid triangles, solid squares and dots denote BSs, uplink cellular transmitters and D2D transmitters, respectively. For clarity we omit plotting D2D receivers, each of which is randomly located on the circle centered at the associated D2D transmitter. 68

3.3 An approximate uplink interference analysis. The typical cellular transmitter is uniformly distributed in $\mathcal{A}$, while cellular interferers form a PPP with density λ_{b} outside the disk $\mathcal{A}$. . . 72

3.4 UE transmit power versus network operating regime SNR_m with $\tilde{N}_0 = -174$ dBm/Hz and $B_w = 1$ MHz. 79
3.5 Validation of the analytical SINR CCDF of cellular links. . . . 84
3.6 Validation of the analytical SINR CCDF of D2D links. 85
3.7 Average rates of cellular and potential D2D UEs in the case of overlay in-band D2D. 87
3.8 Utility value vs. D2D spectrum partition factor η under different values of q, the proportion of potential D2D UEs. 88
3.9 Average rates of cellular and potential D2D UEs in the case of underlay in-band D2D. 91
3.10 Utility value vs. D2D spectrum access factor β under different values of q, the proportion of potential D2D UEs. 92
3.11 A case study on the rate performance of overlay. 94
3.12 A case study on the rate performance of underlay. 95

4.1 A D2D underlaid massive MIMO system consisting of both cellular and D2D links. For clarity, only the central cell is shown. D2D pairs located outside of the cells are out of cellular coverage but still contribute to the total aggregate D2D interference. . . 111
4.2 Simulated cellular spectral efficiency vs. analytical lower bound (4.17) with perfect CSI. 134
4.3 Cellular spectral efficiency with scaled cellular transmit power and perfect CSI. 135
4.4 Simulated D2D spectral efficiency vs. analytical lower bound (4.22) with perfect CSI and $(n_{\mathrm{c}}, n_{\mathrm{d}}) = (0, 2)$. 136
4.5 Effect of multi-user cellular transmission on D2D spectral efficiency with perfect CSI and $(n_{\mathrm{c}}, n_{\mathrm{d}}) = (0, 0)$. 136
4.6 Effect of D2D underlay contamination on asymptotic cellular spectral efficiency of massive MIMO with $(m_{\mathrm{c}}, m_{\mathrm{d}}) = (0, 0)$ and $T_{\mathrm{c}} = 4$. 137

5.1 Received power of an OFDM signal with timing misalignment. $N = 1024$; $N_{cp} = 72$; the used subcarriers are $\{-299, ..., 0, ...300\}$. 162
5.2 Signal and self-interference powers of an OFDM signal received on the central subcarrier with timing misalignment. $N = 1024$; $N_{cp} = 72$; the used subcarriers are $\{-300, ..., 0, ...299\}$. 163
5.3 The upper bound on the mean number of decodable transmitters (c.f. (5.22)) versus pathloss exponent. 166

5.4 Mean number of decodable transmitters versus pathloss exponent in synchronized networks. 168
5.5 Mean number of decodable transmitters versus detection threshold. 170
5.6 Analytical upper bound vs. simulation on the distribution of the number of decodable transmitters: $\lambda = 1/400^2$ m^{-2}. 173
5.7 Analytical upper bound vs. simulation on the distribution of the number of decodable transmitters: $\lambda = 1/800^2$ m^{-2}. 173
5.8 Impact of timing misalignment on the decoding probability of the nearest transmitter. 176
5.9 System throughput versus detection threshold. 177
5.10 Using semi-static receiver timing positioning with multiple timing hypotheses to mitigate the loss of asynchronous transmissions. 181

Chapter 1

Introduction

Device-to-device (D2D) networking enables direct discovery and communication between cellular subscribers that are in proximity. Historically, the idea of integrating D2D networking into cellular networks, or more generally, the concept of hybrid networks consisting of both infrastructure-based and ad hoc networks has long been a topic of considerable interest. It has been shown in earlier studies [92, 143] that cellular coverage and throughput performance can be improved by allowing radio signals to be relayed by mobiles. Researchers also studied ad hoc networks enhanced by infrastructure, particularly from the perspective of transport capacity [50]. It has been demonstrated that better scaling laws of transport capacity can be achieved in a hybrid network than in a purely ad hoc network [81, 93, 148].

This introductory chapter starts with describing the socio-technological trends that have led to the recent surge of interest in D2D networking in Section 1.1. Section 1.2 argues that enabling D2D networking in cellular networks is challenging and requires a rethinking of many of the working assumptions and models used to date for cellular systems, which is the motivation behind this dissertation. It then highlights three fundamental design principles that

will guide the study of D2D networking throughout this dissertation. Section 1.3 provides an overview of the contributions of this dissertation and a brief description of dissertation organization.

1.1 Background

Over the past decade, we have witnessed a mobile revolution, which is driven largely by the invention of smartphones. This revolution is dramatically changing many aspects of our life including learning, playing, shopping and dating. It also gives rise to the popularity of proximity-based services like location-based mobile advertising, local information sharing and mobile social networking. These proximity-based services help discover people and things of interest around us. Existing technologies enabling proximity-based services may be broadly classified into two types: peer-to-peer (P2P) and cloud-based solutions [33].

WiFi Direct [141] and Bluetooth Low Energy [46] are the two most popular P2P technologies being used today. The main limitation of them is that they operate in unlicensed spectrum and thus their transmit powers are low and their communications suffer from uncontrolled interference generated by other devices like microwave ovens and cordless telephones sharing the spectrum. As a result, the range of WiFi Direct is typically limited within tens of meters, while the range of Bluetooth Low Energy is even smaller. The limited ranges of these two technologies make them unsuitable for median-to-long-range proximity-based services. Further, WiFi Direct is not energy efficient;

its energy consumption increases exponentially as the number of devices increases [33].

Cloud-based proximity-aware solutions (e.g. Foursquare and Facebook Places) usually work as follows. Users first register their location information in a central server once launching the application; the central server then distributes the registered location information to other users using the application. Alternatively, the server constantly tracks the user locations using the Global Positioning System (GPS), updates proximity data, and feeds the information back to users. As a result, a device's battery may drain very quickly due to the GPS power consumption and frequent communication session set-up with the network. In addition to battery drain, cloud-based proximity-aware solutions also create privacy concerns as users need to reveal their identities and/or locations to the network.

In view of the limitations of existing proximity-aware technologies, an ideal proximity-aware technology should have the following characteristics. First, it should be *scalable*. Further advancement of proximity-based services hinges on device discovery capability in terms of discovery range and the number of discoverable devices. An ideal proximity-aware technology should provide a much more powerful discovery capability than existing P2P technologies and further be scalable to expand use cases. Second, it should be *energy efficient*. Always-on proximal discovery is highly desirable in proximity-based services [111]. But always-on proximal discovery poses a big challenge to device battery and obviously cannot be met using the power-hungry GPS-based

solutions. An ideal proximity-aware technology should minimize its impact on device battery while providing always-on proximal discovery. Third, it should be *privacy sensitive*. One way to minimizing privacy barriers for consumers is to restrict user identity, location information and any other sensitive private data in local devices rather than revealing them to the cloud.

The need for a better proximity-aware technology has led to a recent surge of interest in enabling D2D networking in licensed cellular spectrum. Efforts have been taken by wireless engineers to meet the socio-technological trend: Qualcomm has built a mobile communication system known as FlashLinQ wherein "wireless sense" is implemented to enable proximity-aware communication among devices [29, 144]. Now the Third Generation Partnership Project (3GPP) is targeting the availability of D2D networking in Long-Term Evolution (LTE) Release 12 and beyond [2].

Another major drive of D2D networking is to enable LTE to become a competitive broadband communication technology for public safety networks [35], used by first responders. Due to the legacy issues and budget constraints, current public safety networks are still mainly based on obsolete 2G technologies like Project 25 (P25) and Terrestrial Trunked Radio (TETRA) while commercial networks are rapidly migrating to LTE. This evolution gap and the desire for enhanced services have led to global attempts to upgrade existing public safety networks. For example, the USA has decided to build an LTE-based public safety network in the 700 MHz band. Compared to commercial networks, public safety networks have much more stringent service require-

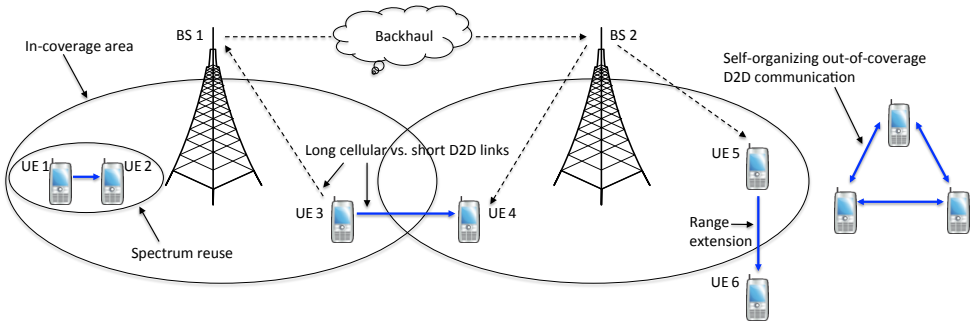

Figure 1.1: Illustration of possible D2D use cases and potential benefits.

ments (e.g. reliability and security) and also require direct communication among mobiles, especially when cellular coverage fails or is not available. This essential direct mode feature is currently missing in LTE.

From a technical perspective, exploiting the natural proximity of communicating devices may provide multiple performance benefits. First, D2D user equipment (UE) may enjoy high data rate and low end-to-end delay due to the short-range direct communication. Second, it is more resource-efficient for proximate UEs to communicate directly with each other, versus routing through base stations (BSs) and possibly the core network. In particular, compared to normal downlink/uplink cellular communication, direct communication saves energy and improves radio resource utilization. Third, switching from an infrastructure path to a direct path offloads cellular traffic, alleviating congestion, and thus benefiting other non-D2D UEs as well. Other benefits may be envisioned such as range extension via UE-to-UE relaying. Figure 1.1 gives an illustration of possible D2D use cases and the potential benefits.

1.2 Motivation

Integrating D2D features into current cellular networks poses many challenges and risks. Cellular networks have existed for several decades. Network operators are likely to resist a technology that dramatically changes the long-standing cellular architecture, which is centered around BSs, unless significant engineering gains and/or big commercial opportunities can be proven. Further, all existing cellular technologies including LTE are mainly designed and optimized for BS-UE links, while the D2D design involves UE-UE links. Also, one has to take into account the impact of D2D on the wide area network (WAN) as a whole. D2D fundamentally alters the cellular architecture, reducing the primacy of BSs and enabling UE devices to transmit directly to nearby UE devices. Such a shift requires a rethinking of many of the working assumptions, models, and analysis used to date for cellular systems.

The goal of this dissertation is to introduce novel models and fundamental analysis to address the challenging engineering aspects of D2D networking in cellular networks. To this end, in the sequel we highlight three design principles that distinguish the engineering of D2D networking from the design of either cellular or ad hoc networking. These fundamental principles will guide our study of D2D networking throughout this dissertation.

The first and foremost design principle is to ensure harmonious coexistence of cellular and D2D networking. A D2D-enabled cellular network, as illustrated in Figure 1.2, is a highly complicated hybrid system. The design of this kind of hybrid systems requires a careful handling of

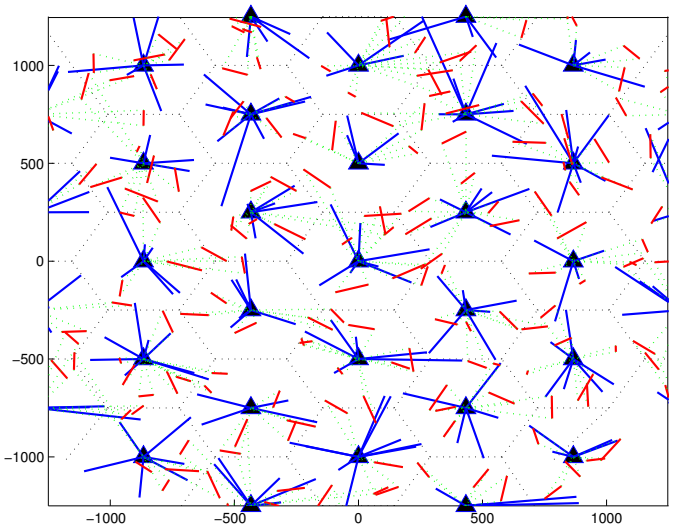

Figure 1.2: A hybrid network consisting of both cellular and D2D links. Solid triangles denote BSs. Blue solid and red solid lines respectively denote cellular and D2D links. The cellular UE-BS association is formed based on maximum received power with pathloss and shadowing considered.

the interaction between cellular and D2D services. D2D networking in licensed cellular spectrum consumes radio resources that may otherwise be utilized for cellular services. This inevitably affects to some extent operators' capability of serving existing cellular customers, and thus careful resource management is essential for the coexistence of cellular and D2D networking. Further, how network resources are managed has a direct impact on the interference environment in a D2D-enabled cellular network. If the underlying cellular network adopts a frequency-division duplexing (FDD) deployment and D2D networking utilizes downlink resources, a transmitting D2D device may cause high interference to nearby cochannel cellular UEs receiving downlink traffic. In contrast, when D2D networking utilizes uplink resources, a receiving D2D de-

vice may experience strong interference from nearby cochannel cellular UEs transmitting uplink traffic. To simplify the interference management, a less efficient approach may be to assign orthogonal radio resources to cellular and D2D links. However, mutual interference between cellular and D2D may persist due to possible adjacent channel power leakage and/or a lack of perfect network synchronization.

Second, D2D networking should leverage infrastructure assistance as much as possible. Mobile ad hoc networks (MANETs) have been studied and developed extensively over about three decades with very limited success [14]. The notoriously difficult design of MANETs is mainly due to lack of infrastructure support. In contrast, D2D networking can typically rely on the assistance from network infrastructure (i.e., BSs) for control functions like synchronization, session setup, resource allocation, routing, and other overhead-consuming functions that are extremely costly in a MANET. The main challenges here are (i) to have a design of network assisted D2D compatible with existing already very complicated cellular control, and (ii) to optimize the design given limited network resources. In the public safety context, D2D must function even when cellular coverage is not available, so it is more like a MANET. D2D in the out-of-cellular-coverage mode however is only required to be rudimentary, and hence is more like a walkie-talkie than a full MANET, which may require streaming video. Further, the wisdom of cellular engineering may facilitate the design of D2D networking even in the out-of-cellular-coverage case. In particular, out-of-cellular-coverage public

safety UEs are often clustered (usually on the order of at most tens of nodes), so a clusterhead may be elected and acts as a *de facto* BS.

Third, energy efficient D2D networking is critical. Battery drain has already been a primary concern for today's mobile devices, especially smartphones. For proximal discovery, D2D networking is targeting at always-on proximal discovery with very long discovery range ($\geq$500 m) and large discovery capacity (1000s devices) [111]. Despite that D2D communication may save UE power by exploiting direct short-range communication, there is a major concern about the UE power consumption involved in the device discovery process, in which a UE may have to periodically broadcast discovery signal and listen to the discovery signals from other UEs. Therefore, it is important to minimize the impact of proximity-aware services on UE battery and optimize the design to save UE power. Equivalently, with a given tolerable level of drain of device battery, D2D networking should be optimized with appropriate PHY techniques (e.g. modulation, coding and power control), MAC design (e.g. scheduling, automatic repeat request (ARQ) and retransmissions), and other upper layers protocols.

1.3 Contributions and Organization

Our main thesis is that integrating D2D into cellular networks poses many challenges to the long-standing cellular architecture. As indicated previously, D2D networking requires a rethinking of models and analysis used to date for cellular networks. This dissertation identifies four outstanding tech-

nical challenges in D2D-enabled cellular networks and addresses them with novel models and fundamental analysis. The technical contributions of this dissertation are covered in Chapters 2 to 5.

Multicast D2D transmissions. Chapter 2 studies multicast D2D transmissions. In D2D-enabled cellular networks, direct multicast transmission, where the same packets from a UE are sent to multiple receivers at the same time, is important for device discovery and applications like location based advertising in commercial networks. Compared to communicating with each receiver separately, one direct multicast transmission reduces overhead and saves resources. Due to the heterogeneous locations of receivers and complicated radio environment, however, link quality may vary significantly over receivers in each multicast cluster; thus retransmissions are often required to cover more or all the receivers. The retransmissions introduce significant correlation among the signals and interference over the multicast processes, making the analysis of multicast performance very challenging.

Chapter 2 proposes a novel hybrid network model consisting of both ad hoc nodes and cellular infrastructure. This model further captures key features of multicast D2D including multicast receiver heterogeneity and retransmissions. The model is applied to analytically characterize several important multicast metrics including:

- the coverage probability;
- the mean number of covered receivers per multicast session;
- the multicast throughput.

The derived expressions allow for efficient numerical evaluation; some of them are even in closed-form. The results indicate that retransmissions may increase the coverage probability and the mean number of covered receivers but may hurt the throughput. Chapter 2 also examines how the multicast performance would be affected by certain factors like network dynamics (due to e.g., UE mobility) and infrastructure assistance (i.e., allowing the network to relay multicast signals). The results show that both may help improve the multicast performance significantly. Chapter 2 further explores how to optimize multicasting by choosing the optimal multicast rate, optimal number of retransmission times, and optimal strategy for the infrastructure assistance.

Spectrum Sharing between Cellular and D2D Communications. Chapter 3 investigates how to share the spectrum resources between cellular and D2D communications, which is a fundamental issue in D2D design. Based on the spectrum sharing manner, D2D can be classified into two types.

- In-band: D2D uses the cellular spectrum.
- Out-of-band: D2D utilizes different bands (e.g., 2.4 GHz ISM band) other than the cellular band.

In-band D2D can be further classified into two categories: *overlay* and *underlay*. Overlay means that cellular and D2D transmitters use orthogonal time/frequency cellular resources, while underlay means that D2D transmitters opportunistically access the time/frequency resources occupied by cellular users. Note that a potential D2D pair can switch between direct and con-

ventional cellular communications. This flexible D2D mode selection feature further complicates the spectrum sharing design.

Chapter 3 develops a unified model and an analytical framework for D2D spectrum sharing. Specifically, Chapter 3 adapts the hybrid network model proposed in Chapter 2 to encompass the above diverse spectrum sharing scenarios and D2D mode selection. Analytical rate expressions are derived for each spectrum sharing scenario. Based on the rate expressions, the design parameters of spectrum sharing and D2D mode selection are optimized. The results indicate that, from an overall mean-rate perspective, both overlay and underlay bring performance improvements (vs. pure cellular).

Massive MIMO Systems with D2D Underlay. In Chapter 4, the focus is shifted from single-antenna transmission to studying multi-antenna transmission, i.e., multi-input multiple-output (MIMO). D2D networking will coexist with MIMO, which has become an indispensable component of current cellular networks. In a D2D underlaid cellular network, the uplink spectrum is reused by the D2D transmissions, causing mutual interference with the on-going cellular transmissions. Massive MIMO is appealing in such a context as the BS's large antenna array can nearly null the D2D-to-BS interference. The multi-user transmission in massive MIMO, however, may lead to increased cellular-to-D2D interference. Further, if cochannel D2D signals are present when estimating massive MIMO channels, the estimated channel state information (CSI) would become less accurate, which may hurt massive MIMO performance.

Chapter 4 extends the baseline single-antenna hybrid network model to multi-antenna transmission and studies the interesting interplay between massive MIMO and underlaid D2D networking in a multi-cell setting. Chapter 4 investigates cellular and D2D spectral efficiency under both perfect and imperfect CSI at the receivers that employ partial zero-forcing. Compared to the case without D2D, there is a loss in cellular spectral efficiency due to D2D underlay. With perfect CSI, the loss can be completely overcome if the number of canceled D2D interfering signals is scaled with the number of BS antennas at an arbitrarily slow rate. With imperfect CSI, in addition to pilot contamination, a new asymptotic effect termed *underlay contamination* arises. In the non-asymptotic regime, simple analytical lower bounds are derived for both the cellular and D2D spectral efficiency.

Asynchronous Multicarrier Wireless Networks. An implicit assumption made in Chapters 2 to 4 is that the networks are synchronized. In D2D discovery, a UE seeks to identify other UEs in its proximity via periodically broadcasting/receiving discovery signals. In this process, devices are usually not or imperfectly synchronized and thus different devices have different notions of timing. From the viewpoint of a typical receiver, the received signals from different transmitters are asynchronous, leading to a loss of orthogonality between subcarriers when an orthogonal frequency-division multiplex (OFDM) waveform is used.

Chapter 5 develops a novel analytical framework for asynchronous wireless networks deploying multicarrier transmission. Based on a detailed link-

level analysis, Chapter 5 proposes a tractable system-level signal-to-interference-plus-noise ratio (SINR) model for asynchronous OFDM networks. The proposed model is used to analytically characterize several important statistics in asynchronous networks with spatially distributed transmitters, including (i) the number of decodable transmitters, (ii) the decoding probability of the nearest transmitter, and (iii) the system throughput. The system-level loss from lack of synchronization is quantified, and to mitigate the loss, Chapter 5 compares and discusses four possible solutions including:

- extended cyclic prefix;
- advanced receiver timing;
- dynamic receiver timing positioning;
- semi-static receiver timing positioning with multiple timing hypotheses.

The model and results are general, and apply not only to D2D networking but also to general ad hoc networks and cellular systems.

Finally, Chapter 6 concludes this dissertation by summarizing key contributions and discussing promising future research directions of D2D networking.

Chapter 2

Multicast D2D Transmissions

In D2D-enabled cellular networks, direct multicast transmission, where the same packets from a UE are sent to multiple receivers, is important for scenarios such as the following.

(i) *Local file transfer/video streaming*: Local UEs may have common packets for nearby receivers; for example, local marketers may send the same advertising messages to people who happen to be in the neighborhood.

(ii) *Device discovery*, referring to the process of discovering surrounding devices, is a basic function for many D2D use cases [18,42]. During device discovery, each device periodically broadcasts beacons to announce its existence, while other devices periodically scan and each may respond to the message once it receives the beacon.

(iii) *Clusterhead selection/coordination*: For out-of-coverage D2D, it is being discussed in 3GPP to have one UE act as a clusterhead within a group of UEs. The clusterhead can help achieve local synchronization, manage radio resources and schedule transmissions. Clusterhead selection normally involves multicast when potential clusterheads send out beacons to announce their roles.

(iv) *Group/broadcast communications*: In public safety networks providing services like police, fire and ambulance, D2D group/broadcast communications are required features [2].

In the aforementioned scenarios, compared to communicating with each receiver separately, one direct multicast transmission reduces overhead and saves resources. However, unlike unicast D2D (see e.g. [89, 147] and references therein), multicast D2D has its own challenges. For example, due to the heterogeneous locations of receivers and complicated radio environment, link quality may vary significantly over receivers in each multicast cluster; thus retransmissions are often required to cover more or all the receivers, which degrades the whole point of multicast vs. unicast. In addition to receiver heterogeneity, it is the UEs rather than BSs that perform multicast; this introduces additional challenge due to the limited capability of UEs. Despite these challenges, compared to multicast in ad hoc networks, multicast D2D has certain conveniences; for example, it may be assisted by the cellular network infrastructure which is not available to ad hoc networks.

It is the significance and distinctive traits of multicast D2D described above that motivate our study in this chapter.

2.1 Related Work

Multicast in cellular networks can be broadly classified into two classes: Single-rate and multi-rate [7]. In single-rate multicast, the transmitter sends the packets to all the receivers at a common rate [47,95,142,145]. For example,

in [47] multicast throughput-delay tradeoff is studied in a single cell system by selecting the median throughput as the multicast rate. In [95], dynamic power and subcarrier allocation is performed to adapt to the receiver with the weakest link. In contrast, receiver heterogeneity is exploited in multi-rate multicast, where different receivers in the same multicast cluster may receive packets at different rates based on e.g. the link qualities [32, 59, 63, 119, 123]. Though being more efficient, multi-rate multicast is much more complex than single-rate multicast in terms of both analysis and implementation.

In parallel with the academic studies, standardization effort in addressing multicast services has been/is being undertaken and mainly focuses on single-rate multicast. For example, multicast services were addressed in GSM/WCDMA and are being addressed in LTE by 3GPP; the 3GPP work item is known as multimedia broadcast and multicast service (MBMS) [3]. Similarly, 3GPP2 addressed multicast services in CDMA2000 with the work item known as broadcast and multicast service (BCMCS) [6].

There also exists much work on multicast in ad hoc networks [26, 85, 94, 118]. For example, in [26] the tradeoff between throughput, stability, and packet loss is studied and a transmission policy is proposed to maximize throughput subject to stability and packet loss constraints. While [85, 118] respectively study transport capacity for single hop and multihop wireless networks, [94] tackles ad hoc multicast from the transmission capacity perspective [139].

Unlike the aforementioned studies, there exists a small set of work on

multicast in hybrid networks consisting of both ad hoc nodes and cellular infrastructure [73, 98, 107]. Though receivers with good channel qualities may relay the multicast traffic to receivers with poor link qualities using ad hoc mode in [107], the multicast transmitter is still the BS. In contrast, [98] studies the multicast transport capacity of a hybrid network, and sheds light on its asymptotic growth rate in the number of network nodes. The more recent work [73] jointly considers coding in caching and multicast delivery in a D2D wireless network, leading to benefits of so-called "coded multicast gain" and spatial reuse. In addition to theoretical analysis, there exist works like [36,117] which rely more on simulations to understand the performance of multicast D2D.

2.2 Contributions and Main Outcomes

The main contributions and outcomes of this chapter are summarized as follows.

A tractable hybrid network model. We introduce a tractable hybrid network model, which consists of both ad hoc nodes and cellular infrastructure and captures the multicast receiver heterogeneity and retransmissions. Specifically, we use independent PPPs to model the spatial positions of the BSs and D2D transmitters. While such a random PPP model is well motivated by the random and unpredictable mobile user locations, using a PPP to model BS locations has been validated in the literature (see e.g. [10, 24, 34, 83, 90]).

Multicast performance analysis. Unlike in one-shot transmission,

there exists significant correlation among the signals and interference over the multicast retransmissions. By tackling this time correlation, we characterize the coverage probability at a particular receiver. Building on the coverage analysis, we derive expressions for the mean number of covered receivers in each multicast cluster. The expressions allow for efficient numerical evaluation; some of them are even in closed-form. Further, we explore multicast throughput and use it as a metric for selecting the optimal multicast rate. These studies reveal a fundamental tradeoff between efficiency (multicast throughput) and reliability (mean number of covered receivers).

Impact of dynamics. Though in our default model multicast transmitters are static, we also explore the impact of dynamics (due to e.g. mobility or bursty transmissions) and analytically show that dynamics *hurt* the performance if one would like to support a target SINR for multiple successive transmissions. In contrast, interestingly, we find that dynamics *improve* the multicast performance in terms of either coverage probability or mean number of covered receivers or multicast throughput.

Network-assisted multicast D2D. We analyze the multicast performance by incorporating network assistance, i.e., allowing the network to relay the multicast signals. It is shown that network assistance can significantly enhance the multicast performance compared to the case of no network assistance. In addition, we formulate a network-assisted multicast D2D optimization problem which minimizes the number of retransmission times subject to a resource constraint at the BSs and a multicast reliability constraint. An

efficient algorithm is also proposed.

2.3 System Model

In this section, we propose a tractable baseline model for studying multicast D2D transmissions.

2.3.1 Distributions of Network Nodes

We consider a hybrid network consisting of both cellular and D2D links. The positions of BSs form an independent Poisson point process (PPP) $\Phi_{\mathrm{b}} = \sum_i \delta_{z_i}$ with intensity λ_b; here δ_z denotes the Dirac measure at position $z \in \mathbb{R}^2$, i.e., for any measurable set $A \subset \mathbb{R}^2$, $\epsilon_z(A) = 1$ if $z \in A$, and 0 otherwise. With a slight abuse of notation, we will also use the position z to indicate the node located at z. Similarly, the positions of multicast D2D transmitters form an independent PPP $\Phi_{\mathrm{m}} = \sum_i \delta_{x_i}$ with intensity λ_m. We further assume that for each D2D transmitter x_i, the positions of its intended receivers form a point process $\Phi_{\mathrm{m},x_i} = \sum_i \delta_{y_i}$ with intensity measure $\Lambda_{x_i}(\cdot) = \lambda_r \nu(\cdot \cap B(x_i, R))$, where $\nu(\cdot)$ is Lebesgue measure in $\mathbb{R}^2$ and $B(x, R)$ denotes the ball centered at x with radius R. Note that we do not assume any specific distribution for the receiver point process Φ_{m,x_i} except the first-order intensity measure; in particular, Φ_{m,x_i} does not have to be Poisson distributed.

Conditioning on Φ_m, $\{\Phi_{\mathrm{m},x_i}\}$ are assumed to be independent. Those familiar with stochastic geometry will immediately recognize that $\{\Phi_{\mathrm{m},x_i}\}$, which are in the space of point processes on $\mathbb{R}^2$, are independent marks of the

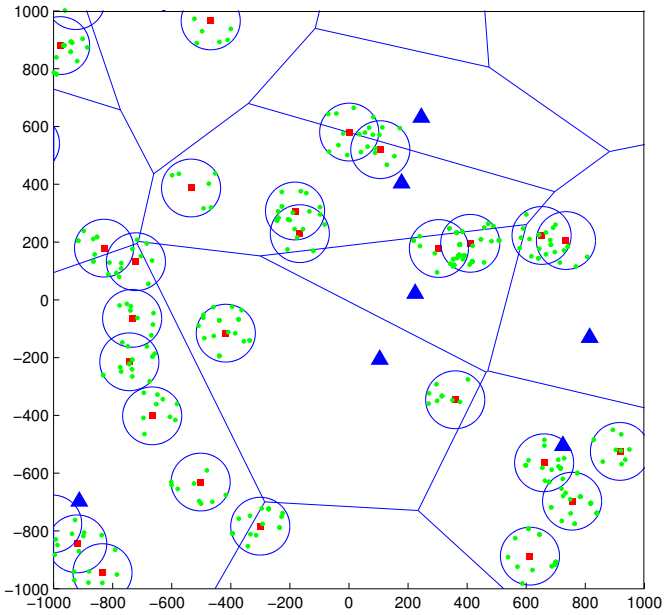

Figure 2.1: A sample realization of the network nodes: Blue solid triangles, red solid squares and green dots denote BSs, multicast D2D transmitters and receivers, respectively.

PPP Φ_{m} [16]. Figure 2.1 illustrates a snapshot of the spatial distribution of network nodes under the above assumptions. Throughout this chapter, the parameters used in plotting numerical results or simulations are summarized in Table 2.1 unless otherwise specified.

2.3.2 Multicast Transmission

Each D2D transmitter x_i has a common message for all the intended receivers in Φ_{m,x_i}; the message can be sent for $\tau_m \in \mathbb{N}$ times, where τ_m is a pre-configured system parameter. Compared to one shot transmission, sending the message $\tau_m > 1$ times enables more intended receivers to successfully decode

BS Density λ_b	$(\pi 500^2)^{-1}$ m^{-2}
D2D Tx Density λ_m	$5 \times (\pi 500^2)^{-1}$ m^{-2}
D2D Rx Density λ	$500 \times (\pi 500^2)^{-1}$ m^{-2}
Path loss Exponent α	3.5
Detection Threshold T	-3 dB
BS Tx Power P_c	40 W
D2D Tx Power P_m	200 mW
Noise PSD	-174 dBm
Noise Figure	9 dB
Channel Bandwidth	10 MHz

Table 2.1: Simulation/Numerical Parameters for Multicast D2D

the message. Further, we assume that multicast transmitters are static during the τ_m transmissions. This fixed-rate repetition multicast scheme will be the focus of this work. The study of the performance of this simple baseline multicast strategy can serve as a benchmark for future work on more efficient D2D multicast schemes, e.g., using advanced coding or adaptive multi-rate multicast.

When D2D UEs are in coverage, the ground cellular network can assist D2D communications. Specifically, each in-coverage multicast D2D transmitter has a serving BS; normally the serving BS is the BS providing the strongest reference signal receiving power (RSRP). In the current set-up, this is equivalent to choosing the nearest BS as the serving BS. We use $z_x \in \Phi_{\mathrm{b}}$ to indicate the nearest BS of D2D transmitter x. Formally, define the Voronoi cell $\mathcal{C}_{z_i}(\Phi_{\mathrm{b}})$ of point z_i with respect to Φ_{b} as

$$\mathcal{C}_{z_i}(\Phi_{\mathrm{b}}) = \{x \in \mathbb{R}^2 : \quad \| x - z_i \| \leq \| x - z_j \|, \ \forall z_j \in \Phi_{\mathrm{b}}\}.$$

Then each BS z can help D2D transmitters located in its Voronoi cell $\mathcal{C}_{z_i}(\Phi_{\mathrm{b}})$ by broadcasting the common messages. Considering the limited time/frequency resource at the BSs, the message of each D2D transmitter x is broadcast by BS z_x at most once.

We assume that D2D is *overlaid* with cellular networks, i.e., D2D transmitters and BSs use orthogonal transmission resources, and thus there is no mutual interference between cellular and D2D transmissions. In addition, we assume the multicast message of each D2D transmitter is known by its serving BS. Note that when cellular network coverage is available, D2D transmissions are under relatively tight network control [88]. So the coordination between cellular and D2D transmissions can be easily achieved by communication through the BS control channels. How the specific coordination is achieved is beyond the scope of this work. One simple coordination strategy may be to multiplex network assistance and D2D multicast transmissions in the time domain: D2D multicast transmissions are scheduled in the first τ_m time slots and the associated BS helps broadcast the message in the (τ_m+1)-th time slot.

2.3.3 Channel Model

Constant transmit powers P_b and P_m are assumed for the BSs and D2D transmitters, respectively. Denote the path loss function as $\ell(r) : \mathbb{R}^+ \mapsto \mathbb{R}^+$, where r denotes the distance; $\ell(r)$ is assumed to be continuous and non-decreasing. When concrete results are desired, we will assume a specific path

loss function $\ell(r) = Ar^{\alpha}$ where $A > 0$ is a constant and $\alpha > 2$ is the path loss exponent.

Focusing on the signal emitted by the typical transmitter x_0 located at the origin, i.e., $x_0 = o$, the received signal $Y_y(n)$ at the receiver $y \in \Phi_{\mathrm{m},o}$ at time n can be written as

$$Y_y(n) = \ell^{-1}(\|y\|)\sqrt{P_m}H_{y,o}(n)X_o + I_y(n) + Z_y(n),$$

where X_x denotes the signal sent by the D2D transmitter x and $\mathbb{E}[\|X_x\|^2] = 1$, $H_{y,x}(n)$ denotes the fading of the link from x to y at time n and is independently distributed as $\mathcal{CN}(0,1)$, $Z_y(n)$ denotes the additive noise at receiver y at time n and is independently distributed as $\mathcal{CN}(0,\sigma^2)$, and $I_y(n)$ denotes the aggregate interference at receiver y at time n and is given by

$$I_y(n) \triangleq \sum_{x \neq o} \ell^{-1}(\|y - x\|)\sqrt{P_m}H_{y,x}X_x.$$

Then the SINR of the link from the typical D2D transmitter $x_0 = o$ to D2D receiver y at time n equals

$$\mathrm{SINR}_{y,x_0}(n) = \frac{F_{y,x_0}(n)/\ell(\|y\|)}{\mathrm{SNR}^{-1} + \sum_{j \neq 0} F_{y,x_j}(n)/\ell(\|x_j - y\|)}, \tag{2.1}$$

where $F_{y,x} = |H_{y,x}|^2 \sim \mathrm{Exp}(1)$, and $\mathrm{SNR}^{-1} = \sigma^2/P_m$.

Similarly, the received downlink signal $Y_y^{(c)}$ at the receiver $y \in \Phi_{\mathrm{m},o}$ can be written as

$$Y_y^{(c)} = \ell^{-1}(\|z_o\|)\sqrt{P_b}H_{y,z_o}X_o + I_y^{(c)} + Z_y,$$

where the aggregate downlink interference

$$I_y^{(c)}(n) \triangleq \sum_{x \neq o} \ell^{-1}(\|y - z_x\|)\sqrt{P_b} H_{y,z_x} X_x.$$

The SINR of the link from the nearest BS z_o of the typical D2D transmitter x_0 to D2D receiver y equals

$$\mathrm{SINR}_{y,z_o}^{(c)} = \frac{F_{y,z_o}/\ell(\|z_o - y\|)}{\mathrm{SNR}_c^{-1} + \sum_{x \neq o} F_{y,z_x}/\ell(\|z_x - y\|)}, \tag{2.2}$$

where $\mathrm{SNR}_c^{-1} = \sigma^2/P_b$.

2.3.4 Performance Metrics

From the perspective of analysis, it suffices to consider the typical multicast cluster with $x_0 = o$ since, as justified by Palm theory [19], its performance indicates the *spatially averaged* performance over all the clusters. Focusing on the typical cluster, we are first interested in the probability that an arbitrary receiver $y \in \Phi_{m,o}$ can decode the multicast message of the typical D2D transmitter x_0; we term this coverage probability. Without network assistance, we say the receiver $y \in \Phi_{m,o}$ is covered if $\exists n \in \{1, 2, ..., \tau_m\}$ such that $\mathrm{SINR}_{y,x_0}(n) \geq T$, where T is the detection threshold of the fixed rate multicast transmission and is normally greater than -6 dB in LTE. Formally, denoting $E_n(y) = \{\mathrm{SINR}_{y,x_o}(n) \geq T\}$, the coverage probability at y without network assistance equals

$$p(y) \triangleq \mathbb{P}^o\left(\cup_{n=1}^{\tau_m} E_n(y)\right), \tag{2.3}$$

where $\mathbb{P}^o(\cdot)$ is the Palm probability associated with the multicast transmitter process Φ_m. For later use, we define $p_n(y) \triangleq \mathbb{P}^o\left(\cap_{m=1}^{n} E_m(y)\right)$.

Similarly, with network assistance, we say the receiver $y \in \Phi_{m,o}$ is covered if either $\exists n \in \{1, 2, ..., \tau_m\}$ such that $\mathrm{SINR}_{y,x_0}(n) \geq T$ or $\mathrm{SINR}_{y,z_o}^{(c)} \geq T$. Formally, denoting $E^{(c)}(y) = \mathrm{SINR}_{y,z_o}^{(c)} \geq T$, the coverage probability at y with network assistance equals

$$\tilde{p}(y) \triangleq \mathbb{P}^o\left(\cup_{n=1}^{\tau_m} E_n(y) \cup E^{(c)}(y)\right). \tag{2.4}$$

While coverage probability characterizes the performance of an individual receiver in the typical cluster, it is also desirable to have a metric to measure the performance of the typical cluster as a whole. Thus, another metric studied in this work is the mean number of covered receivers in the typical cluster. When network assistance is not available, it equals

$$\mathbb{E}^o[N] \triangleq \mathbb{E}^o\left[\sum_{y \in \Phi_{m,o}} \mathbb{I}(\{y \text{ is covered}\})\right], \tag{2.5}$$

where $\mathbb{I}(E)$ is the indicator function which equals 1 if the event E is true and 0 otherwise. We use $\mathbb{E}^o[\tilde{N}]$ to denote the counterpart of $\mathbb{E}^o[N]$ in the case of network assistance.

2.4 Multicast without Network Assistance

In this section we focus on analyzing the multicast performance without network assistance.

2.4.1 Coverage Probability

By the definition of Palm probability, the coverage probability at y without network assistance equals

$$\begin{aligned} p(y) &= \mathbb{E}^o\left[\mathbb{I}(\{\max_{n=1,\ldots,\tau_m} \mathrm{SINR}_{y,x_0}(n) \geq T\})\right] \\ &= \frac{1}{\lambda_m |B|}\mathbb{E}\left[\int_{x\in B} \mathbb{I}(\{\max_{n=1,\ldots,\tau_m} \mathrm{SINR}_{y+x,x}(n) \geq T\})\Phi_m(\mathrm{d}x)\right], \end{aligned}$$

where B is an arbitrarily bounded subset of $\mathbb{R}^2$ and $|B|$ denotes its Lebesgue measure. The last relation clearly demonstrates that the coverage performance of the typical cluster indicates the average coverage performance over the clusters. The coverage probability $p(y)$ is explicitly given in Theorem 2.1.

Theorem 2.1. *The probability that the receiver $y \in \Phi_{m,0}$ is covered by the typical multicast transmitter $x_0 \in \Phi_m$ with τ_m repetitive transmissions is given by*

$$p(y) = \sum_{n=1}^{\tau_m} (-1)^{n+1} \binom{\tau_m}{n} e^{-n\ell(\|y\|)T\cdot SNR^{-1} - 2\pi\lambda_m \int_0^\infty (1-(1+T\ell(\|y\|)/\ell(r))^{-n})r \ \mathrm{d}r}. \quad (2.6)$$

Proof. See Appendix 2.8.1. □

Note that, when τ_m is large, exact calculation of $p(y)$ based on (2.6) may be cumbersome. Instead, one may consider the following bounds of $p(y)$ which follow from Bonferroni inequalities [38]:

$$p^{(k+1)}(y) \leq p(y) \leq p^{(k)}(y),$$

where k is any odd number in $\{1, ..., \tau_m\}$ and $p^{(k)}(y)$ equals the first k summands of the τ_m summands in (2.6). By definition, $p(y) = p^{(\tau_m)}(y)$. In general, one gets tighter bounds by making k larger; $p_k(y)$ reduces to the union bound when $k = 1$.

Based on Theorem 2.1, more specific results can be obtained by plugging explicit path loss functions $\ell(r)$ in (2.6). For example, for the commonly used path loss function $\ell(r) = Ar^{\alpha}$, the following result immediately follows from Theorem 2.1.

Corollary 2.1. *With* $\ell(r) = Ar^{\alpha}$,

$$p(y) = \sum_{n=1}^{\tau_m} (-1)^{n+1} \binom{\tau_m}{n} e^{-nT \cdot SNR^{-1} A \|y\|^{\alpha}} e^{-\lambda_m K(\alpha,n) T^{\frac{2}{\alpha}} \|y\|^2}, \tag{2.7}$$

where

$$K(\alpha, n) = \frac{2\pi}{\alpha} \int_0^{\infty} t^{-\frac{2}{\alpha}-1} \left(1 - \frac{1}{(1+t)^n}\right) \, \mathrm{d}t. \tag{2.8}$$

Figure 2.2 shows the coverage probability as a function of detection threshold. As expected, the farther the potential receiver away from the multicast transmitter, the smaller the coverage probability is. Further, repetitive transmissions are instrumental in improving the coverage probability, especially for far away receivers. But the gain diminishes as τ_m increases.

2.4.2 Spatial Correlation in Multicast D2D Transmissions

Theorem 2.1 and Corollary 2.1 characterize the coverage probability at a particular receiver, which can be treated as first order coverage performance.

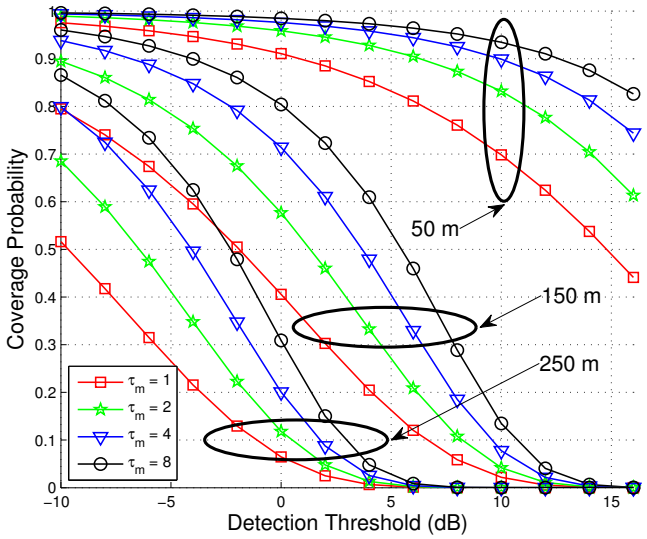

Figure 2.2: Coverage probability versus detection threshold without network assistance: The numbers, 50 m, 150 m, 250 m, indicate three different D2D Tx-Rx distances.

As highlighted in [44, 86], there exist temporal and spatial correlations in the performance at different nodes in a wireless network. Thus, it is of interest to study how the coverage probabilities of different receivers interact, i.e., higher order coverage performance. Intuitively, if some receiver is covered, we may infer that other receivers close to the receiver are also likely to be covered.

Towards a formal understanding of the spatial correlation in D2D multicast transmissions, we define the typical covered receiver process

$$\tilde{\Phi}_{m,o} = \sum_{y \in \Phi_{m,o}} e_y \mathbb{I}(\{y \text{ is covered}\}).$$

Obviously, $\tilde{\Phi}_{m,o}$ is a thinning process "thinned" from the PPP $\Phi_{m,o}$. However, the thinning operations are *not* independent across the points in $\Phi_{m,o}$

because they are correlated through the multicast transmitter process Φ_m, i.e., due to the presence of common randomness in the locations of the multicast transmitters. This dependent thinning makes the thinning process $\Phi_{m,o}$ no longer a PPP. More specifically, let us examine the conditional probability of the event $\{\min_{k=1,\ldots,n} \mathrm{SINR}_{y_1,x_0}(k) \geq T\}$ conditional on the event $\{\min_{k=1,\ldots,n} \mathrm{SINR}_{y_2,x_0}(k) \geq T\}$, i.e.,

$$p_n(y_1|y_2) \triangleq \mathbb{P}^o(\cap_{m=1}^n E_m(y_1)| \cap_{m=1}^n E_m(y_2)). \tag{2.9}$$

Note that this is *not* the conditional coverage probability $p(y_1|y_2)$ which denotes the probability that y_1 is covered conditional on that y_2 is covered. Once we evaluate $p_n(y_1|y_2)$, $p(y_1|y_2)$ can be readily obtained using the inclusion-exclusion principle, as in the proof of Theorem 2.1.

We calculate $p_n(y_1, y_2)$ in two steps: first evaluate $p_n(y_1, y_2|\Phi_m)$ conditioned on Φ_m, and then de-condition on Φ_m to obtain $p_n(y_1, y_2)$. Following similar arguments in the proof of Theorem 2.1, the conditional $p_n(y_1, y_2|\Phi_m)$ can be calculated as

$$p_n(y_1, y_2|\Phi_m) = e^{-n(\ell(\|y_1\|)+\ell(\|y_2\|))T\cdot\mathrm{SNR}^{-1}} \prod_{i=1}^{2}\prod_{j} \frac{1}{(1+\ell(\|y_i\|)T/\ell(\|x_j-y_i\|))^n}.$$

Now de-conditioning on Φ_m yields

$$\begin{aligned} p_n(y_1, y_2) =& \mathbb{E}_{\Phi_m}[p_n(y_1, y_2|\Phi_m)] = e^{-n(\ell(\|y_1\|)+\ell(\|y_2\|))T\cdot\mathrm{SNR}^{-1}} \\ & \times \exp\left(-\lambda_m \int_{\mathbb{R}^2} 1 - \prod_{i=1}^{2} \frac{1}{(1+\ell(\|y_i\|)T/\ell(\|x-y_i\|))^n}\, \mathrm{d}x\right). \end{aligned}$$

Using $p_n(y_1, y_2)$, the conditional probability $p_n(y_1|y_2)$ can be calculated by $p_n(y_1|y_2) = \frac{p_n(y_1,y_2)}{p_n(y_2)}$.

To obtain some insight, let us examine the following expression:

$$\frac{p_n(y_1|y_2)}{p_n(y_1)} = e^{\lambda_m \int_0^\infty \prod_{i=1}^2 \left(1-(1+\ell(\|y_i\|)T/\ell(\|x-y_i\|))^{-n}\right)\, \mathrm{d}x} > 1,$$

which agrees with intuition: Given the event $\{\min_{k=1,\ldots,n} \mathrm{SINR}_{y_2,x_0}(k) \geq T\}$, there is a higher probability that the event $\{\min_{k=1,\ldots,n} \mathrm{SINR}_{y_1,x_0}(k) \geq T\}$ would happen. The following more specific remarks are in order:

- The correlation becomes weaker as λ_m decreases; in particular, when λ_m is asymptotically small, the correlation may be ignored.
- The correlation becomes weaker as n decreases (which leads to reduced temporal correlation).
- The correlation becomes stronger when $\|y_1 - y_2\|$ decreases. In particular, $\lim_{y_2 \to y_1} \frac{p_n(y_1|y_2)}{p_n(y_1)} = \frac{1}{p_n(y_1)}$.
- The correlation becomes stronger as T increases. This is intuitive because with higher T a larger number of interfering nodes come into play. In contrast, when T is small, the outage events at y_1 and y_2 are respectively dominated by a few nearby interferers around them, and the intersection of the two sets of the nearby interferers can be quite small, leading to weak spatial correlation.

2.4.3 Mean Number of Covered Receivers

In this subsection we study the mean number of covered receivers in the typical cluster. For concreteness, we focus on the path loss function $\ell(r) = Ar^\alpha$ in the sequel.

Proposition 2.1. *With $\ell(r) = Ar^\alpha$, the mean number of covered receivers in the typical cluster is given by*

$$\mathbb{E}^o[N] = 2\pi\lambda_r \sum_{n=1}^{\tau_m} (-1)^{n+1} \binom{\tau_m}{n} \int_0^R r e^{-nT \cdot SNR^{-1} A r^\alpha} e^{-\lambda_m K(\alpha,n) T^{\frac{2}{\alpha}} r^2} \, \mathrm{d}r. \quad (2.10)$$

Proof. See Appendix 2.8.2. □

To gain insight from Prop. 2.1, we next focus on a few special cases and/or asymptotic results which have simpler expressions.

No noise. In this case we assume that interference is a dominant issue and thus noise is ignored, i.e., $\sigma^2 \equiv 0$. Then the following corollary follows from Prop. 2.1.

Corollary 2.2. *With $\sigma^2 \equiv 0$ and $\ell(r) = Ar^\alpha$, the mean number of covered receivers in the typical multicast cluster is given by*

$$\mathbb{E}^o[N] = \frac{\pi\lambda_r}{T^{\frac{2}{\alpha}}\lambda_m} \sum_{n=1}^{\tau_m} (-1)^{n+1} \binom{\tau_m}{n} \frac{1}{K(\alpha,n)} \left(1 - e^{-\lambda_m K(\alpha,n) T^{\frac{2}{\alpha}} R^2}\right).$$

In particular, as $\lambda_m \to \infty$, $\mathbb{E}^o[N] \sim \frac{\pi \tilde{K}(\alpha,\tau_m)\lambda_r}{T^{\frac{2}{\alpha}}\lambda_m}$, where

$$\tilde{K}(\alpha,\tau_m) = \sum_{n=1}^{\tau_m} (-1)^{n+1} \binom{\tau_m}{n} \frac{1}{K(\alpha,n)}.$$

It follows from Corollary 2.2 that, when decoding threshold T or cluster size R is small,[1]

$$\mathbb{E}^o[N] \sim \lambda_r \pi R^2 \sum_{n=1}^{\tau_m} (-1)^{n+1} \binom{\tau_m}{n} = \lambda_r \pi R^2,$$

[1]Here we do not consider the case that λ_m is small; small λ_m makes the assumption that the network is interference-limited invalid.

i.e., all the receivers in the typical cluster can be covered in an expectation sense, agreeing with intuition. Further, $\mathbb{E}^o[N]$ is independent of λ_m and τ_m. The last fact implies that a single multicast transmission is optimal when T or R is small enough.

In the extreme case with $\lambda_m \to \infty$, $\mathbb{E}^o[N]$ is inversely proportional to the multicast transmitter density λ_m. Note that the number of repetitions τ_m does not change the scaling law of $\mathbb{E}^o[N]$ (with respect to λ_m); instead, τ_m affects $\mathbb{E}^o[N]$ only up to the multiplicative factor $\tilde{K}(\alpha, n)$.

Pathloss exponent $\alpha = 4$. In this case $\mathbb{E}^o[N]$ in (2.10) reduces to the following:

$$\frac{\pi^{\frac{3}{2}}\lambda_r}{\sqrt{C_1}} e^{\frac{C_2^2}{4C_1}} \sum_{n=1}^{\tau_m} (-1)^{n+1} \binom{\tau_m}{n} \left(Q\left(\frac{C_2}{\sqrt{2C_1}}\right) - Q\left(\sqrt{2C_1}R^2 + \frac{C_2}{\sqrt{2C_1}}\right)\right)$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2}\, \mathrm{d}t$, $C_1 = AT \cdot \mathrm{SNR}^{-1}$ and $C_2 = \lambda_m K(\alpha, n) T^{\frac{2}{\alpha}}$. This gives a quasi-closed form expression for $\mathbb{E}^o[N]$ as $Q(x)$ can be numerically evaluated quite easily.

λ_m is asymptotically small. In this case, using bounded convergence theorem and binomial theorem,

$$\lim_{\lambda_m \to 0} \mathbb{E}^o[N] = \pi\lambda_r \int_0^{R^2} \left(1 - (1 - e^{-T \cdot \mathrm{SNR}^{-1} A t^{\frac{\alpha}{2}}})^{\tau_m}\right)\, \mathrm{d}t.$$

As τ_m increases, the above integrand converges to 1 at a geometric rate and thus the mean number of covered receivers approaches to $\lambda_r \pi R^2$ very quickly. This fact implies that a very small number of repetition transmissions suffices in sparse networks.

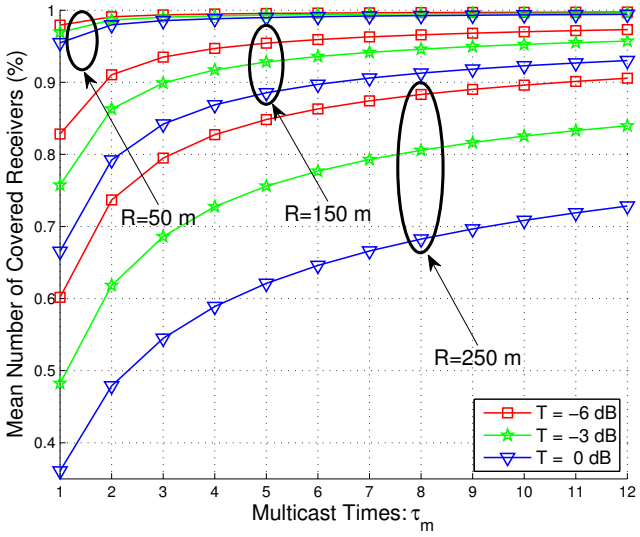

Figure 2.3: Normalized mean number of covered receivers versus multicast times without network assistance.

Figure 2.3 shows the mean number of covered receivers (normalized by all the potential receivers) as a function of multicast times. Again, repetitive transmissions are instrumental but the gain quickly diminishes as τ_m increases. This implies that if a D2D transmitter would like to cover far away receivers, other approaches rather than simple repetitive transmissions are expected; such approaches may include increasing transmit power and interference cancellation.

Thus far we have characterized the mean number of covered receivers in the typical cluster. Other properties may be studied with further assumption on the receiver point processes $\{\Phi_{\mathrm{m},x_i}\}$. For concreteness, assume $\{\Phi_{\mathrm{m},x_i}\}$ are Poisson distributed. Then, due to randomness, all the receivers in some

clusters may be far away from the multicast transmitter; in an extreme case, there may be no receiver at all in some clusters. We term them *null receiver clusters*. It is of interest to quantify the fraction of null receiver clusters. To this end, we first formalize the concept of null receiver cluster.

Definition 2.1. *A multicast cluster is called null receiver cluster if all the receivers have a distance farther than a pre-defined threshold distance R_{th} to the transmitter.*

A possible criterion for threshold distance R_{th} may be as follows.

$$\frac{P_m \ell(r)}{\sigma^2} \geq T, \quad \forall r \leq R_{th}.$$

This criterion implies that, without considering interference and fading, a receiver cannot be covered if its distance from the transmitter is farther than R_{th} due to the weak signal. It follows that $R_{th} = \ell^{(-1)}(\mathrm{SNR}^{-1}T)$ where $\ell^{(-1)}(\cdot)$ denotes the inverse function of $\ell(\cdot)$.

Proposition 2.2. *The fraction of null receiver clusters is $e^{-\lambda_r \pi (\min(R_{th},R))^2}$.*

Proof. See Appendix 2.8.3. □

Note that, conditioning on Φ_m, if $\{\Phi_{\mathrm{m},x_i}\}$ are i.i.d sampled over time, the fraction of null receiver clusters can also be interpreted as the fraction of time that an arbitrary cluster is a null receiver cluster.

2.4.4 Multicast Throughput

Repetition transmission helps improve multicast reliability with increased coverage probability and number of covered receivers. However, repetition consumes more degrees of freedom and thus hurts the throughput. In other words, there exists a fundamental tradeoff between efficiency and reliability. In this section, we explore multicast efficiency. To this end, we define *multicast throughput* (denoted by ξ) as follows.

Definition 2.2. *Multicast throughput is defined as the mean of the sum rate of all the receivers in the typical multicast cluster. Mathematically,*

$$\xi = \mathbb{E}^o[N] \cdot \frac{1}{\tau_m} \log(1+T). \tag{2.11}$$

Multicast throughput may serve as a sensible objective for choosing appropriate multicast rate, i.e., T. On the one hand, with higher T, more sophisticated modulation and coding scheme can be supported and thus higher rate may be achieved. On the other hand, higher T reduces the number of receivers that can be covered by the multicast transmitter. The definition of multicast throughput takes both factors into account by combining $\log(1+T)$ and $\mathbb{E}^o[N]$. So we may optimize multicast rate by maximizing the multicast throughput:

$$\text{maximize}_{T>0} \quad \mathbb{E}^o[N] \cdot \frac{1}{\tau_m} \log(1+T), \tag{2.12}$$

where $\mathbb{E}^o[N]$ is explicitly given in Prop. 2.1. The above optimization is of single variable and thus can be solved efficiently. More explicit results may be

obtained under special cases; for example, the following proposition considers the case with noise ignored and $\lambda_m \to \infty$.

Proposition 2.3. *With $\sigma^2 \equiv 0$ and $\lambda_m \to \infty$, multicast rate optimization (2.12) reduces to*

$$\text{maximize}_{T>0} \quad T^{-\frac{2}{\alpha}} \log(1+T). \tag{2.13}$$

Further, it has a unique optimal point $T^\star > \frac{\alpha}{2} - 1$ that equals the unique solution of the equation: $\frac{x}{1+x} = \frac{2}{\alpha}\log(1+x)$.

Proof. See Appendix 2.8.4. □

To gain some intuition, we show multicast throughput as a function of T in Figure 2.4. It is shown that the optimal rate $T^\star$ is relatively robust to τ_m; for example, with $\alpha = 3.5$, optimal $T^\star$ is around 7 dB for either $\tau_m = 1$ or $\tau_m = 4$. It is also shown that higher multicast throughput is obtained with median path loss exponent, agreeing with intuition: High path loss exponent provides better spatial separation in terms of interference but also leads to high loss of signal power; whereas the converse is true with low path loss exponent.

In (2.11), as τ_m increases, $\mathbb{E}^o[N]$ increases but $\frac{1}{\tau_m}\log(1+T)$ decreases. As the latter typically dominates the former, the defined multicast throughput ξ decreases with τ_m. We illustrate the tradeoff between efficiency – multicast throughput ξ – and reliability – mean number of covered receivers $\mathbb{E}^o[N]$ in Figure 2.5. How to strike a balance between efficiency and reliability depends on the application scenarios. Nevertheless, the bottom line may be that reliability

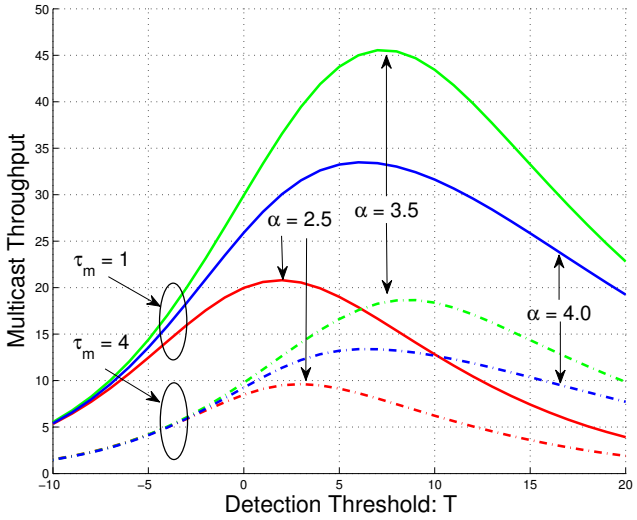

Figure 2.4: Multicast throughput versus detection threshold: $R = 150$ m.

should not be stressed to an extent such that multicast loses its superiority over unicast. For example, ignoring overhead issues, reasonable choice of τ_m should satisfy the follow relation:

$$\xi \geq \mathbb{E}^o \left[\frac{\tau_m}{|\Phi_{m,x_0}|} \sum_{y \in \Phi_{m,x_0}} p(y) \log(1+T) \right],$$

where the right hand side denotes the achievable sum rate if the typical transmitter unicasts to each receiver separately.

2.4.5 Impact of Dynamics

Recall that in our default model multicast transmitters are static, i.e. their positions are fixed and they keep active. Correspondingly, in the previous analysis on coverage probability and mean number of covered receivers, we

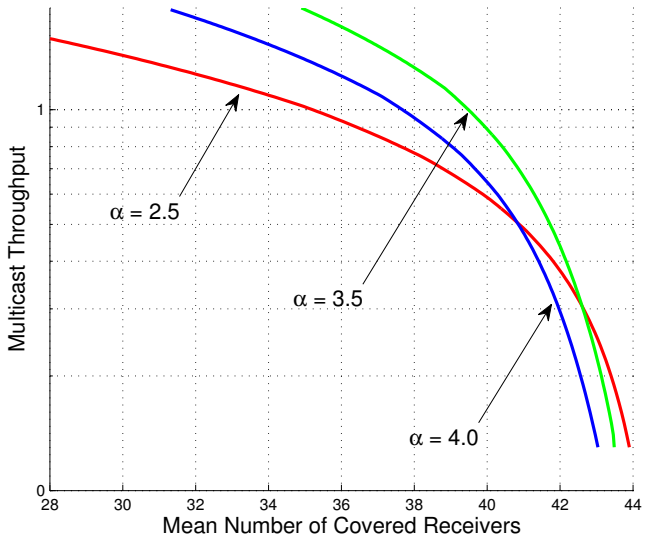

Figure 2.5: Tradeoff between efficiency – multicast throughput – and reliability – mean number of covered receivers.

first perform time average by fixing the spatial realization of Φ_m; then we decondition on Φ_m to average out the spatial randomness. A natural question arises: What is the impact of *dynamics*? To answer this question, we assume in this section that the multicast transmitter process Φ_m is independently resampled at each time slot during the multicast transmissions, i.e., $\{\Phi_m(n)\}$ are independent PPPs. This assumption may model the following scenarios: (i) transmissions are bursty and each transmitter is randomly on/off in each time time slot; (ii) transmitters are of high mobility.

Surprisingly, based on the results for static scenario, the characterization of the performance of dynamic case is quite clean, as given in the following Prop. 2.4.

Proposition 2.4. *With dynamics and path loss function $\ell(r) = Ar^{\alpha}$, the coverage probability $p(y)$ and mean number of covered receivers $\mathbb{E}^o[N]$ are respectively given by (2.7) and (2.10) but with $K(\alpha, n)$ replaced by $nK(\alpha, 1)$.*

Proof. See Appendix 2.8.5. □

To get some insight about how dynamics affect multicast efficiency, let us recall in the static case $\log 1/p_n(y)$ is proportional to $K(\alpha, n)$; in the dynamic case $\log 1/p_n(y)$ is proportional to $nK(\alpha, 1)$. The following Lemma 2.1 shows that $nK(\alpha, 1)$ is greater than $K(\alpha, n)$ except the trivial case $n = 1$. It follows that $p_n(y), n > 1$, in the static case is larger than its counterpart in the dynamic case. In other words, dynamics *hurt* the performance if one would like to support a target SINR for n successive transmissions, agreeing with intuition: Dynamics bring in extra randomness to the received SINR and thus make it harder to successively meet the target SINR.

Lemma 2.1. *For any integer $n > 1$, $nK(\alpha, 1) - K(\alpha, n) > 0$.*

Proof. See Appendix 2.8.6. □

The impact of dynamics on $p(y)$ or $\mathbb{E}^o[N]$ is subtler. Figure 2.6 compares the mean number of covered receivers in static network (i.e., our default model) to that of dynamic network. Interestingly, it shows that dynamics can increase the mean number of covered receivers. Further, the loss due to the static environment can be hardly overcome by increasing the number of retransmissions (at least at the time scale of τ_m). This is because the signal and

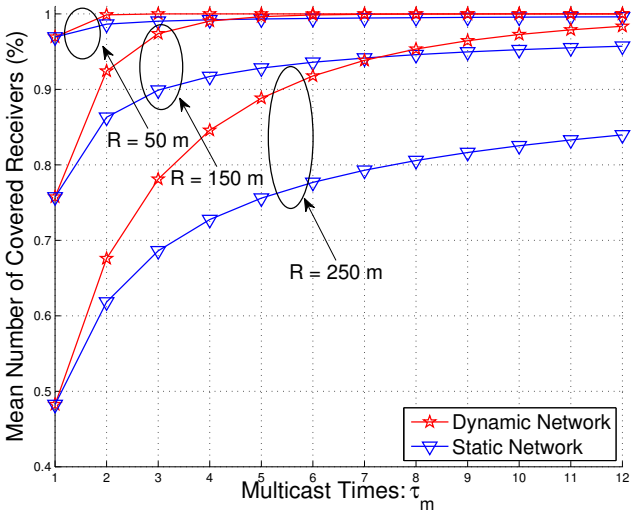

Figure 2.6: Dynamics increase the mean number of covered receivers.

interference powers largely depend on the node locations; multiple transmissions may exploit the fast fading but cannot fundamentally change the signal and interference powers.

The static and dynamic cases considered in this section represent two extremes. The static and high mobility cases represent two extreme mobility patterns; the real mobility pattern lies somewhere in between [90]. The performance of intermediate cases may be obtained by an appropriate combination of the performances of the two extremes. Specifically, we may assume each multicast transmitter is in static status with probability p; otherwise, it is in dynamic status. The statuses are independent across multicast transmitters. Then the performance of a typical static multicast cluster is determined by two independent PPPs: static PPP (with intensity $p\lambda_m$) and dynamic PPP

(with intensity $(1-p)\lambda_m$). The performance of the typical static multicast cluster can be straightforwardly analyzed following our previous analysis on the extreme cases. So is the performance of a typical dynamic multicast cluster. The overall performance then can be obtained by a linear combination of the performances of the typical static and dynamic clusters.

2.5 Multicast with Network Assistance

In this section we analyze the multicast performance by incorporating network assistance, i.e., allowing the network to relay the multicast signals. Recall that $z_{x_0} = z_o$ denotes the BS that is closest to the typical multicast transmitter. We first study the probability that the receiver located at $y \in \Phi_{\mathrm{m},x_0}$ is covered by the BS z_o in the following Lemma 2.2.

Lemma 2.2. *The probability that the receiver located at $y \in \Phi_{m,x_0}$ is covered by the BS z_o is given by*

$$p_c(y) = \int_{\mathbb{R}^2} p_c(y|x) \cdot \lambda_b e^{-\lambda_b \pi \|x\|^2} \, \mathrm{d}x, \tag{2.14}$$

where

$$\begin{aligned} p_c(y|x) =& e^{-T\ell(\|x-y\|)SNR_c^{-1}} \times \\ & \exp\left(- \int_{B^c(0,\|x\|)} (1 - \frac{1}{1 + T\ell(\|x-y\|)/\ell(\|z-y\|)}) \lambda_b \, \mathrm{d}z \right). \end{aligned} \tag{2.15}$$

The proof of Lemma 2.2 follows from [10] and is omitted for brevity. It is noticed that the domains of integrations in Lemma 2.2 are hard to manipulate

to get more explicit results. To overcome this inconvenience, we shall adopt the following approximation:

$$\|z - y\| \approx \|z - x_0\|, \forall z \in \Phi_b. \tag{2.16}$$

The above approximation may be justified when the multicast regions are small compared to the coverage area of each BS. With the above approximation, the following Corollary 2.3 can be obtained.

Corollary 2.3. *With the approximation (2.16) and* $\ell(r) = Ar^\alpha$, $p_c(y) \approx p_c, \forall y \in \Phi_{m,x_0}$, *where*

$$p_c \triangleq 2\pi\lambda_b \int_0^\infty e^{-T\cdot SNR_c^{-1} Ar^\alpha} e^{-2\pi\lambda_b H(T,\alpha) r^2} e^{-\lambda_b \pi r^2} r \, \mathrm{d}r, \tag{2.17}$$

and $H(T,\alpha) = \int_1^\infty \frac{x}{1+x^\alpha/T} \, \mathrm{d}x$. *In particular, when* $\alpha = 4$,

$$p_c = \frac{\pi^{\frac{3}{2}}\lambda_b}{2\sqrt{C_3}} e^{\frac{C_4^2}{4C_3}} Q\left(\frac{C_4}{\sqrt{2C_3}}\right), \tag{2.18}$$

where $C_3 = AT \cdot SNR_c^{-1}$ *and* $C_4 = 2\pi\lambda_b H(T,\alpha) + \pi\lambda_b$; *when there is no noise,* $p_c = \frac{1}{1+2H(T,\alpha)}$.

For simplicity we will use equality instead of an approximation in the sequel. With network assistance, the probability that the receiver $y \in \Phi_{\mathrm{m},x_0}$ is covered as long as either the BS z_o or the multicast transmitter x_0 covers it. Further, these two events are independent. It follows that the coverage probability at $y \in \Phi_{\mathrm{m},x_0}$ with network assistance equals

$$\tilde{p}_0(y) = 1 - (1 - p_c)(1 - p_0(y)).$$

Rearranging the above equality yields the following result.

Proposition 2.5. *With network assistance, the coverage probability of the receiver $y \in \Phi_{m,0}$ equals*

$$\tilde{p}(y) = p(y) + p_c(1 - p(y)), \tag{2.19}$$

where $p(y)$ and p_c are given in Theorem 2.1 and Corollary 2.3, respectively. Accordingly, the mean number of covered receivers equals

$$\mathbb{E}^o[\tilde{N}] = \mathbb{E}^o[N] + p_c(\lambda_r \pi R^2 - \mathbb{E}^o[N]). \tag{2.20}$$

Prop. 2.5 shows that the network assistance is most useful when $\lambda_r \pi R^2 - \mathbb{E}^o[N] \geq 0$ is large. In particular, with moderate to large detection threshold T and cluster range R, network assistance can significantly reduce the number τ_m of transmissions to achieve the same mean number of covered receivers in the absence of network assistance. Figure 2.7 shows the mean number of covered receivers with network assistance as a function of multicast times. As expected, network assistance is very useful; the gain is particularly pronounced in the case of large multicast radius. In addition, Figure 2.7 shows that the analytical results match the empirical results fairly well; in particular, the approximation (2.16) used in the case of network assistance analysis does not lead to noticeable loss of accuracy, at least from the perspective of mean number of covered receivers.

Note that conditioned on $\|z_o\| = r$, $\mathbb{E}^o[\tilde{N}|\|z_0\| = r]$ equals

$$\mathbb{E}^o[N] + e^{-T \cdot \mathrm{SNR}_c^{-1} A r^\alpha - 2\pi\lambda_b H(T,\alpha) r^2}(\lambda_r \pi R^2 - \mathbb{E}^o[N]),$$

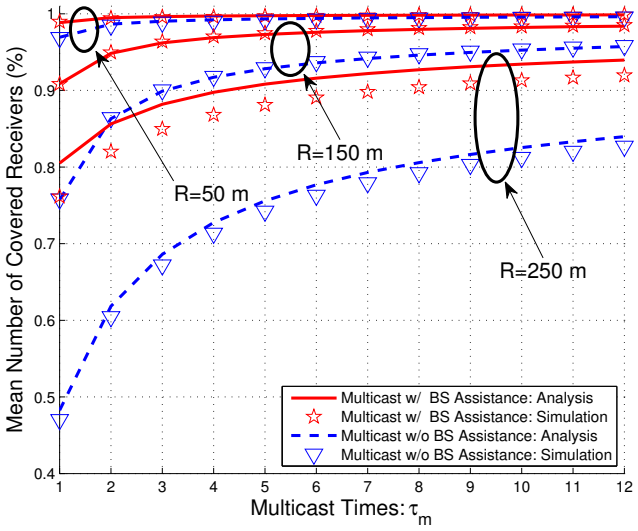

Figure 2.7: Network assistance helps increase the mean number of covered receivers.

from which it is clear that the network assistance is most useful when the distance from the multicast transmitter to its nearest BS is not large. How to optimize this network assistance is the subject of the next section.

Thus far we have seen that both dynamics and network assistance help increase the mean number of covered receivers. Figure 2.8 demonstrates the combined effects of dynamics and network assistance; not surprisingly, their gains accumulate when both dynamics and network assistance are available. An interesting observation from Figure 2.8 is that as τ_m increases, the gain from dynamics can achieve or even exceed the gain from network assistance.

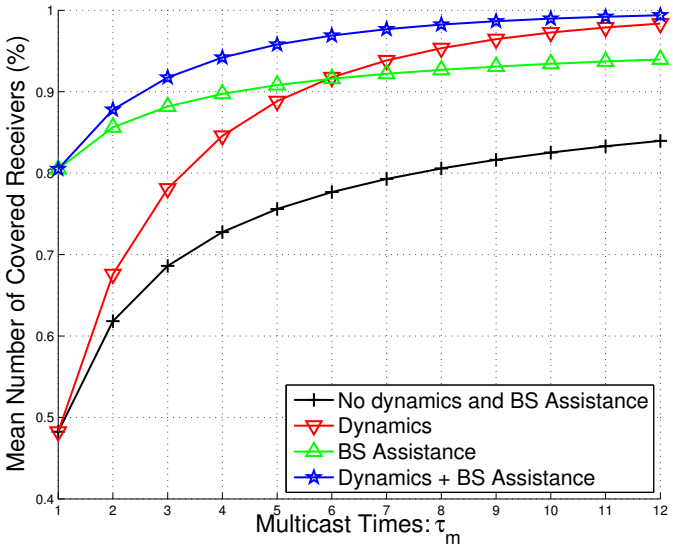

Figure 2.8: Impact of network assistance vs. dynamics: $R = 250$ m.

2.6 Optimizing Multicast Transmissions

In this section we aim to optimize the network assisted multicast transmissions. The overall objective is to seek for optimum network assistance rule to minimize retransmission times while certain network constraints can be satisfied.

For each BS z, let $g_z : \mathbb{R}^+ \to \{0, 1\}$ be a mapping such that $g_z(\|x - z\|) = 1$ if BS z helps D2D transmitter x located in its cell, i.e., $x \in \mathcal{C}_z(\Phi_b)$. As the transmission resources of BSs are limited, we assume each BS z can help at most B multicast sessions in its cell. Mathematically, for $\forall z \in \Phi_b$,

$$\sum_{x \in \Phi_m \cap \mathcal{C}_z(\Phi_b)} g_z(\|x - z\|) \leq B. \tag{2.21}$$

From the spatial average perspective, the following constraint is imposed at

the typical cell.

$$\mathbb{E}^o\left[\sum_{x\in\Phi_m\cap\mathcal{C}_o(\Phi_b)} g_o(\|x\|)\right] \leq B. \tag{2.22}$$

In this section the Palm probability is defined with respect to the BS point process Φ_b instead of D2D transmitter point process Φ_m; the two Palm distributions may be connected with Neveu exchange formula [16]. By definition, $g_o(\cdot) \in \{0,1\}$. However, under the Palm measure, the performance seen by the typical BS is a spatial average; thus with a slight abuse of notation, we allow $g_o(\cdot) \in [0,1]$. In the sequel, we shall refer to (2.21) (resp. (2.22)) as *resource* constraint.

Further, we require that a certain fraction η of the intended receivers associated with the D2D transmitters in each cell should be covered. Mathematically, using Corollary 2.3 and Prop. 2.5, we have the following constraint: $\forall z \in \Phi_b$,

$$\frac{\sum_{x\in\Phi_m\cap\mathcal{C}_z(\Phi_b)} \mathbb{E}[\sum_{y\in\Phi_{m,x}} \mathbb{I}(\{y \text{ is covered}\})]}{\sum_{x\in\Phi_m\cap\mathcal{C}_z(\Phi_b)} \mathbb{E}\left[|\Phi_{m,x}|\right]} \geq \eta, \tag{2.23}$$

where the numerator implicitly depends on $g_z(\cdot)$. From the spatial average perspective, the following constraint is required for the typical cell with the BS located at the origin.

$$\frac{\mathbb{E}^o\left[\sum_{x\in\Phi_m\cap\mathcal{C}_o(\Phi_b)} \sum_{y\in\Phi_{m,x}} \mathbb{I}(\{y \text{ is covered}\})\right]}{\mathbb{E}^o\left[\sum_{x\in\Phi_m\cap\mathcal{C}_z(\Phi_b)} |\Phi_{m,x}|\right]} \geq \eta. \tag{2.24}$$

In the sequel, we shall refer to (2.23) (resp. (2.24)) as *reliability* constraint. The following Prop. 2.6 gives more explicit expressions for the expectation terms involved in (2.22) and (2.24).

Proposition 2.6. *The three expectation terms in (2.22) and (2.24) are respectively given as follows.*

$$\mathbb{E}^o[\sum_{x\in\Phi_m\cap\mathcal{C}_0(\Phi_b)} g_o(\|x\|)] = \frac{\lambda_m}{\lambda_b}\mathbb{E}_D\left[g_o(D)\right]$$

$$\mathbb{E}^o[\sum_{x\in\Phi_m\cap\mathcal{C}_z(\Phi_b)} |\Phi_{m,x}|] = \frac{\lambda_m}{\lambda_b}\bar{N}_{\max}$$

$$\mathbb{E}^o[\sum_{x\in\Phi_m\cap\mathcal{C}_o(\Phi_b)}\sum_{y\in\Phi_{m,x}} \mathbb{I}(\{y \text{ is covered}\})] = \frac{\lambda_m}{\lambda_b}\mathbb{E}_D\left[h(D;\tau_m,g_o(D))\right],$$

where D is a Rayleigh distributed random variable with pdf

$$f_D(r) = 2\pi\lambda_b r e^{-\lambda_b\pi r^2}, \quad r\geq 0;$$

$\bar{N}_{\max} = \lambda_r\pi R^2$*; and* $h:\mathbb{R}^+\mapsto\mathbb{R}^+$ *is given by*

$$h(r;\tau_m,g_o(r)) = \bar{N}(\tau_m) + g_o(r)\cdot q(r)(\bar{N}_{\max}-\bar{N}(\tau_m)),$$

where $\bar{N}(\tau_m)=\mathbb{E}^o[N(\tau_m)]$ *is given in Prop. 2.1,* $q:\mathbb{R}^+\mapsto[0,1]$ *is defined as* $q(r) = e^{-TSNR_c^{-1}Ar^\alpha - 2\pi\lambda_b H(T,\alpha)r^2}$.

Proof. See Appendix 2.8.7. □

Using Prop. 2.6, we can cast the spatial averaged multicast optimization problem as follows.

$$\begin{aligned} \text{minimize} \quad & \tau_m \\ \text{subject to} \quad & \mathbb{E}_D\left[g_o(D)\right] \leq \frac{\lambda_b}{\lambda_m}B \\ & \mathbb{E}_D\left[h(D;\tau_m,g_o(D))\right] \geq \eta\bar{N}_{\max} \\ & 0\leq g_o(r)\leq 1, \forall r\geq 0. \end{aligned} \tag{2.25}$$

This is a mixed integer nonlinear programming which is in general notoriously hard to solve. Worse still, the design space $g_o(\cdot)$ is of infinite dimension; it is not *a priori* clear at all what kind of mapping $g_o(\cdot)$ we ought to pursue. Furthermore, as $g_o(\cdot)$ represents the optimum network assistance rule averaged across the space, it does not lead to readily implementable solution for the network. For these reasons, we are more interested in the following "online" problem: Given a realization of Φ_b and Φ_m, how should each BS z help the D2D transmitters in its cell while satisfying its resource and reliability constraints? Mathematically, each BS z aims to solve the following problem.

$$\begin{aligned} &\text{minimize} \quad \tau_m \\ &\text{subject to} \sum_{x\in\Phi_m\cap\mathcal{C}_z(\Phi_b)} g_z(\|x-z\|) \le B \\ &\qquad \frac{\sum_{x\in\Phi_m\cap\mathcal{C}_z(\Phi_b)} h\left(\|x-z\|;\tau_m, g_z(\|x-z\|)\right)}{|\Phi_m\cap\mathcal{C}_z(\Phi_b)|\cdot\bar{N}_{\max}} \ge \eta \\ &\qquad g_z(\|x-z\|)\in\{0,1\}, \forall x\in\Phi_m\cap\mathcal{C}_z(\Phi_b). \end{aligned} \tag{2.26}$$

Though the above problem is still an integer programming, the design space $g_z(\cdot)$ is of finite dimension. In particular, we only need to determine finite number of binary variables, $g_z(\|x_i - z\|), i = 1, ..., M_z$, where $M_z = |\Phi_m \cap \mathcal{C}_z(\Phi_b)|$. However, with an exhaustive search the complexity is still exponential in M_z. We next analyze the optimality structure of the problem to design an efficient algorithm. To this end, we first note that there always exists a feasible solution; for example, the solution with $g_z(\|x-z\|) \equiv 0$ but large enough τ_m is a feasible one.

For ease of exposition, relabel the D2D transmitters (located in the cell of BS z) in the order of increasing distance to BS z, i.e., $r_1 \leq ... \leq r_{M_z}$ where $r_i = \|x_i - z\|$, and let $\tau_m^\star$ denote the minimum value that can be obtained in the above problem. Then the following result holds.

Proposition 2.7. *There exists an optimal solution such that* $g_z^\star(\|x_1 - z\|) \geq .... \geq g_z^\star(\|x_{M_z} - z\|)$.

Proof. See Appendix 2.8.8. □

For each possible τ_m, Prop. 2.7 implies that BS z can focus on the $\min(M_z, B)$ nearest D2D transmitters and assists as few of them as possible to save resources. Further, as the mapping $h(\cdot)$ is moronically increasing with τ_m, we then can use a binary search for the minimum $\tau_m^\star$ over $\{1, 2, ..., \tau_{\max}\}$, where $\tau_{\max}$ is a large enough integer such that $\frac{1}{M_z}\sum_{i=1}^{M_z} h\left(r_i; \tau_{\max}, 0\right) \geq \eta \bar{N}_{\max}$. The pseudocode of the proposed algorithm can be found in Algorithm 1. The running time of this algorithm is $O(M_z \log \tau_{\max})$; thus for given M_z and $\tau_{\max}$, the proposed algorithm is efficient. However, we need to find a valid but *a priori* unknown $\tau_{\max}$ for initialization purpose. With reasonable η, $\tau_{\max}$'s are usually not large and we can find one quite efficiently. We simulate the proposed algorithm and present the network assistance statistics in Figure 2.9. The abscissa in Figure 2.9 denotes the distance between D2D transmitter and its nearest BS; given the distance, the associated ordinate value gives the corresponding probability that the D2D transmitter is scheduled by the BS

for assistance. As expected, D2D transmitters that are closer to their nearest BSs have higher chance to get network assistance.

Algorithm 1 Multicast D2D with Network Assistance

The following procedure runs on each BS independently. We focus on a particular BS z.

Require: $r_1 \leq ... \leq r_{M_z}$

Choose $\tau_{\max}$ such that $\frac{1}{M_z}\sum_{i=1}^{M_z} h\left(r_i; \tau_{\max}, 0\right) \geq \eta \bar{N}_{\max}$

$\tau_{\min} = 1$

while $\tau_{\min} < \tau_{\max}$ **do**

  for $i = 1$ to M_z **do**

    $g_z(r_i) \Leftarrow 0$

  end for

  $\tau_m \Leftarrow \lfloor \frac{1}{2}(\tau_{\min} + \tau_{\max}) \rfloor$

  $j \Leftarrow 1$

  while $\frac{1}{M_z}\sum_{i=1}^{M_z} h\left(r_i; \tau_m, g_z(r_i)\right) < \eta \bar{N}_{\max}$ and $j \leq \min(B, M_z)$ **do**

    $g_z(r_i) \Leftarrow 1$

    $j{+}{+}$

  end while

  if $j > \min(B, M_z)$ **then**

    $\tau_{\min} \Leftarrow \tau_m$

  else

    $\tau_{\max} \Leftarrow \tau_m$

  end if

end while

Finally, we comment how to construct a reasonably good solution to the original prohibitively difficult network-wise optimization problem (2.25) by collecting and appropriately averaging the network assistance statistics $\{g_z(r)\}$ across the space as follows. We first simulate a large enough network with area e.g. $B(0, R_n)$ where each BS assists the D2D transmitters in its cell using the proposed algorithm. Then, partition $\mathbb{R}^+$ into I non-overlapping intervals $[r_i, r_{i+1}), i = 0, 1, ..., I-1$ with $r_0 = 0, r_I = \infty$, and $|r_{i+1} - r_i| = \Delta$ for

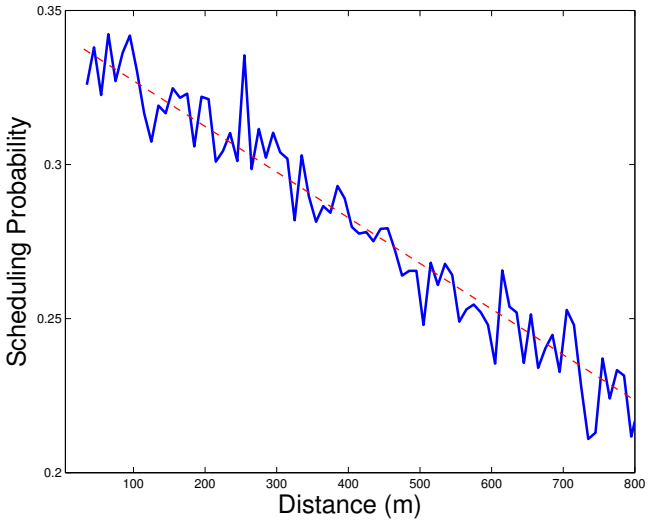

Figure 2.9: Network assistance statistics in the case of optimized multicast transmissions: $\eta = 95\%, B = 2$.

$i = 0, 1, ..., I-2$; and collect the statistics as follows: for $\forall r \in [r_i, r_{i+1})$,

$$\bar{g}_0(r) = \frac{\sum_{z \in B(0,R_n)} \sum_{x \in \Phi_m \cap \mathcal{C}_z(\Phi_b)} 1_{\|x-z\| \in [r_i, r_{i+1})} \cdot g_z(\|x-z\|)}{\sum_{z \in B(0,R_n)} \sum_{x \in \Phi_m \cap \mathcal{C}_z(\Phi_b)} 1_{\|x-z\| \in [r_i, r_{i+1})}}.$$

In this way, we arrive at a piece-wise constant solution $\bar{g}_0(\cdot)$ to the problem (2.25). With $\bar{g}_0(\cdot)$, we can use binary search for the corresponding minimum objective value $\bar{\tau}_m$, which is expected to be approximately equal to the spatial average obtained from simulation, i.e.,

$$\bar{\tau}_m \approx \frac{1}{|\Phi_b(B(0,R_n))|} \sum_{z \in B(0,R_n)} \tau_m^\star(g_z).$$

The above approach leverages the ergodicity of the underlying random processes to relate the statistical average to the spatial average of a sample realization.

2.7 Summary

In this chapter, we propose a tractable analytical model for the analysis and design of multicast D2D. The model has been used to analyze important multicast metrics including the coverage probability, the mean number of covered receivers, and the multicast throughput. We find that retransmissions may increase the coverage probability and mean number of covered receivers but may hurt the throughput. We have also studied how the multicast performance would be affected by certain factors like dynamics and infrastructure assistance. It is found that both may help improve multicast performance significantly. The model and analytical results have been further used to optimize multicasting, i.e., selecting the optimal multicast rate and optimal number of retransmission times. The baseline model developed in this chapter will be adapted or generalized to address other design issues of D2D networking in the following chapters.

2.8 Appendix

2.8.1 Proof of Theorem 2.1

The proof consists of two steps. We first perform time average over fading by fixing the spatial realization of Φ_m; then we de-condition on Φ_m to average out the spatial randomness. This two-step argument used to deal with the temporal correlation of multicast process is motivated by [53] which deals with spatial correlation over multiple receive antennas.

Let $E_I(y) = \bigcap_{n\in I} E_n, I \subset \{1, ..., \tau_m\}$, where we drop the argument of

$E_n(y)$ for notational simplicity. Then conditioned on Φ_m, the probability that $y \in B(0,R)$ is covered by the typical multicast transmitter is given by

$$\begin{aligned} p(y|\Phi_m) &= \mathbb{P}^o\left(\bigcup_{n=1,\ldots,\tau_m} E_n|\Phi_m\right) \\ &= \sum_{n=1}^{\tau_m}(-1)^{n+1}\sum_{I\subset\{1,\ldots,\tau_m\}:|I|=n}\mathbb{P}^o\left(E_I|\Phi_m\right), \end{aligned}$$

where the second equality follows from inclusion-exclusion principle. Note that conditioned on Φ_m, the events E_n, $n = 1, ..., \tau_m$, are independent, because the fading fields are assumed to be independent across both space and time. It follows that $\mathbb{P}^o\left(E_I|\Phi_m\right)$ only depends on the cardinality of I, i.e., $\mathbb{P}^o\left(E_I|\Phi_m\right) \equiv \mathbb{P}^o\left(E_{\{1,\ldots,n\}}|\Phi_m\right), \forall I \subset \{1, ..., \tau_m\}$ with $|I| = n$. Thus, denoting $p_n(y|\Phi_m) = \mathbb{P}^o\left(E_{\{1,\ldots,n\}}|\Phi_m\right)$, $p(y|\Phi_m)$ can be further written as

$$p(y|\Phi_m) = \sum_{n=1}^{\tau_m}(-1)^{n+1}\binom{\tau_m}{n}p_n(y|\Phi_m).$$

Next we focus on computing $p_n(y|\Phi_m)$. Due to the independence of the fading fields across time,

$$\begin{aligned} p_n(y|\Phi_m) &= \mathbb{P}^o(\mathrm{SINR}_{y,0}(k) \geq T, \forall k \in \{1, ..., n\}|\Phi_m) \\ &= \prod_{k=1}^{n}\mathbb{P}^o(\mathrm{SINR}_{y,0}(k) \geq T|\Phi_m), \end{aligned}$$

where $\mathbb{P}^o(\mathrm{SINR}_{y,0}(k) \geq T|\Phi_m)$ equals

$$\begin{aligned} &\mathbb{P}^o\left(\frac{F_{y,x_0}(k)/\ell(\|y\|)}{\mathrm{SNR}^{-1} + \sum_{j\neq i}F_{y,x_j}(k)/\ell(\|x_j - y\|)} \geq T\middle|\Phi_m\right) \\ &= \mathbb{P}^o\left(F_{y,x_0}(k) \geq \ell(\|y\|)T(\mathrm{SNR}^{-1} + \sum_{j\neq 0}F_{y,x_j}(k)/\ell(\|x_j - y\|))\middle|\Phi_m\right) \end{aligned}$$

$$= e^{-\ell(\|y\|)T\cdot \mathrm{SNR}^{-1}} \mathbb{E}_F^o \left[e^{-\ell(\|y\|)T\sum_{j\neq 0} F_{y,x_j}(k)/\ell(\|x_j-y\|)} | \Phi_m \right],$$

where the last equality is due to $F_{y,x_0}(k) \sim \mathrm{Exp}(1)$. Further, by Slyvnyak's theorem [16], the independence of the fading fields across space and the Laplace transform of $F \sim \mathrm{Exp}(1)$ (which is $\mathcal{L}_F(s) = \frac{1}{1+s}$),

$$\begin{aligned} p_n(y|\Phi_m) =& e^{-n\ell(\|y\|)T\cdot \mathrm{SNR}^{-1}} \prod_{k=1}^{n} \mathbb{E}_F \left[e^{-n\ell(\|y\|)T\sum_j F_{y,x_j}(k)/\ell(\|x_j-y\|)} | \Phi_m \right] \\ =& e^{-n\ell(\|y\|)T\cdot \mathrm{SNR}^{-1}} \prod_{j} \prod_{k=1}^{n} \mathbb{E}_F \left[e^{-\ell(\|y\|)TF_{y,x_j}(k)/\ell(\|x_j-y\|)} | \Phi_m \right] \\ =& e^{-n\ell(\|y\|)T\cdot \mathrm{SNR}^{-1}} \prod_{j} \frac{1}{(1+\ell(\|y\|)T/\ell(\|x_j-y\|))^n}. \end{aligned}$$

Now de-conditioning with respect to Φ_m yields

$$\begin{aligned} p_n(y) &= \mathbb{E}_{\Phi_m}[p_0^{(n)}(y|\Phi_m)] \\ &= e^{-n\ell(\|y\|)T\cdot \mathrm{SNR}^{-1}} \mathbb{E}_{\Phi_m} \left[\prod_j \frac{1}{(1+\ell(\|y\|)T/\ell(\|x_j-y\|))^n} \right] \\ &= e^{-n\ell(\|y\|)T\cdot \mathrm{SNR}^{-1}} \mathbb{E}_{\Phi_m} \left[\exp\left(\sum_j \log \frac{1}{(1+\ell(\|y\|)T/\ell(\|x_j-y\|))^n} \right) \right] \\ &= e^{-n\ell(\|y\|)T\cdot \mathrm{SNR}^{-1}} \exp\left(-\lambda_m \int_{\mathbb{R}^2} 1 - \frac{1}{(1+\ell(\|y\|)T/\ell(\|x-y\|))^n} \, \mathrm{d}x \right), \end{aligned}$$

where the last equality follows from the Laplace functional of the PPP Φ_m : $\mathcal{L}_{\Phi_m}(f) = \exp(-\lambda_m \int_{\mathbb{R}^2}(1-e^{-f(x)})\,\mathrm{d}x)$ where $f : \mathbb{R}^2 \to \mathbb{R}^+$ [16]. Further, by the staionarity of the PPP Φ_m and changing Cartesian coordinates to Polar coordinates,

$$p_n(y) = e^{-n\ell(\|y\|)T\cdot \mathrm{SNR}^{-1}} \exp\left(-2\pi\lambda_m \int_0^\infty (1-(1+T\ell(\|y\|)/\ell(r))^{-n}) r \, \mathrm{d}r \right). \tag{2.27}$$

To sum up,

$$p(y) = \mathbb{E}[p(y|\Phi_m)] = \sum_{n=1}^{\tau_m} (-1)^{n+1} \binom{\tau_m}{n} \mathbb{E}[p_n(y|\Phi_m)]$$
$$= \sum_{n=1}^{\tau_m} (-1)^{n+1} \binom{\tau_m}{n} p_n(y).$$

Plugging the explicit expression (2.27) for $p_n(y)$ into the above equality completes the proof.

2.8.2 Proof of Proposition 2.1

Define the typical coverage cell $\mathcal{A}_o$ of the multicast transmitter $x_o \in \Phi_m$ as

$$\mathcal{A}_o = \{y \in B(x_i, R) : \exists n \text{ s.t. } \mathrm{SINR}_{y,o}(n) \geq T\}.$$

We next establish the relation between $\mathbb{E}^o[N]$ and the mean cell volume $\mathbb{E}^o[|\mathcal{A}_o|]$: $\mathbb{E}^o[N] = \lambda_r \mathbb{E}^o[|\mathcal{A}_o|]$. To this end,

$$\mathbb{E}^o[N] = \mathbb{E}^o \left[\sum_{y \in \Phi_{m,o}} \mathbb{I}(\{y \text{ is covered}\}) \right]$$
$$= \mathbb{E}^o \left[\lambda_r \int_{\mathbb{R}^2} \mathbb{I}(\{y \text{ is covered}\}) \, \mathrm{d}y \right]$$
$$= \lambda_r \int_{\mathbb{R}^2} \mathbb{E}^o \left[\mathbb{I}(\{y \text{ is covered}\}) \right] \, \mathrm{d}y$$
$$= \lambda_r \int_{\mathbb{R}^2} p(y) \, \mathrm{d}y,$$

where the second and third equalities follow from Campbell's theorem [16] and Fubini's theorem, respectively. Similarly, we have

$$\mathbb{E}^o[|\mathcal{A}_o|] = \mathbb{E}^o \left[\int_{\mathbb{R}^2} \mathbb{I}(\{y \in \mathcal{C}_0\}) \, \mathrm{d}y \right] = \int_{B(0,R)} p_0(y) \, \mathrm{d}y.$$

It follows that $\mathbb{E}^o[N] = \lambda_r \mathbb{E}^o[|\mathcal{A}_o|]$. The proof will be complete once we compute $\mathbb{E}^o[|\mathcal{A}_0|]$. To this end, using Corollary 2.1, we obtain that

$$\mathbb{E}^o[|\mathcal{A}_0|] = 2\pi \int_0^R \sum_{n=1}^{\tau_m} (-1)^{n+1} \binom{\tau_m}{n} e^{-nT \cdot \mathrm{SNR}^{-1} A r^\alpha - \lambda_m K(\alpha,n) T^{\frac{2}{\alpha}} r^2} r \mathrm{d}r.$$

By Fubini's theorem we can exchange the above summation and integration. Then invoking the established relation $\mathbb{E}^o[N] = \lambda_r \mathbb{E}^o[|\mathcal{A}_o|]$ completes the proof.

2.8.3 Proof of Proposition 2.2

By definition, the fraction of null receiver clusters equals the probability that the typical cluster is null. Denoting this event by E_{null}, we consider the following two cases.

If $R < R_{th}$, then E_{null} is equivalent to the event there exists no point in the typical cluster Φ_{m,x_0}. By Poisson assumption, we have $\Phi_{m,x_0}(B(o,R)) \sim \text{Poisson}(\lambda_r \pi R^2)$. Then,

$$\mathbb{P}^o(E_{null}) = \mathbb{P}^o(\Phi_{m,x_0}(B(o,R)) = 0) = e^{-\lambda_r \pi R^2}.$$

If $R \geq R_{th}$, denoting by $A_n = \{\Phi_{m,x_0}(B(o,R)) = n\}$,

$$\begin{aligned}\mathbb{P}^o(E_{null}) &= \sum_{n=0}^{\infty} \mathbb{P}^o(A_n) \mathbb{P}^o(E_{null}|A_n) \\ &= \sum_{n=0}^{\infty} \frac{(\lambda_r \pi R^2)^n e^{-\lambda_r \pi R^2}}{n!} \left(\frac{\lambda_r \pi R^2 - \lambda_r \pi R_{th}^2}{\lambda_r \pi R^2} \right)^n = e^{-\lambda_r \pi R_{th}^2},\end{aligned}$$

where the second equality follows from that conditioning on $\Phi_{m,x_0}(B(o,R)) = n$ these n points are i.i.d. uniformly distributed in $B(o,R)$. To sum up, we have $\mathbb{P}^o(E_{null}) = e^{-\lambda_r \pi (\min(R_{th},R))^2}$.

2.8.4 Proof of Proposition 2.3

When $\sigma^2 \equiv 0$ and $\lambda_m \to \infty$, by Corollary 2.2

$$\xi = \frac{\pi \tilde{K}(\alpha, \tau_m)\lambda_r}{T^{\frac{2}{\alpha}}\lambda_m} \cdot \frac{1}{\tau_m} \log(1+T).$$

It follows that $\text{maximize}_{T>0}\xi$ is equivalent to $\text{maximize}_{T>0} T^{-\frac{2}{\alpha}} \log(1+T)$. Let $f(x) = x^{-\frac{2}{\alpha}} \log(1+x)$. Direct calculation yields

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \frac{1}{x^{\frac{2}{\alpha}+1}} \left(\frac{x}{1+x} - \frac{2}{\alpha} \log(1+x) \right).$$

Denote by $g(x)$ the term inside the above parentheses. Direct calculation yields $\frac{\mathrm{d}g}{\mathrm{d}x} = \frac{1}{1+x}\left(\frac{1}{1+x} - \frac{2}{\alpha}\right)$. It follows that $\frac{\mathrm{d}g}{\mathrm{d}x} > 0$ when $x \in (0, \frac{\alpha}{2} - 1)$ and $\frac{\mathrm{d}g}{\mathrm{d}x} < 0$ when $x \in (\frac{\alpha}{2} - 1, \infty)$. Correspondingly, $g(x)$ first monotonically increases from 0 to $g(\frac{\alpha}{2} - 1) = 1 - \frac{2}{\alpha}(1 + \log(\frac{\alpha}{2}))$ (which is positive when $\alpha > 2$), and then monotonically decreases from $g(\frac{\alpha}{2} - 1)$ to $-\infty$. Thus, there exists a unique point $x^\star > \frac{\alpha}{2} - 1$ such that $g(x^\star = 0)$, and $f(x)$ monotonically increases when $x \in (0, x^\star)$ and then decreases when $x \in (x^\star, \infty)$. The last fact implies that $x^\star$ is the unique optimal point and this completes the proof.

2.8.5 Proof of Proposition 2.4

The proof is similar to that of Theorem 2.1 except the following arguments:

$$\begin{aligned} p_n(y) &= \mathbb{P}^o\left(\text{SINR}_{y,0}(k) \geq T, \forall k \in \{1, ..., n\}\right) \\ &= \prod_{k=1}^{n} \mathbb{P}^o\left(\text{SINR}_{y,0}(k) \geq T\right), \end{aligned}$$

where the last equality follows from the fact that an independent PPP $\Phi_m(k)$ is drawn at each time slot, and $\mathbb{P}^o(\mathrm{SINR}_{y,0}(k) \geq T)$ equals

$$\begin{aligned}
&\mathbb{P}^o\left(\frac{F_{y,x_0}(k)/\ell(\|y\|)}{\mathrm{SNR}^{-1}+\sum_{x_j\in\Phi_m(k):j\neq 0} F_{y,x_j}(k)/\ell(\|x_j-y\|)} \geq T\right)\\
&= e^{-\ell(\|y\|)T\cdot\mathrm{SNR}^{-1}}\mathbb{E}^o_{\Phi_m(k),F}\left[\exp\left(-\ell(\|y\|)T\sum_{j\in\Phi_m(k)} F_{y,x_j}(k)/\ell(\|x_j-y\|)\right)\right]\\
&= e^{-n\ell(\|y\|)T\cdot\mathrm{SNR}^{-1}}\prod_{k=1}^{n}\exp\left(-\lambda_m\int_{\mathbb{R}^2} 1-\frac{1}{(1+\ell(\|y\|)T/\ell(\|x-y\|))}\,\mathrm{d}x\right)\\
&= e^{-nT\cdot\mathrm{SNR}^{-1}A\|y\|^\alpha}e^{-\lambda_m nK(\alpha,1)T^{\frac{2}{\alpha}}\|y\|^2}.
\end{aligned}$$

Then $p_0(y)$ can be readily obtained by plugging $p_n(y)$ into the equality $p_0(y) = \sum_{n=1}^{\tau_m}(-1)^{n+1}\binom{\tau_m}{n}p_n(y)$. Also, the mean number of covered receivers can be evaluated using the equality $\mathbb{E}^o[N] = \lambda_r\int_{B(0,R)} p_0(y)\,\mathrm{d}y$, which has been established in the proof of Prop. 2.2.

2.8.6 Proof of Lemma 2.1

By definition, we have

$$\begin{aligned}
&nK(\alpha,1)-K(\alpha,n)\\
&= n\frac{2\pi}{\alpha}\int_0^\infty t^{-\frac{2}{\alpha}-1}\left(1-\frac{1}{1+t}\right)\,\mathrm{d}t - \frac{2\pi}{\alpha}\int_0^\infty t^{-\frac{2}{\alpha}-1}\left(1-\frac{1}{(1+t)^n}\right)\,\mathrm{d}t\\
&= \frac{2\pi}{\alpha}\int_0^\infty t^{-\frac{2}{\alpha}-1}\left(n-1-\frac{n}{1+t}+\frac{1}{(1+t)^n}\right)\,\mathrm{d}t.
\end{aligned}$$

Denote by $f(t) = n-1-\frac{n}{1+t}+\frac{1}{(1+t)^n}, t\geq 0$. Note $f'(t) = \frac{(1+t)^{n-1}-1}{(1+t)^{n+1}} \geq 0, \forall t\geq 0$. It follows that $f(t)$ is monotonically increasing on $[0,\infty]$ and $f(t)\geq f(0) = 0$. Thus $nK(\alpha,1)-K(\alpha,n) = \frac{2\pi}{\alpha}\int_0^\infty t^{-\frac{2}{\alpha}-1}f(t)\,\mathrm{d}t \geq 0$ as the integrand is non-negative.

2.8.7 Proof of Proposition 2.6

Under the Palm measure (with respect to Φ_b),

$$\begin{aligned}
&\mathbb{E}^o\left[\sum_{x\in\Phi_m\cap\mathcal{C}_o(\Phi_b)}\sum_{y\in\Phi_{m,x}}\mathbb{I}(\{y \text{ is covered}\})\right]\\
&=\mathbb{E}^o\left[\sum_{x\in\Phi_m\cap\mathcal{C}_o(\Phi_b)}\mathbb{E}[\sum_{y\in\Phi_{m,x}}\mathbb{I}(\{y \text{ is covered}\})]\right]. \qquad (2.28)
\end{aligned}$$

For $x\in\mathcal{C}_o(\Phi_b)\cap\Phi_m$, using Corollary 2.3 and Prop. 2.5, we have

$$\mathbb{E}\left[\sum_{y\in\Phi_{m,x}}\mathbb{I}(\{y \text{ is covered}\})\right]=h(\|x\|;\tau_m,g_o(\|x\|)).$$

It follows that (2.28) equals

$$\begin{aligned}
&\mathbb{E}^o\left[\sum_{x\in\Phi_m\cap\mathcal{C}_o(\Phi_b)}h(\|x\|;\tau_m,g_o(\|x\|))\right]\\
&=\mathbb{E}^o\left[\lambda_m\int_{\mathbb{R}^2}\mathbb{I}(x\in\mathcal{C}_o(\Phi_b))\cdot h(\|x\|;\tau_m,g_o(\|x\|))\,\mathrm{d}x\right]\\
&=\lambda_m\int_{\mathbb{R}^2}\mathbb{E}^o\left[\mathbb{I}(x\in\mathcal{C}_o(\Phi_b))\right]\cdot h(\|x\|;\tau_m,g_o(\|x\|))\,\mathrm{d}x\\
&=\lambda_m\int_{\mathbb{R}^2}\mathbb{P}^o(x\in\mathcal{C}_o(\Phi_b))\cdot h(\|x\|;\tau_m,g_o(\|x\|))\,\mathrm{d}x\\
&=\lambda_m\int_{\mathbb{R}^2}\mathbb{P}^o(\Phi_b(B^0(x,\|x\|))=0)\cdot h(\|x\|;\tau_m,g_o(\|x\|))\,\mathrm{d}x\\
&=\lambda_m\int_{\mathbb{R}^2}e^{-\lambda_b\pi\|x\|^2}\cdot h(\|x\|;\tau_m,g_o(\|x\|))\,\mathrm{d}x\\
&=2\pi\lambda_m\int_{\mathbb{R}^+}e^{-\lambda_b\pi r^2}\cdot h(r;\tau_m,g_o(r))r\,\mathrm{d}r\\
&=\frac{\lambda_m}{\lambda_b}\mathbb{E}_D\left[h(D;\tau_m,g_o(D))\right],
\end{aligned}$$

where we use Campbell's theorem in the first equality and Fubini's theorem in the second equality; the fourth equality follows since $x \in \mathcal{C}_o(\Phi_b)$ if and only if o is the nearest BS in Φ_b, i.e., $\Phi_b(B^0(x, \|x\|)) = 0$; the fifth equality follows from the fact $\Phi_b(B^0(x, \|x\|)) \sim \text{Poisson}(\lambda_m \pi \|x\|^2)$; and we convert from Cartesian to polar coordinates in the penultimate equality. Using similar arguments, we can derive the other two expectation terms; we omit them for brevity.

2.8.8 Proof of Proposition 2.7

Suppose $\mathcal{O} = (\tau_m^\star, \{g_z^\dagger(\|x_i - z\|)\})$ is an optimal solution but does not satisfy $g_z^\dagger(\|x_1 - z\|) \geq .... \geq g_z^\dagger(\|x_{M_z} - z\|)$. Then $\mathcal{O}$ has at least one pair (i, j) such that $1 \leq i < j \leq M_z$ and $0 = g_z^\dagger(\|x_i - z\|) < g_z^\dagger(\|x_j - z\|) = 1$. We will decrease the number of such pairs in $\mathcal{O}$ by swapping the values of the binary decision variables: $g_z^\dagger(\|x_i - z\|) = 1$ and $g_z^\dagger(\|x_j - z\|) = 0$. We denote by $\tilde{\mathcal{O}} = (\tau_m^\star, \{\tilde{g}_z(\|x_i - z\|)\})$ the solution after the swapping. First, we claim $\tilde{\mathcal{O}}$ is feasible; indeed, $\tilde{g}_z(r_k) \in \{0, 1\}$, $\sum_{k=1}^{M_z} \tilde{g}_z(r_k) = \sum_{k=1}^{M_z} g_z^\dagger(r_k) \leq B$, and using the fact that $q(r)$ is strictly decreasing with r,

$$\begin{aligned}
&\frac{1}{M_z} \sum_{k=1}^{M_z} h\left(r_k; \tau_m^\star, \tilde{g}_z(r_k)\right) - \frac{1}{M_z} \sum_{k=1}^{M_z} h\left(r_k; \tau_m^\star, g_z^\dagger(g_k)\right) \\
&= \frac{1}{M_z} (q(r_i) - q(r_j))(\bar{N}_{\max} - \bar{N}(\tau_m)) > 0,
\end{aligned}$$

which shows that $\tilde{\mathcal{O}}$ meets all the constraints. Further, $\tilde{\mathcal{O}}$ gives the optimal objective value $\tau_m^\star$; and thus $\tilde{\mathcal{O}}$ is also an optimal solution. Repeating iteratively the above exchange arguments, we can construct an optimal solution such that $g_z^\star(r_1) \geq .... \geq g_z^\star(r_z)$. This completes the proof.

Chapter 3

Spectrum Sharing between Cellular and D2D Communications

As introduced in Chapter 1, D2D can be classified into two types: in-band and out-of-band; and in-band D2D can be further classified into two categories: overlay and underlay. In this chapter, we extend the hybrid network model proposed in Chapter 2 and develop a unified analytical framework to study spectrum sharing between cellular and D2D communications. We focus on the uplink of cellular networks and study two spectrum sharing models, which are illustrated in Figure 3.1 and described as follows.

Overlay in-band D2D. The uplink spectrum is divided into two orthogonal portions. A fraction η is assigned to D2D communication while the other $1-\eta$ is used for cellular communication. We term η *spectrum partition factor* in the overlay.

Underlay in-band D2D. We assume that each D2D transmitter uses frequency hopping to randomize its interference to other links. Specifically, we divide the uplink channel into B subchannels. Each D2D transmitter may randomly and independently access βB of them, where the factor $\beta \in [0,1]$ measures the aggressiveness of D2D spectrum access. We term β *spectrum*

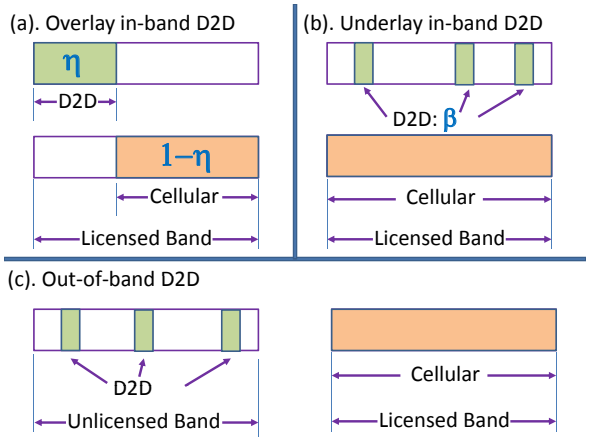

Figure 3.1: Different D2D spectrum sharing scenarios

access factor in the underlay.

How the network performance will be affected by different choices of the spectrum partition factor in the overlay and the spectrum access factor in the underlay? We address this question in this chapter.

3.1 Related Work

Existing research relevant to this work includes spectrum sharing in cognitive radio networks, where secondary cognitive transmitters may access primary spectrum if primary transmitters are not active or they do not cause unacceptable interference [8]. For example, to protect the primary users, secondary transmissions in [77,120] are regulated by sensing the activities of primary transmissions, while [67] imposes stringent secondary access constraints on e.g. collision probability. Multi-antenna techniques are used in [55,150,151]

to minimize secondary interference to primary networks. Auction mechanisms are used in [60, 138] to control the spectrum access of secondary networks. More recently, the economic aspects of spectrum sharing in cognitive radio networks have gained much interest. For example, [104] adopts a dynamical game approach to study the spectrum sharing among a primary user and multiple secondary users. Similarly, a three-stage dynamic game is formulated in [37] to study spectrum leasing and pricing strategies, while [76] designs incentive schemes for spectrum sharing with cooperative communication.

Unlike spectrum sharing in cognitive radio networks, D2D spectrum sharing is controlled by cellular networks. How D2D should access the spectrum is a largely open question, though some initial results exist (see e.g. [65, 78, 146, 147]). D2D spectrum sharing is further complicated by D2D *mode selection* which means that a potential D2D pair can switch between direct and conventional cellular communications [42, 88]. Determining an optimum D2D mode selection threshold – which we define as the Tx-Rx distance under which D2D communication should occur – is another objective of this chapter.

Note that, as highlighted in Chapter 1, D2D is different from ad hoc networks whose analysis and design are notoriously difficult (see e.g. [45, 50, 139]). A key difference is that D2D networking can be assisted by the cellular network infrastructure which is not available to a typical ad hoc network [88, 91]. Nevertheless, D2D networking introduces its own challenges. For example, the interference situation in the underlay in-band D2D is more complicated than in a purely ad hoc network. Further, enabling D2D communication requires

a lot of new functionalities [2, 4, 88] and significantly complicates the cellular network design. To sum up, spectrum sharing in cellular networks with D2D can be quite different from those of either ad hoc networks or traditional cellular networks.

3.2 Contributions and Main Outcomes

The main contributions and outcomes of this chapter are summarized as follows.

A refined tractable hybrid network model. We refine the hybrid network model previously proposed in Chapter 2 by further incorporating many important characteristics of D2D-enabled cellular networks including D2D mode selection, transmit power control and orthogonal scheduling of cellular users within a cell.

A unified performance analysis approach. We present a general analytical framework and conduct a unified performance analysis of two D2D spectrum sharing scenarios: overlay and underlay in-band D2D. In particular, we derive analytical rate expressions and apply them to optimize spectrum sharing parameters.

Design insights. The following observations are made from the derived analytical and/or numerical results under the model studied in this chapter and may be informative for system design.

Overlay vs. underlay. We evaluate the rate performance in both over-

lay and underlay scenarios. We observe that D2D mobiles can enjoy much higher data rate than regular cellular mobiles in both scenarios. As for cellular mobiles in the overlay, their rate performance also improves due to the offloading capability of D2D communication. In contrast, the rate performance of cellular mobiles in the underlay does not improve or even slightly degrades with D2D communication.[1] This is because cellular mobiles suffer from interference caused by the underlaid D2D transmissions, which offsets the D2D offloading gain. From an overall mean-rate (averaged across both cellular and D2D mobiles) perspective, both overlay and underlay provide performance improvements (vs. pure cellular).

D2D mode selection. We derive the optimal D2D mode selection threshold that minimizes the transmit power of a potential D2D transmitter. We find that the optimal threshold is inversely proportional to the square root of BS density and monotonically increases with the pathloss exponent. Moreover, it is invariant with the distance distribution of potential D2D pairs. D2D mode selection and spectrum sharing may be jointly optimized from e.g. the rate perspective. From a coverage perspective, we reveal a tradeoff between the D2D spectrum access and mode selection threshold in the underlay: as more D2D links are allowed (due to a more relaxed mode selection threshold), the network should actually make less spectrum available to them to limit their interference.

[1]Note that the underlay study in this chapter assumes that D2D randomly accesses the cellular spectrum. With carefully designed dynamic scheduling in the underlay, the rate of cellular mobiles may also increase, and the rate of D2D mobiles may further increase.

3.3 System Model

3.3.1 Network Model

As shown in Figure 3.2, we consider a hybrid network consisting of both cellular and D2D links and focus on the uplink. The BSs are regularly placed according to a hexagonal grid. Denoting by $1/\lambda_{\rm b}$ the area of a hexagonal cell, $\lambda_{\rm b}$ can be regarded as the average number of BSs per unit area. The transmit UEs are randomly distributed and modeled by an independently *marked* PPP denoted as

$$\tilde{\Phi} = \{(X_i, \delta_i, L_i, P_i)\}. \tag{3.1}$$

Here $\{X_i\}$ denote the spatial locations of the UEs. Denote by $\Phi \in \mathbb{R}^2$ the *unmarked* PPP $\{X_i\}$ with λ being its intensity. $\{\delta_i\}$ denote the types of the UEs and are assumed to be i.i.d. Bernoulli random variables with $\mathbb{P}(\delta_i = 1) = q \in [0, 1]$. In particular, UE i is called a *potential* D2D UE[2] if $\delta_i = 1$; otherwise, it is called a *cellular* UE. So, q is a simple indicator of the load of potential D2D traffic. $\{L_i\}$ denote the lengths of radio links. For notational simplicity, denote by $L_{\rm c}$ (resp. $L_{\rm d}$) the generic random variable for the link length of a typical cellular UE (resp. potential D2D UE). $\{P_i\}$ denote the transmit powers of UEs. In this work we use *channel inversion* for power control, i.e., $P_i = L_i^{\alpha}$, where $\alpha > 2$ denotes the pathloss exponent; extension to distance-proportional fractional power control (i.e., $P_i = L_i^{\alpha\epsilon}$ where $\epsilon \in [0, 1]$) is straightforward.

[2]It is called *potential* D2D UE as a UE with D2D traffic can use either cellular or D2D mode.

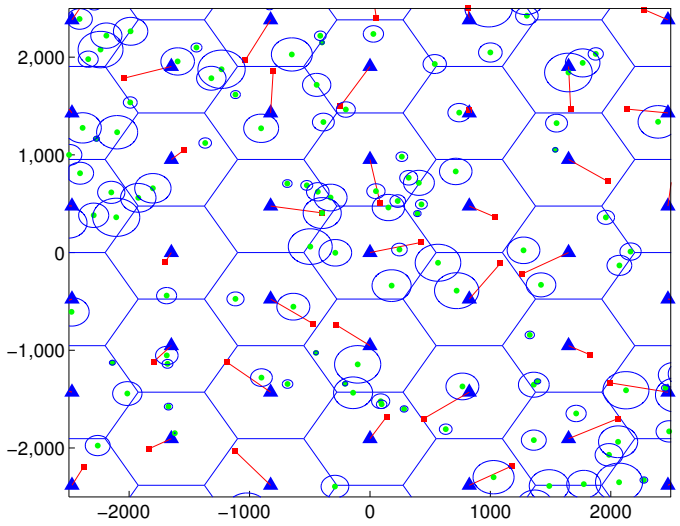

Figure 3.2: A hybrid network consisting of both cellular and D2D links. Solid triangles, solid squares and dots denote BSs, uplink cellular transmitters and D2D transmitters, respectively. For clarity we omit plotting D2D receivers, each of which is randomly located on the circle centered at the associated D2D transmitter.

Similarly, we use P_{c} and P_{d} to denote the generic random variables for the transmit powers of cellular and potential D2D UEs, respectively.

Remark on channel inversion. Note that channel inversion in this work only compensates for the large-scale pathloss. In particular, it does *not* compensate for the small-scale fading. This channel inversion scheme has two advantages: 1) it does not lead to excessively large transmit power when the link is poor (due to the small-scale fading), and 2) the transmitter only needs a long-term statistic (i.e. pathloss) to decide its transmit power, i.e., instantaneous channel state information is not required to be available at the transmitter as small-scale fading is not compensated for. Note that for ease of

exposition, we assume in the analysis that the average received power is 1 due to channel inversion, i.e., $P_i = L_i^{\alpha}$. In other words, P_i should be considered as *virtual* transmit power, and should be scaled appropriately to map to the *actual* transmit power $\tilde{P}_i$, say, $\tilde{P}_i = \rho P_i$, where ρ is the coefficient of proportionality. Normally, $\rho \ll 1$ since the practical transmit power of wireless devices is far less than the pathloss.

Next, let us introduce the notation SNR_m to denote the average *received* signal power normalized by noise power, i.e.,

$$\mathrm{SNR}_m = \frac{\tilde{P}L^{-\alpha}}{\tilde{N}_0 B_w} = \frac{\rho P L^{-\alpha}}{\tilde{N}_0 B_w} = \frac{1}{\rho^{-1}\tilde{N}_0 B_w}, \tag{3.2}$$

where $\tilde{N}_0$ denotes the one-sided power spectral density of the additive white Gaussian noise, and B_w denotes the channel bandwidth. In the rest of this chapter, if the average received power is normalized to 1, we use N_0 to denote the *equivalent* noise power $\rho^{-1}\tilde{N}_0 B_w$. By choosing the operating regime SNR_m (or equivalently, the coefficient ρ) appropriately, we can make sure that the UE power constraints are satisfied and thus there is no need to truncate UE transmit power to meet the peak power constraint. We will give more detailed results in Section 3.4.2 to illustrate the above argument.

The potential of D2D will largely depend on the amount of local traffic that may be routed through local direct paths, instead of infrastructure paths. One possible approach to model "data localization" would be based on current user traffic statistics. However, it appears very challenging to acquire such traffic data, which is typically owned by operators and contains

sensitive and proprietary information. Even if the current traffic data could be obtained from the operators, it might not be too useful, since presumably D2D's availability could change future traffic patterns. For example, once users realize high D2D speeds are possible, more local sharing is likely to occur. So far, no commonly agreed upon D2D distance distribution has appeared in the literature. In the absence of such an accepted model, we assume that each potential D2D receiver is randomly and independently placed around its associated potential D2D transmitter with isotropic direction and Rayleigh distributed distance D with probability density function (PDF) given by

$$f_D(x) = 2\pi\xi x e^{-\xi\pi x^2}, \quad x \geq 0. \tag{3.3}$$

In other words, the potential D2D receiver is randomly placed around its associated potential D2D transmitter according to a two-dimensional Gaussian distribution, which results in (3.3). A similar Gaussian assumption has also been used in [20] to analyze the performance of FlashLinQ. The analysis and calculations in this work can be used to study other D2D distance distributions as well.

In this work, we consider *distance-based D2D mode selection*: cellular mode is used if $D \geq \mu$; otherwise, D2D mode is selected. If we assume that the received signal power (averaged over fast fading) is only a function of distance and pathloss exponent, distance-based D2D mode selection is equivalent to the average received-signal-power or SNR-based mode selection, to which the results in this work can be directly applied.

3.3.2 Transmission Scheduling

Cellular transmitters including cellular UEs and potential D2D UEs in cellular mode form a PPP Φ_c with intensity $\lambda_{\mathrm{c}} = (1-q)\lambda + q\lambda\mathbb{P}(D \geq \mu)$. We assume an orthogonal multiple access technique and that uplink transmitters are scheduled in a round-robin fashion. It follows that only one uplink transmitter in each macrocell can be active at a time. Generally speaking, scheduling cellular transmitters in an orthogonal manner leads to *dependent* thinning of PPP Φ_{c}. This makes the analysis intractable and some simplified assumptions are needed (see e.g. [105]). In this chapter, denoting by $\mathcal{A}$ the coverage region of a hexagonal macrocell, we approximate $\mathcal{A}$ by a disk that has the same area as the hexagonal cell, i.e., $\mathcal{A} = \mathcal{B}(0, R)$ where $\mathcal{B}(x, r)$ denotes the ball centered at x with radius r and $R = \sqrt{\frac{1}{\pi\lambda_{\mathrm{b}}}}$. To avoid triviality, we assume $\lambda_{\mathrm{c}} \geq \lambda_{\mathrm{b}}$, which is reasonable as the uplink transmitter density is usually larger than the BS density. Further, we assume that the typical active cellular transmitter is uniformly distributed in the coverage region $\mathcal{A}$, and that the locations of cellular interferers form a PPP $\Phi_{\mathrm{c,a}}$ with intensity λ_{b}. For the typical uplink transmission, cellular interferers are located outside the region $\mathcal{A}$. Figure 3.3 illustrates the proposed approximate interference analysis for a typical uplink transmission. Due to the use of this approximation, the analytical results about cellular performance derived in this work are approximations; for notation simplicity, we will present the results as equalities instead of the more cumbersome approximations in the sequel. The approximation will be numerically validated later.

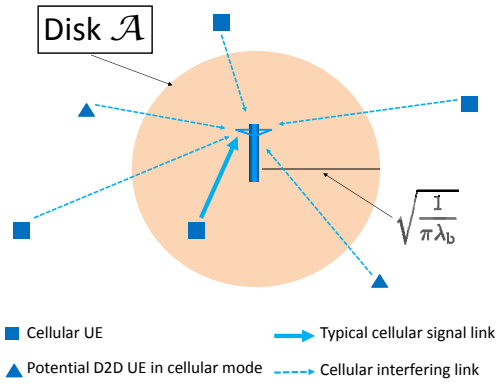

Figure 3.3: An approximate uplink interference analysis. The typical cellular transmitter is uniformly distributed in $\mathcal{A}$, while cellular interferers form a PPP with density λ_b outside the disk $\mathcal{A}$.

As for potential D2D UEs in D2D mode, they form a PPP Φ_d with intensity $\lambda_d = q\lambda\mathbb{P}(D < \mu)$. For D2D medium access control, we consider a simple spatial Aloha access scheme: in each time slot each potential D2D UE in D2D mode transmits with probability κ and is silent with probability $1-\kappa$; the activity decisions are independently made over different time slots and different transmitters. The study of this simple baseline medium access scheme can serve as a benchmark for future work on more sophisticated D2D scheduling schemes, e.g., carrier sense multiple access (CSMA) or centralized scheduling.

3.3.3 Performance Metrics

We will analyze the average rates of cellular and potential D2D UEs, T_c and T_d. Recall that potential D2D UEs can use either cellular or D2D

mode. Denote by $\hat{T}_{\mathrm{d}}$ the average rate of potential D2D UEs in D2D mode. Conditioning on using cellular mode, the rate T_{d} of potential D2D UEs equals the rate T_{c} of cellular UEs; conditioning on using D2D mode, the rate T_{d} of potential D2D UEs equals $\hat{T}_{\mathrm{d}}$ by definition. Under the assumed distance-based D2D mode selection, a typical potential D2D UE uses cellular mode with probability $\mathbb{P}(D \geq \mu)$ and D2D mode with probability $\mathbb{P}(D < \mu)$. To sum up, the average rate of potential D2D UEs can be written as

$$T_{\mathrm{d}} = \mathbb{P}(D \geq \mu) \cdot T_{\mathrm{c}} + \mathbb{P}(D < \mu) \cdot \hat{T}_{\mathrm{d}}. \tag{3.4}$$

3.4 Preliminary Analysis

In this section we present preliminary analytical results, which lay the foundation for the study of overlay and underlay in-band D2D in the next two sections.

3.4.1 A Unified Analytical Approach

Consider a typical transmitter and receiver pair interfered by K types of heterogeneous interferers. The set of the k-th type of interferers is denoted as $\mathcal{M}_k$. In this work, we focus on frequency-flat narrowband channels; the results can be readily extended to OFDM-based frequency-selective wideband channels, each of which can conceptually be regarded as a set of parallel frequency-flat narrowband channels.

The baseband received signal $Y_0[n]$ at the typical receiver located at

the origin can be written as

$$Y_0[n] = \sqrt{P_0 L_0^{-\alpha} G_0} S_0[n] + \sum_{k=1}^{K} \sum_{i \in \mathcal{M}_k} \sqrt{P_i \|X_i\|^{-\alpha} G_i} S_i[n] + Z[n], \tag{3.5}$$

where P_0, L_0, G_0 and $S_0[n]$ are associated with the typical link and denote the typical link's transmit power, length, channel fading and unit-variance signal, respectively; $P_i, \|X_i\|, G_i$ and $S_i[n]$ are associated with the interfering link from transmitter i to the typical receiver and denote the interfering link's transmit power, length, channel fading and unit-variance signal, respectively; $Z[n]$ is additive white Gaussian noise with constant PSD $\tilde{N}_0$ Watts/Hz. It follows that the received SINR is given by $\text{SINR} = \frac{W}{I + \tilde{N}_0 B_w}$, where signal power $W = P_0 L_0^{-\alpha} G_0$, and interference power $I = \sum_{k=1}^{K} \sum_{i \in \mathcal{M}_k} P_i \|X_i\|^{-\alpha} G_i$.

In this chapter, recall we assume channel inversion, i.e., $P_0 L_0^{-\alpha} = 1$, and thus $W = G_0$. For simplicity we consider Rayleigh fading, i.e., $G_0 \sim \text{Exp}(1)$, and assume independent fading over space. Then the following corollary will be particularly useful.

Corollary 3.1. *Suppose* $SINR = \frac{W}{I + N_0}$, *where* $W \sim Exp(1)$ *and* I *respectively denote the (random) signal and interference powers, and* N_0 *denotes the equivalent noise power. If* W *and* I *are independent,*

$$\mathbb{E}\left[\log(1 + SINR)\right] = \int_0^\infty \frac{e^{-N_0 x}}{1 + x} \mathcal{L}_I(x) \, \mathrm{d}x, \tag{3.6}$$

where $\mathcal{L}_I(s) = \mathbb{E}[e^{-sI}]$ *denotes the Laplace transform of* I.

Corollary 3.1 follows from a more general result given in [87] and its proof may also be directly found in [16]. Note that in this work, interference

is not Gaussian but may be considered as *conditionally Gaussian*. Specifically, assuming all the transmitters use Gaussian signaling, the interference is Gaussian conditioned on the fading and node locations in the network. Then treating the interference as noise, we may invoke Shannon's formula to determine the maximum achievable spectral efficiency of a typical link. If the random fading and node locations are furthered averaged out, we arrive at the expression (3.6). Though not optimal in an information-theoretical sense, (3.6) serves as a good performance metric and has been widely adopted in literature [16].

Next we define the ergodic link spectral efficiency R, which combines modulation and coding schemes in the physical layer and multiple access protocols in the medium access control layer, as follows.

$$R = \mathbb{E}\left[\Delta \cdot \log(1 + \mathrm{SINR})\right], \tag{3.7}$$

where Δ denotes the time and/or frequency resources accessed by the typical link. For example, in the overlay with spectrum partition factor η, a typical D2D link with random Aloha access probability κ can effectively access $\kappa\eta$ time-frequency resources. We will analyze ergodic link spectral efficiency R in detail in Sections 3.5 and 3.6.

3.4.2 Transmit Power Analysis

In this subsection, we analyze the transmit power distributions, particularly $\mathbb{E}[P_{\mathrm{c}}]$ and $\mathbb{E}[P_{\mathrm{d}}]$, the average transmit powers of cellular and potential

D2D UEs. The derived results are not only interesting in its own right but also are extensively used in the analysis of rate performance later. To this end, denote by $\hat{P}_{\mathrm{d}}$ the generic random variable for the transmit power of potential D2D UEs in D2D mode.

Lemma 3.1. *The average transmit powers of a typical cellular UE, a potential D2D UE and, a D2D-mode potential D2D UE are respectively given by*

$$\mathbb{E}[P_c] = \frac{1}{(1+\frac{\alpha}{2})\pi^{\frac{\alpha}{2}}\lambda_b^{\frac{\alpha}{2}}} \tag{3.8}$$

$$\mathbb{E}[P_d] = e^{-\xi\pi\mu^2}\mathbb{E}[P_c] + (\xi\pi)^{-\frac{\alpha}{2}}\gamma(\frac{\alpha}{2}+1, \xi\pi\mu^2) \tag{3.9}$$

$$\mathbb{E}[\hat{P}_d] = \frac{(\xi\pi)^{-\frac{\alpha}{2}}}{1-e^{-\xi\pi\mu^2}}\gamma(\frac{\alpha}{2}+1, \xi\pi\mu^2), \tag{3.10}$$

where $\gamma(s,x) = \int_0^x z^{s-1}e^{-z}\,\mathrm{d}z$ *is the lower incomplete Gamma function.*

Proof. See Appendix 3.9.1. □

Note that both $\mathbb{E}[P_{\mathrm{c}}]$ and $\mathbb{E}[P_{\mathrm{d}}]$ increase as pathloss exponent α increases and are inversely proportional to the square root of BS density. Next we examine how to choose D2D mode selection threshold μ to minimize the average transmit power $\mathbb{E}[P_{\mathrm{d}}]$ of potential D2D UEs.

Proposition 3.1. *For any distribution function* $f_D(x)$ *of the nonnegative random distance* D, $\mathbb{E}[P_d]$ *is minimized when*

$$\mu^\star = (\mathbb{E}[P_c])^{\frac{1}{\alpha}} = \left(\frac{1}{1+\frac{\alpha}{2}}\right)^{\frac{1}{\alpha}}\sqrt{\frac{1}{\pi\lambda_b}}. \tag{3.11}$$

Proof. See Appendix 3.9.2. □

Prop. 3.1 shows that $\mu^\star$ is only a function of the average transmit power $\mathbb{E}[P_{\rm c}]$ of cellular UEs and is independent of the distribution of D; in particular, the Rayleigh assumption made in (3.3) is not needed here. Specifically, $\mu^\star$ is inversely proportional to the square root of BS density $\lambda_{\rm b}$, which is intuitive: cellular mode becomes more favorable when more BSs are available. In addition, $(\frac{1}{1+\frac{\alpha}{2}})^{\frac{1}{\alpha}}$ monotonically increases as α increases. This implies that $\mu^\star$ increases in α, agreeing with intuition: local transmission with D2D mode is more favorable for saving transmit power when the pathloss exponent increases.

Before ending this section, we give a numerical example in Figure 3.4 showing that UE power constraint can be satisfied by choosing the right operating regime SNR_m. Throughout this chapter, the parameters used in plotting numerical or simulation results are summarized in Table 3.1 unless otherwise specified. In Figure 3.4, the cellular peak power $P_{\rm c,max}$ is defined as the minimum transmit power used by a cell-edge cellular transmitter to meet the target SNR_m, i.e., $P_{\rm c,max}$ is determined by $\text{SNR}_m = \frac{P_{\rm c,max}R^{-\alpha}}{\tilde{N}_0 B_w}$. Similar, the D2D peak power $P_{\rm d,max}$ is determined by $\text{SNR}_m = \frac{P_{\rm d,max}\mu^{-\alpha}}{\tilde{N}_0 B_w}$. The average power of a cellular (resp. D2D) transmitter can be obtained by multiplying (3.8) (resp. (3.10)) with the scaling factor $\tilde{N}_0 B_w \cdot \text{SNR}_m$. As shown in Figure 3.4, the typical UE power constraint 23 dBm (i.e. 200 mW) is well respected even when SNR_m=10 dB, a relatively high average received SNR in the uplink. Besides, Figure 3.4 also shows that compared to cellular transmitters, D2D transmitters can save about 15 dB transmit power in achieving the same SNR_m target,

Density of macrocells λ_{b}	$(\pi 500^2)^{-1}$ m^{-2}
Density of UEs λ	$10 \times (\pi 500^2)^{-1}$ m^{-2}
D2D distance parameter ξ	$10 \times (\pi 500^2)^{-1}$ m^{-2}
Proportion of potential D2D UEs q	0.2
Pathloss exponent α	3.5
SNR_m	10 dB
D2D mode selection threshold μ	200 m
D2D Aloha access probability κ	1
Spectrum partition factor η	0.2
UE weights $(w_{\mathrm{c}}, w_{\mathrm{d}})$	$(0.6, 0.4)$
Spectrum access factor β	1
Number of subchannels B	1

Table 3.1: Simulation/Numerical Parameters for Spectrum Sharing between Cellular and D2D

demonstrating the energy efficiency of D2D communication. In other words, with the same power budget, the D2D links can enjoy about 15 dB higher SNR_m than the cellular links.

3.5 Analysis of Overlay In-Band D2D

3.5.1 Link Spectral Efficiency

Let us consider a typical D2D link. As the underlying PPP is stationary, without loss of generality we assume that the typical receiver is located at the origin. Note that the analysis carried out for a typical link indicates the spatially averaged performance of the network by Slivnyak's theorem [122]. Henceforth, we focus on characterizing the performance of a typical link, which may be either a D2D or cellular link depending on the context.

With overlay in-band D2D, the interferers are cochannel D2D trans-

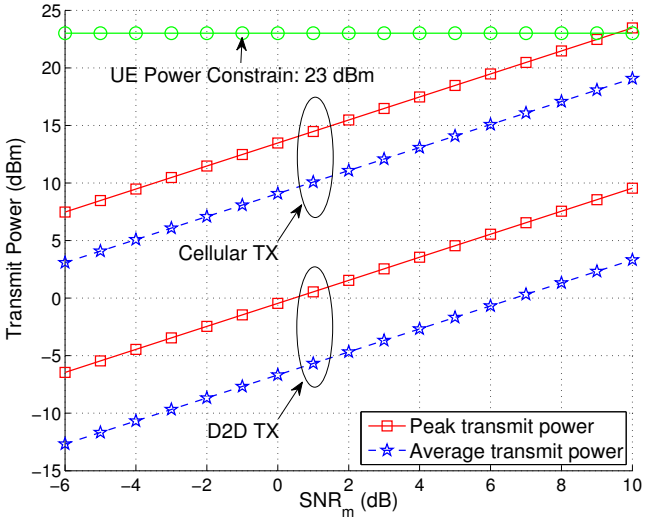

Figure 3.4: UE transmit power versus network operating regime SNR_m with $\tilde{N}_0 = -174$ dBm/Hz and $B_w = 1$ MHz.

mitters. Due to the random Aloha access, the effective interferers constitute a homogeneous PPP with density $\kappa\lambda_{\mathrm{d}}$, a PPP thinned from Φ_{d} with thinning probability κ. Denoting this thinned PPP by $\kappa\Phi_{\mathrm{d}}$, the interference at the typical D2D receiver is given by

$$I_{\mathrm{d}} = \sum_{X_i \in \kappa\Phi_{\mathrm{d}} \setminus \{o\}} \hat{P}_{\mathrm{d},i} G_i \|X_i\|^{-\alpha}. \tag{3.12}$$

Proposition 3.2. *With overlay in-band D2D, the complementary cumulative distribution function (CCDF) of the SINR of D2D links is given by*

$$\mathbb{P}(SINR \geq x) = \exp\left(-N_0 x - c x^{\frac{2}{\alpha}}\right), \quad x \geq 0, \tag{3.13}$$

where c is a non-negative constant given by

$$c = \frac{\kappa q(\frac{\lambda}{\xi} - (\frac{\lambda}{\xi} + \lambda\pi\mu^2)e^{-\xi\pi\mu^2})}{sinc(\frac{2}{\alpha})}. \tag{3.14}$$

Further, the spectral efficiency R_d of D2D links is given by

$$R_d = \kappa \int_0^\infty \frac{e^{-N_0 x}}{1+x} e^{-c x^{\frac{2}{\alpha}}} \, \mathrm{d}x. \tag{3.15}$$

Proof. This proposition follows by evaluating the Laplace transform of I_d and using (3.6). See Appendix 3.9.3 for details. □

Note that in Prop. 3.2, as μ increases, c monotonically increases, which in turn results in monotonically decreasing R_d. This agrees with intuition: the spectral efficiency of D2D links decreases when more potential D2D UEs choose D2D mode (leading to increased interference). In particular,

$$R_{\mathrm{d,min}} = \lim_{\mu\to\infty} R_\mathrm{d} = \kappa \int_0^\infty \frac{e^{-N_0 x}}{1+x} e^{-\kappa q \frac{\lambda}{\xi} (\mathrm{sinc}(\frac{2}{\alpha}))^{-1} x^{\frac{2}{\alpha}}} \, \mathrm{d}x. \tag{3.16}$$

The typical D2D link experiences the most severe interference in this case and thus has the minimum spectral efficiency. On the contrary,

$$R_{\mathrm{d,max}} = \lim_{\mu\to 0} R_\mathrm{d} = \kappa \int_0^\infty \frac{1}{1+x} e^{-N_0 x} \, \mathrm{d}x = \kappa e^{N_0} E_1(N_0), \tag{3.17}$$

where $E_1(z) = \int_z^\infty \frac{e^{-x}}{x} \, \mathrm{d}x$ is the exponential integral. The typical D2D link is free of interference in this case and thus has the maximum spectral efficiency.

Now let us consider a typical uplink. With overlay in-band D2D, the interferers are cellular transmitters in the other cells. The interference at the typical BS is given by

$$I_\mathrm{c} = \sum_{X_i \in \Phi_{\mathrm{c,a}} \cap \mathcal{A}^c} P_{\mathrm{c},i} G_i \|X_i\|^{-\alpha}.$$

Proposition 3.3. *With overlay in-band D2D, the CCDF of the SINR of cellular links is given by*

$$\mathbb{P}(SINR \geq x) = \exp(-N_0 x) \times \exp\left(-2\pi\lambda_b \int_R^\infty (1 - {}_2F_1(1, \frac{2}{\alpha}; 1 + \frac{2}{\alpha}; -\frac{x}{(R/r)^\alpha}))r \ \mathrm{d}r\right), x \geq 0, \quad (3.18)$$

where ${}_2F_1(a, b; c; x)$ *denotes the hypergeometric function, and* $R = \sqrt{\frac{1}{\pi\lambda_b}}$. *Further, the spectral efficiency* R_c *of cellular links is given by*

$$R_c = \frac{\lambda_b}{\lambda_c}(1 - e^{-\frac{\lambda_c}{\lambda_b}}) \int_0^\infty \frac{e^{-N_0 x}}{1+x} \times \exp\left(-2\pi\lambda_b \int_R^\infty (1 - {}_2F_1(1, \frac{2}{\alpha}; 1 + \frac{2}{\alpha}; -\frac{x}{(R/r)^\alpha}))r \ \mathrm{d}r\right) \ \mathrm{d}x, \quad (3.19)$$

where $\lambda_c = (1-q)\lambda + q\lambda e^{-\xi\pi\mu^2}$.

Proof. See Appendix 3.9.4. □

Unlike the closed form expression (3.13) for the CCDF of the SINR of D2D links, the expression (3.18) for the CCDF of the uplink SINR involves an integration. To have some insights, we consider sparse (i.e. $\lambda_\mathrm{b} \to 0$) and dense (i.e. $\lambda_\mathrm{b} \to \infty$) networks in the following corollary.

Corollary 3.2. *In a sparse network with* $\lambda_b \to 0$, *the CCDF of the SINR of cellular links is given by*

$$\mathbb{P}(SINR \geq x) = \exp\left(-\left(N_0 + \frac{4}{\alpha^2 - 4}\right)x\right), \quad x \geq 0, \quad (3.20)$$

In a dense network with $\lambda_b \to \infty$, the CCDF of the SINR of cellular links is given by

$$\mathbb{P}(SINR \geq x) = \exp\left(-N_0 x - \frac{1}{2sinc(\frac{2}{\alpha})} x^{\frac{2}{\alpha}}\right), \quad x \geq 0. \tag{3.21}$$

Proof. See Appendix 3.9.5. □

In a sparse network, Corollary 3.2 implies that interference and noise have the same impact on the SINR coverage performance (in the order sense). From this perspective, we may simply consider interference as an extra source of noise. Thus, the sparse network is noise-limited. In contrast, in a dense network, the impact of interference behaves differently for UEs with different SINR targets. For users with low SINR target (i.e. $x \to 0$), interference has a more pronounced impact on the SINR coverage performance than the noise. The converse is true for users with high SINR target (i.e. $x \to \infty$).

If we denote by θ_{c} the SINR threshold for successful uplink transmissions and consider the outage probability $\mathbb{P}(\mathrm{SINR} < \theta_{\mathrm{c}})$, Corollary 3.2 implies that, as $\theta_{\mathrm{c}} \to 0$, the outage probability of a sparse network scales as $\Theta(\theta_{\mathrm{c}})$. In this case, the outage performance is noise-limited. In contrast, the outage probability of a dense network scales as $\Theta(\theta_{\mathrm{c}}^{\frac{2}{\alpha}})$. In this case, the outage performance is interference-limited.

Before ending this section, we validate the analytical results, particularly the CCDF of the uplink SINR (since we adopt an approximate approach for the analysis on the uplink SINR). As all the major analytical results presented in this work are functions of SINR, it suffices to validate the analytical

SINR distributions by simulation rather than repetitively validating each analytical expression which in turn is a function of SINR. In Figure 3.5, we compare the analytical uplink SINR CCDF (given in Prop. 3.3) to the corresponding empirical distribution obtained from simulation using the hexagonal model. The simulation steps are described as follows. (1) Place the BSs according to a hexagonal grid in a large area C; the area of a hexagon equals $1/\lambda_{\mathrm{b}}$. (2) Generate a random number N such that $N \sim \mathrm{Poisson}(\lambda_{\mathrm{c}}|C|)$. (3) Generate N points that are uniformly distributed in C; these N points represent the cellular transmitters. (4) For each BS, it randomly schedules a cellular transmitter if there is at least one in its coverage region. (5) Determine the transmit power of each scheduled transmitter. (6) Generate independently the fading gains from each scheduled cellular transmitter to each BS. (7) Collect the SINR statistics of the cellular links located in the central hexagonal cells (to avoid boundary effect). (8) Repeat the above steps $10,000$ times.

Figure 3.5 shows that the analytical results match the empirical results fairly well; the small gaps arise as we approximate the hexagonal model using a PPP model with a guard radius. Recall that in Figure 3.5, $\lambda_{\mathrm{b}} = (\pi 500^2)^{-1}$ m^{-2}, and thus the network is sparse. In the sparse regime, as established in Corollary 2, SNR_m (or equivalently, N_0) has a considerable impact on the uplink SINR CCDF; Figure 3.5 confirms the analytical result.

We next compare the SINR distribution of a typical D2D link (given in Prop. 3.2) to the corresponding empirical distribution obtained from Monte Carlo simulation. The results are shown in Figure 3.6, from which we can see

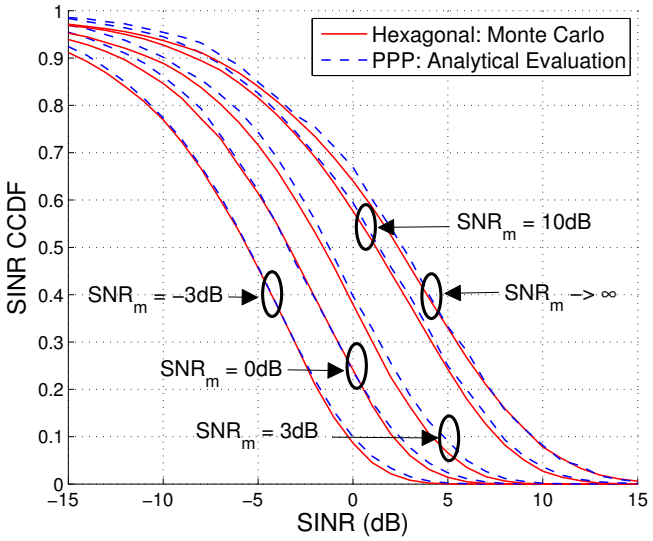

Figure 3.5: Validation of the analytical SINR CCDF of cellular links.

that the analytical results closely match the empirical results as in this case no approximation is made in the analysis.

3.5.2 Optimizing Spectrum Partition

In this section we study how to choose the optimal spectrum partition factor $\eta^\star$ such that

$$\eta^\star = \arg \max_{\eta \in [0,1]} \quad u(T_\mathrm{c}, T_\mathrm{d}), \tag{3.22}$$

where $u(T_c, T_d)$ is a utility function that can take different forms under different design metrics. In this work we use the popular weighted proportional fair function: $w_\mathrm{c} \log T_\mathrm{c} + w_\mathrm{d} \log T_\mathrm{d}$, where $w_\mathrm{c}, w_\mathrm{d} > 0$ are weight factors such that $w_\mathrm{c} + w_\mathrm{d} = 1$.

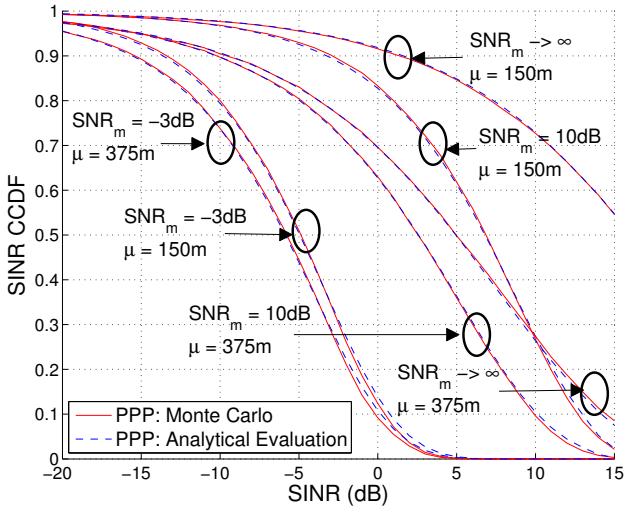

Figure 3.6: Validation of the analytical SINR CCDF of D2D links.

To optimize the spectrum partition, we first need the rate expressions for $T_{\rm c}$ and $T_{\rm d}$. For a given spectrum partition η, the (normalized) rate (bit/s/Hz) $T_{\rm c}$ of cellular UEs equals $R_{\rm c}$ multiplied by the available spectrum resource $1-\eta$. In contrast, the rate $T_{\rm d}$ of potential D2D UEs equals $(1-\eta)R_{\rm c}$ if cellular mode is used; otherwise, i.e., D2D mode is used, it equals $\eta R_{\rm d}$. To summarize,

$$T_{\rm c} = (1-\eta)R_{\rm c}, \quad T_{\rm d} = (1-\eta)e^{-\xi\pi\mu^2}R_{\rm c} + \eta(1-e^{-\xi\pi\mu^2})R_{\rm d}. \tag{3.23}$$

Figure 3.7 shows the average rates of cellular and potential D2D UEs as a function of D2D mode selection threshold μ.[3] As expected, the average

[3]Note that for fair comparison, we normalize the transmit powers of cellular and D2D transmitters by taking into account their transmission bandwidths when plotting numerical results in this chapter. For example, in Figure 3.7 with $\eta = 0.2$, when $\text{SNR}_m = 10$ dB, the

rate of cellular UEs increases as μ increases. This is because with increasing μ, less potential D2D UEs choose cellular mode and correspondingly cellular UEs can be scheduled more often. In contrast, the average rate of potential D2D UEs first increases and then decreases as μ increases. This is because the average rate of potential D2D UEs is co-determined by its cellular-mode rate and D2D-mode rate: cellular-mode rate increases with μ while D2D-mode rate decreases with μ (due to the increased intra-tier interference). Figure 3.7 also shows that with appropriate choice of μ, potentials D2D UEs can enjoy much higher rate than cellular UEs. Meanwhile, cellular UEs also benefit from offloading the traffic by D2D communication.

We are now in a position to derive the optimal weighted proportional fair spectrum partition, which reads as follows.

Proposition 3.4. *The optimal weighted proportional fair spectrum partition* $\eta^\star$ *is given by*

$$\eta^\star = 1 - \frac{w_c}{w_c + w_d} \cdot \frac{1}{1 - (e^{\xi\pi\mu^2} - 1)^{-1} \frac{R_c}{R_d}} \tag{3.24}$$

if $R_d > \frac{w_c + w_d}{w_d} \frac{1}{e^{\xi\pi\mu^2} - 1} R_c$*; otherwise,* $\eta^\star = 0$*. In particular,* $\lim_{\mu\to\infty} \eta^\star = w_d$*.*

Proof. See Appendix 3.9.6. □

From Prop. 3.4, we can see that if $\mu \to \infty$, i.e., potential D2D UEs are restricted to use D2D mode, the optimal partition $\eta^\star$ converges to w_{d}. So

average received SNR of cellular links and of D2D links are given by $\mathrm{SNR}_m + 10\log_{10}\frac{1}{1-\eta} = 10.97$ dB, $\mathrm{SNR}_m + 10\log_{10}\frac{1}{\eta} = 16.99$ dB, respectively.

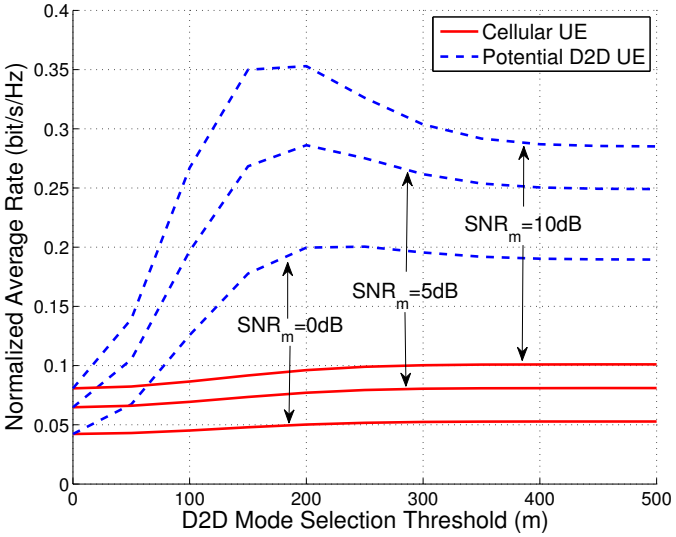

Figure 3.7: Average rates of cellular and potential D2D UEs in the case of overlay in-band D2D.

the optimal partition $\eta^\star$ simply equals the weight we assign to the potential D2D UEs. In Figure 3.8, we plot the utility value vs. η under different values of q, the proportion of potential D2D UEs. It can be seen that the optimal $\eta^\star = 0.4 = w_{\mathrm{d}}$, which is independent of q. This plot validates Prop. 3.4.

3.6 Analysis of Underlay In-Band D2D

3.6.1 Link Spectral Efficiency

Considering the random access of D2D in both frequency and time domains, the effective D2D interferers constitute a homogeneous PPP with density $\kappa\beta\lambda_{\mathrm{d}}$, a PPP thinned from Φ_{d} with thinning probability $\kappa\beta$. Denote this thinned PPP by $\kappa\beta\Phi_{\mathrm{d}}$. Considering further the interference from cellular transmitters $\Phi_{\mathrm{c,a}}$ with density λ_{b}, the interference at the typical D2D receiver

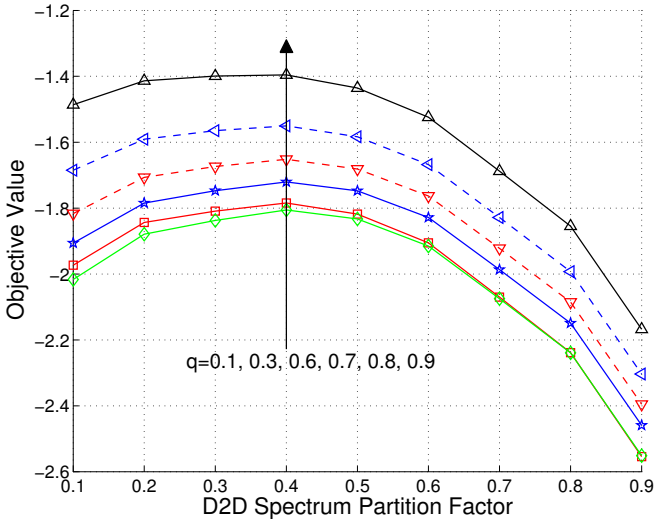

Figure 3.8: Utility value vs. D2D spectrum partition factor η under different values of q, the proportion of potential D2D UEs.

is given by

$$I_{\mathrm{d}} = \sum_{X_i \in \kappa\beta\Phi_{\mathrm{d}} \setminus \{o\}} \hat{P}_{\mathrm{d},i} G_i \|X_i\|^{-\alpha} + \sum_{X_i \in \Phi_{\mathrm{c,a}}} P_{\mathrm{c},i} G_i \|X_i\|^{-\alpha}. \tag{3.25}$$

Proposition 3.5. *With underlay in-band D2D, the CCDF of the SINR of D2D links is given by*

$$\mathbb{P}(SINR \geq x) = \exp\left(-N_0 x - c\beta x^{\frac{2}{\alpha}} - \frac{1}{2sinc(\frac{2}{\alpha})}(\beta x)^{\frac{2}{\alpha}}\right), \quad x \geq 0. \tag{3.26}$$

Further, the spectral efficiency R_d of D2D links is given by

$$R_d = \kappa \int_0^{\infty} \frac{e^{-N_0 x}}{1+x} \exp\left(-c\beta x^{\frac{2}{\alpha}} - \frac{1}{2sinc(\frac{2}{\alpha})}(\beta x)^{\frac{2}{\alpha}}\right) \mathrm{d}x. \tag{3.27}$$

Proof. See Appendix 3.9.7. □

Now let us consider a typical uplink. With underlay in-band D2D, the interferers are out-of-cell cellular transmitters $\Phi_{\mathrm{c,a}} \cap \mathcal{A}^c$ and D2D transmitters in $\kappa\beta\Phi_{\mathrm{d}}$. The interference at the typical BS is given by

$$I_{\mathrm{c}} = \sum_{X_i \in \Phi_{\mathrm{c,a}} \cap \mathcal{A}^c} P_{\mathrm{c},i} G_i \|X_i\|^{-\alpha} + \sum_{X_i \in \kappa\beta\Phi_{\mathrm{d}}} \hat{P}_{\mathrm{d},i} G_i \|X_i\|^{-\alpha}.$$

Proposition 3.6. *With underlay in-band D2D, the CCDF of the SINR of cellular links is given by*

$$\mathbb{P}(SINR \geq x) = \exp\left(-N_0 x - c\beta^{1-\frac{2}{\alpha}} x^{\frac{2}{\alpha}}\right) \times \exp\left(-2\pi\lambda_b \int_R^\infty (1 - {}_2F_1(1, \frac{2}{\alpha}; 1 + \frac{2}{\alpha}; -\frac{x}{(R/r)^\alpha}))r \ \mathrm{d}r\right), x \geq 0. \quad (3.28)$$

Further, the spectral efficiency R_c of cellular links is given by

$$R_c = \frac{\lambda_b}{\lambda_c}(1 - e^{-\frac{\lambda_c}{\lambda_b}}) \int_0^\infty \frac{e^{-N_0 x}}{1+x} \cdot e^{-c\beta^{1-\frac{2}{\alpha}} x^{\frac{2}{\alpha}}} \cdot \times \exp\left(-2\pi\lambda_b \int_R^\infty (1 - {}_2F_1(1, \frac{2}{\alpha}; 1 + \frac{2}{\alpha}; -\frac{x}{(R/r)^\alpha}))r \ \mathrm{d}r\right) \ \mathrm{d}x. \quad (3.29)$$

Proof. See Appendix 3.9.8. □

Prop. 3.5 (resp. Prop. 3.6) implies that the spectral efficiency R_{d} of D2D links (resp. R_{c} of cellular links) decreases as β increases. In other words, with larger β, the increased D2D interferer density in each subchannel has a more significant impact than the decreased transmit power per subchannel of D2D interferers. To sum up, from the perspective of maximizing the spectral efficiency of either D2D or cellular links, the design insight here is that underlay

in-band D2D should access small bandwidth with high power density rather than spreading the power over large bandwidth. However, small β limits the spectrum resource available to the D2D transmissions, which in turn limits the D2D throughput or rate.

3.6.2 Optimizing Spectrum Access

As in the case of overlay, we choose an optimal spectrum access factor $\beta^\star$ in underlay case such that

$$\beta^\star = \arg \max_{\beta \in [0,1]} \quad w_{\mathrm{c}} \log T_{\mathrm{c}} + w_{\mathrm{d}} \log T_{\mathrm{d}}. \tag{3.30}$$

To this end, we first need the rate expressions for T_{c} and T_{d}. By definition, it is easy to see that

$$T_{\mathrm{c}} = R_{\mathrm{c}}, \quad T_{\mathrm{d}} = e^{-\xi\pi\mu^2} R_{\mathrm{c}} + \beta(1 - e^{-\xi\pi\mu^2}) R_{\mathrm{d}}, \tag{3.31}$$

where R_{c} and R_{d} are given in (3.29) and (3.27), respectively.

Figure 3.9 shows the average rates of cellular and potential D2D UEs as a function of μ in the underlay scenario. Recall in the overlay case, the rate of cellular UEs increases with μ due to D2D offloading gain. In contrast, here the rate of cellular UEs stays almost constant or even slightly decreases with μ. This is because cellular UEs now suffer from the interference caused by the underlaid D2D transmissions; this offsets the offloading gain. Figure 3.9 also shows that larger β leads to higher rate of potential D2D UEs but lower rate of cellular UEs, which is pretty intuitive.

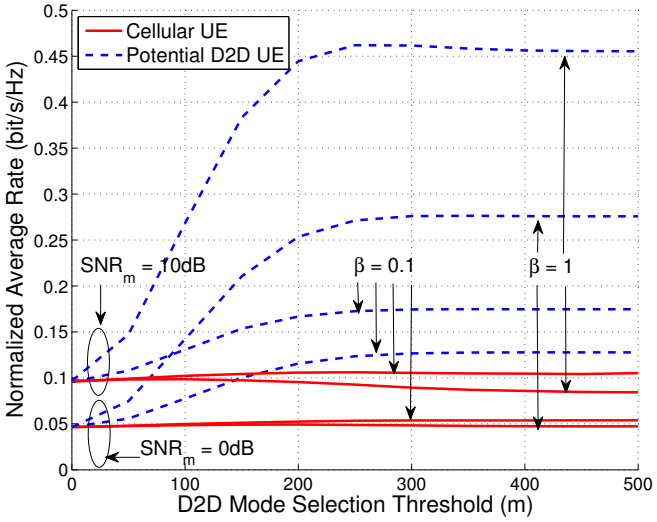

Figure 3.9: Average rates of cellular and potential D2D UEs in the case of underlay in-band D2D.

Next we optimize the spectrum access. From (3.31), we can see that the spectrum access factor β in the underlay scenario has a much more complicated impact on T_{c} and T_{d} than η does in the overlay scenario. As a result, a closed-form solution for $\beta^{\star}$ is hard to obtain. Nevertheless, the optimization problem is of single variable and can be numerically solved. In Figure 3.10, we plot the utility value vs. β under different values of q, the proportion of potential D2D UEs. It can be seen that the optimal $\eta^{\star}$ decreases as q increases.

Thus far, spectrum sharing is optimized from the rate perspective. In practice, D2D spectrum sharing may have to take into account other factors. Take the underlay scenario for example. In order to protect the cellular transmissions, we may have to limit the proportion of the spectrum that can be accessed by D2D. Specifically, assume that D2D transmissions have a target

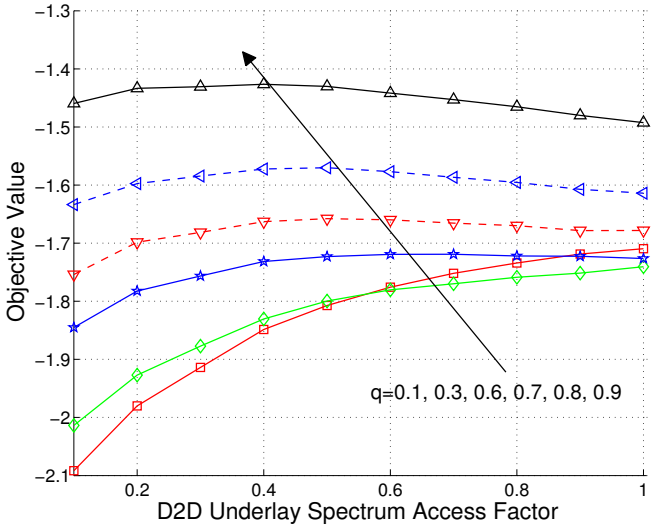

Figure 3.10: Utility value vs. D2D spectrum access factor β under different values of q, the proportion of potential D2D UEs.

outage probability $\epsilon_{\rm d}$. Then D2D coverage probability must satisfy

$$\mathbb{P}(\frac{W}{I_{\rm d}+N_0} \geq \theta_{\rm d}) = \mathbb{E}[e^{-\beta B\theta_{\rm d}(I_{\rm d}+N_0)}] \geq 1-\epsilon_{\rm d}$$

$$\Rightarrow \quad \left(N_0 B\theta_{\rm d} + \theta_{\rm d}^{\frac{2}{\alpha}} c(\mu)\right)\beta + \frac{\theta_{\rm d}^{\frac{2}{\alpha}}}{2\mathrm{sinc}(\frac{2}{\alpha})}\beta^{\frac{2}{\alpha}} \leq \log(\frac{1}{1-\epsilon_{\rm d}}), \tag{3.32}$$

where $\theta_{\rm d}$ is the SINR threshold for successful D2D transmissions, and $c(\mu) = c$ (given in Prop. 3.2) monotonically increases as μ increases. Inequality (3.32) reveals the tradeoff between β and μ. In particular, if each D2D transmission has access to more spectrum, i.e. larger β, the signal power is spread over wider channel bandwidth and thus the effective SINR gets "thinner" in each subchannel. This in turn implies that given link reliability requirement on $\theta_{\rm d}$ and $\epsilon_{\rm d}$, less cochannel D2D transmissions can be supported, i.e., μ has to

be decreased to make more potential D2D UEs use cellular mode rather than D2D mode.

Similarly, if cellular transmissions have a target outage probability $\epsilon_{\rm c}$, the cellular coverage probability must satisfy

$$\begin{aligned}\mathbb{P}(\frac{W}{I_{\rm c}+N_0} \geq \theta_{\rm c}) =& \mathbb{E}[e^{-\theta_{\rm c}(I_{\rm c}+N_0)}] \geq 1-\epsilon_{\rm c} \\ \Rightarrow\ \theta_{\rm c}^{\frac{2}{\alpha}} c(\mu)\beta^{1-\frac{2}{\alpha}} \leq& \log(\frac{1}{1-\epsilon_{\rm c}}) - N_0 B\theta_{\rm c} \\ & - 2\pi\lambda_{\rm b}\int_R^\infty (1-{}_2F_1(1,\frac{2}{\alpha};1+\frac{2}{\alpha};-\frac{\theta_{\rm c}}{(R/r)^\alpha}))r\ {\rm d}r. \quad (3.33)\end{aligned}$$

As in (3.32), a joint constraint on β and μ is imposed by (3.33); a tradeoff between β and μ exists. Incorporating the constraints (3.32) and (3.33) into the D2D underlay spectrum access optimization problem is an interesting topic for future work.

3.7 Overlay vs. Underlay: A Case Study

In this section we provide a case study to compare the rate performance of overlay with that of underlay. The results are shown in Figure 3.11 and Figure 3.12, in which the label "Overall" indicates the rate performance averaged across both cellular UEs and potential D2D UEs. In this case study, the percentage of D2D links equals $q(1-e^{-\xi\pi\mu^2}) = 0.2(1-e^{-\frac{10}{\pi 500^2}\pi 200^2}) = 16\%$. Even with only 16% of the links being D2D links, the overall rate increases remarkably in both overlay and underlay due to the high rates of D2D links. Figure 3.11 and Figure 3.12 further show that a small η (say, 0.3) in overlay leads to as good rate performance of potential D2D UEs as its counterpart

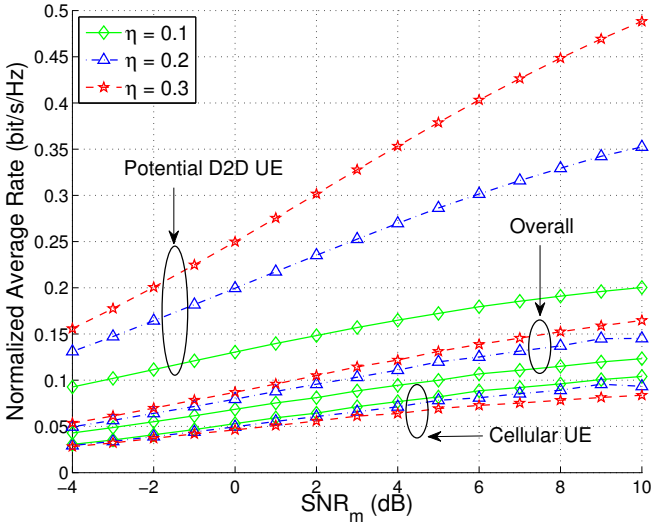

Figure 3.11: A case study on the rate performance of overlay.

in underlay with a large β (say, 0.9) because D2D UEs in overlay are free of cellular interference.

It can be observed from Figure 3.11 that the rate of potential D2D UEs in overlay increases almost linearly as η increases. In contrast, Figure 3.12 shows that, as β increases, the rate of potential D2D UEs in underlay increases in a diminishing way. This is because as β increases, the interference from cellular transmissions and the mutual interference of D2D links increase and thus the received SINR degrades. Meanwhile, as η (resp. β) increases, the rate of cellular UEs decreases due to the less spectrum resource (resp. more D2D interference). Further, the rate performance of cellular UEs is relatively sensitive to η in overlay but is robust to β in the underlay.

Recall that the higher the SNR_m, the higher the transmit powers. As

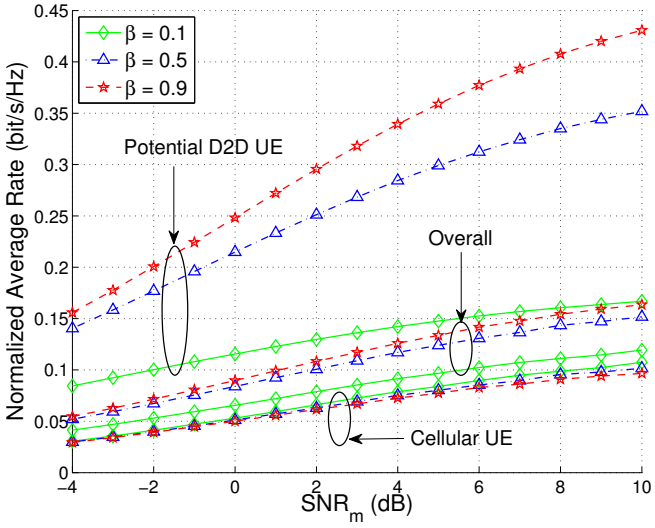

Figure 3.12: A case study on the rate performance of underlay.

SNR_m increases from -4 dB to 10 dB, the rate of cellular UEs in both overlay and underlay and the rate of potential D2D UEs in overlay increase linearly, implying that the transmit powers are not high enough to make the performance interference-limited. In contrast, the rate of potential D2D UEs in underlay increases in a diminishing way, especially when SNR_m exceeds 4 dB, implying that the performance is gradually limited by the interference caused by the increased transmit powers.

We summarize the main lessons drawn from the above discussion as follows. In underlay, the rate of potential D2D UEs is more resource-limited; a linear increase in the spectrum resource can generally lead to a linear rate increase. In contrast, the rate of potential D2D UEs is more interference-limited, mainly due to the cochannel cellular interference. As for the rate of

cellular UEs, it is sensitive to the reduction of spectrum in overlay but is more robust to the cochannel D2D interference in underlay.

3.8 Summary

In this chapter we have jointly studied D2D spectrum sharing and mode selection using a hybrid network model and a unified analytical approach. Two scenarios, overlay and underlay, have been investigated. Though spectrum sharing has been mainly studied from a rate perspective, we also show in the underlay case how to apply the derived results to study spectrum sharing from a coverage perspective, leading to the discovery of the tradeoff between the underlay D2D spectrum access and mode selection.

Note that, though we examine D2D spectrum sharing from a frequency domain perspective, it is straightforward to interpret the derived results in this chapter from a two-dimensional time-frequency perspective. Take the overlay for example: the equivalent interpretation is that a proportion η of OFDMA resource blocks is assigned to D2D while the remaining resource blocks are used by cellular.

3.9 Appendix

3.9.1 Proof of Lemma 3.1

Recall that we adopt an approximate approach on the uplink analysis. In particular, we approximate the coverage region of a hexagonal macrocell as a ball with radius $R = \sqrt{\frac{1}{\pi\lambda_{\mathrm{b}}}}$ and assume that the typical active cellular

transmitter is uniformly distributed in $\mathcal{A}$. Thus, $\mathbb{P}(L_{\rm c} \le x) = x^2/R^2, 0 \le x \le R$. Taking the derivative and using $R = \sqrt{\frac{1}{\pi\lambda_{\rm b}}}$, we obtain that the PDF of the length of a typical cellular link is $f_{L_{\rm c}}(x) = 2\pi\lambda_{\rm b} x \cdot \mathbb{I}_{x\in[0,\frac{1}{\sqrt{\pi\lambda_{\rm b}}}]}$. Then the average transmit power of a cellular transmitter equals

$$\mathbb{E}[P_{\rm c}] = \mathbb{E}[L_{\rm c}^{\alpha}] = \int_0^{1/\sqrt{\pi\lambda_{\rm b}}} 2\pi\lambda_{\rm b} x^{\alpha+1}\, {\rm d}x = \frac{1}{(1+\frac{\alpha}{2})\pi^{\frac{\alpha}{2}}\lambda_{\rm b}^{\frac{\alpha}{2}}}. \tag{3.34}$$

The PDF of the length of a typical D2D link is

$$f_{L_{\rm d}|D<\mu}(x) = \begin{cases} \dfrac{f_D(x)}{\mathbb{P}(D<\mu)} = \dfrac{2\pi\xi x e^{-\xi\pi x^2}}{1-e^{-\xi\pi\mu^2}} & \text{if } x \in [0,\mu); \\ 0 & \text{otherwise.} \end{cases}$$

Correspondingly, its α-th moment can be computed as follows:

$$\begin{aligned} \mathbb{E}[L_{\rm d}^{\alpha}|D<\mu] =& \frac{1}{\mathbb{P}(D<\mu)} \int_0^{\mu} 2\pi\xi x^{\alpha+1} e^{-\xi\pi x^2}\, {\rm d}x \\ =& \frac{(\xi\pi)^{-\frac{\alpha}{2}}}{1-e^{-\xi\pi\mu^2}} \gamma(\frac{\alpha}{2}+1, \xi\pi\mu^2), \end{aligned} \tag{3.35}$$

where we have used $\mathbb{P}(D<\mu) = 1-e^{-\xi\pi\mu^2}$ in the last equality. By definition,

$$\mathbb{E}[P_{\rm d}] = \mathbb{P}(D \ge \mu)\mathbb{E}[L_{\rm c}^{\alpha}] + (1-\mathbb{P}(D\ge\mu))\mathbb{E}[L_{\rm d}^{\alpha}|D<\mu]. \tag{3.36}$$

Substituting $\mathbb{E}[L_{\rm c}^{\alpha}]$ and $\mathbb{E}[L_{\rm d}^{\alpha}|D<\mu]$ into the above equation completes the proof.

3.9.2 Proof of Proposition 3.1

As in the proof of Lemma 3.1, the average transmit power of a potential D2D UE equals $\mathbb{E}[P_{\rm c}]$ conditioned on cellular mode is used; otherwise, i.e., conditioned on D2D mode is used, it equals $\mathbb{E}[\hat{P}_{\rm d}]$. Conditioning on D2D

mode, the D2D link length L_{d} is distributed as $\frac{1}{\mathbb{P}(D<\mu)} f_D(x), 0 \leq L_{\mathrm{d}} < \mu$. It follows that $\mathbb{E}[\hat{P}_{\mathrm{d}}] = \mathbb{E}[L_{\mathrm{d}}^{\alpha} | D < \mu] = \int_0^{\mu} x^{\alpha} \frac{f_D(x)}{\mathbb{P}(D<\mu)} \, \mathrm{d}x$ and thus

$$\begin{aligned}\mathbb{E}[P_{\mathrm{d}}] &= (1 - \int_0^{\mu} f_D(x) \, \mathrm{d}x) \cdot \mathbb{E}[P_{\mathrm{c}}] + \int_0^{\mu} f_D(x) \, \mathrm{d}x \cdot \int_0^{\mu} x^{\alpha} \frac{f_D(x)}{\mathbb{P}(D<\mu)} \, \mathrm{d}x \\ &= (1 - \int_0^{\mu} f_D(x) \, \mathrm{d}x) \cdot \mathbb{E}[P_{\mathrm{c}}] + \int_0^{\mu} x^{\alpha} f_D(x) \, \mathrm{d}x. \end{aligned} \tag{3.37}$$

Taking the derivative of $\mathbb{E}[P_{\mathrm{d}}]$ with respect to μ, $\frac{d}{d\mu}\mathbb{E}[P_{\mathrm{d}}] = f_D(\mu)(\mu^{\alpha} - \mathbb{E}[P_{\mathrm{c}}])$. Setting the derivative to zero, we obtain the stationary point $(\mathbb{E}[P_{\mathrm{c}}])^{1/\alpha}$. It is easy to see that $\mathbb{E}[P_{\mathrm{d}}]$ is decreasing when $\mu \in [0, (\mathbb{E}[P_{\mathrm{c}}])^{1/\alpha})$ and is increasing when $\mu \in [(\mathbb{E}[P_{\mathrm{c}}])^{1/\alpha}, \infty)$. Hence, $\mathbb{E}[P_{\mathrm{d}}]$ is minimized at $\mu^{\star} = (\mathbb{E}[P_{\mathrm{c}}])^{1/\alpha}$. The proof is complete by plugging the explicit expression for $\mathbb{E}[P_{\mathrm{c}}]$ (given in (3.8)).

3.9.3 Proof of Proposition 3.2

Consider the conditional Laplace transform

$$\begin{aligned}\mathcal{L}_{I_{\mathrm{d}}}(s) &= \mathbb{E}[\exp(-s \sum_{X_i \in \Phi_{\mathrm{d}} \setminus \{o\}} \hat{P}_{\mathrm{d},i} G_i \|X_i\|^{-\alpha}) | o \in \Phi_{\mathrm{d}}] \\ &= \mathbb{E}^{!o}[\exp(-s \sum_{X_i \in \Phi_{\mathrm{d}}} \hat{P}_{\mathrm{d},i} G_i \|X_i\|^{-\alpha})] \\ &= \mathbb{E}[\exp(-s \sum_{X_i \in \Phi_{\mathrm{d}}} \hat{P}_{\mathrm{d},i} G_i \|X_i\|^{-\alpha})] \\ &= \exp\left(-2\pi\kappa\lambda_{\mathrm{d}} \int_0^{\infty} (1 - \mathbb{E}[\exp(-s\hat{P}_{\mathrm{d}} G r^{-\alpha})]) r \, \mathrm{d}r\right) \\ &= \exp\left(-\frac{\pi\kappa\lambda_{\mathrm{d}}}{\mathrm{sinc}(\frac{2}{\alpha})} \mathbb{E}[\hat{P}_{\mathrm{d}}^{\frac{2}{\alpha}}] s^{\frac{2}{\alpha}}\right), \end{aligned} \tag{3.38}$$

where $\mathbb{E}^{!o}[\cdot]$ denotes the expectation with respect to the reduced Palm distribution, the third equality follows from Slivnyak's theorem [122], the fourth

equality is due to the probability generating functional of PPP [16], and we have used $G \sim \text{Exp}(1)$ in the last equality. Note that $\lambda_{\text{d}} = q\lambda(1-e^{-\xi\pi\mu^2})$, and from the proof of Lemma 3.1,

$$\mathbb{E}[\hat{P}_{\text{d}}^{\frac{2}{\alpha}}] = \mathbb{E}[L_{\text{d}}^2] = \frac{1}{\xi\pi} - \frac{\mu^2 e^{-\xi\pi\mu^2}}{1-e^{-\xi\pi\mu^2}}. \tag{3.39}$$

Plugging λ_{d} and $\mathbb{E}[\hat{P}_{\text{d}}^{\frac{2}{\alpha}}]$ into $\mathcal{L}_{I_{\text{d}}}(s)$ yields $\mathcal{L}_{I_{\text{d}}}(s) = e^{-cs^{\frac{2}{\alpha}}}$, where c is given in Prop. 3.2. Invoking (3.6) yields the spectral efficiency $R_{\text{d}} = \kappa\mathbb{E}[\log(1+\text{SINR})]$ of D2D links.

The CCDF of the SINR of D2D links can be obtained as follows:

$$\mathbb{P}(\text{SINR} \geq x) = \mathbb{P}\left(G_o \geq x(I_{\text{d}} + N_0)\right) = \mathbb{E}[e^{-x(I_{\text{d}}+N_0)}] = e^{-xN_0}\mathcal{L}_{I_{\text{d}}}(x). \tag{3.40}$$

Plugging $\mathcal{L}_{I_{\text{d}}}(x)$ into (3.40) completes the proof.

3.9.4 Proof of Proposition 3.3

The spectral efficiency R_{c} of cellular links is given by

$$R_{\text{c}} = \mathbb{E}^o[\frac{1}{N}\log(1+\text{SINR})] = \mathbb{E}^o[\frac{1}{N}] \cdot \mathbb{E}[\log(1+\text{SINR})], \tag{3.41}$$

where N is the random number of potential uplink transmitters located in the cell. Due to the PPP assumption, N is a Poisson random variable with parameter $\lambda_{\text{c}}/\lambda_{\text{b}}$. Denoting by $\tilde{N}$ the number of *other* potential uplink transmitters located in the cell except the one under consideration, i.e., $N = 1 + \tilde{N}$. Thus,

$$\mathbb{E}^o[\frac{1}{N}] = \sum_{n=1}^{\infty} \frac{1}{n} \cdot \mathbb{P}^o(N=n) = \sum_{n=1}^{\infty} \frac{1}{n} \cdot \mathbb{P}^o(\tilde{N} = n-1)$$

$$= \sum_{n=1}^{\infty} \frac{1}{n} \cdot \frac{(\frac{\lambda_c}{\lambda_b})^{n-1}}{(n-1)!} e^{-\frac{\lambda_c}{\lambda_b}} = \frac{\lambda_b}{\lambda_c} \sum_{n=1}^{\infty} \frac{(\frac{\lambda_c}{\lambda_b})^{n}}{n!} e^{-\frac{\lambda_c}{\lambda_b}} = \frac{\lambda_b}{\lambda_c}(1 - e^{-\frac{\lambda_c}{\lambda_b}}), \quad (3.42)$$

where the third equality is due to Slivnyak's theorem [122]: conditioning on the transmitter under consideration, the other potential uplink transmitters are still PPP distributed and thus $\tilde{N} \sim \text{Poisson}(\lambda_c/\lambda_b)$ under the Palm probability $\mathbb{P}^o$.

Next we calculate $\mathbb{E}[\log(1 + \text{SINR})]$. To this end, we first calculate the Laplace transform

$$\begin{aligned}
\mathcal{L}_{I_c}(s) &= \mathbb{E}[\exp(-s \sum_{X_i \in \Phi_{c,a} \cap \mathcal{A}^c} P_{c,i} G_i \|X_i\|^{-\alpha})] \\
&= \mathbb{E}[\exp(-s \sum_{X_i \in \Phi_{c,a}} P_{c,i} G_i \|X_i\|^{-\alpha} \mathbb{I}_{\{\|X_i\| \geq R\}})] \\
&= \mathbb{E}[\prod_{X_i \in \Phi_{c,a}} \exp(-s P_{c,i} G_i \|X_i\|^{-\alpha} \mathbb{I}_{\{\|X_i\| \geq R\}})] \\
&= \exp(-\int_0^{2\pi} \int_0^{\infty} 1 - \mathbb{E}[e^{-s P_c G r^{-\alpha} \mathbb{I}_{\{r \geq R\}}}] \lambda_b r \, \mathrm{d}r \, \mathrm{d}\theta) \\
&= \exp(-2\pi\lambda_b \int_R^{\infty} (1 - \mathbb{E}[e^{-s P_c G r^{-\alpha}}]) r \, \mathrm{d}r) \\
&= \exp(-2\pi\lambda_b \int_R^{\infty} (1 - \mathbb{E}[e^{-s L_c^{\alpha} G r^{-\alpha}}]) r \, \mathrm{d}r), \qquad (3.43)
\end{aligned}$$

where we have used the probability generating functional of PPP in the fourth equality, and $P_c = L_c^{\alpha}$ in the last equality. Using that L_c is distributed as $f_{L_c}(x) = 2\pi\lambda_b x \mathbb{I}(x \in [0, 1/\sqrt{\pi\lambda_b}])$, and $G \sim \text{Exp}(1)$,

$$\begin{aligned}
\mathbb{E}_{G,L_c}[e^{-x L_c^{\alpha} G r^{-\alpha}}] &= \mathbb{E}_{L_c}\left[\frac{1}{1 + x L_c^{\alpha} r^{-\alpha}}\right] = 2\pi\lambda_b \int_0^{\sqrt{\frac{1}{\pi\lambda_b}}} \frac{t}{1 + x r^{-\alpha} t^{\alpha}} \, \mathrm{d}t \\
&= {}_2F_1(1, \frac{2}{\alpha}; 1 + \frac{2}{\alpha}; -\frac{x}{(r\sqrt{\pi\lambda_b})^{\alpha}}).
\end{aligned}$$

Plugging $\mathcal{L}_{I_\mathrm{c}}(s)$ into (3.6) yields

$$\mathbb{E}[\log(1+\mathrm{SINR})] = \int_0^\infty \frac{e^{-N_0 x}}{1+x} \times$$
$$\exp(-2\pi\lambda_\mathrm{b} \int_R^\infty (1 - {}_2F_1(1, \frac{2}{\alpha}; 1+\frac{2}{\alpha}; -\frac{x}{(r\sqrt{\pi\lambda_\mathrm{b}})^\alpha}))r \ \mathrm{d}r) \ \mathrm{d}x. \tag{3.44}$$

Plugging (3.42) and (3.44) into (3.41) gives the spectral efficiency R_c of cellular links. The CCDF of the SINR of cellular links can be similarly obtained as in (3.40).

3.9.5 Proof of Corollary 3.2

For a sparse cellular network with small λ_b, R is large and thus for $r \in [R, \infty)$,

$$1 - \mathbb{E}[e^{-xGL_\mathrm{c}^\alpha r^{-\alpha}}] \approx xr^{-\alpha}\mathbb{E}[GL_\mathrm{c}^\alpha] = xr^{-\alpha}\mathbb{E}[G]\mathbb{E}[P_\mathrm{c}] = \frac{xr^{-\alpha}}{(\frac{\alpha}{2}+1)(\pi\lambda_\mathrm{b})^{\frac{\alpha}{2}}}, \tag{3.45}$$

where we have used the approximation $1 - e^{-y} \approx y$ for small value y, the independence of fading G and link length L_c and $\mathbb{E}[P_\mathrm{c}] = \mathbb{E}[L_\mathrm{c}^\alpha]$ in the first equality, and have plugged in $\mathbb{E}[G] = 1$ and $\mathbb{E}[L_\mathrm{c}^\alpha] = \frac{1}{(\frac{\alpha}{2}+1)(\pi\lambda_\mathrm{b})^{\frac{\alpha}{2}}}$ (c.f. Lemma 3.1) in the last equality. Accordingly, the Laplace transform of the uplink interference is given by

$$\mathcal{L}_{I_\mathrm{c}}(s) = \exp(-2\pi\lambda_\mathrm{b} \int_R^\infty (1 - \mathbb{E}[e^{-sGL_\mathrm{c}^\alpha r^{-\alpha}}])r \ \mathrm{d}r)$$
$$\approx \exp(-2\pi\lambda_\mathrm{b} s \int_R^\infty \frac{r^{-\alpha+1}}{(\frac{\alpha}{2}+1)(\pi\lambda_\mathrm{b})^{\frac{\alpha}{2}}} \ \mathrm{d}r) = \exp(-\frac{4}{\alpha^2-4}s). \tag{3.46}$$

For a dense network with large λ_b, R is small and thus

$$\mathcal{L}_{I_\mathrm{c}}(s) \approx \exp(-2\pi\lambda_\mathrm{b} \int_0^\infty (1 - \mathbb{E}[e^{-sGL_\mathrm{c}^\alpha r^{-\alpha}}])r \ \mathrm{d}r)$$

$$= \exp(-\frac{\pi\lambda_{\mathrm{b}}}{\mathrm{sinc}(\frac{2}{\alpha})}\mathbb{E}[P_{\mathrm{c}}^{\frac{2}{\alpha}}]s^{\frac{2}{\alpha}}) = \exp(-\frac{1}{2\mathrm{sinc}(\frac{2}{\alpha})}s^{\frac{2}{\alpha}}), \tag{3.47}$$

where we have plugged in $\mathbb{E}[P_{\mathrm{c}}^{\frac{2}{\alpha}}] = \mathbb{E}[L_{\mathrm{c}}^2] = \frac{1}{2\pi\lambda_{\mathrm{b}}}$ (c.f. (3.34)) in the last equality. Combining the above asymptotic results with Prop. 3.3 completes the proof.

3.9.6 Proof of Proposition 3.4

Note that

$$\begin{aligned}\arg\max_{\eta\in[0,1]} w_c \log T_{\mathrm{c}} + w_d \log T_{\mathrm{d}} &= \arg\max_{\eta\in[0,1]} \log T_{\mathrm{c}}^{w_c} \cdot T_{\mathrm{d}}^{w_d} \\ &= \arg\max_{\eta\in[0,1]} g(\eta),\end{aligned} \tag{3.48}$$

where $g(\eta) = \log(1-\eta)^{w_c}((1-\eta)e^{-\xi\pi\mu^2}R_c + \eta(1-e^{-\xi\pi\mu^2})R_d)^{w_d}$. For ease of notation we let $Q_c = e^{-\xi\pi\mu^2}R_c$ and $Q_d = (1-e^{-\xi\pi\mu^2})R_d$. Taking derivative of $g(\eta)$ with respect to η,

$$\begin{aligned}\frac{d}{d\eta}g(\eta) =& \frac{(1-\eta)^{w_c-1}(1-\eta)Q_c + \eta Q_d)^{w_d-1}}{(1-\eta)^{w_c}((1-\eta)Q_c + \eta Q_d)^{w_d}} \\ &\times (w_c+w_d)\left((Q_c-Q_d)\eta - (Q_c - \frac{w_d}{w_c+w_d}Q_d)\right).\end{aligned} \tag{3.49}$$

If $Q_c - Q_d > 0$, the stationary point $\eta^\dagger = \frac{Q_c - \frac{w_d}{w_c+w_d}Q_d}{Q_c-Q_d} = 1 - \frac{w_c}{w_c+w_d}\frac{Q_d}{Q_d-Q_c} \geq 1$, and $g(\eta)$ monotonically decreases on $\eta \leq \eta^\dagger$ and monotonically increases on $\eta > \eta^\dagger$. Hence, $\eta^\star = 0$. If $Q_c - Q_d = 0$, $\frac{d}{d\eta}g(\eta) < 0$ and $g(\eta)$ monotonically decreases. Thus, $\eta^\star = 0$. If $Q_c - Q_d < 0$, $g(\eta)$ monotonically increases on $\eta \leq \eta^\dagger$ and monotonically decreases on $\eta > \eta^\dagger$. Note the stationary point $\eta^\dagger = 1 - \frac{w_c}{w_c+w_d}\frac{Q_d}{Q_d-Q_c} < 1$. Hence, $\eta^\star = \max(0, \eta^\dagger)$. Also, $\eta^\dagger \leq 0$ if and only if

$(w_c + w_d)Q_c \geq w_d Q_d$, which implies $\eta^\star = 0$ if and only if $(w_c + w_d)Q_c \geq w_d Q_d$. To sum up,

$$\eta^\star = \begin{cases} 1 - \frac{w_c}{w_c+w_d}\frac{Q_d}{Q_d - Q_c} & \text{if } Q_c < \frac{w_d}{w_c+w_d} Q_d; \\ 0 & \text{otherwise.} \end{cases}$$

Plugging the explicit expressions of Q_c and Q_d complete the proof.

3.9.7 Proof of Proposition 3.5

Note that the average D2D transmit power $\mathbb{E}[\hat{P}_{\mathrm{d}}]$ here is only $1/\beta B$ of the one given in Lemma 3.1 as each D2D transmitter accesses βB subchannels and needs to split its power accordingly, i.e., $\hat{P}_{\mathrm{d}} = \frac{1}{\beta B} L_{\mathrm{d}}^{\alpha}$. Hence, $\mathbb{E}[\hat{P}_{\mathrm{d}}^{\frac{2}{\alpha}}] = (\frac{1}{\beta B})^{\frac{2}{\alpha}} \mathbb{E}[L_{\mathrm{d}}^2]$. The corresponding spectral efficiency (normalized by bandwidth B) is given by

$$\begin{aligned} R_{\mathrm{d}} =& \kappa \mathbb{E}[\log(1 + \mathrm{SINR})] = \kappa \int_0^\infty \frac{e^{-N_0 x}}{1+x} \cdot \mathcal{L}_{I_{\mathrm{d}}}(\beta B x)\, \mathrm{d}x \\ =& \kappa \int_0^\infty \frac{e^{-N_0 x}}{1+x} e^{-\frac{\pi \kappa \lambda_{\mathrm{d}}}{\mathrm{sinc}(\frac{2}{\alpha})} \mathbb{E}[\hat{P}_{\mathrm{d}}^{\frac{2}{\alpha}}] (\beta B x)^{\frac{2}{\alpha}}} e^{-\frac{\pi \lambda_{\mathrm{b}}}{\mathrm{sinc}(\frac{2}{\alpha})} \mathbb{E}[P_{\mathrm{c}}^{\frac{2}{\alpha}}] (\beta B x)^{\frac{2}{\alpha}}}\, \mathrm{d}x. \end{aligned} \tag{3.50}$$

Here $\lambda_{\mathrm{d}} = \beta \cdot q\lambda(1 - e^{-\xi\pi\mu^2})$, the density of D2D transmitters that is "seen" from each subchannel, and from the proof in Appendix 3.9.1,

$$\begin{aligned} \mathbb{E}[\hat{P}_{\mathrm{d}}^{\frac{2}{\alpha}}] &= (\frac{1}{\beta B})^{\frac{2}{\alpha}} \mathbb{E}[L_{\mathrm{d}}^2] = (\frac{1}{\beta B})^{\frac{2}{\alpha}} \left(\frac{1}{\xi\pi} - \frac{\mu^2 e^{-\xi\pi\mu^2}}{1 - e^{-\xi\pi\mu^2}} \right) \\ \mathbb{E}[P_{\mathrm{c}}^{\frac{2}{\alpha}}] &= (\frac{1}{B})^{\frac{2}{\alpha}} \mathbb{E}[L_{\mathrm{c}}^2] = (\frac{1}{B})^{\frac{2}{\alpha}} \frac{1}{2\pi\lambda_{\mathrm{b}}}. \end{aligned}$$

Plugging λ_{d}, $\mathbb{E}[\hat{P}_{\mathrm{d}}^{\frac{2}{\alpha}}]$ and $\mathbb{E}[P_{\mathrm{c}}^{\frac{2}{\alpha}}]$ into (3.50) establishes the expression for the D2D link spectral efficiency. The CCDF of the SINR of D2D links can be similarly obtained as in (3.40).

3.9.8 Proof of Proposition 3.6

As in the proof of Prop. 3.5, the Laplace transform of the interference can be calculated as follows.

$$\mathcal{L}_{I_{\mathrm{c}}}(s) = e^{-\frac{\pi\kappa\lambda_{\mathrm{d}}}{\mathrm{sinc}(\frac{2}{\alpha})}\mathbb{E}[\hat{P}_{\mathrm{d}}^{\frac{2}{\alpha}}]s^{\frac{2}{\alpha}} - 2\pi\lambda_{\mathrm{b}}\int_R^\infty (1 - {}_2F_1(1,\frac{2}{\alpha};1+\frac{2}{\alpha};-\frac{s/B}{(r\sqrt{\pi\lambda_{\mathrm{b}}})^\alpha}))r\ \mathrm{d}r}. \tag{3.51}$$

Then as in the proof of Prop. 3.3, the spectral efficiency R_{c} of cellular links is given by

$$R_{\mathrm{c}} = \mathbb{E}^o[\frac{1}{N}\log(1+\mathrm{SINR})] = \frac{\lambda_{\mathrm{b}}}{\lambda_{\mathrm{c}}}(1 - e^{-\frac{\lambda_{\mathrm{c}}}{\lambda_{\mathrm{b}}}})\int_0^\infty \frac{e^{-N_0 x}}{1+x}\mathcal{L}_{I_{\mathrm{c}}}(Bx)\ \mathrm{d}x. \tag{3.52}$$

Plugging $\mathcal{L}_{I_{\mathrm{c}}}(s)$ (3.51) into the above equation establishes the expression for the cellular link spectral efficiency. The CCDF of the SINR of D2D links can be similarly obtained as in (3.40).

Chapter 4

Massive MIMO Systems with D2D Underlay

Recently, there has been a surge of interest in massive MIMO, mostly due to the work [99]. In a massive MIMO system, each BS uses a very large antenna array to serve multiple users in each time-frequency resource block [99]. If the number of antennas at the BS is significantly larger than the number of served users, the channel of each user to/from the BS is nearly orthogonal to that of any other user. This allows for very simple transmit or receive processing techniques like matched filtering to be nearly optimal with enough antennas even in the presence of interference.

In this chapter, we study the interplay between massive MIMO and underlaid D2D networking. In a D2D underlaid cellular network, the uplink spectrum is reused by the D2D transmissions, causing mutual interference with the ongoing cellular transmissions. Massive MIMO is appealing in such a context as the BS's large antenna array can nearly null the D2D-to-BS interference. The multi-user transmission in massive MIMO, however, may lead to increased cellular-to-D2D interference. To protect D2D links, the number of simultaneously active uplink users might have to be limited, eating into massive MIMO gain. It is not *a priori* clear to what extent the D2D signals would

be affected by the multiuser transmission and the tradeoff between supporting D2D communication and scaling up the uplink capacity in a massive MIMO system. Further, if cochannel D2D signals are present when estimating massive MIMO channels, the estimated CSI would become less accurate, which may hurt massive MIMO performance. It is not *a priori* clear however to what extent the D2D signals would affect the channel estimation and consequently the performance of the massive MIMO system.

4.1 Related Work

Extensive investigation on MIMO has been carried out since the point-to-point MIMO channel capacity was analyzed in [43, 125]. A natural shift from the point-to-point MIMO channel is to study the role of MIMO in multi-user channels. Much progress has been made, especially in multiple access channel and broadcast channel whose information theoretical capacities have been characterized [25, 135, 136, 140]. In the context of ad hoc networks, early work like [27, 113] mainly used simulation to study the performance gains of MIMO and drew implications accordingly, while analytical studies explicitly considering random node distribution may be found in [9, 48]. Notably, based on the assumption that nodes are distributed according to a PPP, stochastic geometry has been widely used to characterize the transmission capacity of ad hoc networks under different MIMO techniques: Single stream transmission [66, 69, 74], multi-stream transmission [96, 131] and space division multiple access (SDMA) [80].

In the context of cellular networks, many MIMO techniques have been proposed and investigated. These include antenna selection [62], interference cancellation [11], mulituser detection [30, 133], networked MIMO [132, 149], distributed antenna architectures [28, 61], and massive MIMO [99]. Massive MIMO, the focus of this chapter, is not a brand new concept. Historically, there has been much academic interest in applying random matrix theory to investigate the asymptotic MIMO performance [43, 128] or the isomorphic multi-user detection problem in code division multiple access (CDMA) systems with random spreading sequences [127, 134]. The recent widespread interest in massive MIMO is mostly due to [99], where several practically important system aspects such as pilot contamination have been identified. The work [99] has stimulated a surge of interest in massive MIMO (see [21, 23, 64, 68, 75, 97, 103, 115] and references therein).

In the context of cellular networks with D2D networking, existing research is mainly focused on single-antenna networks (see e.g. [20, 73, 78, 89, 91, 146, 147]) while research on the use of antenna arrays has just begun [71, 72, 84, 101, 124]. To mitigate or avoid mutual interference between cellular and D2D transmissions, [71, 124] considered precoding while [72, 84] studied various relaying strategies. In contrast, [101] proposed not to schedule uplink users that may generate excessive interference to D2D users. How D2D MIMO and cellular MIMO interact, especially in the massive MIMO context, is still largely open.

4.2 Contributions and Main Outcomes

The main contributions and outcomes of this chapter are summarized as follows.

A tractable multi-antenna hybrid network model. We extend our previous single-antenna hybrid network model used in Chapters 2 and 3 to multi-antenna transmission. We consider a multi-cell setting and focus on the uplink. The spatial positions of the underlaid D2D transmitters are modeled by a PPP. All the transmissions (both cellular and D2D) in this model are SIMO (i.e., single-input multiple-output) with each BS having a very large antenna array. For the receive processing, we extend the partial zero-forcing (PZF) receiver studied in ad hoc networks [74] to the hybrid network in question. Spectral efficiency is used as the sole metric throughout this chapter.

Spectral efficiency with perfect CSI. In the asymptotic regime where the number of BS antennas $M \to \infty$ and with perfect CSI, we find that the received signal-to-interference-plus-noise ratio (SINR) of any cellular user increases unboundedly and the effects of noise, fast fading, and the interfering signals from the other co-channel cellular users and the *infinite* D2D transmitters vanish completely. Equivalently, it is possible to reduce cellular transmit power as $\Theta(1/M)$ but still achieve a non-vanishing cellular spectral efficiency, as in the case without D2D underlay [103]. Compared to the case without D2D, with scaled cellular transmit power $\Theta(1/M)$, there is a loss in cellular spectral efficiency if a constant number of D2D interfering signals is canceled. The loss can be overcome if the number of canceled D2D interfering signals is

scaled appropriately (e.g., $\Theta(\sqrt{M})$). In the non-asymptotic regime, we derive simple analytical lower bounds for both cellular and D2D spectral efficiency; the derived bounds allow for very efficient numerical evaluation.

Spectral efficiency with imperfect CSI. We study pilot-based CSI estimation in which known training sequences are transmitted and the receivers use minimum mean squared error (MMSE) estimators for channel estimation. In the asymptotic regime with the estimated CSI, it is known that the received SINR of any cellular user is bounded due to pilot contamination [99]. With D2D underlay, the bounded SINR is further degraded due to a new asymptotic effect which we term *underlay contamination*. Due to the underlay contamination, we find that scaling down cellular transmit power results in vanishing cellular spectral efficiency, no matter how slow the scaling rate is. This is dramatically different from the case without D2D underlay, for which [103] shows that cellular transmit power can be scaled down as $\Theta(1/\sqrt{M})$. To recover the power scaling law $\Theta(1/\sqrt{M})$, one possible approach is to deactivate the D2D links in the training phase of massive MIMO; however, compared to the case without D2D, there is a loss in cellular spectral efficiency due to D2D-to-cellular interference in the data transmission phase. Instead, if the cellular transmit power is not scaled down and D2D links are deactivated in the training phase, massive MIMO automatically eliminates the effect of D2D-to-cellular interference in the data transmission phase.

4.3 Mathematical Models

4.3.1 Network Model

Consider a multi-cell D2D underlaid massive MIMO system shown in Fig. 4.1. In this system, there are $B+1$ cells; in each cell $b, b = 0, 1, ..., B$, K cellular user equipments (UEs) transmit to the BS b. We assume that the K cellular UEs are uniformly distributed in each cell; this assumption is not essential in the analysis but will be used in the simulation. We denote by $\mathcal{K}_b$ the set of the K cellular UEs in the cell b, and $\mathcal{C}_b$ the coverage area of the cell b. We assume that $\mathcal{C}_b \cap \mathcal{C}_{b'} = \emptyset, \forall b \neq b'$.

The cellular system is underlaid with D2D UEs. The locations of D2D transmitters are distributed as a homogeneous PPP Φ with density λ. We partition Φ into $B+2$ disjoint PPPs $\Phi_0, ..., \Phi_{B+1}$, where $\Phi_b = \Phi \cup \mathcal{C}_b, \forall b = 0, ..., B$, and $\Phi_{B+1} = \Phi \backslash \cup_{i=0}^{B} \Phi_i$. Each D2D receiver is located at a random distance of D meters from its associated D2D transmitter with uniformly distributed direction.

We focus on SIMO in this chapter, i.e., a transmitter (either cellular or D2D) uses one antenna for transmission, while a BS and a D2D receiver respectively use M and N antennas for receiving. Note that we are interested in the performance regime where M is large and thus the assumption $M \gg K$ is made throughout this chapter.

In this system, all the transmitters use the same time-frequency resource block, leading to cochannel interference. We assume that cellular and D2D UEs transmit at constant powers P_{c} and P_{d} respectively.

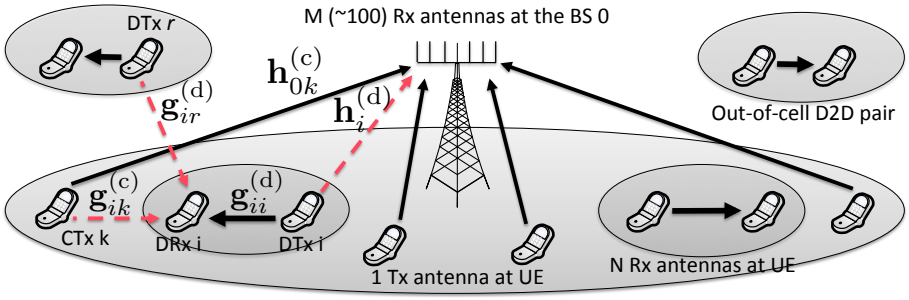

Figure 4.1: A D2D underlaid massive MIMO system consisting of both cellular and D2D links. For clarity, only the central cell is shown. D2D pairs located outside of the cells are out of cellular coverage but still contribute to the total aggregate D2D interference.

4.3.2 Baseband Channel Models

Without loss of generality, we focus on the central cell, whose BS is indexed by $b = 0$ and located at the origin. This helps simplify the notation. The $M \times 1$ dimensional baseband received signal at the central BS is

$$\begin{aligned}\mathbf{y}_0^{(c)} = &\sum_{b=0}^{B} \sum_{k \in \mathcal{K}_b} \sqrt{P_c \Xi_{bk}^{(c)}} \|x_{bk}^{(c)}\|^{-\frac{\alpha_c}{2}} \mathbf{h}_{bk}^{(c)} u_{bk}^{(c)} \\ &+ \sum_{i \in \Phi} \sqrt{P_d \Xi_i^{(d)}} \|x_i^{(d)}\|^{-\frac{\alpha_c}{2}} \mathbf{h}_i^{(d)} u_i^{(d)} + \mathbf{v}_0^{(c)}, \end{aligned} \tag{4.1}$$

where $\Xi_{bk}^{(c)}$ denotes the shadowing from cellular transmitter k in the cell b to the BS 0, $x_{bk}^{(c)}$ denotes the position of cellular transmitter k in the cell b, $\alpha_c > 2$ denotes the pathloss exponent of UE-BS links, $\mathbf{h}_{bk}^{(c)} \in \mathcal{C}^{M \times 1}$ is the vector channel from cellular transmitter k in the cell b to the BS 0, $u_{bk}^{(c)}$ denotes the zero-mean unit-variance transmit symbol of cellular transmitter k in the cell b, $\Xi_i^{(d)}, x_i^{(d)}, \mathbf{h}_i^{(d)} \in \mathcal{C}^{M \times 1}$ and $u_i^{(d)}$ are similarly defined for D2D transmitter i,

and $\mathbf{v}_0^{(\mathrm{c})} \in \mathcal{C}^{M\times 1}$ is complex Gaussian noise with covariance $N_0\mathbf{I}_M$ at the BS 0 with $\mathbf{I}_M$ denoting the M dimensional identity matrix.

Similarly, the $N \times 1$ dimensional baseband received signal at the D2D receiver r is

$$\begin{aligned}\mathbf{y}_r^{(\mathrm{d})} =& \sum_{b=0}^{B}\sum_{k\in\mathcal{K}_b}\sqrt{P_{\mathrm{c}}\Xi_{rbk}^{(\mathrm{c})}}(d_{rbk}^{(\mathrm{c})})^{-\frac{\alpha_{\mathrm{d}}}{2}}\mathbf{g}_{rbk}^{(\mathrm{c})}u_{bk}^{(\mathrm{c})} \\ &+\sum_{i\in\Phi}\sqrt{P_{\mathrm{d}}\Xi_{ri}^{(\mathrm{d})}}(d_{ri}^{(\mathrm{d})})^{-\frac{\alpha_{\mathrm{d}}}{2}}\mathbf{g}_{ri}^{(\mathrm{d})}u_i^{(\mathrm{d})}+\mathbf{v}_r^{(\mathrm{d})}, \end{aligned} \tag{4.2}$$

where $\Xi_{rbk}^{(\mathrm{c})}, \Xi_{ri}^{(\mathrm{d})}$ are the shadowing from cellular transmitter k in the cell b to D2D receiver r and from D2D transmitter i to D2D receiver r respectively, $d_{rbk}^{(\mathrm{c})} = \|x_{bk}^{(\mathrm{c})} - z_r^{(\mathrm{d})}\|$ and $d_{ri}^{(\mathrm{d})} = \|x_i^{(\mathrm{d})} - z_r^{(\mathrm{d})}\|$ with $z_r^{(\mathrm{d})}$ denoting the position of D2D receiver r, $\alpha_{\mathrm{d}} > 2$ denotes the pathloss exponent of UE-UE links, $\mathbf{g}_{rbk}^{(\mathrm{c})}, \mathbf{g}_{ri}^{(\mathrm{d})} \in \mathcal{C}^{N\times 1}$ are the vector channels from cellular transmitter k in the cell b to D2D receiver r and from D2D transmitter i to D2D receiver r respectively, and $\mathbf{v}_r^{(\mathrm{d})} \in \mathcal{C}^{N\times 1}$ is complex Gaussian noise with covariance $N_0\mathbf{I}_N$.

Note that we have used different pathloss exponents α_{c} and α_{d} for UE-BS and UE-UE links (cf. (4.1) and (4.2)) due to their different propagation characteristics. Specifically, the antenna height of a macro BS is tens of meters, while the typical antenna height at a UE is under 2 m. As a result, both terminals of a UE-UE link are low and see similar near street scattering environment, which is different from the radio environment around a macro BS [88].

In this chapter, we assume Gaussian signaling, i.e., $\{u_{bk}^{(\mathrm{c})}\}, \{u_i^{(\mathrm{d})}\}$ are

i.i.d. complex Gaussian $\mathcal{CN}(0,1)$, and i.i.d. shadowing with mean $\bar{\Xi}$. We also assume that all the vector channels have i.i.d. $\mathcal{CN}(0,1)$ elements, independent across transmitters. It follows that the *favorable propagation* condition [115] desired in massive MIMO systems holds in our model:

$$\frac{1}{M}\mathbf{h}_{br}^{(s)*}\mathbf{h}_{b'\ell}^{(s')} \xrightarrow{a.s.} \begin{cases} 1 & \text{if } s=s',\ b=b' \text{ and } r=\ell; \\ 0 & \text{otherwise,} \end{cases}$$

where $s \in \{\mathrm{c},\mathrm{d}\}$, $\xrightarrow{a.s.}$ denotes the almost sure convergence as $M \to \infty$, and when $s = \mathrm{d}$ the first subindex b in $\mathbf{h}_{br}^{(s)}$ should be understood as null. Recent measurement campaigns have given evidence to validate favorable propagation for massive MIMO in practice [82].

4.3.3 Receive Filters

Denote by $\mathbf{w}_k^{(\mathrm{c})}$ the filter used by the central BS for receiving the signal of cellular transmitter k in the central cell, i.e., the central BS detects the symbol $u_{0k}^{(\mathrm{c})}$ based on $\mathbf{w}_k^{(\mathrm{c})*}\mathbf{y}_0^{(\mathrm{c})}$. Similarly, D2D receiver r detects the symbol $u_r^{(\mathrm{d})}$ based on $\mathbf{w}_r^{(\mathrm{d})*}\mathbf{y}_r^{(\mathrm{d})}$, where $\mathbf{w}_r^{(\mathrm{d})}$ denotes the filter used by D2D receiver r. The performance of the D2D underlaid massive MIMO system depends on the receive filters. In general, either the receive filters can be designed to boost desired signal power or they can be used to cancel undesired interference. In this chapter, we focus on a particular type of linear filters: the PZF receiver, which uses a subset of the degrees of freedom for boosting received signal power and the remainder for interference cancellation.

The central BS uses m_{c} and m_{d} degrees of freedom to cancel the interference from the *nearest* m_{c} cellular interferers and the nearest m_{d} D2D

interferers. A feasible choice of $(m_{\mathrm{c}}, m_{\mathrm{d}})$ needs to be in the following set:

$$\mathcal{Z}_{\mathrm{c}} = \{(m_{\mathrm{c}}, m_{\mathrm{d}}) \in \mathbb{N} \times \mathbb{N} : m_{\mathrm{c}} \leq (B+1)K - 1, m_{\mathrm{c}} + m_{\mathrm{d}} \leq M - 1\}. \quad (4.3)$$

The PZF filter $\mathbf{w}_k^{(\mathrm{c})}$ is the projection of the channel vector $\mathbf{h}_{0k}^{(\mathrm{c})}$ onto the subspace orthogonal to the one spanned by the channel vectors of canceled interferers. For ease of reference, we denote by $\mathcal{K}_{bk}^{(\mathrm{c})}$ the set of *uncanceled* cellular interferers in the cell b and $\Phi_k^{(\mathrm{c})}$ the set of uncanceled D2D interferers when detecting the symbol $u_{0k}^{(\mathrm{c})}$ of cellular transmitter k in the central cell.

Similarly, each D2D receiver uses n_{c} and n_{d} degrees of freedom to cancel the interference from the nearest n_{c} cellular interferers and the nearest n_{d} D2D interferers, and $(n_{\mathrm{c}}, n_{\mathrm{d}})$ needs to be in the following set:

$$\mathcal{Z}_{\mathrm{d}} = \{(n_{\mathrm{c}}, n_{\mathrm{d}}) \in \mathbb{N} \times \mathbb{N} : n_{\mathrm{c}} \leq (B+1)K, n_{\mathrm{c}} + n_{\mathrm{d}} \leq N - 1\}. \quad (4.4)$$

The PZF filter $\mathbf{w}_r^{(\mathrm{d})}$ of D2D receiver r is the projection of the channel vector $\mathbf{g}_{rr}^{(\mathrm{d})}$ onto the subspace orthogonal to the one spanned by the channel vectors of canceled interferers. For ease of reference, we denote by $\mathcal{K}_{br}^{(\mathrm{d})}$ the set of uncanceled cellular interferers in the cell b and $\Phi_r^{(\mathrm{d})}$ the set of uncanceled D2D interferers at D2D receiver r.

Remark on PZF receiver. Although suboptimal, PZF receivers have several advantages that motivate us to focus on them in this chapter. On the one hand, PZF receivers are relatively general: they reduce to maximum ratio combining (MRC) receivers when $m_{\mathrm{c}} + m_{\mathrm{d}} = 0$ and $n_{\mathrm{c}} + n_{\mathrm{d}} = 0$ and to conventional fully ZF receivers when $m_{\mathrm{c}} + m_{\mathrm{d}} = M - 1$ and $n_{\mathrm{c}} + n_{\mathrm{d}} = N - 1$.

It has been shown that PZF receivers can achieve the same scaling law in terms of transmission capacity as MMSE receivers [74], which is not true for either MRC or fully ZF receivers. On the other hand, PZF receivers are analytically more tractable than other more sophisticated receivers like MMSE receivers from a system point of view. This analytical tractability allows us to develop an explicit characterization on the performance of the massive MIMO system with D2D underlay. Nevertheless, we would like to point out that, as noted in [74], MMSE fitlers should be used in practice because they have less stringent CSI requirement while being the best linear filters.

4.4 Spectral Efficiency with Perfect Channel State Information

In this section, we derive the spectral efficiency of cellular and D2D links under the assumption of perfect CSI; the case of imperfect CSI will be treated in the next section.

4.4.1 Asymptotic Cellular Spectral Efficiency

For cellular UE k in the central cell, the post-processing SINR with the PZF filter $\mathbf{w}_k^{(\mathrm{c})}$ is

$$\mathrm{SINR}_k^{(\mathrm{c})} = \frac{S_k^{(\mathrm{c})}}{I_k^{(\mathrm{c}\to\mathrm{c})} + I_k^{(\mathrm{d}\to\mathrm{c})} + \|\mathbf{w}_k^{(\mathrm{c})}\|^2 N_0}, \tag{4.5}$$

where $S_k^{(\mathrm{c})} = P_{\mathrm{c}} \Xi_{0k}^{(\mathrm{c})} \|x_{0k}^{(\mathrm{c})}\|^{-\alpha_{\mathrm{c}}} \|\mathbf{w}_k^{(\mathrm{c})*} \mathbf{h}_{0k}^{(\mathrm{c})}\|^2$ denotes the desired signal power of cellular UE k, $I_k^{(\mathrm{c}\to\mathrm{c})}$ and $I_k^{(\mathrm{d}\to\mathrm{c})}$ respectively denote the cochannel cellular and

D2D interference powers experienced by cellular UE k and are given by

$$I_k^{(\mathrm{c}\to\mathrm{c})} = \sum_{b=0}^{B} \sum_{\ell\in\mathcal{K}_{bk}^{(\mathrm{c})}} P_{\mathrm{c}} \Xi_{b\ell}^{(\mathrm{c})} \|x_{b\ell}^{(\mathrm{c})}\|^{-\alpha_{\mathrm{c}}} |\mathbf{w}_k^{(\mathrm{c})*} \mathbf{h}_{b\ell}^{(\mathrm{c})}|^2 \tag{4.6}$$

$$I_k^{(\mathrm{d}\to\mathrm{c})} = \sum_{i\in\Phi_k^{(\mathrm{c})}} P_{\mathrm{d}} \Xi_i^{(\mathrm{d})} \|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}} |\mathbf{w}_k^{(\mathrm{c})*} \mathbf{h}_i^{(\mathrm{d})}|^2. \tag{4.7}$$

The spectral efficiency of cellular UE k in the central cell is defined as

$$R_k^{(\mathrm{c})} = \mathbb{E}\left[\log(1+\mathrm{SINR}_k^{(\mathrm{c})})\right], \tag{4.8}$$

where the expectation is taken with respect to the fast fading, shadowing and random locations of UEs.

Proposition 4.1. *With perfect CSI, as $M\to\infty$, the desired signal power $S_k^{(c)}$ when normalized by M^2 and conditioned on $\Xi_{0k}^{(c)}$ and $x_{0k}^{(c)}$ converges to*

$$\lim_{M\to\infty} \frac{1}{M^2} S_k^{(c)} \xrightarrow{a.s.} P_c \Xi_{0k}^{(c)} \|x_{0k}^{(c)}\|^{-\alpha_c}, \tag{4.9}$$

and the cellular interference power $I_k^{(c\to c)}$, the D2D interference power $I_k^{(d\to c)}$, and the noise power $\|\mathbf{w}_k^{(c)}\|^2 N_0$ when normalized by M^2 converge as follows.

$$\lim_{M\to\infty} \frac{1}{M^2} I_k^{(c\to c)} \xrightarrow{a.s.} 0, \quad \lim_{M\to\infty} \frac{1}{M^2} I_k^{(d\to c)} \xrightarrow{p.} 0, \quad \lim_{M\to\infty} \frac{1}{M^2} \|\mathbf{w}_k^{(c)}\|^2 N_0 \xrightarrow{a.s.} 0, \tag{4.10}$$

where $\xrightarrow{p.}$ denotes the convergence in probability.

Proof. See Appendix 4.8.1. □

Prop. 4.1 shows that with perfect CSI, as $M\to\infty$, the post-processing $\mathrm{SINR}_k^{(\mathrm{c})}$ increases unboundedly in probability (as almost sure convergence implies convergence in probability). More specifically, a deterministic received

power of the desired signal from cellular UE k (conditioned on its pathloss and shadowing) can be achieved and the effects of noise, fast fading, and the interfering signals from the other $K-1$ cellular UEs and the *infinite* D2D transmitters vanish completely. Therefore, Prop. 4.1 validates the intuition that with perfect CSI, D2D-to-cellular interference can be made arbitrarily small with a large enough antenna array at the BS. Perhaps the most interesting observation from Prop. 4.1 is that the D2D-to-cellular interference can be completely nulled out, though (1) the number of the PPP distributed D2D interferers is infinite and (2) the mean of the aggregate D2D interference is infinite. Further, the proof of Prop. 4.1 shows that a simple MRC filter with $m_{\mathrm{c}} = m_{\mathrm{d}} = 0$ suffices.

Though Prop. 4.1 shows that arbitrarily large received SINR and thus arbitrarily large rate (at least in theory) can be achieved with massive MIMO, it is not possible to fully exploit a very high SINR due to practical constraints such as the highest order of modulation and coding schemes. Nevertheless, the large array gains may be translated into power savings for cellular UEs: with a given SNR target we can lower the transmit powers of cellular UEs and thus improve their energy efficiency, as shown in the following proposition.

Proposition 4.2. *With perfect CSI, fixed PZF parameters (m_c, m_d), scaled cellular transmit power P_c/M, and conditioned on $\Xi_{0k}^{(c)}$ and $x_{0k}^{(c)}$, as $M \to \infty$, the spectral efficiency $R_k^{(c)}$ of cellular UE k in the central cell converges to*

$$R_k^{(c)} \to \mathbb{E}_{\Phi,\eta}\left[\log\left(1 + \frac{SNR_{0k}^{(c)}}{\sum_{i\in\Phi_k^{(c)}} \frac{P_d}{N_0}\|x_i^{(d)}\|^{-\alpha_c}\eta_i + 1}\right)\right], \tag{4.11}$$

where $SNR_{bk}^{(c)} = P_c \Xi_{bk}^{(c)} \|x_{bk}^{(c)}\|^{-\alpha_c}/N_0$, $\{\eta_i\}$ *are i.i.d. random variables distributed as* $\eta_i \sim Exp(1)$. *Further, if* $m_d + 1 > \frac{\alpha_c}{2}$,

$$\lim_{M\to\infty} R_k^{(c)} \geq \log\left(1 + \frac{SNR_{0k}^{(c)}}{\rho(m_d, \alpha_c) + 1}\right), \tag{4.12}$$

where

$$\rho(m, \alpha) = \frac{2(\pi\lambda)^{\frac{\alpha}{2}} P_d \bar{\Xi} \Gamma(m + 1 - \frac{\alpha}{2})}{(\alpha - 2) N_0 \Gamma(m)}, \tag{4.13}$$

where the Gamma function $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\, \mathrm{d}t$.

Proof. See Appendix 4.8.2. □

Note that in Prop. 4.2, if the underlaid D2D transmitters did not exist, the spectral efficiency $R_k^{(\mathrm{c})}$ of cellular UE k (conditioned on its pathloss and shadowing) in the central cell would converge to $\log\left(1 + \mathrm{SNR}_{0k}^{(\mathrm{c})}\right)$, the maximum achievable spectral efficiency of a point-to-point SISO (single-input single-output) Gaussian channel. It is as if massive MIMO could simultaneously support K interference-free SISO links while reducing the power of each cellular UE by $10 \log_{10} M$ dB. This result is consistent with Prop. 1 in [103] without D2D underlay.

With D2D underlay, the asymptotic result (4.11) shows that there is a loss in cellular spectral efficiency due to the uncanceled interfering signals from the D2D transmitters in $\Phi_k^{(\mathrm{c})}$, i.e., D2D transmitters in Φ except the nearest m_{d} ones whose signals are canceled by the PZF filter. Though it is possible to derive an exact analytical expression (involving integrals) for (4.11), we give a

more intuitive lower bound (4.12), which succinctly characterizes the loss due to the D2D underlay through a single term $\rho(m_{\mathrm{d}}, \alpha_{\mathrm{c}})$. Several remarks are in order.

Remark 1. The term $\rho(m_{\mathrm{d}}, \alpha_{\mathrm{c}})$ corresponding to the uncanceled D2D interference increases with P_{d} and λ and decreases with m_{d}, agreeing with intuition: larger transmit power or larger density of D2D interferers or smaller number of canceled D2D interferers leads to higher D2D-to-cellular interference, thus lowering the cellular spectral efficiency. Further, $\rho(m_{\mathrm{d}}, \alpha_{\mathrm{c}}) \sim \lambda^{\frac{\alpha_{\mathrm{c}}}{2}}$ because a linear increase in λ implies that the distances of the PPP distributed D2D transmitters to the BS decrease as $\lambda^{\frac{1}{2}}$ and thus the D2D-to-cellular interference power increases as $\lambda^{\frac{\alpha_{\mathrm{c}}}{2}}$.

Remark 2. Note that the lower bound (4.12) is meaningful only if $m_{\mathrm{d}} + 1 > \frac{\alpha_{\mathrm{c}}}{2}$. As $m_{\mathrm{d}} \to \frac{\alpha_{\mathrm{c}}}{2} - 1$, $\Gamma(m_{\mathrm{d}} + 1 - \frac{\alpha_{\mathrm{c}}}{2}) \to \infty$ and thus $\rho(m_{\mathrm{d}}, \alpha_{\mathrm{c}}) \to \infty$. In fact, from the proof of Prop. 4.2, we can see that if this condition is violated, the expected D2D-to-cellular interference would be infinite.

Next we show that the loss of cellular spectral efficiency due to D2D underlay can be recovered if we scale the number m_{d} of canceled D2D interferers to infinity as $M \to \infty$. Further, the growth rate of m_{d} can be arbitrarily slow.

Proposition 4.3. *With perfect CSI, arbitrary but fixed m_c, scaled cellular transmit power P_c/M, and conditioned on $\Xi_{0k}^{(c)}$ and $x_{0k}^{(c)}$, if m_d increases to infinity at a rate slower than $\Theta(M)$, the spectral efficiency $R_k^{(c)}$ of cellular UE*

k in the central cell converges as follows.

$$R_k^{(c)} \to \log\left(1 + SNR_{0k}^{(c)}\right), \quad \text{as } M \to \infty. \tag{4.14}$$

Proof. According to Stirling's formula, $\Gamma(t+1) \sim \sqrt{2\pi t}(\frac{t}{e})^t$ when t is large. It follows that

$$\begin{aligned}\frac{\Gamma(m_{\rm d}+1-\frac{\alpha_{\rm c}}{2})}{\Gamma(m_{\rm d})} &\sim \frac{\sqrt{2\pi(m_{\rm d}-\frac{\alpha_{\rm c}}{2})}(\frac{m_{\rm d}-\frac{\alpha_{\rm c}}{2}}{e})^{m_{\rm d}-\frac{\alpha_{\rm c}}{2}}}{\sqrt{2\pi(m_{\rm d}-1)}(\frac{m_{\rm d}-1}{e})^{m_{\rm d}-1}} \\ &= \left(\frac{e}{m_{\rm d}-\frac{\alpha_{\rm c}}{2}}\right)^{\frac{\alpha_{\rm c}}{2}-1}\left(\frac{m_{\rm d}-\frac{\alpha_{\rm c}}{2}}{m_{\rm d}-1}\right)^{m_{\rm d}-\frac{1}{2}} \sim \left(\frac{e}{m_{\rm d}-\frac{\alpha_{\rm c}}{2}}\right)^{\frac{\alpha_{\rm c}}{2}-1}. \end{aligned} \tag{4.15}$$

Therefore, as $m_{\rm d} \to \infty$, $\rho(m_{\rm d}, \alpha_{\rm c}) \to 0$ and thus

$$\begin{aligned}\log\left(1+\mathrm{SNR}_{0k}^{(\mathrm{c})}\right) &\geq \lim_{M\to\infty} R_k^{(\mathrm{c})} \\ &\geq \log\left(1+\frac{\mathrm{SNR}_{0k}^{(\mathrm{c})}}{\rho(m_{\rm d},\alpha_{\rm c})+1}\right) \to \log\left(1+\mathrm{SNR}_{0k}^{(\mathrm{c})}\right). \end{aligned} \tag{4.16}$$

This completes the proof. □

Prop. 4.3 implies that massive MIMO can asymptotically null out all the interfering signals from the infinite D2D transmitters but still maintain a linear scaling in the desired signal power, i.e., we can reduce cellular transmit power as $\Theta(1/M)$ but still achieve the spectral efficiency of an interference-free cellular link.

4.4.2 Non-asymptotic Cellular Spectral Efficiency

Next we analyze the cellular spectral efficiency in the non-asymptotic regime to generate more insights into the impact of the various system param-

eters. To this end, using Jensen's inequality we derive a lower bound for $R_k^{(c)}$ in the following proposition.

Proposition 4.4. *With perfect CSI, $M \geq m_c + m_d + 1$ and $m_d > \frac{\alpha_c}{2} - 1$, and conditioned on $\{\Xi_{bk}^{(c)}\}$ and $\{x_{bk}^{(c)}\}$, the spectral efficiency $R_k^{(c)}$ of cellular UE k in the central cell is lower bounded as*

$$R_k^{(c)} \geq R_k^{(c,lb)} = \log\left(1 + \frac{(M - m_c - m_d - 1)SNR_{0k}^{(c)}}{\sum_{b=0}^{B}\sum_{\ell \in \mathcal{K}_{bk}^{(c)}} SNR_{b\ell}^{(c)} + \rho(m_d, \alpha_c) + 1}\right), \quad (4.17)$$

where $\rho(m, \alpha)$ is defined in (4.13).

Proof. See Appendix 4.8.3. □

Note that the first term in the denominator of (4.17) corresponds to the uncanceled cellular interference; it decreases as m_{c} increases. Similarly, the second term in the denominator of (4.17) corresponds to the uncanceled D2D interference; it decreases as m_{d} increases. In contrast, the numerator of (4.17) corresponds to the desired signal power; it decreases as m_{c} and/or m_{d} increase. The lower bound (4.17) demonstrates the various tradeoffs when choosing the PZF parameters m_{c} and m_{d}. Note that such tradeoffs disappear in the asymptotic regime (cf. Prop. 4.2 and 4.3). If the PZF parameter $m_{\mathrm{c}} = (B+1)K - 1$, then all the cochannel cellular interference will be nulled out, leading to the following specialized lower bound.

$$R_k^{(\mathrm{c,lb})} = \log\left(1 + \frac{(M - (B+1)K - m_{\mathrm{d}})\mathrm{SNR}_{0k}^{(\mathrm{c})}}{\rho(m_{\mathrm{d}}, \alpha_{\mathrm{c}}) + 1}\right). \quad (4.18)$$

We point out that the received signal power gain is only proportional to $M - m_{\mathrm{c}} - m_{\mathrm{d}} - 1$ in the lower bound (4.17). One might think the power gain should be proportional to $M - m_{\mathrm{c}} - m_{\mathrm{d}}$, the number of degrees of freedom left for power boosting after using $m_{\mathrm{c}} + m_{\mathrm{d}}$ degrees of freedom for interference cancellation. The fallacy of the above argument is that it ignores the effect of fading, which makes a power gain proportional to $M - m_{\mathrm{c}} - m_{\mathrm{d}}$ unachievable.

4.4.3 D2D Spectral Efficiency

For D2D receiver r, the post-processing SINR with PZF filter is

$$\mathrm{SINR}_r^{(\mathrm{d})} = \frac{S_r^{(\mathrm{d})}}{I_r^{(\mathrm{c}\to\mathrm{d})} + I_r^{(\mathrm{d}\to\mathrm{d})} + \|\mathbf{w}_r^{(\mathrm{d})}\|^2 N_0}, \tag{4.19}$$

where $S_r^{(\mathrm{d})} = P_{\mathrm{d}} \Xi_{rr}^{(\mathrm{d})} (d_{rr}^{(\mathrm{d})})^{-\alpha_{\mathrm{d}}} \|\mathbf{w}_r^{(\mathrm{d})*} \mathbf{g}_{rr}^{(\mathrm{d})}\|^2$ denotes the desired signal power of D2D Tx-Rx pair r, $I_r^{(\mathrm{c}\to\mathrm{d})}$ and $I_r^{(\mathrm{d}\to\mathrm{d})}$ respectively denote the cochannel cellular and D2D interference powers experienced by D2D receiver r and are given by

$$\begin{aligned} I_r^{(\mathrm{c}\to\mathrm{d})} &= \sum_{b=0}^{B} \sum_{k \in \mathcal{K}_{br}^{(\mathrm{d})}} P_{\mathrm{c}} \Xi_{rbk}^{(\mathrm{c})} (d_{rbk}^{(\mathrm{c})})^{-\alpha_{\mathrm{d}}} |\mathbf{w}_r^{(\mathrm{d})*} \mathbf{g}_{rbk}^{(\mathrm{c})}|^2 \\ I_r^{(\mathrm{d}\to\mathrm{d})} &= \sum_{i \in \Phi_r^{(\mathrm{d})}} P_{\mathrm{d}} \Xi_{ri}^{(\mathrm{d})} (d_{ri}^{(\mathrm{d})})^{-\alpha_{\mathrm{d}}} |\mathbf{w}_r^{(\mathrm{d})*} \mathbf{g}_{ri}^{(\mathrm{d})}|^2. \end{aligned} \tag{4.20}$$

The spectral efficiency of the D2D Tx-Rx pair r is defined as

$$R_r^{(\mathrm{d})} = \mathbb{E}\left[\log(1 + \mathrm{SINR}_r^{(\mathrm{d})})\right], \tag{4.21}$$

where the expectation is taken with respect to the fast fading, shadowing and random locations of UEs.

As the number N of antennas at the UE is often limited due to hardware constraints, it is not very meaningful to study the asymptotic performance with $N \to \infty$. Instead, as in the case of cellular spectral efficiency, we provide a lower bound for $R_r^{(\mathrm{d})}$ in the non-asymptotic regime, which characterizes the impact of the various system parameters on the D2D spectral efficiency.

Proposition 4.5. *With perfect CSI, $N \geq n_c + n_d + 1$ and $n_d > \frac{\alpha_d}{2} - 1$, and conditioned on $\Xi_{rr}^{(d)}$, $d_{rr}^{(d)}$, $\{\Xi_{bk}^{(c)}\}$ and $\{x_{bk}^{(c)}\}$, the spectral efficiency $R_r^{(d)}$ of D2D Tx-Rx pair r is lower bounded as*

$$R_r^{(d)} \geq R_r^{(d,lb)} = \log\left(1 + \frac{(N - n_c - n_d - 1)SNR_r^{(d)}}{\sum_{b=0}^{B}\sum_{k\in\mathcal{K}_{br}^{(d)}} \frac{P_c}{N_0}\Xi_{rbk}^{(c)}(d_{rbk}^{(c)})^{-\alpha_d} + \rho(n_d, \alpha_d) + 1}\right), \tag{4.22}$$

where $SNR_r^{(d)} = P_d\Xi_{rr}^{(d)}(d_{rr}^{(d)})^{-\alpha_d}/N_0$, and $\rho(m, \alpha)$ is defined in (4.13).

Proof. The proof is similar to that of Prop. 4.4 and is omitted for brevity. □

Many of the remarks on Prop. 4.4 apply to Prop. 4.5 as well and are not repeated here. One additional remark is that the cellular-to-D2D interference is not homogeneous: the D2D receivers located in the boundary of the cellular network experience less cellular interference than the D2D receivers located in the central cell. But if we focus on the D2D performance in the central cell and choose the number of cellular cells large enough, this heterogeneity can be made negligible.

4.5 Spectral Efficiency with Imperfect Channel State Information

4.5.1 Estimating UE-BS Channels

We consider pilot-based CSI estimation in which known training sequences are transmitted and used for estimation purpose. To alleviate the training overhead and coordination complexity, we assume that each BS b does not estimate the channels from other-cell transmitters (either cellular or D2D). Note that as the number $|\Phi_b|$ of D2D transmitters in the cell b is Poisson distributed, there may be less than m_{d} D2D transmitters in the cell b. Therefore, during the training phase, each BS b requires the K cellular UEs and the $m_{\mathrm{d},b} \triangleq \min(m_{\mathrm{d}}, |\Phi_b|)$ nearest D2D transmitters (w.r.t. the BS b) in its cell to simultaneously transmit orthogonal training sequences. The BSs do not coordinate the other D2D transmitters, which can send independent symbols during the training phase.

Unlike the perfect CSI case, other-cell transmissions (both cellular and D2D) now have a more delicate impact on the performance of the central cell. To accommodate this, in this subsection we extend the previous notation as follows. We add an additional subscript b to $x_i^{(\mathrm{d})}, \Xi_i^{(\mathrm{d})}$ and $\mathbf{h}_i^{(\mathrm{d})}$, and obtain $x_{bi}^{(\mathrm{d})}, \Xi_{bi}^{(\mathrm{d})}$ and $\mathbf{h}_{bi}^{(\mathrm{d})}$, indicating that they are associated with D2D transmitter i in the cell b. Similarly, we use $\Phi_{bk}^{(\mathrm{c})}$ to denote the set of uncanceled D2D interferers in the cell $b, b = 0, ..., B+1$. Note that the coverage of the "cell" $B+1$ is simply the complement (w.r.t. $\mathbb{R}^2$) of the coverage areas of the cells $0, ..., B$, and the "cell" $B+1$ does not contain a BS.

Denoting by $T_{\rm c} \geq K + m_{\rm d}$ the length of a training sequence, we can represent the training sequences as a $T_{\rm c} \times (K + m_{\rm d})$ dimensional matrix $\sqrt{T_{\rm c}}\mathbf{Q}^{({\rm c})} = \sqrt{T_{\rm c}}(\mathbf{q}_1^{({\rm c})}, ..., \mathbf{q}_{K+m_{\rm d}}^{({\rm c})})$ satisfying $\mathbf{Q}^{({\rm c})*}\mathbf{Q}^{({\rm c})} = \mathbf{I}_{K+m_{\rm d}}$. These pilots are reused over different cells. In the training phase, the $M \times T_{\rm c}$ dimensional baseband received signal $\mathbf{Y}_0^{({\rm c})}$ at the central BS is

$$\begin{aligned}\mathbf{Y}_0^{({\rm c})} =& \sum_{b=0}^{B} \sum_{k \in \mathcal{K}_b} \sqrt{T_{\rm c} P_{\rm c} \Xi_{bk}^{({\rm c})}} \|x_{bk}^{({\rm c})}\|^{-\frac{\alpha_{\rm c}}{2}} \mathbf{h}_{bk}^{({\rm c})} \mathbf{q}_k^{({\rm c})*} \\ &+ \sum_{b=0}^{B} \sum_{i=1}^{m_{{\rm d},b}} \sqrt{T_{\rm c} P_{\rm c} \Xi_{bi}^{({\rm d})}} \|x_{bi}^{({\rm d})}\|^{-\frac{\alpha_{\rm c}}{2}} \mathbf{h}_{bi}^{({\rm d})} \mathbf{q}_{K+i}^{({\rm c})*} \\ &+ \sum_{b=0}^{B+1} \sum_{r \in \Phi_{bk}^{({\rm c})}} \sqrt{P_{\rm d} \Xi_{br}^{({\rm d})}} \|x_{br}^{({\rm d})}\|^{-\frac{\alpha_{\rm c}}{2}} \mathbf{h}_{br}^{({\rm d})} \mathbf{u}_{br}^{({\rm d})*} + \mathbf{V}_0^{({\rm c})}, \end{aligned} \tag{4.23}$$

where the $T_{\rm c} \times 1$ dimensional vector $\mathbf{u}_{br}^{({\rm d})}$ contains the data symbols sent by D2D interferer r in the cell b, and the $M \times T_{\rm c}$ dimensional noise matrix $\mathbf{V}_0^{({\rm c})}$ consists of i.i.d. $\mathcal{CN}(0, N_0)$ elements. Note that the coordinated D2D transmitters also use power $P_{\rm c}$ during the training phase since they now transmit to their associated BSs.

We assume that the central BS uses linear MMSE estimator for the channel estimation. To this end, we first project the received signal $\mathbf{Y}_0^{({\rm c})}$ in the direction of $\mathbf{q}_{\tilde{k}}^{({\rm c})}$ and normalize it to obtain

$$\begin{aligned}\tilde{\mathbf{y}}_k^{(s)} &= \frac{1}{\sqrt{T_{\rm c} P_{\rm c} \Xi_{0k}^{(s)}} \|x_{0k}^{(s)}\|^{-\frac{\alpha_{\rm c}}{2}}} \mathbf{Y}_0^{({\rm c})} \mathbf{q}_{\tilde{k}}^{({\rm c})} \\ &= \mathbf{h}_{0k}^{(s)} + \sum_{b=1}^{B} \sqrt{\beta_{bk}^{(s)}} \mathbf{h}_{bk}^{(s)} + \tilde{\mathbf{v}}_k^{(s)}, \quad (s, \tilde{k}) \in \{({\rm c}, k), ({\rm d}, K+k)\}, \end{aligned} \tag{4.24}$$

where

$$\beta_{bk}^{(s)} \triangleq \begin{cases} 0 & \text{if } s = \mathrm{d} \text{ and } k > m_{\mathrm{d},b} \text{ ;} \\ \frac{\Xi_{bk}^{(s)} \|x_{bk}^{(s)}\|^{-\alpha_{\mathrm{c}}}}{\Xi_{0k}^{(s)} \|x_{0k}^{(s)}\|^{-\alpha_{\mathrm{c}}}} & \text{otherwise,} \end{cases}$$

and $\tilde{\mathbf{v}}_k^{(s)}$ denotes the equivalent channel estimation "noise" and is given by

$$\tilde{\mathbf{v}}_k^{(s)} = \frac{1}{\sqrt{T_{\mathrm{c}} P_{\mathrm{c}} \Xi_{0k}^{(s)}} \|x_{0k}^{(s)}\|^{-\frac{\alpha_{\mathrm{c}}}{2}}} \left(\sum_{b=0}^{B+1} \sum_{r \in \Phi_{bk}^{(\mathrm{c})}} \sqrt{P_{\mathrm{d}} \Xi_{br}^{(\mathrm{d})}} \|x_{br}^{(\mathrm{d})}\|^{-\frac{\alpha_{\mathrm{c}}}{2}} \mathbf{h}_{br}^{(\mathrm{d})} \bar{u}_{br}^{(\mathrm{d})} + \bar{\mathbf{v}}_k^{(\mathrm{c})} \right). \tag{4.25}$$

where $\bar{u}_{br}^{(\mathrm{d})} = \mathbf{u}_{br}^{(\mathrm{d})*} \mathbf{q}_{\tilde{k}}^{(\mathrm{c})}$ and $\bar{\mathbf{v}}_k^{(\mathrm{c})} = \mathbf{V}_0^{(\mathrm{c})} \mathbf{q}_{\tilde{k}}^{(\mathrm{c})}$.

Lemma 4.1. *The linear MMSE estimate of* $\mathbf{h}_{0k}^{(s)}, s \in \{c, d\}$*, is given by* $\hat{\mathbf{h}}_{0k}^{(s)} = \xi_k^{(s)} \tilde{\mathbf{y}}_k^{(s)}$*, where*

$$\xi_k^{(s)} = \left(1 + \sum_{b=1}^{B} \beta_{bk}^{(s)} + \frac{\sum_{b=0}^{B+1} \sum_{r \in \Phi_{bk}^{(c)}} P_d \Xi_{br}^{(d)} \|x_{br}^{(d)}\|^{-\alpha_c} + N_0}{T_c P_c \Xi_{0k}^{(s)} \|x_{0k}^{(s)}\|^{-\alpha_c}} \right)^{-1}. \tag{4.26}$$

Further, $\mathbb{E}[\hat{\mathbf{h}}_{0k}^{(s)}] = 0$ *and* $\mathbb{E}[\hat{\mathbf{h}}_{0k}^{(s)} \hat{\mathbf{h}}_{0k}^{(s)*}] = \xi_k^{(s)} \mathbf{I}_M$*. As for the estimation error* $\boldsymbol{\epsilon}_k^{(s)} = \mathbf{h}_{0k}^{(s)} - \hat{\mathbf{h}}_{0k}^{(s)}$*,* $\mathbb{E}[\boldsymbol{\epsilon}_k^{(s)}] = 0$ *and* $\mathbb{E}[\boldsymbol{\epsilon}_k^{(s)} \boldsymbol{\epsilon}_k^{(s)*}] = (1 - \xi_k^{(s)}) \mathbf{I}_M$*.*

Proof. See Appendix 4.8.4. □

Lemma 4.1 shows that the longer the length T_{c} of a training sequence, the smaller the covariance of the estimation error $\boldsymbol{\epsilon}_k^{(s)}$ and thus the more accurate the channel estimation $\hat{\mathbf{h}}_{0k}^{(s)}$, agreeing with intuition. In particular, $\mathbb{E}[\boldsymbol{\epsilon}_k^{(s)} \boldsymbol{\epsilon}_k^{(s)*}] \to \frac{\sum_{b=1}^{B} \beta_{bk}^{(s)}}{1 + \sum_{b=1}^{B} \beta_{bk}^{(s)}} \mathbf{I}_M$, as $T_{\mathrm{c}} \to \infty$. This shows that even with infinitely long training sequences, the channel estimation cannot be perfect due to pilot contamination.

4.5.2 Asymptotic Cellular Spectral Efficiency

In this subsection, we examine the asymptotic performance of the cellular links as $M \to \infty$. For simplicity, we focus on $m_{\mathrm{c}} = m_{\mathrm{d}} = 0$. Then $\mathbf{w}_k^{(\mathrm{c})} = \hat{\mathbf{h}}_{0k}^{(\mathrm{c})}$ is the MRC filter. Since multiplying the filter by a constant does not affect the post-processing SINR, we may choose $\mathbf{w}_k^{(\mathrm{c})} = \mathbf{Y}_0^{(\mathrm{c})}\mathbf{q}_k^{(\mathrm{c})}$. It follows that $\lim_{M\to\infty} \frac{1}{M}\mathbf{w}_k^{(\mathrm{c})*}\mathbf{y}_0^{(\mathrm{c})}$ equals

$$\lim_{M\to\infty} \frac{1}{M}\Bigg(\sum_{b=0}^{B} \sqrt{T_{\mathrm{c}} P_{\mathrm{c}} \Xi_{bk}^{(\mathrm{c})}} \|x_{bk}^{(\mathrm{c})}\|^{-\frac{\alpha_{\mathrm{c}}}{2}} \mathbf{h}_{bk}^{(\mathrm{c})} + \Bigg(\sum_{i\in\Phi} \sqrt{P_{\mathrm{d}} \Xi_i^{(\mathrm{d})}} \|x_i^{(\mathrm{d})}\|^{-\frac{\alpha_{\mathrm{c}}}{2}} \mathbf{h}_i^{(\mathrm{d})} \bar{u}_i^{(\mathrm{d})} + \bar{\mathbf{v}}_0^{(\mathrm{c})} \Bigg) \Bigg)^{*} \mathbf{y}_0^{(\mathrm{c})} = \sum_{b=0}^{B} \sqrt{T_{\mathrm{c}}} P_{\mathrm{c}} \Xi_{bk}^{(\mathrm{c})} \|x_{bk}^{(\mathrm{c})}\|^{-\alpha_{\mathrm{c}}} u_{bk}^{(\mathrm{c})} + \sum_{i\in\Phi} P_{\mathrm{d}} \Xi_i^{(\mathrm{d})} \|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}} \bar{u}_i^{(\mathrm{d})*} u_i^{(\mathrm{d})}. \tag{4.27}$$

The first term in (4.27) is the usual phenomenon appearing in massive MIMO [99]. In particular, it indicates that asymptotically the effects of uncorrelated receiver noise and fast fading vanish, and there is no intra-cell interference. The remaining effect is the residual other-cell interference due to pilot reuse across the cells [99]. With D2D underlay, we observe that a new effect (i.e., the last term in (4.27)) indicating the residual D2D-to-cellular interference arises. The reason why the effect of D2D underlay does not vanish can be explained as follows. The interfering signal of D2D transmitter i in the training phase correlates with the interfering signal of D2D transmitter i in the data transmission phase through the common channel vector $\mathbf{h}_i^{(\mathrm{d})}$. Therefore, unlike the uncorrelated receiver noises in the estimation phase and in the data transmission phase, when multiplying the estimated channel $\hat{\mathbf{h}}_{0k}^{(\mathrm{c})}$

with the received signal $\mathbf{y}_0^{(\mathrm{c})}$, the effect of D2D underlay cannot be eliminated even with infinitely many antennas at the BS. We term this effect *underlay contamination*.

Note that the D2D underlay contamination term in (4.27) involves the products of complex Gaussian random variables $\bar{u}_i^{(\mathrm{d})*} u_i^{(\mathrm{d})}$, the D2D interfering signals are not Gaussian distributed. It has been proven, however, that given a covariance constraint Gaussian noise is the worst-case noise for additive noise channels. Therefore, treating the D2D interfering signals as Gaussian noises, we obtain the following Lemma 4.2.

Lemma 4.2. *With imperfect CSI at the central BS and $m_c = m_d = 0$, i.e., the MRC receiver $\mathbf{w}_k^{(c)} = \hat{\mathbf{h}}_{0k}^{(c)}$, the following spectral efficiency $\hat{R}_k^{(c)}$ can be achieved for cellular UE k in the central cell.*

$$\hat{R}_k^{(c)} = \mathbb{E}\left[\log\left(1 + \frac{\hat{S}_k^{(c)}}{\hat{I}_k^{(c\to c)} + \hat{I}_k^{(d\to c)}}\right)\right], \tag{4.28}$$

where $\hat{S}_k^{(c)} = T_c P_c^2 |\Xi_{0k}^{(c)}|^2 \|x_{0k}^{(c)}\|^{-2\alpha_c}$, and

$$\hat{I}_k^{(c\to c)} = \sum_{b=1}^{B} T_c P_c^2 |\Xi_{bk}^{(c)}|^2 \|x_{bk}^{(c)}\|^{-2\alpha_c}, \quad \hat{I}_k^{(d\to c)} = \sum_{i\in\Phi} P_d^2 |\Xi_i^{(d)}|^2 \|x_i^{(d)}\|^{-2\alpha_c}. \tag{4.29}$$

Unlike the perfect CSI case in which the SINR of a cellular link can be made arbitrarily large (c.f. Prop. 4.1), Lemma 4.2 shows that with imperfect CSI there is a limit on the received SINR in massive MIMO due to the pilot contamination and D2D underlay contamination. With D2D underlay, conditioned on UE positions and shadowing, the loss of SINR (in dB) of cellular

UE k in the central cell is $10\log_{10}\left(1+\hat{I}_k^{(\mathrm{d}\to\mathrm{c})}/\hat{I}_k^{(\mathrm{c}\to\mathrm{c})}\right)$. There are four possible approaches to mitigate the loss. First, we can decrease the D2D transmit power. This approach reduces the link budgets of D2D links, limiting the range of D2D communication. Second, we can increase the cellular transmit power. This approach increases the energy consumption of cellular UEs and also results in more cellular-to-D2D interference. Third, we can increase the length of training sequence. But longer training sequence consumes more cellular transmission resources in terms of both power and bandwidth. Fourth, we can deactivate the D2D links in the training phase of massive MIMO. Then we retain the usual asymptotic cellular spectral efficiency in massive MIMO:

$$\hat{R}_k^{(\mathrm{c})} = \mathbb{E}\left[\log\left(1+\frac{|\Xi_{0k}^{(\mathrm{c})}|^2\|x_{0k}^{(\mathrm{c})}\|^{-2\alpha_{\mathrm{c}}}}{\sum_{b=1}^{B}|\Xi_{bk}^{(\mathrm{c})}|^2\|x_{bk}^{(\mathrm{c})}\|^{-2\alpha_{\mathrm{c}}}}\right)\right]. \tag{4.30}$$

Certainly, the last approach reduces time resources for D2D communication.

The following Corollary 4.1 shows that with D2D underlay contamination it is impossible to scale down cellular transmit powers, and thus D2D underlay hurts the energy efficiency of cellular UEs in massive MIMO.

Corollary 4.1. *Scaling down cellular transmit powers results in vanishing cellular spectral efficiency, i.e.,* $\hat{R}_k^{(c)} \to 0$, *as* $P_c \to 0$.

To achieve a non-vanishing cellular spectral efficiency while scaling down cellular transmit powers, one approach would be to schedule two independent sets of active D2D transmitters in the estimation phase and in the data transmission phase of massive MIMO. This solves underlay contamination. The disadvantage is that the BSs cannot use the estimated D2D UE-BS

channels in the estimation phase to cancel the interference from the other set of D2D transmitters in the data transmission phase. Therefore, its performance is not clear in a non-asymptotic regime. Another simpler approach would be to deactivate the D2D links in the training phase of massive MIMO. Then we can scale down cellular transmit powers as in the following Prop. 4.6.

Proposition 4.6. *With D2D links deactivated in the training phase of massive MIMO and scaled cellular transmit power $P_c/\sqrt{M}$, as $M \to \infty$, the achievable spectral efficiency $\hat{R}_k^{(c)}$ of cellular UE k in the central cell converges as follows.*

$$\hat{R}_k^{(c)} \to \mathbb{E}\left[\log\left(1 + \frac{T_c (SNR_{0k}^{(c)})^2}{\sum_{b=1}^{B} T_c (SNR_{bk}^{(c)})^2 + \sum_{i\in\Phi} \frac{P_d}{N_0} \Xi_i^{(d)} \|x_i^{(d)}\|^{-\alpha_c} + 1}\right)\right]. \tag{4.31}$$

Proof. See Appendix 4.8.5. □

Finally, we give a more explicit expression for the asymptotic cellular spectral efficiency to allow for efficient numerical evaluation.

Proposition 4.7. *The achievable spectral efficiency $\hat{R}_k^{(c)}$ of cellular UE k in the central cell equals*

$$\hat{R}_k^{(c)} = \int_0^\infty \frac{1}{z}(1 - \mathbb{E}[e^{-z\hat{S}_k^{(c)}}])\mathbb{E}[e^{-z\hat{I}_k^{(c\to c)}}]\mathbb{E}[e^{-z\hat{I}_k^{(d\to c)}}]\, \mathrm{d}z, \tag{4.32}$$

where

$$\mathbb{E}[e^{-z\hat{S}_k^{(c)}}] = \mathbb{E}[e^{-zT_c P_c^2 |\Xi_{0k}^{(c)}|^2 \|x_{0k}^{(c)}\|^{-2\alpha_c}}] \tag{4.33}$$

$$\mathbb{E}[e^{-z\hat{I}_k^{(c\to c)}}] = \prod_{b=1}^{B} \mathbb{E}[e^{-zT_c P_c^2 |\Xi_{bk}^{(c)}|^2 \|x_{bk}^{(c)}\|^{-2\alpha_c}}] \tag{4.34}$$

$$\mathbb{E}[e^{-z\hat{I}_k^{(d\to c)}}] = \exp\left(-\pi\lambda\Gamma(1-1/\alpha_c) P_d^{2/\alpha_c} \mathbb{E}[\Xi^{2/\alpha_c}] z^{1/\alpha_c}\right). \tag{4.35}$$

Proof. For any $x > 0$, $\log(1+x) = \int_0^\infty \frac{1}{z}(1-e^{-xz})e^{-z}\,\mathrm{d}z$ [56]. Therefore,

$$\begin{aligned}\mathbb{E}\left[\log\left(1+\frac{X}{Y}\right)\right] &= \mathbb{E}\left[\int_0^\infty \frac{1}{z}(1-e^{-z\frac{X}{Y}})e^{-z}\,\mathrm{d}z\right] \\ &= \mathbb{E}\left[\int_0^\infty \frac{1}{z}(1-e^{-zX})e^{-zY}\,\mathrm{d}z\right]. \end{aligned} \tag{4.36}$$

Using the above equality, the linearity of expectation, and the independence of $\hat{S}_k^{(\mathrm{c})}$, $\hat{I}_k^{(\mathrm{c}\to\mathrm{c})}$ and $\hat{I}_k^{(\mathrm{d}\to\mathrm{c})}$,

$$\begin{aligned}\hat{R}_k^{(\mathrm{c})} &= \mathbb{E}\left[\int_0^\infty \frac{1}{z}(1-e^{-z\hat{S}_k^{(\mathrm{c})}})e^{-z(\hat{I}_k^{(\mathrm{c}\to\mathrm{c})}+\hat{I}_k^{(\mathrm{d}\to\mathrm{c})})}\,\mathrm{d}z\right] \\ &= \int_0^\infty \frac{1}{z}(1-\mathbb{E}[e^{-z\hat{S}_k^{(\mathrm{c})}}])\mathbb{E}[e^{-z\hat{I}_k^{(\mathrm{c}\to\mathrm{c})}}]\mathbb{E}[e^{-z\hat{I}_k^{(\mathrm{d}\to\mathrm{c})}}]\,\mathrm{d}z. \end{aligned} \tag{4.37}$$

The expressions for $\mathbb{E}[e^{-z\hat{S}_k^{(\mathrm{c})}}]$ and $\mathbb{E}[e^{-z\hat{I}_k^{(\mathrm{c}\to\mathrm{c})}}]$ follow by definitions. Using the Laplace functional of the PPP Φ [19], we have

$$\mathbb{E}[e^{-z\hat{I}_k^{(\mathrm{d}\to\mathrm{c})}}] = \exp\left(-2\pi\lambda\int_0^\infty \left(1-\mathbb{E}[\exp(-zP_{\mathrm{d}}^2\Xi^2 r^{-2\alpha_{\mathrm{c}}})]\right) r\,\mathrm{d}r\right), \tag{4.38}$$

which equals (4.35). □

4.6 Simulation and Numerical Results

In this section, we provide simulation and numerical results to demonstrate the analytical results and obtain insights into how the various system parameters affect the cellular and D2D spectral efficiency. The specific parameters used are summarized in Table 5.1 unless otherwise specified. The cellular network consists of 19 hexagonal cells; the side length of each cell is R_c. There are K uniformly distributed cellular UEs in each cell, while D2D UEs are distributed as a PPP. The shadowing is lognormal with deviation

BS coverage radius R_c	500 m
# cellular UEs K	4
Density of D2D UEs λ	$\frac{12}{\pi R_c^2}$ m^{-2}
# BS antennas M	100
# UE Rx antennas N	4
UE-BS PL exponent α_{c}	3.76
UE-UE PL exponent α_{d}	4.37
UE-BS PL reference $C_{\mathrm{c},0}$	15.3 dB
UE-UE PL reference $C_{\mathrm{d},0}$	38.5 dB
Cellular Tx power P_{c}	23 dBm
D2D Tx power P_{d}	13 dBm
Channel bandwidth	10 MHz
Noise PSD	-174 dBm/Hz
BS noise figure	6 dB
UE noise figure	9 dB
Lognormal shadowing σ	7 dB

Table 4.1: Simulation/Numerical Parameters for Massive MIMO with D2D Underlay

σ (dB). The pathloss parameters given in Table 5.1 correspond to a carrier frequency of 2 GHz. Specifically, we use the 3GPP macrocell propagation model (urban area) for UE-BS channels [1] and the revised *Winner + B1* model (non-light-of-sight with -5 dB offset) for UE-UE channels [5]. Note that different pathloss reference values $C_{\mathrm{c},0}$ and $C_{\mathrm{d},0}$ are used in the UE-BS and UE-UE channels. Therefore, when evaluating the analytical expressions using the parameters in Table 5.1, $P_{\mathrm{c}} = 23 - C_{\mathrm{c},0}$ (dBm) and $P_{\mathrm{d}} = 13 - C_{\mathrm{c},0}$ (dBm) for the UE-BS channels while $P_{\mathrm{c}} = 23 - C_{\mathrm{d},0}$ (dBm) and $P_{\mathrm{d}} = 13 - C_{\mathrm{d},0}$ (dBm) for the UE-UE channels.

We first compare the simulated cellular spectral efficiency to the corre-

sponding analytical lower bound (4.17) under various PZF parameters $(m_{\rm c}, m_{\rm d})$ in Fig. 4.2. The conditioned random variables in (4.17) are averaged out in Fig. 4.2. We can see that the analytical lower bound (4.17) closely matches the simulation results. The larger $m_{\rm d}$, the better match between the simulation and the analytical lower bound (4.17). This is because larger $m_{\rm d}$ implies less D2D-to-cellular interferers and thus smaller interference variance. As a result, the lower bound based on Jensen's inequality becomes more accurate with larger $m_{\rm d}$. Note that, with $K = 4$, $m_{\rm c} = 0$ and $m_{\rm c} = 3$ correspond to MRC and ZF (w.r.t. intra-cell cellular UEs), respectively. Comparing the spectral efficiency with $(m_{\rm c}, m_{\rm d}) = (0, 2)$ to that of $(m_{\rm c}, m_{\rm d}) = (3, 2)$, we can see that ZF has better performance and the spectral efficiency gain is about 1.6 bps/Hz. This observation implies that although asymptotically ZF and MRC have similar performance, it is still quite beneficial to appropriately suppress the co-channel cellular interference in practical non-asymptotic regime.

Since the lower bound (4.17) is accurate, next we use it to demonstrate the cellular spectral efficiency with scaled cellular transmit power (i.e., $P_{\rm c} \to P_{\rm c}/M$) in Fig. 4.3. We consider two PZF choices: PZF with constant $m_{\rm d}$ and PZF with scaled $m_{\rm d} = \Theta(\sqrt{M})$. As a benchmark, we also include the curves corresponding the scenarios without D2D underlay. Also, D2D transmit power is decreased by 10 times to accelerate the convergence. Several observations from Fig. 4.3 are in order. First, unlike the case with unscaled cellular transmit power, Fig. 4.3 shows that ZF and MRC have similar performance. Second, adopting a constant $m_{\rm d}$ results in a fixed loss in the cellular spectral

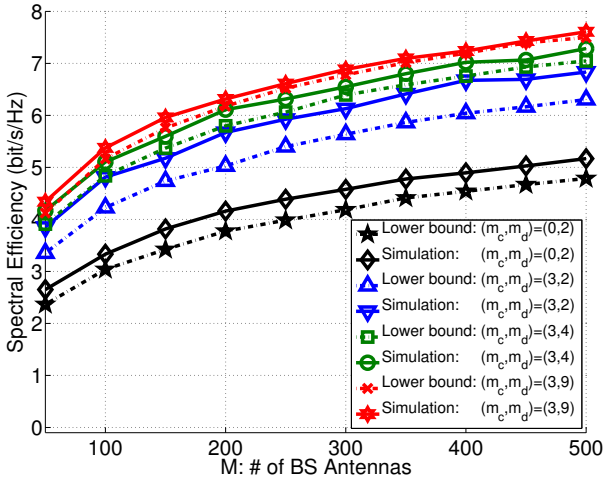

Figure 4.2: Simulated cellular spectral efficiency vs. analytical lower bound (4.17) with perfect CSI.

efficiency due to the underlaid D2D interference; this loss cannot be overcome by increasing the number of BS antennas when the cellular transmit power is also scaled down as $\Theta(1/M)$. This observation confirms the analytical results in Prop. 4.2. Third, the loss in the cellular spectral efficiency due to D2D underlay can be overcome by scaling $m_{\rm d}$ as $\Theta(\sqrt{M})$, validating the theoretical finding in Prop. 4.3. But the convergence rate is relatively slow.

Fig. 4.4 compares the simulated D2D spectral efficiency to the corresponding analytical lower bound (4.22) under different D2D distances and $(n_{\rm c}, n_{\rm d}) = (0, 2)$. The conditioned random variables in (4.22) are averaged out in Fig. 4.4. We can see that the analytical lower bound (4.22) closely matches the simulation results when $N \geq 6$ while being a bit loose when $N < 6$. The accuracy of the lower bound obtained from Jensen's inequality implies that after canceling 2 nearest D2D interferers, the variance of the residual D2D-to-

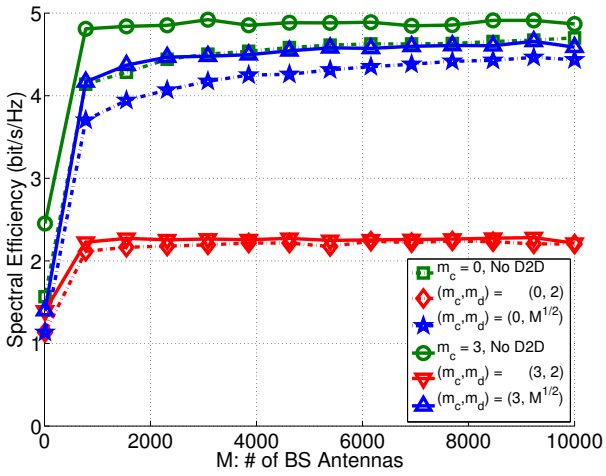

Figure 4.3: Cellular spectral efficiency with scaled cellular transmit power and perfect CSI.

cellular interference is relatively small. Fig. 4.4 also shows that D2D spectral efficiency is quite sensitive to its communication range: there is a loss of about 3 bps/Hz in spectral efficiency if D2D range is increased from 20 m to 35 m.

Next we evaluate the effect of multi-user cellular transmission on D2D spectral efficiency. Fig. 4.5 shows the D2D spectral efficiency as a function of the number K of co-channel cellular UEs per cell. Not surprisingly, as K increases, D2D spectral efficiency decreases due to the increased cellular-to-D2D interference. The interesting observation from Fig. 4.5 is that even with $(n_{\mathrm{c}}, n_{\mathrm{d}}) = (0, 0)$ (i.e., the MRC receiver) the average D2D spectral efficiency is not severely affected by scaling up the number of cellular UEs. For example, when K increases from 10 to 20, the loss in D2D spectral efficiency is less than 0.5 bps/Hz. This implies that we can scale up the uplink capacity in a massive MIMO system without much loss in the average D2D spectral efficiency.

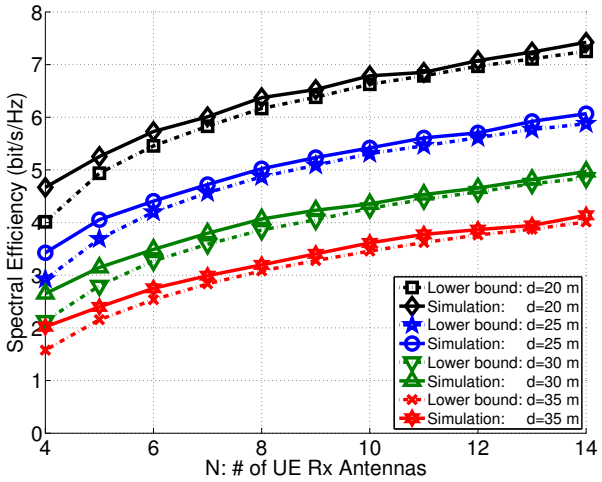

Figure 4.4: Simulated D2D spectral efficiency vs. analytical lower bound (4.22) with perfect CSI and $(n_{\mathrm{c}}, n_{\mathrm{d}}) = (0, 2)$.

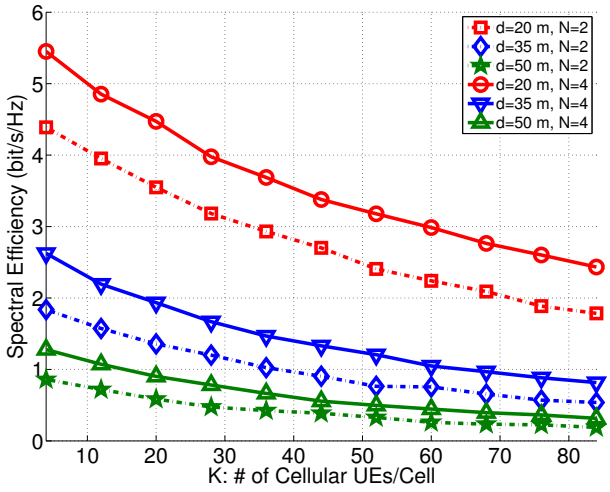

Figure 4.5: Effect of multi-user cellular transmission on D2D spectral efficiency with perfect CSI and $(n_{\mathrm{c}}, n_{\mathrm{d}}) = (0, 0)$.

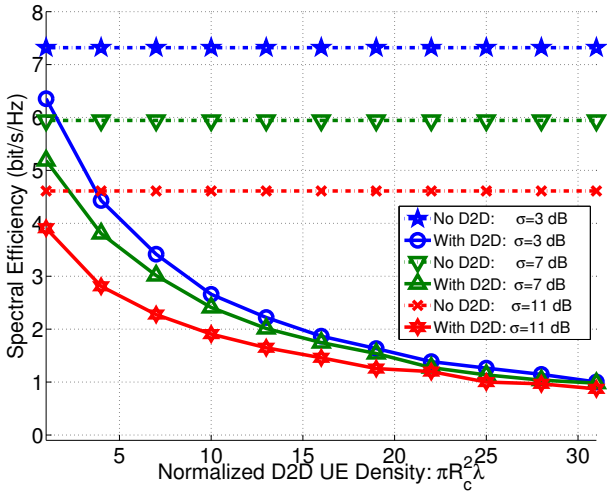

Figure 4.6: Effect of D2D underlay contamination on asymptotic cellular spectral efficiency of massive MIMO with $(m_{\mathrm{c}}, m_{\mathrm{d}}) = (0, 0)$ and $T_{\mathrm{c}} = 4$.

Fig. 4.6 illustrates the effect of D2D underlay contamination on the asymptotic cellular spectral efficiency of massive MIMO. Compared to the case without D2D, where only pilot contamination exists, D2D underlay contamination degrades the achievable asymptotic massive MIMO spectral efficiency. For example, with shadowing deviation $\sigma = 7$ dB and $\pi R_c^2 \lambda = 4$, the spectral efficiency is reduced from 6 bps/Hz to about 3.8 bps/Hz. Further, the more the underlaid D2D UEs, the smaller the asymptotic cellular spectral efficiency. Fig. 4.6 shows that when $\pi R_c^2 \lambda \geq 22$ the effect of D2D underlay contamination dominates in the overall effect of pilot and underlay contamination.

4.7 Summary

In this chapter, we have studied the spectral efficiency of a D2D underlaid massive MIMO system under perfect and imperfect CSI. We have found that massive MIMO can efficiently handle the D2D-to-cellular interference. Meanwhile, from an average perspective, D2D links are relatively robust to the cellular-to-D2D interference even if there are quite many cochannel cellular users. D2D interference does make the estimated CSI in massive MIMO less accurate and thus in turn hurts the cellular spectral efficiency. One simple approach to alleviating this effect is to deactivate D2D links in the cellular training phase. Overall, our study suggests that D2D may be much simpler in massive MIMO cellular systems than in current cellular systems.

4.8 Appendix

4.8.1 Proof of Proposition 4.1

We show that a PZF receiver with $m_{\mathrm{c}} = m_{\mathrm{d}} = 0$, i.e., the MRC receiver, at the BS suffices. With $m_{\mathrm{c}} = m_{\mathrm{d}} = 0$, the PZF receiver $\mathbf{w}_k^{(\mathrm{c})} = \mathbf{h}_{0k}^{(\mathrm{c})}$. By the law of large numbers, $\frac{1}{M}\|\mathbf{h}_{0k}^{(\mathrm{c})}\|^2 \xrightarrow{a.s.} 1$, $\frac{1}{M}\mathbf{h}_{0k}^{(\mathrm{c})*}\mathbf{h}_{b\ell}^{(\mathrm{c})} \xrightarrow{a.s.} 0, \ell \neq k$ or $b \neq 0$. It follows that when conditioned on $\Xi_{0k}^{(\mathrm{c})}$ and $x_{0k}^{(\mathrm{c})}$,

$$\frac{1}{M^2} P_{\mathrm{c}} \Xi_{0k}^{(\mathrm{c})} \|x_{0k}^{(\mathrm{c})}\|^{-\alpha_{\mathrm{c}}} \|\mathbf{h}_{0k}^{(\mathrm{c})}\|^4 \xrightarrow{a.s.} P_{\mathrm{c}} \Xi_{0k}^{(\mathrm{c})} \|x_{0k}^{(\mathrm{c})}\|^{-\alpha_{\mathrm{c}}}. \tag{4.39}$$

Also, the noise term normalized by M^2 converges as $\frac{1}{M^2} N_0 \|\mathbf{h}_{0k}^{(\mathrm{c})}\|^2 \xrightarrow{a.s.} 0$. Further, interchanging the order of the limit and the finite sum, the cellular

interference normalized by M^2 converges as

$$\lim_{M\to\infty}\frac{1}{M^2}\sum_{b=0}^{B}\sum_{\ell\in\mathcal{K}_{bk}^{(\mathrm{c})}}P_{\mathrm{c}}\Xi_{b\ell}^{(\mathrm{c})}\|x_{b\ell}^{(\mathrm{c})}\|^{-\alpha_{\mathrm{c}}}|\mathbf{h}_{0k}^{(\mathrm{c})*}\mathbf{h}_{b\ell}^{(\mathrm{c})}|^2$$

$$=\sum_{b=0}^{B}\sum_{\ell\in\mathcal{K}_{bk}^{(\mathrm{c})}}P_{\mathrm{c}}\Xi_{b\ell}^{(\mathrm{c})}\|x_{b\ell}^{(\mathrm{c})}\|^{-\alpha_{\mathrm{c}}}\left(\lim_{M\to\infty}\frac{1}{M^2}|\mathbf{h}_{0k}^{(\mathrm{c})*}\mathbf{h}_{b\ell}^{(\mathrm{c})}|^2\right)\xrightarrow{a.s.}0. \tag{4.40}$$

Next we show that the D2D interference normalized by M^2 converges to 0 as $M\to\infty$. Note that in this case *we cannot directly interchange the order of the limit and the infinite sum* to conclude that it converges to 0 almost surely. Instead, we can prove its convergence in probability, i.e., for any $\epsilon>0$,

$$\lim_{M\to\infty}\mathbb{P}\left(\frac{1}{M^2}\sum_{i\in\Phi}P_{\mathrm{d}}\Xi_i^{(\mathrm{d})}\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}|\mathbf{h}_{0k}^{(\mathrm{c})*}\mathbf{h}_i^{(\mathrm{d})}|^2<\epsilon\right)=1. \tag{4.41}$$

To this end, we partition the D2D transmitters into two groups: one group is composed of those transmitters located within distance r_o from the BS and the other group is composed of those transmitters located with distance grater than r_o from the BS. Then using the inequalities

$$\mathbb{P}(X+Y\geq\epsilon)\leq\mathbb{P}\left(X\geq\frac{\epsilon}{2}\text{ or }Y\geq\frac{\epsilon}{2}\right)\leq\mathbb{P}\left(X\geq\frac{\epsilon}{2}\right)+\mathbb{P}\left(Y\geq\frac{\epsilon}{2}\right),$$

where X and Y are two arbitrary random variables, we have

$$\mathbb{P}\left(\frac{1}{M^2}\sum_{i\in\Phi}P_{\mathrm{d}}\Xi_i^{(\mathrm{d})}\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}|\mathbf{h}_{0k}^{(\mathrm{c})*}\mathbf{h}_i^{(\mathrm{d})}|^2\geq\epsilon\right)$$

$$\leq\mathbb{P}\left(\frac{1}{M^2}\sum_{i\in\Phi\cap\mathcal{B}^c(o,r_o)}P_{\mathrm{d}}\Xi_i^{(\mathrm{d})}\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}|\mathbf{h}_{0k}^{(\mathrm{c})*}\mathbf{h}_i^{(\mathrm{d})}|^2\geq\frac{\epsilon}{2}\right)$$

$$+\mathbb{P}\left(\frac{1}{M^2}\sum_{i\in\Phi\cap\mathcal{B}(o,r_o)}P_{\mathrm{d}}\Xi_i^{(\mathrm{d})}\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}|\mathbf{h}_{0k}^{(\mathrm{c})*}\mathbf{h}_i^{(\mathrm{d})}|^2\geq\frac{\epsilon}{2}\right). \tag{4.42}$$

Next we show in two steps that the two terms on the right hand side of (4.42) can be made arbitrarily small by choosing M large enough.

Step 1. For the first term on the right hand side of (4.42), we have

$$\begin{aligned}
&\mathbb{P}\left(\frac{1}{M^2}\sum_{i\in\Phi\cap\mathcal{B}^c(o,r_o)} P_{\mathrm{d}}\Xi_i^{(\mathrm{d})}\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}|\mathbf{h}_{0k}^{(\mathrm{c})*}\mathbf{h}_i^{(\mathrm{d})}|^2 \geq \frac{\epsilon}{2}\right)\\
&\leq \frac{2}{\epsilon}\mathbb{E}\left[\frac{1}{M^2}\sum_{i\in\Phi\cap\mathcal{B}^c(o,r_o)} P_{\mathrm{d}}\Xi_i^{(\mathrm{d})}\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}|\mathbf{h}_{0k}^{(\mathrm{c})*}\mathbf{h}_i^{(\mathrm{d})}|^2\right] &(4.43)\\
&= \frac{2P_{\mathrm{d}}\bar{\Xi}}{\epsilon M}\mathbb{E}[\sum_{i\in\Phi\cap\mathcal{B}^c(o,r_o)}\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}] &(4.44)\\
&= \frac{4\pi\lambda P_{\mathrm{d}}\bar{\Xi}}{\epsilon M}\int_{r_o}^{\infty} r^{1-\alpha_{\mathrm{c}}}\,\mathrm{d}r &(4.45)\\
&= \frac{4\pi\lambda P_{\mathrm{d}}\bar{\Xi}}{\epsilon M}\frac{1}{(\alpha_{\mathrm{c}}-2)r_o^{\alpha_{\mathrm{c}}-2}}, &(4.46)
\end{aligned}$$

where (4.43) is due to the Markov inequality, (4.44) is due to $\mathbb{E}[\Xi_i^{(\mathrm{d})}] = \bar{\Xi}$ and $\mathbb{E}[|\mathbf{h}_{0k}^{(\mathrm{c})*}\mathbf{h}_i^{(\mathrm{d})}|^2] = M$, and (4.45) is due to Campbell's formula [19], and we use the assumption that $\alpha_{\mathrm{c}} > 2$ in (4.46). It follows that that there exists M_1 large enough such that for all $M \geq M_1$,

$$\mathbb{P}\left(\frac{1}{M^2}\sum_{i\in\Phi\cap\mathcal{B}^c(o,r_o)} P_{\mathrm{d}}\Xi_i^{(\mathrm{d})}\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}|\mathbf{h}_{0k}^{(\mathrm{c})*}\mathbf{h}_i^{(\mathrm{d})}|^2 \geq \frac{\epsilon}{2}\right) < \frac{\delta}{2}, \tag{4.47}$$

where δ is an arbitrary small positive constant.

Step 2. For the second term on the right hand side of (4.42),

$$\mathbb{P}\left(\frac{1}{M^2}\sum_{i\in\Phi\cap\mathcal{B}(o,r_o)} P_{\mathrm{d}}\Xi_i^{(\mathrm{d})}\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}|\mathbf{h}_{0k}^{(\mathrm{c})*}\mathbf{h}_i^{(\mathrm{d})}|^2 \geq \frac{\epsilon}{2}\right)$$

$$= \mathbb{P}\left(\frac{1}{M^2} \sum_{i \in \Phi \cap \mathcal{B}(o,r_o)} P_{\mathrm{d}} \Xi_i^{(\mathrm{d})} \|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}} |\mathbf{h}_{0k}^{(\mathrm{c})*} \mathbf{h}_i^{(\mathrm{d})}|^2 \geq \frac{\epsilon}{2} \middle| E \right) \mathbb{P}(E)$$
$$+ \mathbb{P}\left(\frac{1}{M^2} \sum_{i \in \Phi \cap \mathcal{B}(o,r_o)} P_{\mathrm{d}} \Xi_i^{(\mathrm{d})} \|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}} |\mathbf{h}_{0k}^{(\mathrm{c})*} \mathbf{h}_i^{(\mathrm{d})}|^2 \geq \frac{\epsilon}{2} \middle| E^c \right) \mathbb{P}(E^c).$$

where $E = \{|\Phi \cap \mathcal{B}(o, r_o)| \leq C\}$ and E^c is the complement of E.

Step 2(a). Note that the number of D2D transmitters in $\mathcal{B}(o, r_o)$, denoted as $|\Phi \cap \mathcal{B}(o, r_o)|$, is Poisson distributed with mean $\lambda \pi r_o^2$. We can choose C large enough such that

$$\mathbb{P}\left(\frac{1}{M^2} \sum_{i \in \Phi \cap \mathcal{B}(o,r_o)} P_{\mathrm{d}} \Xi_i^{(\mathrm{d})} \|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}} |\mathbf{h}_{0k}^{(\mathrm{c})*} \mathbf{h}_i^{(\mathrm{d})}|^2 \geq \frac{\epsilon}{2} \middle| E^c \right) \mathbb{P}(E^c)$$
$$\leq \mathbb{P}(E^c) = 1 - \sum_{n=0}^{C} \frac{(\lambda \pi r_o^2)^n}{n!} e^{-\lambda \pi r_o^2} < \frac{\delta}{4}. \tag{4.48}$$

Note that the choice of C depends on δ.

Step 2(b). Since $\frac{1}{M} \mathbf{h}_{0k}^{(\mathrm{c})*} \mathbf{h}_i^{(\mathrm{d})} \xrightarrow{a.s.} 0$, conditioning on $|\Phi \cap \mathcal{B}(o, r_o)| \leq C$, we have

$$\frac{1}{M^2} \sum_{i \in \Phi \cap \mathcal{B}(o,r_o)} P_{\mathrm{d}} \Xi_i^{(\mathrm{d})} \|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}} |\mathbf{h}_{0k}^{(\mathrm{c})*} \mathbf{h}_i^{(\mathrm{d})}|^2 \xrightarrow{a.s.} 0.$$

It follows that there exists M_2 large enough such that for all $M \geq M_2$,

$$\mathbb{P}\left(\frac{1}{M^2} \sum_{i \in \Phi \cap \mathcal{B}(o,r_o)} P_{\mathrm{d}} \Xi_i^{(\mathrm{d})} \|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}} |\mathbf{h}_{0k}^{(\mathrm{c})*} \mathbf{h}_i^{(\mathrm{d})}|^2 \geq \frac{\epsilon}{2} \middle| E \right) \mathbb{P}(E)$$
$$\leq \mathbb{P}\left(\frac{1}{M^2} \sum_{i \in \Phi \cap \mathcal{B}(o,r_o)} P_{\mathrm{d}} \Xi_i^{(\mathrm{d})} \|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}} |\mathbf{h}_{0k}^{(\mathrm{c})*} \mathbf{h}_i^{(\mathrm{d})}|^2 \geq \frac{\epsilon}{2} \middle| E \right) < \frac{\delta}{4}. \tag{4.49}$$

Combining (4.47), (4.48) and (4.49) obtained in Steps 1, 2(a) and 2(b) respectively, we have for all $M \geq \max\{M_1, M_2\}$,

$$\mathbb{P}\left(\frac{1}{M^2}\sum_{i\in\Phi} P_{\mathrm{d}}\Xi_i^{(\mathrm{d})}\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}|\mathbf{h}_{0k}^{(\mathrm{c})*}\mathbf{h}_i^{(\mathrm{d})}|^2 \geq \epsilon\right) \leq \frac{\delta}{2}+\frac{\delta}{4}+\frac{\delta}{4}=\delta.$$

As δ is an arbitrary positive constant, we conclude that (4.41) holds. This completes the proof.

4.8.2 Proof of Proposition 4.2

When the transmit powers of cellular UEs scale as P_{c}/M, as in the proof of Prop. 4.1, we can show that as $M \to \infty$, the desired signal power $S_k^{(\mathrm{c})}$, the cellular interference power $I_k^{(\mathrm{c}\to\mathrm{c})}$, and the noise power $\|\mathbf{w}_k^{(\mathrm{c})}\|^2 N_0$ normalized by M converge as follows.

$$\begin{aligned}
&\lim_{M\to\infty}\frac{1}{M}S_k^{(\mathrm{c})} \xrightarrow{a.s.} P_{\mathrm{c}}\Xi_{0k}^{(\mathrm{c})}\|x_{0k}^{(\mathrm{c})}\|^{-\alpha_{\mathrm{c}}}\\
&\lim_{M\to\infty}\frac{1}{M}I_k^{(\mathrm{c}\to\mathrm{c})} \xrightarrow{a.s.} 0\\
&\lim_{M\to\infty}\frac{1}{M}\|\mathbf{w}_k^{(\mathrm{c})}\|^2 N_0 \xrightarrow{a.s.} N_0.
\end{aligned} \tag{4.50}$$

For any D2D interferer $i \in \Phi_k^{(\mathrm{c})}$, it generates interference

$$P_{\mathrm{d}}\Xi_i^{(\mathrm{d})}\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}|\mathbf{w}_k^{(\mathrm{c})*}\mathbf{h}_i^{(\mathrm{d})}|^2.$$

By Central Limit Theorem, $\frac{1}{\sqrt{M}}\mathbf{w}_k^{(\mathrm{c})*}\mathbf{h}_i^{(\mathrm{d})} \xrightarrow{d.} \mathcal{CN}(0,1)$, where $\xrightarrow{d.}$ denotes the convergence in distribution. It follows that $\frac{1}{M}|\mathbf{w}_k^{(\mathrm{c})*}\mathbf{h}_i^{(\mathrm{d})}|^2 \xrightarrow{d.} \eta_i \sim \mathrm{Exp}(1)$. Therefore, D2D interferer i's interference power normalized by M converges

to $P_{\mathrm{d}}\Xi_i^{(\mathrm{d})}\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}\eta_i$ in distribution. Summing over all the D2D interferers, we have

$$\lim_{M\to\infty}\frac{1}{M}I_k^{(\mathrm{d}\to\mathrm{c})}\xrightarrow{d.}\sum_{i\in\Phi_k^{(\mathrm{c})}}P_{\mathrm{d}}\Xi_i^{(\mathrm{d})}\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}\eta_i. \tag{4.51}$$

Therefore, the spectral efficiency of cellular UE k converges as in (4.11).

With a slight abuse of notation, we also denote by $I_k^{(\mathrm{d}\to\mathrm{c})}$ the asymptotic interference power $\sum_{i\in\Phi_k^{(\mathrm{c})}}P_{\mathrm{d}}\Xi_i^{(\mathrm{d})}\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}\eta_i$. The lower bound (4.12) is due to Jensen's inequality:

$$\mathbb{E}\left[\log\left(1+\frac{P_{\mathrm{c}}\Xi_{0k}^{(\mathrm{c})}\|x_{0k}^{(\mathrm{c})}\|^{-\alpha_{\mathrm{c}}}}{I_k^{(\mathrm{d}\to\mathrm{c})}+N_0}\right)\right]\geq\log\left(1+\frac{P_{\mathrm{c}}\Xi_{0k}^{(\mathrm{c})}\|x_{0k}^{(\mathrm{c})}\|^{-\alpha_{\mathrm{c}}}}{\mathbb{E}[I_k^{(\mathrm{d}\to\mathrm{c})}]+N_0}\right). \tag{4.52}$$

As the BS uses m_{d} degrees of freedom to cancel the interference from the m_{d} nearest D2D transmitters when detecting the signal of cellular UE k, $\Phi_k^{(\mathrm{c})}$ consists of the points from the original PPP Φ except the nearest m_{d} points to the origin. Let us order the points in Φ based on their distances to the BS in an ascending manner, i.e., $\|x_1^{(\mathrm{d})}\|\leq\|x_2^{(\mathrm{d})}\|\leq....$ Then

$$\begin{aligned}\mathbb{E}[I_k^{(\mathrm{d}\to\mathrm{c})}]&=\mathbb{E}_\Phi[\sum_{i\in\Phi_k^{(\mathrm{c})}}P_{\mathrm{d}}\mathbb{E}[\Xi_i^{(\mathrm{d})}]\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}\mathbb{E}[\eta_i]]=P_{\mathrm{d}}\bar{\Xi}\mathbb{E}_\Phi[\sum_{i\in\Phi_k^{(\mathrm{c})}}\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}]\\&=P_{\mathrm{d}}\bar{\Xi}\mathbb{E}_\Phi[\sum_{i=m_{\mathrm{d}}+1}^{\infty}\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}].\end{aligned} \tag{4.53}$$

Conditioning on the location $x_{m_{\mathrm{d}}}^{(\mathrm{d})}=(r,\theta)$ of the m_{d}-th nearest point in Φ,

$$\begin{aligned}\mathbb{E}[I_k^{(\mathrm{d}\to\mathrm{c})}|x_{m_{\mathrm{d}}}^{(\mathrm{d})}=(r,\theta)]&=P_{\mathrm{d}}\bar{\Xi}\mathbb{E}_\Phi\left[\sum_{i=m_{\mathrm{d}}+1}^{\infty}\|x_i^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}}\big|x_{m_{\mathrm{d}}}^{(\mathrm{d})}=(r,\theta)\right]\\&=P_{\mathrm{d}}\bar{\Xi}2\pi\lambda\int_r^{\infty}t^{1-\alpha_{\mathrm{c}}}\,\mathrm{d}t=\frac{P_{\mathrm{d}}\bar{\Xi}2\pi\lambda}{\alpha_{\mathrm{c}}-2}r^{2-\alpha_{\mathrm{c}}},\end{aligned} \tag{4.54}$$

where the second equality is due to Campbell formula [19]. To decondition on $x_{m_\mathrm{d}}^{(\mathrm{d})} = (r, \theta)$, we need the PDF of $\|x_{m_\mathrm{d}}^{(\mathrm{d})}\|$ derived in [52]:

$$f_{\|x_{m_\mathrm{d}}^{(\mathrm{d})}\|}(r) = \frac{2(\lambda\pi r^2)^{m_\mathrm{d}}}{r(m_\mathrm{d}-1)!} e^{-\lambda\pi r^2}, \quad r \geq 0. \tag{4.55}$$

Using the fact that $x_{m_\mathrm{d}}^{(\mathrm{d})}$ is uniform in direction and $f_{\|x_{m_\mathrm{d}}^{(\mathrm{d})}\|}(r)$, we decondition on $x_{m_\mathrm{d}}^{(\mathrm{d})}$ in (4.54) and obtain

$$\begin{aligned}\mathbb{E}[I_k^{(\mathrm{d}\to\mathrm{c})}] &= \frac{P_\mathrm{d}\bar{\Xi}2\pi\lambda}{\alpha_\mathrm{c}-2} \int_0^\infty r^{2-\alpha_\mathrm{c}} f_{\|x_{m_\mathrm{d}}^{(\mathrm{d})}\|}(r)\, \mathrm{d}r \\ &= \frac{P_\mathrm{d}\bar{\Xi}2\pi\lambda}{\alpha_\mathrm{c}-2} \cdot \frac{1}{(m_\mathrm{d}-1)!} (\lambda\pi)^{\frac{\alpha_\mathrm{c}}{2}-1} \int_0^\infty t^{m_\mathrm{d}-\frac{\alpha_\mathrm{c}}{2}} e^{-t}\, \mathrm{d}t, \end{aligned} \tag{4.56}$$

where we have changed variable $t = \lambda\pi r^2$ in (4.56). By the definition of the Gamma function,

$$\mathbb{E}[I_k^{(\mathrm{d}\to\mathrm{c})}] = \frac{2P_\mathrm{d}\bar{\Xi}}{\alpha_\mathrm{c}-2} (\pi\lambda)^{\frac{\alpha_\mathrm{c}}{2}} \frac{\Gamma(m_\mathrm{d}+1-\frac{\alpha_\mathrm{c}}{2})}{\Gamma(m_\mathrm{d})}, \tag{4.57}$$

Plugging (4.57) into (4.52) yields the desired lower bound (4.12).

4.8.3 Proof of Proposition 4.4

Using the convexity of the function $\log(1+\frac{1}{x})$ and applying Jensen's inequality [103],

$$\begin{aligned} R_k^{(\mathrm{c})} \geq R_k^{(\mathrm{c,lb})} &= \log\left(1 + \left(\mathbb{E}\left[\frac{1}{\mathrm{SINR}_k^{(\mathrm{c})}}\right]\right)^{-1}\right) \\ &= \log\left(1 + \left(\mathbb{E}\left[\frac{1}{S_k^{(\mathrm{c})}}\right] \cdot (\mathbb{E}[I_k^{(\mathrm{c}\to\mathrm{c})}] + \mathbb{E}[I_k^{(\mathrm{d}\to\mathrm{c})}] + N_0)\right)^{-1}\right). \end{aligned} \tag{4.58}$$

In the following three steps, we calculate $\mathbb{E}\left[\frac{1}{S_k^{(\mathrm{c})}}\right]$, $\mathbb{E}[I_k^{(\mathrm{c}\to\mathrm{c})}]$, and $\mathbb{E}[I_k^{(\mathrm{d}\to\mathrm{c})}]$, respectively. Without loss of generality, we assume that $\mathbf{w}_k^{(\mathrm{c})}$ is normalized, i.e., $\|\mathbf{w}_k^{(\mathrm{c})}\| = 1$.

Step 1: calculating $\mathbb{E}\left[\frac{1}{S_k^{(\mathrm{c})}}\right]$. By definition $\|\mathbf{w}_k^{(\mathrm{c})*}\mathbf{h}_{0k}^{(\mathrm{c})}\|^2$ is the squared norm of the projection of the vector $\mathbf{h}_{0k}^{(\mathrm{c})}$ onto the subspace orthogonal to the one spanned by the channel vectors of canceled interferers. The space is of $M - m_{\mathrm{c}} - m_{\mathrm{d}}$ dimensions and is independent of $\mathbf{h}_{0k}^{(\mathrm{c})}$. It follows that $\|\mathbf{w}_k^{(\mathrm{c})*}\mathbf{h}_{0k}^{(\mathrm{c})}\|^2 \sim \chi^2_{2(M-m_{\mathrm{c}}-m_{\mathrm{d}})}$, i.e., $\|\mathbf{w}_k^{(\mathrm{c})*}\mathbf{h}_{0k}^{(\mathrm{c})}\|^2 \sim \Gamma(M-m_{\mathrm{c}}-m_{\mathrm{d}},1)$. Therefore, conditioned on $\Xi_{0k}^{(\mathrm{c})}$ and $x_{0k}^{(\mathrm{c})}$, $\frac{1}{S_k^{(\mathrm{c})}}$ is inverse-Gamma distributed and its mean equals

$$\mathbb{E}\left[\frac{1}{S_k^{(\mathrm{c})}}\right] = \frac{1}{P_{\mathrm{c}}\Xi_{0k}^{(\mathrm{c})}\|x_{0k}^{(\mathrm{c})}\|^{-\alpha_{\mathrm{c}}}(M - m_{\mathrm{c}} - m_{\mathrm{d}} - 1)}. \tag{4.59}$$

Step 2: calculating $\mathbb{E}[I_k^{(c\to c)}]$. Since $\|\mathbf{w}_k^{(\mathrm{c})}\| = 1$ and $\mathbf{w}_k^{(\mathrm{c})}$ is independent of $\mathbf{h}_{b\ell}^{(\mathrm{c})}, \forall \ell \in \mathcal{K}_{bk}^{(\mathrm{c})}, \forall b$, $\mathbf{w}_k^{(\mathrm{c})*}\mathbf{h}_{b\ell}^{(\mathrm{c})}$ is a linear combination of complex Gaussian random variables and thus is distributed as $\mathcal{CN}(0,1)$. It follows that $|\mathbf{w}_k^{(\mathrm{c})*}\mathbf{h}_{b\ell}^{(\mathrm{c})}|^2 \sim \mathrm{Exp}(1)$ and

$$\begin{aligned}\mathbb{E}[I_k^{(\mathrm{c}\to\mathrm{c})}] &= \mathbb{E}[\sum_{b=0}^{B}\sum_{\ell\in\mathcal{K}_{bk}^{(\mathrm{c})}} P_{\mathrm{c}}\Xi_{b\ell}^{(\mathrm{c})}\|x_{b\ell}^{(\mathrm{c})}\|^{-\alpha_{\mathrm{c}}}|\mathbf{w}_k^{(\mathrm{c})*}\mathbf{h}_{b\ell}^{(\mathrm{c})}|^2] \\ &= \sum_{b=0}^{B}\sum_{\ell\in\mathcal{K}_{bk}^{(\mathrm{c})}} P_{\mathrm{c}}\bar{\Xi}\|x_{b\ell}^{(\mathrm{c})}\|^{-\alpha_{\mathrm{c}}}. \end{aligned}\tag{4.60}$$

Step 3: calculating $\mathbb{E}[I_k^{(d\to c)}]$. With a similar argument as in Step 2, we

have $|\mathbf{w}_k^{(\mathrm{c})*}\mathbf{h}_i^{(\mathrm{d})}|^2 \sim \mathrm{Exp}(1)$ and

$$\begin{aligned}\mathbb{E}[I_k^{(\mathrm{d}\to\mathrm{c})}] &= \mathbb{E}_\Phi[\sum_{i\in\Phi_k^{(\mathrm{c})}} P_\mathrm{d}\Xi_i^{(\mathrm{d})}\|x_i^{(\mathrm{d})}\|^{-\alpha_\mathrm{c}}\mathbb{E}_\mathbf{h}[|\mathbf{w}_k^{(\mathrm{c})*}\mathbf{h}_i^{(\mathrm{d})}|^2]] \\ &= P_\mathrm{d}\bar{\Xi}\mathbb{E}_\Phi[\sum_{i\in\Phi_k^{(\mathrm{c})}} \|x_i^{(\mathrm{d})}\|^{-\alpha_\mathrm{c}}]. \end{aligned} \tag{4.61}$$

The remaining steps for calculating $\mathbb{E}[I_k^{(\mathrm{d}\to\mathrm{c})}]$ follow the same steps in the proof of Prop. 4.2, i.e., the steps after (4.53), and $\mathbb{E}[I_k^{(\mathrm{d}\to\mathrm{c})}]$ is given in (4.57).

Finally, plugging (4.59), (4.60) and (4.57) into (4.58) completes the proof.

4.8.4 Proof of Lemma 4.1

Since the $M\times T_\mathrm{c}$ dimensional noise matrix $\mathbf{V}_0^{(\mathrm{c})}$ consists of i.i.d. $\mathcal{CN}(0,N_0)$ elements and $\mathbf{q}_{\tilde{k}}^{(\mathrm{c})}$ is an orthonomal vector, $\mathbf{V}_0^{(\mathrm{c})}\mathbf{q}_{\tilde{k}}^{(\mathrm{c})}$ also consists of i.i.d. $\mathcal{CN}(0,N_0)$ elements. Similarly, $\mathbf{u}_{br}^{(\mathrm{d})*}\mathbf{q}_{\tilde{k}}^{(\mathrm{c})} \sim \mathcal{CN}(0,1)$. Using the independence of $\mathbf{h}_{br}^{(\mathrm{d})}$, $\mathbf{u}_{br}^{(\mathrm{d})*}\mathbf{q}_k^{(\mathrm{c})}$ and $\mathbf{V}_0^{(\mathrm{c})}\mathbf{q}_{\tilde{k}}^{(\mathrm{c})}$, we have $\mathbb{E}[\tilde{\mathbf{v}}_k^{(s)}]=0$, and

$$\begin{aligned}\mathbb{E}[\tilde{\mathbf{v}}_k^{(s)}\tilde{\mathbf{v}}_k^{(s)*}] &= \frac{1}{T_\mathrm{c}P_\mathrm{c}\Xi_{0k}^{(s)}\|x_{0k}^{(s)}\|^{-\alpha_\mathrm{c}}}\Bigg(\sum_{b=0}^{B+1}\sum_{r\in\Phi_{bk}^{(\mathrm{c})}} P_\mathrm{d}\Xi_{br}^{(\mathrm{d})}\|x_{br}^{(\mathrm{d})}\|^{-\alpha_\mathrm{c}}\times \\ &\qquad \mathbb{E}[\mathbf{h}_{br}^{(\mathrm{d})}\mathbf{u}_{br}^{(\mathrm{d})*}\mathbf{q}_{\tilde{k}}^{(\mathrm{c})}\mathbf{q}_{\tilde{k}}^{(\mathrm{c})*}\mathbf{u}_{br}^{(\mathrm{d})}\mathbf{h}_{br}^{(\mathrm{d})*}] + \mathbb{E}[\mathbf{V}^{(\mathrm{c})}\mathbf{q}_{\tilde{k}}^{(\mathrm{c})}\mathbf{q}_{\tilde{k}}^{(\mathrm{c})*}\mathbf{V}^{(\mathrm{c})*}]\Bigg) \\ &= \frac{1}{T_\mathrm{c}P_\mathrm{c}\Xi_{0k}^{(s)}\|x_{0k}^{(s)}\|^{-\alpha_\mathrm{c}}}\left(\sum_{b=0}^{B+1}\sum_{r\in\Phi_{bk}^{(\mathrm{c})}} P_\mathrm{d}\Xi_{br}^{(\mathrm{d})}\|x_{br}^{(\mathrm{d})}\|^{-\alpha_\mathrm{c}}\mathbb{E}[\mathbf{h}_{br}^{(\mathrm{d})}\mathbf{h}_{br}^{(\mathrm{d})*}] + N_0\mathbf{I}_M\right) \\ &= \frac{\sum_{b=0}^{B+1}\sum_{r\in\Phi_{bk}^{(\mathrm{c})}} P_\mathrm{d}\Xi_{br}^{(\mathrm{d})}\|x_{br}^{(\mathrm{d})}\|^{-\alpha_\mathrm{c}} + N_0}{T_\mathrm{c}P_\mathrm{c}\Xi_{0k}^{(s)}\|x_{0k}^{(s)}\|^{-\alpha_\mathrm{c}}}\mathbf{I}_M. \end{aligned} \tag{4.62}$$

Further, using that $\mathbf{h}_{bk}^{(s)}$ consists of i.i.d. $\mathcal{CN}(0,1)$ elements and the independence of $\mathbf{h}_{bk}^{(s)}$ and $\tilde{\mathbf{v}}_k^{(s)}$,

$$\mathbb{E}[\mathbf{h}_{0k}^{(s)}\tilde{\mathbf{y}}_k^{(s)*}] = \mathbb{E}[\mathbf{h}_{0k}^{(s)}\mathbf{h}_{0k}^{(s)*} + \sum_{b=1}^{B}\sqrt{\beta_{bk}^{(s)}}\mathbf{h}_{0k}^{(s)}\mathbf{h}_{bk}^{(s)*} + \mathbf{h}_{0k}^{(s)}\tilde{\mathbf{v}}_k^{(s)*}] = \mathbf{I}_M. \tag{4.63}$$

Similarly, we have

$$\begin{aligned}\mathbb{E}[\tilde{\mathbf{y}}_k^{(s)}\tilde{\mathbf{y}}_k^{(s)*}] &= \mathbb{E}[\mathbf{h}_{0k}^{(s)}\mathbf{h}_{0k}^{(s)*} + \sum_{b=1}^{B}\beta_{bk}^{(s)}\mathbf{h}_{bk}^{(s)}\mathbf{h}_{bk}^{(s)*} + \tilde{\mathbf{v}}_k^{(s)}\tilde{\mathbf{v}}_k^{(s)*}] \\ &= \left(1 + \sum_{b=1}^{B}\beta_{bk}^{(s)} + \frac{\sum_{b=0}^{B+1}\sum_{r\in\Phi_{bk}^{(\mathrm{c})}}P_{\mathrm{d}}\Xi_{br}^{(\mathrm{d})}\|x_{br}^{(\mathrm{d})}\|^{-\alpha_{\mathrm{c}}} + N_0}{T_{\mathrm{c}}P_{\mathrm{c}}\Xi_{0k}^{(s)}\|x_{0k}^{(s)}\|^{-\alpha_{\mathrm{c}}}}\right)\mathbf{I}_M \\ &= \frac{1}{\xi_k^{(s)}}\mathbf{I}_M. \end{aligned}\tag{4.64}$$

Therefore, the MMSE estimate of $\mathbf{h}_{0k}^{(s)}$ is

$$\hat{\mathbf{h}}_{0k}^{(s)} = \mathbb{E}[\mathbf{h}_{0k}^{(s)}\tilde{\mathbf{y}}_k^{(s)*}](\mathbb{E}[\tilde{\mathbf{y}}_k^{(s)}\tilde{\mathbf{y}}_k^{(s)*}])^{-1}\tilde{\mathbf{y}}_k^{(s)} = \xi_k^{(s)}\tilde{\mathbf{y}}_k^{(s)}. \tag{4.65}$$

Clearly, $\hat{\mathbf{h}}_{0k}^{(s)}$ is zero mean and its covariance is $\mathbb{E}[\hat{\mathbf{h}}_{0k}^{(s)}\hat{\mathbf{h}}_{0k}^{(s)*}] = \xi_k^{(s)}\mathbf{I}_M$. As for the estimation error $\boldsymbol{\epsilon}_k^{(s)} = \mathbf{h}_{0k}^{(s)} - \hat{\mathbf{h}}_{0k}^{(s)}$, it is clearly zero mean and its covariance is

$$\mathbb{E}[\boldsymbol{\epsilon}_k^{(s)}\boldsymbol{\epsilon}_k^{(s)*}] = \mathbb{E}[\mathbf{h}_{0k}^{(s)}\mathbf{h}_{0k}^{(s)*}] - \mathbb{E}[\hat{\mathbf{h}}_{0k}^{(s)}\hat{\mathbf{h}}_{0k}^{(s)*}] = (1-\xi_k^{(s)})\mathbf{I}_M. \tag{4.66}$$

4.8.5 Proof of Proposition 4.6

With D2D links deactivated in the training phase, we have

$$\begin{aligned}&\lim_{M\to\infty}\frac{1}{\sqrt{M}}\left(\frac{1}{M^{1/4}}\sum_{b=0}^{B}\sqrt{T_{\mathrm{c}}P_{\mathrm{c}}\Xi_{bk}^{(\mathrm{c})}}\|x_{bk}^{(\mathrm{c})}\|^{-\frac{\alpha_{\mathrm{c}}}{2}}\mathbf{h}_{bk}^{(\mathrm{c})} + \bar{\mathbf{v}}_0^{(\mathrm{c})}\right)^{*}\mathbf{y}_0^{(\mathrm{c})} \\ &= \sum_{b=0}^{B}\sqrt{T_{\mathrm{c}}}P_{\mathrm{c}}\Xi_{bk}^{(\mathrm{c})}\|x_{bk}^{(\mathrm{c})}\|^{-\alpha_{\mathrm{c}}}u_{bk}^{(\mathrm{c})}\end{aligned}$$

$$+ \lim_{M\to\infty} \frac{1}{M^{3/4}} \sum_{b=0}^{B} \sqrt{T_{\rm c} P_{\rm c} \Xi_{bk}^{(\rm c)}} \|x_{bk}^{(\rm c)}\|^{-\frac{\alpha_{\rm c}}{2}} \sum_{i\in\Phi} \sqrt{P_{\rm d} \Xi_i^{(\rm d)}} \|x_i^{(\rm d)}\|^{-\frac{\alpha_{\rm c}}{2}} \mathbf{h}_{bk}^{(\rm c)*} \mathbf{h}_i^{(\rm d)} u_i^{(\rm d)}$$

$$+ \lim_{M\to\infty} \frac{1}{M^{3/4}} \sum_{b=0}^{B} \sqrt{T_{\rm c} P_{\rm c} \Xi_{bk}^{(\rm c)}} \|x_{bk}^{(\rm c)}\|^{-\frac{\alpha_{\rm c}}{2}} \mathbf{h}_{bk}^{(\rm c)*} \mathbf{v}_0^{(\rm c)}$$

$$+ \lim_{M\to\infty} \frac{1}{M^{3/4}} \sum_{b=0}^{B} \sqrt{P_{\rm c} \Xi_{bk}^{(\rm c)}} \|x_{bk}^{(\rm c)}\|^{-\frac{\alpha_{\rm c}}{2}} \bar{\mathbf{v}}_0^{(\rm c)*} \mathbf{h}_{bk}^{(\rm c)} u_{bk}^{(\rm c)}$$

$$+ \lim_{M\to\infty} \frac{1}{\sqrt{M}} \sum_{i\in\Phi} \sqrt{P_{\rm d} \Xi_i^{(\rm d)}} \|x_i^{(\rm d)}\|^{-\frac{\alpha_{\rm c}}{2}} \bar{\mathbf{v}}_0^{(\rm c)*} \mathbf{h}_i^{(\rm d)} u_i^{(\rm d)} + \lim_{M\to\infty} \frac{1}{\sqrt{M}} \bar{\mathbf{v}}_0^{(\rm c)*} \mathbf{v}_0^{(\rm c)}. \quad (4.67)$$

For the second term on the right hand side of (4.67), we can show that it converges to 0 in probability by following the same arguments of the proof of Prop. 4.1. For the third and fourth terms on the right hand side of (4.67), it is clear that they converge to 0 almost surely. The last term on the right hand side of (4.67) converges in distribution to a zero-mean complex Gaussian random variable of variance N_0^2. The fifth term is zero mean and has variance $\sum_{i\in\Phi} P_{\rm d} \Xi_i^{(\rm d)} \|x_i^{(\rm d)}\|^{-\alpha_{\rm c}} N_0$ but not Gaussian. Using the worst-case noise argument, we conclude (4.31) is achievable.

Chapter 5

Asynchronous Multicarrier Wireless Networks

An implicit assumption made in Chapters 2 to 4 is that the networks are synchronized. In D2D discovery, a UE seeks to identify other UEs in its proximity via periodically broadcasting/receiving discovery signals. A listening UE can decode the discovery signal broadcast by a transmitting UE if the SINR exceeds some detection threshold, which depends on the used modulation and coding scheme. The number of transmitting UEs that can be decoded is an important metric of discovery effectiveness. In the discovery process, devices are usually not or imperfectly synchronized and thus different devices have different notions of timing. Therefore, this chapter investigates the following question: if we take a snapshot of the network at a randomly selected time-frequency slot and randomly select a receiving UE, then how many (if any) transmitting UEs can be discovered or decoded by the selected receiving UE given that the UEs each have different notions of timing?

The question posed above is of interest not only in D2D discovery but also in general cellular systems and ad hoc networks. For example, in the downlink of a cellular network, how many BSs can be decoded by a typical

UE at a given SINR? This is a key consideration for soft handover or multiple-BS coverage or offloading in dense heterogeneous networks [79]. A similar metric can be used for neighbor discovery in wireless ad hoc networks [130].

5.1 Background and Related Work

The answer to the posed question obviously depends on how the transmitting nodes are spatially distributed. We assume that the transmitting nodes are distributed according to a PPP, which has been recently shown to accurately model (with small modifications or shifts) a very large class of wireless networks, including even regular grids (with sufficient shadowing) [22] and most random spatial distributions with a small and constant SINR shift [49]. It is therefore reasonable to assume that the conclusions in this chapter also will hold for most plausible network topologies. Because of its excellent analytical properties, the PPP has found numerous applications in various types of wireless networks [10, 17, 40, 54].

Despite this encouraging progress in applying the PPP to wireless networking, existing works nearly universally assume that the networks are perfectly synchronized. In cellular networks, BSs in different cells may not be synchronized in a Frequency Division Duplex (FDD) deployment, or have synchronization errors in a Time Division Duplex (TDD) deployment. These facts also lead to synchronization issues in D2D discovery. In particular, UEs participating in the discovery are synchronized with their associated BSs and thus may not be synchronized or at best imperfectly synchronized among them-

selves even when other factors like propagation delays are not considered, let alone the UEs that are out of cellular coverage [88]. The synchronous assumption becomes even more questionable when it comes to an ad hoc network in which network-wide synchronization is almost impossible. In such contexts, different transmitters have different notions of timing. From the viewpoint of a typical receiver, which also has its own notion of timing, the multicarrier OFDM signals from the transmitters are asynchronous and also do not align with the receiver's timing, leading to a loss of orthogonality between subcarriers.

The impact of synchronization errors on single-user OFDM has been extensively investigated in the literature (see e.g. [102, 110, 116, 121, 137]). Extension of the analysis in single-user OFDM to multiuser OFDM, however, is not straightforward as the latter involves a much larger set of random variables. Analysis of asynchronous OFDM in the uplink of cellular systems includes [39,108,112,126], while the downlink counterpart may be found in [58,100] and ad hoc networks in [57]. The works [39, 108, 112, 126] are focused on a single-cell setting and do not consider other-cell interference that plays a key role in system-level performance. In contrast, cochannel interference is modeled and studied in [57, 58, 100]. But [57, 58, 100] do not consider or leverage the randomness inherent in the positions of network nodes, and the system-level studies therein are mainly based on Monte Carlo simulations.

5.2 Contributions and Main Outcomes

The main goal of this chapter is to incorporate the effect of asynchronous OFDM transmissions in the system-level study of wireless networks in which the positions of transmitting nodes are modeled by a PPP. The main contributions and outcomes of this chapter are summarized as follows.

A tractable SINR model for asynchronous OFDM networks. We carry out a detailed link-level analysis on the impact of timing misalignment in OFDM transmission. Based on the link-level analysis, we propose a tractable first-order SINR model, which can be conveniently used in system-level studies.

System-level analysis of asynchronous PPP networks. We apply the proposed SINR model to study the system-level performance of asynchronous networks where the locations of transmitting nodes are modeled by a PPP and an OFDM waveform is used. Taking from a typical receiver's point of view, we derive analytical results for the average number of decodable transmitters, the decoding probability of the nearest transmitter, and system throughput. Further, we derive an upper bound on the distribution of the number of decodable transmitters. Note that, according to Palm theory [16], the statistical performance experienced by a typical receiver is equivalently the spatially averaged performance over all receivers. The analysis of perfectly synchronized networks can be treated as a special case of this work. For example, the result on the decoding probability of the nearest transmitter reduces to [10] that studies a perfectly synchronized cellular network.

Solutions for mitigating the impact of asynchronous transmissions. We compare and discuss four possible solutions including extended cyclic prefix, advanced receiver timing, dynamic receiver timing positioning, and semi-static receiver timing positioning with multiple timing hypotheses. These solutions, detailed in Section 5.6, differ in complexity and may be applicable in different scenarios for mitigating the loss due to asynchronous transmissions.

5.3 System Model

We consider a network in which transmitters use an OFDM waveform. The baseband equivalent time-domain signal $s_i(t)$ emitted by transmitter i can be written as

$$s_i(t) = \sqrt{E_i} \sum_{m=-\infty}^{\infty} \frac{1}{N} \sum_{k} S_i[k;m] e^{j2\pi \frac{k}{T}(t-mT_s)} \mathbb{I}_{[-T_{cp},T_d)}(t-mT_s), \qquad (5.1)$$

where E_i denotes the transmit energy per sample of transmitter i, m is the OFDM symbol index, N denotes the total number of subcarriers, k is the subcarrier index, $S_i[k;m]$ denotes transmitter i's data symbol on the k-th subcarrier during the m-th OFDM symbol, $T_s = T_d + T_{cp}$ denotes the duration of an OFDM symbol with T_d denoting the duration of the data part and T_{cp} the duration of the cyclic prefix, and $\mathbb{I}_A(t)$ is an indicator function: it equals 1 if $t \in A$ and zero otherwise.. The data symbols $\{S_i[k;m]\}$ are complex and assumed to be independent and identically distributed (i.i.d.) with zero mean and unit variance.

We are interested in asynchronous scenarios where different transmitters have different notions of timing and so do the receivers. The more commonly studied synchronous scenarios where all the nodes are synchronized is a special case of this model. In an asynchronous network, we are interested in what a typical receiver "sees" at a random time-frequency resource unit. Note that the spectral width can be arbitrary. It can be a complete OFDM channel or a subband of an OFDM channel. In the latter case, transmitter i simply puts zero-valued data symbols $S_i[k;m]$ on the unused subcarriers, as in OFDMA.

The active transmitters at the time-frequency resource unit in question are assumed to be randomly distributed according to a PPP Φ with density λ. The location of transmitter $i \in \Phi$ is denoted by X_i. Note that our model does not preclude the possibility that there may be other transmitters active at some other time-frequency resource units. For example, we may consider a super PPP $\Phi' \supseteq \Phi$, where Φ' denotes the set of all the nodes in the network, and a time-frequency grid composed of orthogonal time-frequency resource units.[1] Each node randomly selects a time-frequency resource unit and transmits an OFDM waveform. Then the active transmitters at a randomly selected time-frequency resource unit constitute a PPP Φ, thinned from the super PPP Φ'. This described random access scheme is in fact part of the D2D discovery design used in LTE Direct [5].

[1]We ignore possible leakages from other time-frequency resource units when considering a particular time-frequency resource unit.

In this asynchronous network, we will study system-level questions such as the number of transmitting nodes that can be decoded by a typical receiver. To this end, since the transmitter process is stationary, we may assume without loss of generality that the typical receiver is located at the origin. Further, we consider flat-fading OFDM channels, i.e., the multipath spreads are small (w.r.t. sampling period). The last assumption holds for example in the following three scenarios: (1) there are not many obstacles in the radio environment and the arrival times of the multipaths are not resolvable at the receiver; (2) the received signal power is dominated by a single path, e.g. the line-of-sight path if it exists; and (3) the transmit signal is restricted to a flat-fading subband of a frequency-selective channel, as in OFDMA. We leave the important extension to frequency-selective OFDM channels as future work.

More specific modeling assumptions related to the system-level study will be given in Section 5.5.

5.4 Tractable SINR Model for Asynchronous Networks

5.4.1 Link-Level Timing Misalignment Analysis

In this subsection, we analyze the impact of timing misalignment from a link-level perspective. Though similar analysis may be found in the rich OFDM literature (see e.g. [121]), we briefly revisit this analysis to motivate our proposed SINR model that captures the impact of asynchronous transmissions. To this end, we shall focus on the link between transmitter i and the typical receiver and ignore the signals from the other transmitters for now.

Note that the n-th time-domain sample of the m-th OFDM symbol from the signal $s_i(t)$ is given by

$$\begin{aligned} s_i[n;m] &= s_i\left(mT_s + n\frac{T_d}{N}\right) \\ &= \frac{\sqrt{E_i}}{N}\sum_k S_i[k;m]e^{j2\pi\frac{k}{N}n}, \quad n = -N_{cp}, ..., N-1, \end{aligned} \tag{5.2}$$

where $N_{cp} = NT_{cp}/T_d$ is the number of cyclic prefix samples. Denote by D_i the timing misalignment between transmitter i and the typical receiver. Without loss of generality, we assume $D_i \in \mathbb{D} \triangleq [-(N+N_{cp}), N+N_{cp})$.[2]

In each OFDM symbol m, the typical receiver would like to decode the m-th OFDM symbol sent by transmitter i. To this end, it discards the first N_{cp} samples falling in the current receiving window and performs a fast Fourier transform (FFT) on the remaining N samples. We consider the following four cases, in which for notational simplicity we drop the additive noise term and assume that the channel gain is 1 unless otherwise noted.

Case 1: $-(N+N_{cp}) \le D_i < -N$**.** The N samples used for the FFT of the m-th OFDM symbol are

$$y[n;m] = s_i[n - D_i - N - N_{cp}; m+1], \quad n = 0, ..., N-1. \tag{5.3}$$

The received signal on the ℓ-th subcarrier during the m-th OFDM symbol is given by

$$Y[\ell;m] = \sqrt{E_i}e^{j2\pi\frac{\ell}{N}(-D_i - N_{cp})}S_i[\ell; m+1], \tag{5.4}$$

[2]This assumption can be easily relaxed by using different notations m and m' to respectively index OFDM symbols at the transmitter and at the receiver in the following analysis.

which is derived in Appendix 5.8.1. Thus, the received symbol on the ℓ-th subcarrier during OFDM symbol time m is just a phase rotated version of the transmitted symbol on the ℓ-th subcarrier during OFDM symbol time $m+1$. If $S_i[\ell;m]$ is desired, the useful signal power is 0. Otherwise, transmitter i's signal appears as interference and its interference power (energy/symbol) on the ℓ-th subcarrier during the m-th OFDM symbol equals

$$P_i[\ell;m] = \mathbb{E}[|Y[\ell;m]|^2] = G_i[m]E_i, \tag{5.5}$$

where we have included the effect of channel gain $G_i[m]$ from transmitter i to the typical receiver during OFDM symbol time m. Note that $G_i[m]$ is independent of subcarrier ℓ as we assume that the channel is flat-fading.

Case 2: $-N \leq D_i < 0$**.** The N samples used for the FFT of the m-th OFDM symbol are

$$y[n;m] = \begin{cases} s_i[-D_i+n;m], & 0 \leq n \leq N-1+D_i; \\ s_i[n-(N+D_i)-N_{cp};m+1], & N+D_i \leq n \leq N-1. \end{cases} \tag{5.6}$$

The received signal on the ℓ-th subcarrier during the m-th OFDM symbol is given by

$$\begin{aligned} Y[\ell;m] =& \sqrt{E_i}\frac{N+D_i}{N}S_i[\ell;m]e^{-j2\pi\frac{\ell}{N}D_i} - \sqrt{E_i}\frac{D_i}{N}S_i[\ell;m+1]e^{j2\pi\frac{\ell}{N}(-D_i-N_{cp})} \\ &+ \sqrt{E_i}\frac{1}{N}\sum_{k\neq\ell}\left(\frac{1-e^{j2\pi\frac{k-\ell}{N}(N+D_i)}}{1-e^{j2\pi\frac{k-\ell}{N}}}\right) \\ &\times\left(S_i[k;m]e^{-j2\pi\frac{k}{N}D_i} - S_i[k;m+1]e^{j2\pi\frac{k}{N}(-D_i-N_{cp})}\right), \end{aligned} \tag{5.7}$$

which is derived in Appendix 5.8.1. Thus, the total received power on the ℓ-th

subcarrier during the m-th OFDM symbol from transmitter i is

$$P_i[\ell;m] = G_i[m]E_i\left(\frac{1}{N^2}\left((N+D_i)^2+D_i^2\right)+\frac{2}{N^2}\sum_{k\neq\ell}\frac{\sin^2\left(\pi\frac{N+D_i}{N}(k-\ell)\right)}{\sin^2\left(\pi\frac{1}{N}(k-\ell)\right)}\right), \tag{5.8}$$

where we have used the assumption that $\{S_i[k;m]\}$ are i.i.d. and have zero mean and unit variance. If $S_i[\ell;m]$ is desired, the useful signal power is $\frac{(N+D_i)^2}{N^2}G_i[m]E_i$; the remaining terms in (5.8) contribute to self-interference including both inter-carrier interference (ICI) and inter-symbol interference (ISI). Otherwise, transmitter i's signal appears as interference whose power is characterized by (5.8).

Case 3: $0 \leq D_i < N_{cp}$**.** The N samples used for the FFT of the m-th OFDM symbol are

$$y[n;m] = s_i[n-D_i;m], \quad 0 \leq n \leq N-1. \tag{5.9}$$

As in Case 1, we can show that the received signal on the ℓ-th subcarrier during the m-th OFDM symbol is given by

$$Y[\ell;m] = \sqrt{E_i}S_i[\ell;m]e^{-j2\pi\frac{\ell}{N}D_i}. \tag{5.10}$$

If $S_i[\ell;m]$ is desired, the useful signal power is $G_i[m]E_i$, and there is no self-interference. Otherwise, transmitter i's signal appears as interference with power $G_i[m]E_i$.

Case 4: $N_{cp} \leq D_i < N + N_{cp}$**.** The N samples used for the FFT of the m-th OFDM symbol are

$$y[n;m] = \begin{cases} s_i[n+N+N_{cp}-D_i;m-1], & 0 \leq n \leq D_i - N_{cp} - 1; \\ s_i[n-D_i;m], & D_i - N_{cp} \leq n \leq N-1. \end{cases} \tag{5.11}$$

As in Case 2, we can show that the received signal on the ℓ-th subcarrier during the m-th OFDM symbol is given by

$$\begin{aligned} Y[\ell;m] =& \sqrt{E_i}\frac{N-D_i+N_{cp}}{N}S_i[\ell;m]e^{-j2\pi\frac{\ell}{N}D_i} + \sqrt{E_i}\frac{D_i-N_{cp}}{N}S_i[\ell;m-1] \\ & \times e^{-j2\pi\frac{\ell}{N}(D_i-N_{cp})} + \sqrt{E_i}\frac{1}{N}\sum_{k\neq\ell}\left(\frac{1-e^{j2\pi\frac{k-\ell}{N}(D_i-N_{cp})}}{1-e^{j2\pi\frac{k-\ell}{N}}}\right) \\ & \times \left(-S_i[k;m]e^{-j2\pi\frac{k}{N}D_i} + S_i[k;m-1]e^{-j2\pi\frac{k}{N}(D_i-N_{cp})}\right). \end{aligned} \tag{5.12}$$

Thus, the total received power on the ℓ-th subcarrier during the m-th OFDM symbol from transmitter i is

$$\begin{aligned} P_i[\ell;m] =& G_i[m]E_i\Bigg(\frac{1}{N^2}\left((N-D_i+N_{cp})^2+(D_i-N_{cp})^2\right) \\ & + \frac{2}{N^2}\sum_{k\neq\ell}\frac{\sin^2\left(\pi\frac{D_i-N_{cp}}{N}(k-\ell)\right)}{\sin^2\left(\pi\frac{1}{N}(k-\ell)\right)}\Bigg). \end{aligned} \tag{5.13}$$

If $S_i[\ell;m]$ is desired, the useful signal power is $\frac{(N-D_i+N_{cp})^2}{N^2}G_i[m]E_i$; the remaining terms in (5.13) contribute to self-interference including both ICI and ISI. Otherwise, transmitter i's signal appears as interference whose power is characterized by (5.13).

5.4.2 From Link-Level to System-Level Studies

In this subsection, we discuss how to apply the previous link-level analysis on the impact of timing misalignment to OFDM transmission in system-level studies. In an OFDM system without adaptive modulation and coding per subcarrier, a transmitter simultaneously sends a block of coded bits on the used subcarriers. The probability that the receiver can decode the block

sent by transmitter i depends on all the SINR values of the used subcarriers. Transmitter i's SINR of subcarrier ℓ is given by

$$\mathrm{SINR}_i[\ell] = \frac{g(D_i)G_iE_i}{P_i[\ell] - g(D_i)G_iE_i + \sum_{j\neq i} P_j[\ell] + N_0}, \tag{5.14}$$

where we have dropped the OFDM symbol index m, N_0 denotes the noise power, and

$$g(d) = \begin{cases} 0 & -(N+N_{cp}) \leq d < -N; \\ \frac{(N+d)^2}{N^2} & -N \leq d < 0; \\ 1 & 0 \leq d < N_{cp}; \\ \frac{(N+N_{cp}-d)^2}{N^2} & N_{cp} \leq d < N+N_{cp}. \end{cases} \tag{5.15}$$

In a system-level study, the subcarrier SINR values are usually mapped to a unique SINR, based on which the decision on whether the block is decodable is made. For example, the exponential effective SINR mapping (EESM) is a popular mapping method [129]. In an asynchronous network with timing misalignment, the calculation of $\mathrm{SINR}_i[\ell]$ can be difficult because the detailed modeling of timing errors in a system-level study can be cumbersome. Further, the received power $P_i[\ell]$ depends on timing misalignment in a delicate way (c.f. (5.8) and (5.13)), which makes the analytical evaluation of system-level performance even more challenging.

To solve the above mentioned difficulties, we propose a simple first-order model, which can be conveniently used in system-level studies.

System-Level Abstraction. *In a system-level study of the asynchronous network with timing misalignment, the subcarrier $SINR_i[\ell]$ may be approximately calculated as follows.*

1. *Model and calculate the timing misalignment D_i between transmitter i and the typical receiver.*
2. *Calculate the useful signal power as $g(D_i)G_iE_i$, where $g(d)$ is defined in (5.15).*
3. *Approximate the total received signal power from transmitter j as $P_j[\ell] = G_jE_j, j = 1, 2, ...$*
4. *Calculate $SINR_i[\ell]$ according to (5.14).*

The proposed system-level abstraction has two main advantages: (1) when evaluating $\text{SINR}_i[\ell]$ it only needs to consider the timing misalignment of the receiver with respect to transmitters i; and (2) compared to the original complicated expressions (c.f. (5.8) and (5.13)), the total received signal power from transmitter j is simply approximated as $P_j[\ell] = G_jE_j$. These two facts greatly simplify system-level studies.

The validness of the proposed system-level abstraction hinges on the condition that the total received signal power from transmitter j can be well approximately as $P_j[\ell] = G_jE_j$, regardless of the timing misalignment D_j. As shown in a numerical example in Fig. 5.1, this approximation is quite accurate: the received powers are almost uniform on the used subcarriers except a few edge subcarriers under various timing misalignment cases. Fig. 5.2 further shows how the timing misalignment in OFDM transmission affects the power of useful signal as well as the power of self-interference. For example, the received SNR of the central subcarrier would be limited to less than 20 dB when the receiving window is later than the actual timing of the received

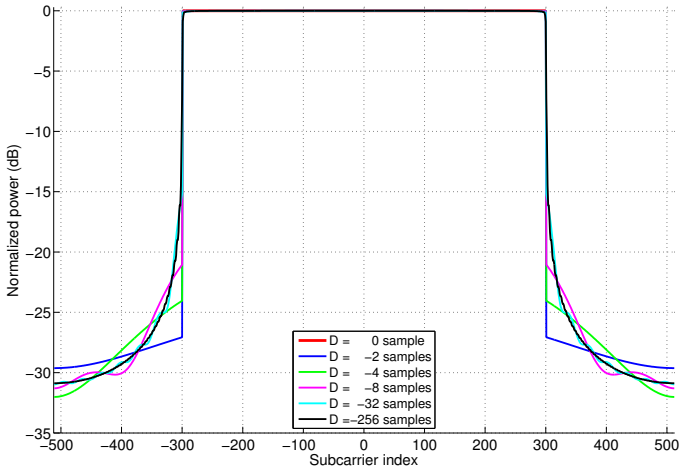

Figure 5.1: Received power of an OFDM signal with timing misalignment. $N = 1024$; $N_{cp} = 72$; the used subcarriers are $\{-299, ..., 0, ...300\}$.

signal by 6 samples (mainly due to the self-interference).

5.5 On the Decodable Transmitters of a Typical Receiver

In this section, we apply the proposed system-level abstraction to study several important statistics about the transmitters whose packets can be decoded by the typical receiver in the asynchronous network. Such statistics include the average number of decodable transmitters, the decoding probability of the nearest transmitter, the distribution of the number of decodable transmitters, and system throughput.

To this end, we first notice that with the proposed system-level ab-

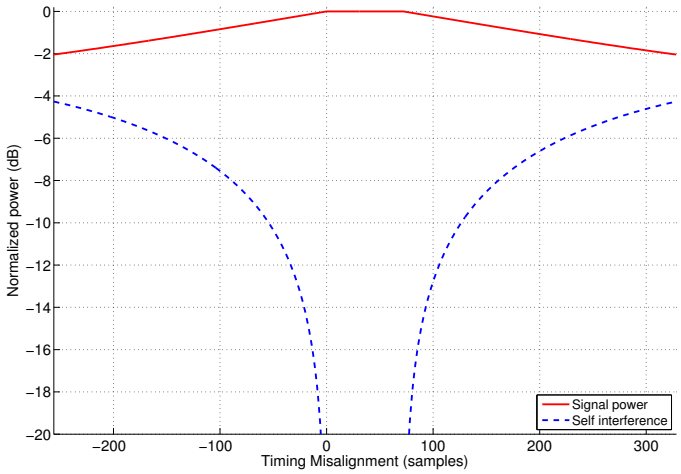

Figure 5.2: Signal and self-interference powers of an OFDM signal received on the central subcarrier with timing misalignment. $N = 1024$; $N_{cp} = 72$; the used subcarriers are $\{-300, ..., 0, ...299\}$.

straction, the subcarrier $\text{SINR}_i[\ell]$ now can be written as

$$\text{SINR}_i[\ell] = \frac{g(D_i)G_iE_i}{(1 - g(D_i))G_iE_i + \sum_{j \neq i} G_jE_j + N_0}. \tag{5.16}$$

Noting that the right hand side of (5.16) is independent of ℓ, we can simply use the subcarrier $\text{SINR}_i[\ell]$ as the block SINR_i, based on which the decision on whether a packet is decodable can be made. Therefore, in the sequel we drop the subcarrier index ℓ in (5.16) and treat it as a block SINR.

In the following system-level study we assume that (i) transmitters use constant transmit power E, (ii) the timing mismatches $\{D_i\}$ are i.i.d. with cumulative distribution function (CDF) $F_D(\cdot)$, and (iii) the channel gain G_i is modeled as

$$G_i = \|X_i\|^{-\alpha} F_i, \tag{5.17}$$

where $\alpha > 2$ is the pathloss exponent, and F_i denotes the fading of the link from transmitter i to the typical receiver. For simplicity, we consider independent Rayleigh fading, i.e., $F_i \sim \text{Exp}(1)$; more general fading and/or the effect of shadowing may be treated by further applying Displacement theorem for the PPP [22], which is not the focus of this chapter. With these assumptions, the SINR_i now can be written as

$$\text{SINR}_i = \frac{g(D_i)\|X_i\|^{-\alpha}F_i}{(1-g(D_i))\|X_i\|^{-\alpha}F_i + \sum_{j\neq i}\|X_j\|^{-\alpha}F_j + N_0/E}. \tag{5.18}$$

We let $\mathcal{E}_i$ be the event that a packet from transmitter i is decodable. Then the event $\mathcal{E}_i$ occurs if and only if the received SINR_i is above some detection threshold T, which is a function of the used modulation and coding scheme. Mathematically, the number Υ of decodable transmitters is given by

$$\Upsilon = \sum_i \mathbb{I}(\mathcal{E}_i) = \sum_i \mathbb{I}(\text{SINR}_i \geq T), \tag{5.19}$$

where $\mathbb{I}(\mathcal{E})$ is an indicator function which equals 1 if the event $\mathcal{E}$ is true and 0 otherwise. Clearly, Υ is a random variable and will be the central object studied in the sequel.

5.5.1 Mean Number of Decodable Transmitters

We first consider the average number of decodable transmitters $\mathbb{E}[\Upsilon]$.

Proposition 5.1. *The mean number of decodable transmitters is given by*

$$\mathbb{E}[\Upsilon] = \pi\lambda \int_{\mathbb{D}} \int_0^\infty \mathbb{I}\left(g(\tau) > \frac{T}{1+T}\right)$$

$$\times e^{-h(\tau,T)SNR^{-1}v^{\frac{\alpha}{2}}} e^{-\lambda\pi sinc^{-1}\left(\frac{2}{\alpha}\right)(h(\tau,T))^{\frac{2}{\alpha}}v} \, \mathrm{d}v F_D(\, \mathrm{d}\tau), \quad (5.20)$$

where $h(\tau,T) = \frac{T}{(1+T)g(\tau)-T}$, $SNR = E/N_0$, *and* $sinc(x) = \frac{\sin(\pi x)}{\pi x}$.

Proof. See Appendix 5.8.2. □

To gain some insights from Prop. 5.1, we next focus on the special case that the network is interference-limited, i.e., $N_0 \to 0$.

Corollary 5.1. *In the interference-limited case with* $N_0 \to 0$, *(5.20) reduces to a simpler form:*

$$\mathbb{E}[\Upsilon] = \mathbb{E}_D\left[\frac{\mathbb{I}\left(g(D) > \frac{T}{1+T}\right) sinc\left(\frac{2}{\alpha}\right)}{(h(D,T))^{\frac{2}{\alpha}}}\right], \quad (5.21)$$

which can be upper bounded as

$$\mathbb{E}[\Upsilon] \leq \frac{sinc\left(\frac{2}{\alpha}\right)}{T^{\frac{2}{\alpha}}}. \quad (5.22)$$

The upper bound (5.22) follows because by definition $g(\tau) \leq 1$ (c.f. (5.15)) and thus $h(\tau,T) \geq T$ for all $\tau \in \mathbb{D}$ satisfying $g(\tau) > T/(1+T)$. The upper bound is attained when timing misalignment D is restricted within the range of cyclic prefix. This simple upper bound only depends on two network parameters: α and T. In particular, the upper bound decreases as the detection threshold T increases, agreeing with intuition: the mean number of decodable transmitters decreases when the modulation and coding rate are chosen such that T is higher.

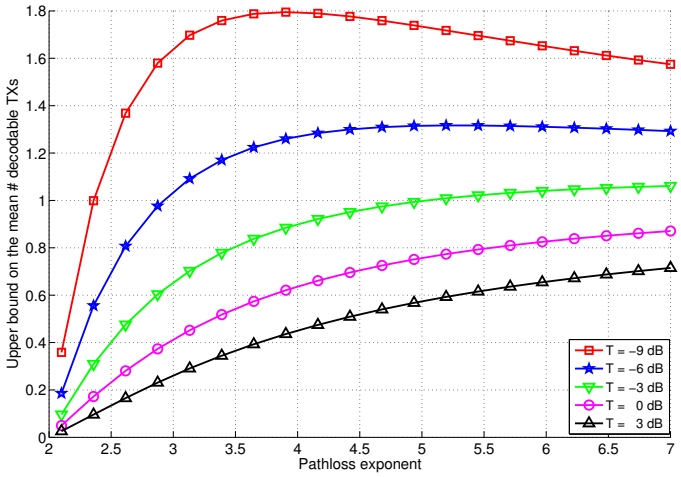

Figure 5.3: The upper bound on the mean number of decodable transmitters (c.f. (5.22)) versus pathloss exponent.

The dependency of the upper bound on the pathloss exponent α is more complicated and is illustrated in Fig. 5.3. Note that $\text{sinc}\left(\frac{2}{\alpha}\right)$ is increasing with $\alpha \in (2, \infty)$. In contrast, when $0 < T < 1$, $T^{\frac{2}{\alpha}}$ is increasing with $\alpha \in (2, \infty)$, but when $T \geq 1$, $T^{\frac{2}{\alpha}}$ is decreasing with $\alpha \in (2, \infty)$. Therefore, when $T \geq 1$, the upper bound increases with $\alpha \in (2, \infty)$. The intuition is that in order to decode packets from more transmitters in the median-to-high modulation and coding rate regime, it is important to reduce the interference power in the interference-limited scenario and thus high pathloss exponent is favorable. When $0 < T < 1$, it is possible that the upper bound first increases and then decreases as the pathloss exponent increases. This is because in the low modulation and coding rate regime, it is also important to preserve the useful signal power while reducing the interference power. In particular, for very low T, as α increases beyond some point, the loss of the useful signal

power will outweigh the gain of interference reduction and thus the mean number of decodable transmitters will eventually decrease. Another interesting observation from Fig. 5.3 is that the mean number of decodable transmitters is very small: it is less than 2 even when T is as low as -9 dB. We will explore this fact more in later sections.

Though the above discussion is carried out in the interference-limited case, the overall insights still hold when noise is taken into account. For example, Fig. 5.4 considers noise (whose power is given in Table 5.1) and shows the performance under two two transmitter densities. The dense case with $\lambda = 1/20^2$ m^{-2} is interference-limited; in this case, we can see that the upper bound shown in Fig. 5.3 is quite close to the true values shown in Fig. 5.4. In the sparse case with $\lambda = 1/400^2$ m^{-2} where the noise has a more pronounced effect, Fig. 5.4 shows that a moderate pathloss exponent (around 3.3) is preferred as it strikes a balance between interference reduction and preserving the useful signal power.

Next let us turn to the impact of timing misalignment. As expected and shown in (5.22), there is a loss in the mean number of decodable transmitters due to the timing misalignment. However, if the timing misalignment is restricted within the range of cyclic prefix, i.e., $D \in [0, N_{cp})$, then $g(D) \equiv 1$ and thus the upper bound in (5.22) is attained. In this case, there is no loss due to the timing misalignment. Otherwise, the loss exists and depends on the distribution of the timing misalignment. Note that the integrand in (5.20) is zero if $g(\tau) \leq T/(1+T)$. The physical interpretation is that when $g(\tau) \leq T/(1+T)$,

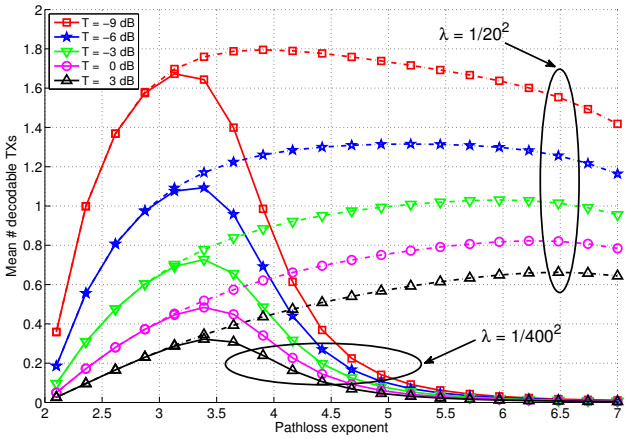

Figure 5.4: Mean number of decodable transmitters versus pathloss exponent in synchronized networks.

the self interference caused by timing misalignment is already large enough to cause the decoding failure.

To obtain a more concrete understanding of the impact of timing misalignment, we show some numerical results in the sequel. As a null hypothesis, we assume that the distribution of the timing misalignment is Gaussian with mean 0 and standard deviation σ but is truncated within the range $[-(N+N_{cp}), N+N_{cp})$. The specific parameters used in plotting numerical or simulation results in this chapter are summarized in Table 5.1 unless otherwise specified. Note that, with the OFDM sampling period normalized to 1, N denotes the duration of the data part of an OFDM symbol. Accordingly, we normalize timing error deviation σ and measure it in terms of N, as indicated in Table 5.1.

Fig. 5.5 shows the mean number of decodable transmitters versus the

Tx density λ	$1/400^2$ m^{-2}
PL exponent α	3.8
Tx power	23 dBm
Channel bandwidth	10 MHz
Noise PSD	-174 dBm
Rx noise figure	9 dB
Detection threshold T	-12 dB
(N, N_{cp})	(1024, 72)
Timing error deviation σ	$0.2N$

Table 5.1: Simulation/Numerical Parameters for Asynchronous OFDM Networks

detection threshold. From Fig. 5.5, we can see that asynchronous transmissions have a remarkable effect on the performance; for example, when aiming at decoding one transmitter on average and $\lambda = 1/20^2$ m^{-2}, the loss in the supported detection threshold is about 2 dB (resp. 4 dB) with $\sigma = 0.2N$ (resp. $\sigma = 0.4N$). Similarly, with the detection threshold $T = -4$ dB, the loss in the mean number of decodable transmitters is 21% (resp. 44%) when $\sigma = 0.2N$ (resp. $\sigma = 0.4N$). Fig. 5.5 also shows that the relative loss in the mean number of decodable transmitters due to asynchronous transmissions increases as the detection threshold increases, implying that asynchronous transmissions have a more significant impact on high-rate communication. Similar observations hold when $\lambda = 1/400^2$ m^{-2}. Note that the simulation results clearly match the analysis in Fig. 5.5; this provides a sanity check for the derived analytical results.

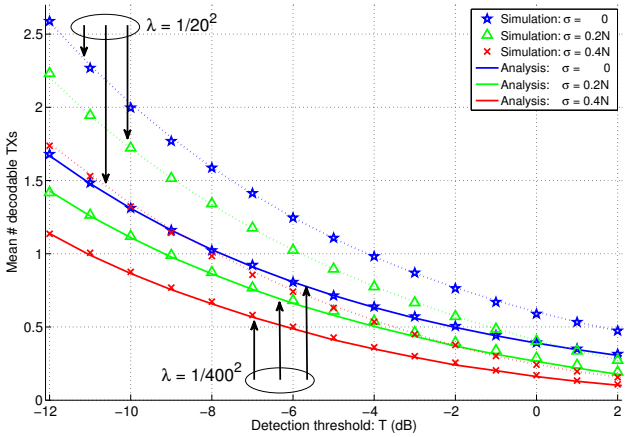

Figure 5.5: Mean number of decodable transmitters versus detection threshold.

5.5.2 An Upper Bound on the Distribution of the Number of Decodable Transmitters

In the previous subsection, we studied the first order statistic of the number Υ of decodable transmitters. In this subsection, we take a broader view and study the distribution of the number Υ of decodable transmitters. Though an exact characterization is possible, the resulting expressions involve very high dimensional integrals even in the case of perfectly synchronized networks [79]. Instead, we provide a simple upper bound on the distribution of Υ in the following proposition.

Proposition 5.2. *The number Υ of decodable transmitters is (first order) stochastically dominated by a truncated Poisson random variable $\Upsilon^{(u)}$, i.e., $\mathbb{P}(\Upsilon^{(u)} \geq n) \geq \mathbb{P}(\Upsilon \geq n), n = 0, 1, \ldots$. The distribution of $\Upsilon^{(u)}$ is given as*

follows: $\mathbb{P}(\Upsilon^{(u)} = n) = \frac{1}{C}\frac{\tilde{\lambda}^n}{n!}, n = 0, ..., \lfloor \frac{1+T}{T} \rfloor$, *where*

$$\tilde{\lambda} = \pi\lambda \int_0^\infty \mathbb{E}_D \left[\mathbb{I}\left(g(D) > \frac{T}{1+T} \right) e^{-\frac{Tv^{\alpha/2}}{g(D)SNR}} \right] \mathrm{d}v, \tag{5.23}$$

and C *is a normalization constant such that* $\sum_{n=0}^{\lfloor \frac{1+T}{T} \rfloor} \mathbb{P}(\Upsilon^{(u)} = n) = 1$.

Proof. See Appendix 5.8.3. □

To gain some insights from Prop. 5.2, we next focus on the special case with $T > 1$. Then Prop. 5.1 implies that $\Upsilon^{(\mathrm{u})}$ is a Bernoulli random variable: it equals 1 with probability $\tilde{\lambda}/(1+\tilde{\lambda})$ and 0 otherwise. The mean of $\Upsilon^{(\mathrm{u})}$ is $\tilde{\lambda}/(1+\tilde{\lambda})$. If the network is very sparse such that $\lambda \sim o(1)$, then $\tilde{\lambda}/(1+\tilde{\lambda}) \sim \tilde{\lambda} = \Theta(\lambda)$. When the transmit power is fixed, the performance of sparse networks is noise-limited. This indicates that in the noise-limited case the probability that the receiver can decode a packet from some transmitter is $O(\lambda)$. So is the mean number of decodable transmitters. In the next subsection, we will show that the probability is $\Omega(\lambda)$ as $\lambda \to 0$, and thus the probability actually scales as $\Theta(\lambda)$.

If the network is very dense, i.e., $\lambda \to \infty$, then $\tilde{\lambda}/(1+\tilde{\lambda}) \sim 1$. Clearly, the performance of dense networks is interference-limited. As a result, one might think that in the interference-limited case the probability that the receiver can decode a packet from some transmitter is close to 1. The fallacy of the above argument is that $\tilde{\lambda}/(1+\tilde{\lambda})$ is an upper bound and may not be tight as $\lambda \to \infty$. In fact, the right intuition should be that the received SINR from any transmitter in the interference-limited case would not be large and thus

the probability that no transmitter can be decoded can be relatively high if the detection threshold T is large. The last intuition can be further confirmed by examining Fig. 5.3. For example, Fig. 5.3 shows that the mean number of decodable transmitters is less than 0.5 at $\alpha = 4$ and $T = 3$ dB, implying that the probability that no transmitter can be decoded is greater than 0.5.

Note that the parameter $\tilde{\lambda}$ may take more explicit form in some special cases. For example, when $\alpha = 4$,

$$\tilde{\lambda} = \frac{\pi^{\frac{3}{2}}\lambda}{2}\sqrt{\frac{\mathrm{SNR}}{T}}\mathbb{E}_D\left[\mathbb{I}\left(g(D) > \frac{T}{1+T}\right)\sqrt{g(D)}\right]. \tag{5.24}$$

Therefore, if $T > 1$ and $\alpha = 4$, the probability that the receiver can decode a packet from some transmitter is proportional to the square root of SNR in the noise-limited case, agreeing with intuition: the radio link length is proportional to $\mathrm{SNR}^{1/4}$ when $\alpha = 4$ and thus the decoding probability should be proportional to $\mathrm{SNR}^{1/2}$ in $\mathbb{R}^2$. Similar intuition may be used to explain why the probability is inversely proportional to the square root of the detection threshold T.

Figs. 5.6 and 5.7 compare the analytical upper bound on the distribution of the number Υ of decodable transmitters to the corresponding true distribution obtained from simulation under two different transmitter densities. It can be seen that the analytical upper bound is more accurate when the network is sparser (i.e. less interference-limited).

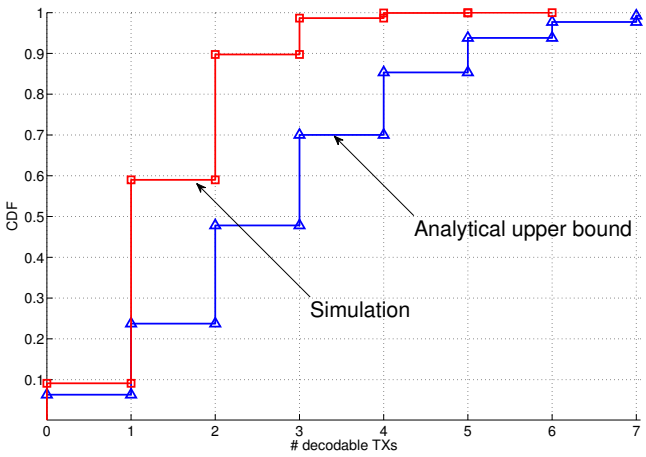

Figure 5.6: Analytical upper bound vs. simulation on the distribution of the number of decodable transmitters: $\lambda = 1/400^2$ m^{-2}.

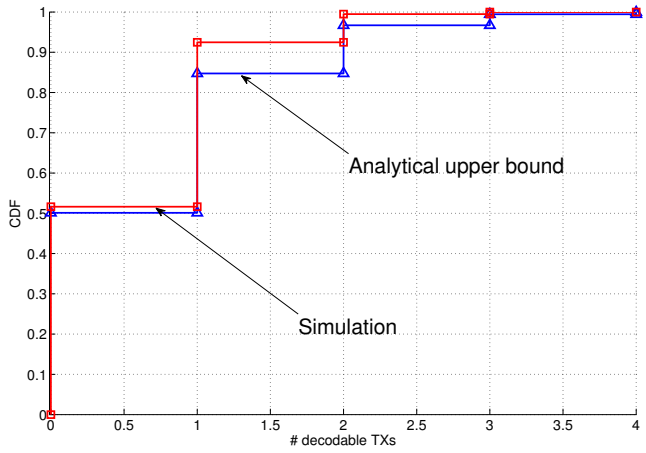

Figure 5.7: Analytical upper bound vs. simulation on the distribution of the number of decodable transmitters: $\lambda = 1/800^2$ m^{-2}.

5.5.3 On Decoding the Nearest Transmitter

According to Prop. 5.2, the receiver can decode the packet from at most one transmitter if $T > 1$. The decodable transmitter is typically the nearest one, though fading and timing misalignment may affect the result. Further, the probability of decoding the nearest transmitter indicates the coverage performance of cellular networks where the positions of BSs are modeled by a PPP [10]. Therefore, it is of particularly interest to study the probability that the receiver can decode a packet sent by the nearest transmitter. We answer this question in the following Proposition 5.3.

Proposition 5.3. *The probability that the receiver can decode a packet sent by the nearest transmitter X_0 is given by*

$$\begin{aligned}\mathbb{P}(SINR_0 \geq T) =& \pi\lambda \int_{\mathbb{D}} \int_0^{\infty} \mathbb{I}\left(g(\tau) > \frac{T}{1+T}\right) \\ & \times e^{-h(\tau,T)SNR^{-1}v^{\frac{\alpha}{2}}} e^{-\pi\lambda(1+\rho(h(\tau,T),\alpha))v} \, \mathrm{d}v F_D(\, \mathrm{d}\tau), \end{aligned} \tag{5.25}$$

where $\rho(x,\alpha) = x^{\frac{2}{\alpha}} \int_{x^{-\frac{2}{\alpha}}}^{\infty} \frac{1}{1+v^{\frac{\alpha}{2}}} \, \mathrm{d}v$, and $h(\tau,T)$ is defined in Prop. 5.1.

Proof. See Appendix 5.8.4. □

From Prop. 5.3, it is easy to see the probability that the receiver can decode a packet sent by the nearest transmitter is $\Theta(\lambda)$ as $\lambda \to 0$. Thus, the probability that the receiver can decode a packet sent by at least one transmitter is $\Omega(\lambda)$ as $\lambda \to 0$. The last fact has been used in the previous section when stating that with $T > 1$ the probability that the receiver can decode a packet sent by some transmitter scales as $\Theta(\lambda)$.

When the network is interference-limited, i.e., $N_0 \to 0$, (5.25) reduces to

$$\mathbb{P}(\mathrm{SINR}_0 \geq T) = \mathbb{E}_D \left[\frac{\mathbb{I}\left(g(D) > \frac{T}{1+T}\right)}{1+\rho(h(D,T),\alpha)} \right] \leq \frac{1}{1+\rho(T,\alpha)}, \tag{5.26}$$

where we have used the fact that $h(\tau, T) \geq T$, for all $\tau \in \mathbb{D}$ satisfying $g(\tau) > T/(1+T)$, in the last inequality. The above upper bound is attained when $D \equiv 0$, i.e., the network is perfectly synchronized, which has been studied in [10]. In fact, as long as the timing misalignment D is restricted within the range of cyclic prefix, the upper bound can be attained. As in the case of the mean number of decodable transmitters, there is a loss in the probability of decoding the nearest transmitter due to the timing misalignment, and the loss depends on the distribution of the timing misalignment.

Fig. 5.8 shows the decoding probability of the nearest transmitter versus the detection threshold. From Fig. 5.8, we can see that, when aiming at decoding probability 0.5 and $\lambda = 1/20^2$ m^{-2}, the loss in the supported detection threshold is about 3 dB (resp. 6 dB) with $\sigma = 0.2N$ (resp. $\sigma = 0.4N$). Fig. 5.8 also shows that the impact of asynchronous transmissions becomes more significant as the detection threshold increases. Similar observations hold when $\lambda = 1/400^2$ m^{-2}.

5.5.4 Optimizing System Throughput

The average number of decodable transmitters characterized in Prop. 5.1 is monotonically increasing as the detection threshold T decreases. How-

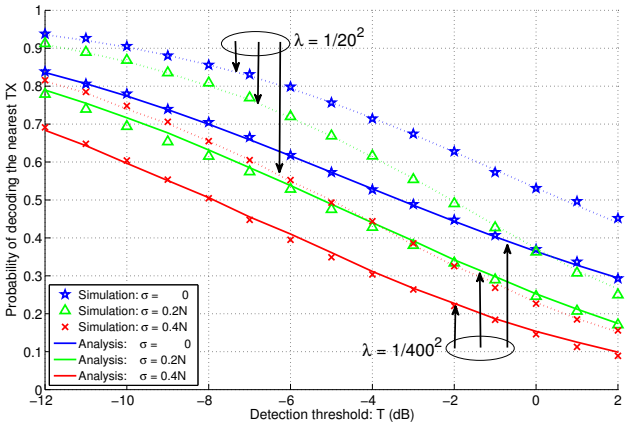

Figure 5.8: Impact of timing misalignment on the decoding probability of the nearest transmitter.

ever, reducing the detection threshold T implies that we adopt lower modulation order and/or coding rate. This may be undesirable from a throughput point of view. In order to take into account this tradeoff, we define *system throughput* ξ as the mean of the sum rate of all the transmitters to the typical receiver. Mathematically,

$$\xi = \mathbb{E}\left[\sum_i \mathbb{I}(\mathrm{SINR}_i \geq T)\log(1+T)\right]. \tag{5.27}$$

With this definition, the following result follows immediately.

Corollary 5.2. *The system throughput equals* $\xi = \log(1+T)\mathbb{E}[\Upsilon]$ *with* $\mathbb{E}[\Upsilon]$ *given in Prop. 5.1.*

Now we may optimize the detection threshold T by maximizing the system throughput ξ. This optimization is of single variable and thus can be solved efficiently. To gain some intuition, we show the system throughput

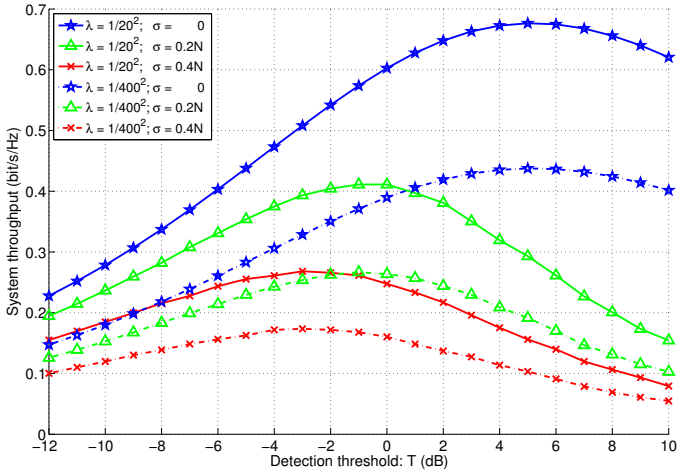

Figure 5.9: System throughput versus detection threshold.

as a function of T in Fig. 5.9. From Fig. 5.9, we can see that the optimal detection thresholds are respectively 5 dB, -1 dB and -3 dB when $\sigma = 0$, $0.2N$ and $0.4N$. This implies that we have to be more conservative in setting the detection threshold when the networks are asynchronous (vs. synchronized networks). Another interesting observation from Fig. 5.9 is that the optimal detection thresholds are nearly unaffected by the transmitter density.

5.6 Solutions for Mitigating the Loss of Asynchronous Transmissions

In the previous section, we have seen that asynchronous transmissions may have a remarkable effect on the system-level performance. In this section we discuss four possible solutions, which differ in complexity and may be applicable in different scenarios, to mitigate the loss due to asynchronous transmissions.

Extended cyclic prefix. If the timing mismatches are concentrated in the range $[0, N_x)$ where $N_x > N_{cp}$, we can solve the timing misalignment problem by simply extending the length of the cyclic prefix beyond N_x. However, using cyclic prefix of extended length comes at the cost of more power and time spent in sending the cyclic prefix instead of being used to communicate data. This is a tradeoff, the characterization of which is beyond the scope of this chapter. The general conclusion is that this approach is applicable to the scenarios where N_x is not too large.

Advanced receiver timing. If the timing mismatches are concentrated in the range $[-N_x, N_y)$ where $N_x, N_y > 0$ and $N_x + N_y \leq N_{cp}$, we can solve the timing misalignment problem by simply advancing the receiver timing by N_x. Then the timing mismatches will be concentrated in the range $[0, N_x + N_y)$. As $N_x + N_y \leq N_{cp}$, there will be no loss due to the timing misalignment after shifting the receiver's timing earlier. This approach is very simple but is only applicable to the scenarios where $N_x + N_y \leq N_{cp}$, and it also requires knowledge of N_x.

Dynamic receiver timing positioning. The receiver may estimate the timings used by each transmitter through either pilot-based or non pilot-based synchronization methods. Once a transmitter's timing is obtained, the receiver can adaptively adjust its receiving window to decode the transmitter's packet. Compared to the previous two approaches, dynamic receiver timing positioning is applicable to many more scenarios but at the cost of higher complexity. In particular, as the transmitters have i.i.d. timing mismatches, the

typical receiver needs to estimate every transmitter's timing and accordingly positions its receiving window to decode a transmitter's packet.

Semi-static receiver timing positioning with multiple timing hypotheses. Instead of estimating each transmitter's timing, the receiver may simply adopt multiple timing hypotheses: $-n_1\Delta, ..., 0, ..., n_2\Delta$, where $n\Delta$ denotes the timing difference between the hypothesis n and the receiver's timing. For every timing hypothesis, the receiver accordingly adjusts its receiving window and performs decoding; the packets from the transmitters whose timings happen to be around the current timing hypothesis may be decoded. This semi-static receiver timing positioning approach reduces the complexity of dynamic receiver timing positioning but still requires the receiver to use multiple timing windows. Further, a careful choice of n_1, n_2 and Δ is important for the design. In general, the more the used timing hypotheses, the smaller the loss due to timing misalignment but the higher the complexity.

The above proposed solutions may be combined depending on the application scenarios. For example, advanced receiver timing may be jointly used with extended cyclic prefix to make the condition $N_x + N_y \leq N_{cp}$ hold. In practice, the design decision on which solution should be used or how they should be combined is best made based on the specific scenario under consideration. Note that if our target is not to decode as many transmitters as possible but for example is to decode the nearest transmitter, synchronizing directly with the nearest transmitter is of reasonable complexity and recovers the loss.

Let us consider the solution of semi-static receiver timing positioning with multiple timing hypotheses since it can be applied to many scenarios while having reasonable complexity. We take the mean number of decodable transmitters as the metric to evaluate its effectiveness. The following corollary immediately follows from Prop. 5.1.

Corollary 5.3. *Denote by* $\mathcal{H} = \{-n_1\Delta, ..., 0, ..., n_2\Delta\}$ *the set of timing hypotheses. The mean number of decodable transmitters is given by (5.20) but with* $g(x)$ *substituted by* $\tilde{g}(x) \triangleq \max_{\tau\in\mathcal{H}} g(x-\tau)$.

The rationale of Corollary 5.3 is straightforward: a transmitter is decodable as long as it is decodable under any of the used timing hypotheses. Fig. 5.10 shows the effectiveness of using multiple timing hypotheses. As expected, the more the used timing hypotheses, the more the mean number of decodable transmitters. Also, we can see from Fig. 5.10 that using 3 timing hypotheses helps recover the majority of the loss.

5.7 Summary

In view of the lack of network-wide synchronization in many wireless networks, this chapter has presented a baseline SINR model for asynchronous OFDM networks, which can be conveniently used in system-level studies. The model is then applied to characterize several important statistics in asynchronous PPP networks including the number of decodable transmitters, the decoding probability of the nearest transmitter, and system throughput. The

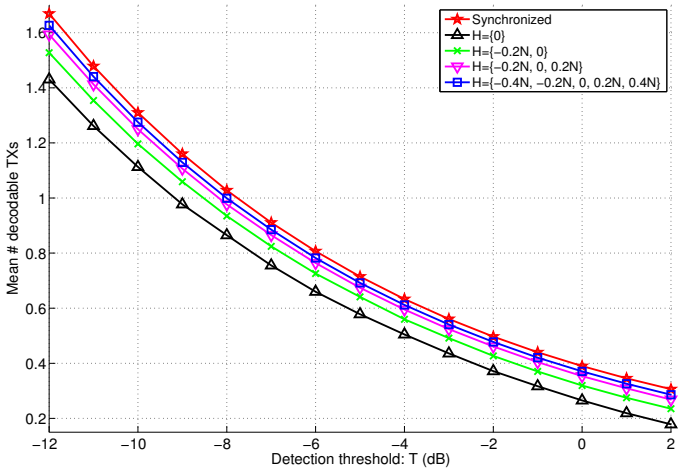

Figure 5.10: Using semi-static receiver timing positioning with multiple timing hypotheses to mitigate the loss of asynchronous transmissions.

derived results complement existing analysis of synchronized networks using stochastic geometry. Further, this chapter has compared and discussed four possible solutions for mitigating the loss of asynchronous transmissions. The model and results are general, and apply to ad hoc networks, cellular systems, and neighbor discovery in D2D networks.

5.8 Appendix

5.8.1 Derivation of Equations (5.4) and (5.7)

We first derive (5.4). By the definition of discrete-time Fourier transform,

$$Y[\ell; m] = \sum_{n=0}^{N-1} y[n; m] e^{-j2\pi \frac{\ell}{N} n}$$

$$
\begin{aligned}
&= \sum_{n=0}^{N-1} s_i[n - D_i - N - N_{cp}; m+1] e^{-j2\pi \frac{\ell}{N} n} \\
&= \sum_{n=0}^{N-1} \sqrt{E_i} \frac{1}{N} \sum_k S_i[k; m+1] e^{j2\pi \frac{k}{N}(n - D_i - N - N_{cp})} e^{-j2\pi \frac{\ell}{N} n} \\
&= \sqrt{E_i} \sum_k S_i[k; m+1] e^{j2\pi \frac{k}{N}(-D_i - N_{cp})} \frac{1}{N} \sum_{n=0}^{N-1} e^{j2\pi \frac{k-\ell}{N} n} \\
&= \sqrt{E_i} e^{j2\pi \frac{\ell}{N}(-D_i - N_{cp})} S_i[\ell; m+1], \qquad (5.28)
\end{aligned}
$$

where we have plugged (5.3) in the second equality and used (5.2) in the third equality, and the last equality follows from the fact that $\frac{1}{N} \sum_{n=0}^{N-1} e^{j2\pi \frac{k-\ell}{N} n} = \delta[k - \ell]$.

We next derive (5.7). Using the definition of discrete-time Fourier transform, (5.6) and (5.2) yields

$$
\begin{aligned}
Y[\ell; m] &= \sum_{n=0}^{N-1} y[n; m] e^{-j2\pi \frac{\ell}{N} n} = \sum_{n=0}^{N-1+D_i} s[-D_i + n; m] e^{-j2\pi \frac{\ell}{N} n} \\
&\quad + \sum_{n=N+D_i}^{N-1} s[n - (N + D_i) - N_{cp}; m+1] e^{-j2\pi \frac{\ell}{N} n} \\
&= \sum_{n=0}^{N-1+D_i} \sqrt{E_i} \cdot \frac{1}{N} \sum_k S_i[k; m] e^{j2\pi \frac{k}{N}(-D_i + n)} e^{-j2\pi \frac{\ell}{N} n} + \\
&\quad \sum_{n=N+D_i}^{N-1} \sqrt{E_i} \cdot \frac{1}{N} \sum_k S_i[k; m+1] e^{j2\pi \frac{k}{N}(n - (N+D_i) - N_{cp})} e^{-j2\pi \frac{\ell}{N} n}. \qquad (5.29)
\end{aligned}
$$

The first sum in (5.29) equals

$$
\sqrt{E_i} \frac{N + D_i}{N} S_i[\ell; m] e^{-j2\pi \frac{\ell}{N} D_i} + \sqrt{E_i} \frac{1}{N} \sum_{k \neq \ell} S_i[k; m] e^{-j2\pi \frac{k}{N} D_i} \sum_{n=0}^{N-1+D_i} e^{j2\pi \frac{k-\ell}{N} n},
$$

and the second sum in (5.29) equals

$$
- \sqrt{E_i} \frac{D_i}{N} S_i[\ell; m+1] e^{j2\pi \frac{\ell}{N}(-D_i - N_{cp})}
$$

$$+\sqrt{E_i}\frac{1}{N}\sum_{k\neq\ell}S_i[k;m+1]e^{j2\pi\frac{k}{N}(-D_i-N_{cp})}\sum_{n=N+D_i}^{N-1}e^{j2\pi\frac{k-\ell}{N}n}.$$

Combining the above two results, and plugging in the following two equations

$$\begin{aligned}\sum_{n=0}^{N-1+D_i}e^{j2\pi\frac{k-\ell}{N}n}&=\frac{1-e^{j2\pi\frac{k-\ell}{N}(N+D_i)}}{1-e^{j2\pi\frac{k-\ell}{N}}}\\ \sum_{n=N+D_i}^{N-1}e^{j2\pi\frac{k-\ell}{N}n}&=\frac{e^{j2\pi\frac{k-\ell}{N}(N+D_i)}(1-e^{j2\pi\frac{\ell-k}{N}D_i})}{1-e^{j2\pi\frac{k-\ell}{N}}},\end{aligned}\tag{5.30}$$

we obtain (5.7).

5.8.2 Proof of Proposition 5.1

For notational simplicity, denote by $I_\Phi=\sum_{X_j\in\Phi}\|X_j\|^{-\alpha}F_j$. Then by definition,

$$\begin{aligned}\mathbb{E}[\Upsilon]&=\mathbb{E}\left[\sum_{X_i\in\Phi}\mathbb{I}\left(\frac{g(D_i)\|X_i\|^{-\alpha}F_i}{(1-g(D_i))\|X_i\|^{-\alpha}F_i+I_{\Phi-\delta_{X_i}}+N_0/E}\geq T\right)\right]\\&=\int_{\mathbb{R}^2}\int_{\mathbb{D}}\int_{\mathbb{R}}\mathbb{E}\left[\mathbb{I}\left(\frac{g(\tau)\|x\|^{-\alpha}f}{(1-g(\tau))\|x\|^{-\alpha}f+I_\Phi+N_0/E}\geq T\right)\right]\\&\quad F_F(\,\mathrm{d}f)F_D(\,\mathrm{d}\tau)M(\,\mathrm{d}x),\end{aligned}\tag{5.31}$$

where $M(\cdot)$ is the mean measure of the PPP Φ, i.e., $M(A)=\mathbb{E}[\Phi(A)]$ for any measurable set $A\subset\mathbb{R}^2$, and we have used the reduced Campbell formula for the PPP [16] in the last equality. Noting that F_i's are i.i.d. Rayleigh fading, $F_F(\,\mathrm{d}f)=e^{-f}\,\mathrm{d}f, f\geq 0$. For the homogeneous PPP Φ, $M(\,\mathrm{d}x)=\lambda\,\mathrm{d}x$. Using these two facts and changing the integral with respect to $x\in\mathbb{R}^2$ into polar coordinates, we have

$$\mathbb{E}[\Upsilon]=2\pi\lambda\int_0^\infty\int_{\mathbb{D}}\int_0^\infty\mathbb{E}\left[\mathbb{I}\left(\frac{g(\tau)r^{-\alpha}f}{(1-g(\tau))r^{-\alpha}f+I_\Phi+N_0/E}\geq T\right)\right]$$

$$
\begin{aligned}
&\times e^{-f}\ \mathrm{d}f F_D(\ \mathrm{d}\tau) r\ \mathrm{d}r \\
&= 2\pi\lambda \int_0^\infty \int_{\mathbb{D}} \int_0^\infty \mathbb{E}\left[\mathbb{I}\left(g(\tau) > \frac{T}{1+T}\right)\mathbb{I}\left(f \geq r^\alpha h(\tau,T)(I_\Phi + N_0/E)\right)\right] \\
&\quad \times e^{-f}\ \mathrm{d}f F_D(\ \mathrm{d}\tau) r\ \mathrm{d}r \\
&= 2\pi\lambda \int_0^\infty \int_{\mathbb{D}} \mathbb{I}\left(g(\tau) > \frac{T}{1+T}\right) e^{-r^\alpha h(\tau,T) N_0/E} \mathbb{E}\left[e^{-r^\alpha h(\tau,T) I_\Phi}\right] F_D(\ \mathrm{d}\tau) r\ \mathrm{d}r \\
&= 2\pi\lambda \int_{\mathbb{D}} \int_0^\infty \mathbb{I}\left(g(\tau) > \frac{T}{1+T}\right) e^{-r^\alpha h(\tau,T) N_0/E} e^{-\lambda\pi \mathrm{sinc}^{-1}\left(\frac{2}{\alpha}\right)(h(\tau,T))^{\frac{2}{\alpha}} r^2} \\
&\quad \times r\ \mathrm{d}r F_D(\ \mathrm{d}\tau),
\end{aligned} \tag{5.32}
$$

where we have used the shorthand function $h(\tau, T)$ in the second equality and applied in the last equality the Laplace transform of the interference generated by a Poisson field of interferers with Rayleigh fading [51]:

$$
\mathcal{L}_{I_\Phi}(s) \triangleq \mathbb{E}[e^{-sI_\Phi}] = \exp\left(-\frac{\lambda\pi s^{\frac{2}{\alpha}}}{\mathrm{sinc}\left(\frac{2}{\alpha}\right)}\right). \tag{5.33}
$$

With a change of variables $r^2 \to v$ in (5.32), we obtain (5.20) and complete the proof.

5.8.3 Proof of Proposition 5.2

The set of transmitters in Φ whose packets can be decoded can be upper bounded as

$$
\begin{aligned}
\tilde{\Phi} &= \sum_{X_i \in \Phi} \delta_{X_i} \mathbb{I}(\mathrm{SINR}_i \geq T) \\
&\leq \sum_{X_i \in \Phi} \delta_{X_i} \mathbb{I}\left(\frac{g(D_i) F_i \|X_i\|^{-\alpha}}{N_0/E} \geq T\right) \mathbb{I}\left(g(D_i) > \frac{T}{1+T}\right) \triangleq \tilde{\Phi}^{(\mathrm{u})}.
\end{aligned} \tag{5.34}
$$

Note that given Φ the Bernoulli random variables

$$
\mathbb{I}\left(\frac{g(D_i) F_i \|X_i\|^{-\alpha}}{N_0/E} \geq T\right) \mathbb{I}\left(g(D_i) > \frac{T}{1+T}\right), i = 1, 2, ...,
$$

are independent. It follows that $\tilde{\Phi}^{(\mathrm{u})}$ is an independent thinning of Φ with thinning probability

$$\begin{aligned} p(x) &= \mathbb{E}\left[\mathbb{I}\left(\frac{g(D_i)F_i\|X_i\|^{-\alpha}}{N_0/E} \geq T\right)\mathbb{I}\left(g(D_i) > \frac{T}{1+T}\right)\right] \\ &= \mathbb{E}_D\left[\mathbb{I}\left(g(D) > \frac{T}{1+T}\right)\exp\left(-\frac{T\|x\|^{\alpha}}{g(D)\mathrm{SNR}}\right)\right], \end{aligned} \tag{5.35}$$

where we have used the independence of D and F, and $F \sim \exp(1)$. Therefore, $\tilde{\Phi}^{(\mathrm{u})}$ is a PPP with intensity measure $\Lambda(A) = \int_A p(x)\lambda \, \mathrm{d}x$. Further, $\Upsilon^{(\mathrm{u})} = \tilde{\Phi}^{(\mathrm{u})}(\mathbb{R}^2)$ is Poisson with parameter

$$\begin{aligned} \Lambda(\mathbb{R}^2) &= \int_{\mathbb{R}^2} p(x)\lambda \, \mathrm{d}x \\ &= \int_{\mathbb{R}^2} \mathbb{E}_D\left[\mathbb{I}\left(g(D) > \frac{T}{1+T}\right)\exp\left(-\frac{T\|x\|^{\alpha}}{g(D)\mathrm{SNR}}\right)\right]\lambda \, \mathrm{d}x \\ &= \pi\lambda\int_0^{\infty} \mathbb{E}_D\left[\mathbb{I}\left(g(D) > \frac{T}{1+T}\right)\exp\left(-\frac{Tv^{\frac{\alpha}{2}}}{g(D)\mathrm{SNR}}\right)\right] \mathrm{d}v. \end{aligned} \tag{5.36}$$

Next we show that $\Upsilon^{(\mathrm{u})}$ can be truncated at $\lfloor\frac{1+T}{T}\rfloor$, following a similar argument as in [16]. To this end, suppose there are n decodable transmitters, without loss of generality assumed to be $X_0, ..., X_{n-1}$. Then we have

$$\frac{g(D_i)\|X_i\|^{-\alpha}F_i}{-g(D_i)\|X_i\|^{-\alpha}F_i + \sum_{j=0}^{n-1}\|X_j\|^{-\alpha}F_j + I_{\Phi-\cup_{j=0}^{n-1}\delta_{X_j}} + N_0/E} \geq T, \tag{5.37}$$

for $i = 0, ..., n-1$, which implies that

$$\frac{\|X_i\|^{-\alpha}F_i}{\sum_{j=0, j\neq i}^{n-1}\|X_j\|^{-\alpha}F_j + I_{\Phi-\cup_{j=0}^{n-1}\delta_{X_j}} + N_0/E} \geq T, \quad i = 0, ..., n-1. \tag{5.38}$$

With some algebraic manipulations, we have the following set of inequalities:

$$(1+T)\|X_i\|^{-\alpha}F_i \geq T(\sum_{j=0}^{n-1}\|X_j\|^{-\alpha}F_j + I_{\Phi-\cup_{j=0}^{n-1}\delta_{X_j}} + N_0/E), i = 0, ..., n-1.$$

Summing the above set of inequalities,

$$(1+T)\sum_{j=0}^{n-1}\|X_j\|^{-\alpha}F_j \geq nT(\sum_{j=0}^{n-1}\|X_j\|^{-\alpha}F_j + I_{\Phi-\cup_{j=0}^{n-1}\delta_{X_j}} + N_0/E)$$
$$> nT\sum_{j=0}^{n-1}\|X_j\|^{-\alpha}F_j. \tag{5.39}$$

It follows that $n \leq \lfloor \frac{1+T}{T} \rfloor$, and thus the proposition has been proven.

5.8.4 Proof of Proposition 5.3

To begin with, we condition on the location of the nearest transmitter $X_0 = x = (r, \theta)$ and its associated fading $F_0 = f$ and timing misalignment $D_0 = \tau$. Then

$$\begin{aligned}
&\mathbb{P}(\mathrm{SINR}_0 \geq T | X_0 = x, F_0 = f, D_0 = \tau) \\
&= \mathbb{P}\left(\frac{g(\tau)\|x\|^{-\alpha}f}{(1-g(\tau))\|x\|^{-\alpha}f + I_{\Phi-\delta_x} + N_0/E} \geq T \Big| X_0 = x, F_0 = f, D_0 = \tau\right) \\
&= \mathbb{P}(f \geq r^{\alpha}h(\tau,T)(I_{\Phi-\delta_x} + N_0/E) \big| X_0 = x, F_0 = f, D_0 = \tau)\mathbb{I}\left(g(\tau) > \frac{T}{1+T}\right) \\
&= \mathbb{P}^{x,f,\tau}(f \geq r^{\alpha}h(\tau,T)(I_{\Phi-\delta_x} + N_0/E) \big| \Phi(B(o,r)) = 0)\mathbb{I}\left(g(\tau) > \frac{T}{1+T}\right),
\end{aligned} \tag{5.40}$$

where $\mathbb{P}^{x,f,\tau}(\cdot)$ denotes the Palm distribution with respect to Φ, i.e., the probability law conditioned on that there exists a point at location x with the marks f and τ. Note that, conditioned on that the nearest point is located in x, there are no other points in Φ located in the ball $B(o,r)$ centered at o with radius r, i.e., $\Phi(B(o,r)) = 0$. This condition has been made explicitly in (5.40). Further, the first term in (5.40) equals

$$\mathbb{P}^{x,f,\tau}(f \geq r^{\alpha}h(\tau,T)(I_{\Phi\cap B^c(o,r)-\delta_x} + N_0/E) \big| \Phi(B(o,r)) = 0)$$

$$= \mathbb{P}^{x,f,\tau}(f \geq r^{\alpha} h(\tau, T)(I_{\Phi \cap B^c(o,r)-\delta_x} + N_0/E))$$
$$= \mathbb{P}(f \geq r^{\alpha} h(\tau, T)(I_{\Phi \cap B^c(o,r)} + N_0/E)). \quad (5.41)$$

The first equality in (5.41) is due to the independence of $I_{\Phi \cap B^c(o,r)-\delta_x}$ and $\Phi(B(o,r)) = 0$, which follows from the complete independence property of PPP. The second equality in (5.41) is due to Slivnyak-Mecke Theorem [16].

Following a similar derivation as in [10], we can uncondition on $F_0 = f$ and $X_0 = x$ to obtain

$$\mathbb{P}(\mathrm{SINR}_0 \geq T | D_0 = \tau) = \pi\lambda \mathbb{I}\left(g(\tau) > \frac{T}{1+T}\right) \times \int_0^{\infty} e^{-v^{\frac{\alpha}{2}} h(\tau,T) N_0/E} e^{-\pi\lambda v(1+\rho(h(\tau,T),\alpha))} \, \mathrm{d}v, \quad (5.42)$$

where $\rho(t, \alpha)$ is defined in Prop. 5.3. Unconditioning further on $D = \tau$ yields (5.25). This completes the proof.

Chapter 6

Conclusions

6.1 Summary

D2D is an exciting and innovative feature that will be present in forthcoming cellular networks. A scalable, energy efficient, and privacy sensitive D2D design will create big opportunities for mobile industry, while facilitating the inter-operability between critical public safety networks and ubiquitous commercial networks. D2D fundamentally alters the cellular architecture, reducing the primacy of BSs and enabling UEs to discover and directly communicate with proximate UEs. Unlike mobile ad hoc networks, D2D can usually rely on the assistance from network infrastructure for control functions like synchronization, session setup, resource allocation, routing, and other overhead-consuming functions that are extremely costly in a mobile ad hoc network.

Although simpler than a mobile ad hoc network, adding D2D features to cellular networks poses many challenges. A D2D-enabled cellular network is a highly complicated hybrid system. The design of this kind of hybrid systems requires a careful handling of the interaction between cellular and D2D services including resource management and interference control. As Chapter 1 dis-

cussed in detail, D2D networking requires a rethinking of many of the working assumptions and models used to date for cellular systems. This dissertation has identified four outstanding technical challenges in D2D-enabled cellular networks and addressed them with novel models and fundamental analysis.

In Chapter 2, we proposed a novel hybrid network model consisting of both ad hoc nodes and cellular infrastructure. This model captures key features of multicast D2D including multicast receiver heterogeneity and retransmissions, while being tractable for analytical purpose. Under the proposed model, we carried out a comprehensive analysis and optimization of multicast D2D. In Chapter 3, we adapted the hybrid network model and further incorporated D2D mode selection to study spectrum sharing between cellular and D2D communications. We derived analytical rate expressions and applied them to optimize the design of spectrum sharing. In Chapter 4, we extended the baseline single-antenna hybrid network model to multi-antenna transmission and focused on the interplay between massive MIMO and underlaid D2D networking. We investigated the spectral efficiency of such multi-antenna hybrid networks under both perfect CSI and imperfect CSI assumptions, and derived novel asymptotic and non-asymptotic results. Unlike Chapters 2 to 4 that assume the networks are synchronized, Chapter 5 is focused on the effect of asynchronous multicarrier transmission. Based on a detailed link-level analysis, we proposed a tractable SINR model and applied it to study system-level performance of asynchronous wireless networks.

6.2 Future Directions

The models and analysis in this dissertation can be used as a stepstone for a wide range of interesting topics including hybrid scheduling (i.e., a combination of centralized and distributed scheduling), feedback design, multihop and cooperative D2D communications. Please refer to [88] for a detailed discussion. Here we propose two more forward-looking extensions, which are interesting in the context of the 5th generation mobile networks (5G). The first is to enable D2D in millimeter wave (mmWave) bands, and the second is to study the interplay between D2D and small cells.

6.2.1 Millimeter Wave D2D

The focus of this dissertation is on enabling D2D networking in microwave frequencies that are below 5 GHz. To support multi-Gbps data rates in 5G, it is important to leverage the huge chunks of contiguous spectrum available at high frequencies, particularly mmWave bands [13, 23, 41, 106]. There exist many challenges regarding the use of mmWave bands in cellular networks due to their different propagation characteristics (vs. microwave frequency bands like the 900 MHz band). Specifically, with much smaller wavelength mmWave signals suffer from high pathloss. To overcome the high pathloss as well as other losses due to rain and oxygen absorption and higher noise floor associated with larger bandwidth, mmWave transmission requires a large array gain provided by appropriate beamforming techniques [70]. Further, mmWave signals cannot penetrate most solid materials and have very limited diffraction

ability [109, 114]. Therefore, signal paths may be easily blocked by obstacles, and only a few of them may arrive at the receiver. As a result, direct mmWave transmission appears restricted to short-range communication (say within 200 m).

Since mmWave transmission and D2D will likely to coexist in 5G, it is of interest to explore the opportunities of mmWave D2D transmission. To this end, the first step would be to develop appropriate channel models for mmWave D2D. Note that existing channel measurement results on asymmetric BS-UE radio access channels [114] may not be directly applicable to symmetric UE-UE channels [88], and thus additional measurement campaigns may be required. Once the channel models are established, the network models in this dissertation can be adapted to incorporate the distinctive traits of mmWave D2D transmission. Then specific design questions on mmWave D2D can be addressed; for example, system-level performance analysis and optimization may be carried out along the lines of this dissertation.

It is also of interest to explore how D2D may help mmWave cellular networks. One possibility is to employ D2D relaying to extend mmWave cellular coverage and network connectivity. For example, suppose a BS would like to communicate to a UE using mmWave transmission. The communication would be unsuccessful if the mmWave signal arriving at the UE is not strong enough due to the various blockages in the environment. Instead, D2D relaying may help the mmWave signal turn around obstacles and set up the connection provided there exists a feasible path from the BS to the UE.

6.2.2 D2D in Heterogeneous Networks

The focus of this dissertation is on enabling D2D networking in cellular networks that consist of only tower-mounted macro BSs. Network densification via deploying small cells has greatly increased area spectral efficiency due to the reduced lengths of radio access links, increased spatial spectrum reuse, and lightened load per cell. Networks having both macro and small cells are known as heterogeneous networks (HetNets) [12, 31]. Small cells bring BSs closer to UEs, resulting in shorter Tx-Rx distances. This idea is somewhat similar to that of D2D; the difference is that D2D exploits the natural proximity of nearby devices while small cells create the proximity.

Since small cells will be a key aspect of 5G, it is interesting to extend the work in this dissertation to HetNets and study the interplay between D2D and small cells. To this end, the first step would be to model the various types of BSs in HetNets. Random PPP models, used throughout this dissertation, can capture the randomness in the deployments of small cells and can be applied to model HetNets [34]. Then specific design issues on D2D networking in HetNets can be addressed; for example, all the design questions of D2D networking studied in this dissertation can be re-examined in a HetNet setting. Further, one can study how D2D networking affects the performance and design of HetNets.

One particularly interesting interplay between D2D and small cells is in load balancing. A major concern about deploying small cells is that they have very small coverage areas due to their low transmit powers. As a result,

small cells are often lightly loaded and do not accomplish as much as desired, while macrocells are still heavily loaded. To alleviate this issue, biasing allows small cells to expand their coverage areas, enabling more UEs to connect to small cells and thus helping to balance the load distribution [15]. Aggressive biasing, however, can lead to low SNR for the UEs located in the cell edges of small cells due to the limited power budgets of small cells. This problem can be solved by D2D relaying, which breaks a long link into two or more short links so that each link can have a sufficient power budget. Theoretically, the coverage of small cells can be unlimited with dense enough relaying devices. Therefore, in addition to offloading local traffic from cellular networks, D2D can help small cells accomplish more by extending their coverage to achieve more balanced load distribution in cellular networks. The design challenge here are (i) to dynamically discover and set up the optimal routing paths between BSs and end UEs, and (ii) to effectively and efficiently schedule the radio links including both D2D links and BS-UE links.

Bibliography

[1] 3GPP, "LTE; evolved universal terrestrial radio access (E-UTRA); radio frequency (RF) system scenarios," *TR 36.942 V9.0.1*, April 2010.

[2] ——, "3rd generation partnership project; technical specification group SA; feasibility study for proximity services (ProSe) (release 12)," *TR 22.803 V1.0.0*, August 2012.

[3] ——, "Multimedia broadcast/multicast service (MBMS); architecture and functional description," *TS 23.246 V11.1.0*, March 2012.

[4] ——, "Study on LTE device to device proximity services," December 2012.

[5] ——, "3rd generation partnership project; technical specification group radio access network; study on LTE device to device proximity services; radio aspects (release 12)," *TR 36.843 V12.0.1*, March 2014.

[6] 3GPP2, "CDMA2000 high rate broadcast-multicast packet data air interface specification," *C.S0054-0 V1.0*, February 2004.

[7] R. O. Afolabi, A. Dadlani, and K. Kim, "Multicast scheduling and resource allocation algorithms for OFDMA-based systems: A survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 240–254, 2013.

[8] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "A survey on spectrum management in cognitive radio networks," *IEEE Communications Magazine*, vol. 46, no. 4, pp. 40–48, April 2008.

[9] O. B. S. Ali, C. Cardinal, and F. Gagnon, "Performance of optimum combining in a Poisson field of interferers and Rayleigh fading channels," *IEEE Transactions on Wireless Communications*, vol. 9, no. 8, pp. 2461–2467, August 2010.

[10] J. G. Andrews, F. Baccelli, and R. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122–3134, November 2011.

[11] J. G. Andrews, "Interference cancellation for cellular systems: A contemporary overview," *IEEE Wireless Communications*, vol. 12, no. 2, pp. 19–29, April 2005.

[12] ——, "Seven ways that HetNets are a cellular paradigm shift," *IEEE Communications Magazine*, vol. 51, no. 3, pp. 136–144, March 2013.

[13] J. G. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, to appear, 2014.

[14] J. G. Andrews, S. Shakkottai, R. Heath, N. Jindal, M. Haenggi, R. Berry, D. Guo, M. Neely, S. Weber, S. Jafar *et al.*, "Rethinking information theory for mobile ad hoc networks," *IEEE Communications Magazine*, vol. 46, no. 12, pp. 94–101, December 2008.

[15] J. G. Andrews, S. Singh, Q. Ye, X. Lin, and H. S. Dhillon, "An overview of load balancing in HetNets: Old myths and open problems," *IEEE Wireless Communications*, vol. 21, no. 2, pp. 18–25, April 2014.

[16] F. Baccelli and B. Blaszczyszyn, "Stochastic geometry and wireless networks - Part I: Theory," *Foundations and Trends in Networking*, vol. 3, no. 3-4, pp. 249–449, 2009.

[17] F. Baccelli, B. Blaszczyszyn, and P. Muhlethaler, "An Aloha protocol for multihop mobile wireless networks," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 421–436, February 2006.

[18] F. Baccelli, N. Khude, R. Laroia, J. Li, T. Richardson, S. Shakkottai, S. Tavildar, and X. Wu, "On the design of device-to-device autonomous discovery," in *Proceedings of Fourth International Conference on Communication Systems and Networks*, 2012, pp. 1–9.

[19] F. Baccelli and P. Brémaud, *Elements of queueing theory: Palm Martingale calculus and stochastic recurrences*. Springer Verlag, 2003, vol. 26.

[20] F. Baccelli, J. Li, T. Richardson, S. Shakkottai, S. Subramanian, and X. Wu, "On optimizing CSMA for wide area ad hoc networks," *Queueing Systems*, vol. 72, no. 1-2, pp. 31–68, October 2012.

[21] T. Bai and R. W. Heath Jr, "Asymptotic coverage probability and rate in massive MIMO networks," in *Proceedings of IEEE Global Conference on Signal and Information Processing*, December 2014.

[22] B. Blaszczyszyn, M. K. Karray, and H. P. Keeler, "Using Poisson processes to model lattice cellular networks," in *Proceedings of IEEE INFOCOM*, 2013, pp. 773–781.

[23] F. Boccardi, R. W. Heath Jr, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, February 2014.

[24] T. X. Brown, "Cellular performance bounds via shotgun cellular systems," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 11, pp. 2443–2455, November

2000.

[25] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, July 2003.

[26] P. Chaporkar and S. Sarkar, "Wireless multicast: Theory and approaches," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 1954–1972, June 2005.

[27] B. Chen and M. J. Gans, "MIMO communications in ad hoc networks," *IEEE Transactions on Signal Processing*, vol. 54, no. 7, pp. 2773–2783, July 2006.

[28] W. Choi and J. G. Andrews, "Downlink performance and capacity of distributed antenna systems in a multicell environment," *IEEE Transactions on Wireless Communications*, vol. 6, no. 1, pp. 69–73, January 2007.

[29] M. Corson, R. Laroia, J. Li, V. Park, T. Richardson, and G. Tsirtsis, "Toward proximity-aware internetworking," *IEEE Wireless Communications*, vol. 17, no. 6, pp. 26–33, December 2010.

[30] H. Dai, A. F. Molisch, and H. V. Poor, "Downlink capacity of interference-limited MIMO systems with joint detection," *IEEE Transactions on Wireless Communications*, vol. 3, no. 2, pp. 442–453, March 2004.

[31] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Communications*, vol. 18, no. 3, pp. 10–21, June 2011.

[32] S. Deb, S. Jaiswal, and K. Nagaraj, "Real-time video multicast in WiMAX networks," in *Proceedings of IEEE INFOCOM*, 2008, pp. 1579–1587.

[33] Deutsche Telekom AG, Orange Silicon Valley, Qualcomm Technologies Incorporated, Tagged Incorporated, and Samsung Electronics, "LTE Direct workshop white paper," May 2013. [Online]. Available: https://www.qualcomm.com/media/documents/files/lte-direct-whitepaper.pdf

[34] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and analysis of K-tier downlink heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 550 –560, April 2012.

[35] T. Doumi, M. F. Dolan, S. Tatesh, A. Casati, G. Tsirtsis, K. Anchan, and D. Flore, "LTE for public safety networks," *IEEE Communications Magazine*, vol. 51, no. 2, pp. 106–112, February 2013.

[36] J. Du, W. Zhu, J. Xu, Z. Li, and H. Wang, "A compressed HARQ feedback for device-to-device multicast communications," in *Proceedings of IEEE Vehicular Technology Conference*, 2012, pp. 1–5.

[37] L. Duan, J. Huang, and B. Shou, "Duopoly competition in dynamic spectrum leasing and pricing," *IEEE Transactions on Mobile Computing*, vol. 11, no. 11, pp. 1706–1719, November 2012.

[38] R. Durrett, *Probability: Theory and examples*. Cambridge University Press, 2010.

[39] M. S. El-Tanany, Y. Wu, and L. Házy, "OFDM uplink for interactive broadband wireless: Analysis and simulation in the presence of carrier, clock and timing errors," *IEEE Transactions on Broadcasting*, vol. 47, no. 1, pp. 3–19, March 2001.

[40] H. ElSawy, E. Hossain, and M. Haenggi, "Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: A survey," *IEEE*

Communications Surveys & Tutorials, vol. 15, no. 3, pp. 996–1019, July 2013.

[41] Ericsson, "5G radio access: Research and vision," *white paper*, June 2013.

[42] G. Fodor, E. Dahlman, G. Mildh, S. Parkvall, N. Reider, G. Miklós, and Z. Turányi, "Design aspects of network assisted device-to-device communications," *IEEE Communications Magazine*, vol. 50, no. 3, pp. 170–177, March 2012.

[43] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Autumn 1996.

[44] R. K. Ganti and M. Haenggi, "Spatial and temporal correlation of the interference in aloha ad hoc networks," *IEEE Communications Letters*, vol. 13, no. 9, pp. 631–633, September 2009.

[45] A. J. Goldsmith and S. B. Wicker, "Design challenges for energy constrained ad hoc wireless networks," *IEEE Wireless Communications*, vol. 9, no. 4, pp. 8–27, August 2002.

[46] C. Gomez, J. Oller, and J. Paradells, "Overview and evaluation of bluetooth low energy: An emerging low-power wireless technology," *Sensors*, vol. 12, no. 9, pp. 11 734–11 753, 2012.

[47] P. K. Gopala and H. El Gamal, "On the throughput-delay tradeoff in cellular multicast," in *Proceedings of International Conference on Wireless Networks, Communications and Mobile Computing*, vol. 2, 2005, pp. 1401–1406.

[48] S. Govindasamy, D. W. Bliss, and D. H. Staelin, "Spectral efficiency in single-hop ad-hoc wireless networks with interference using adaptive antenna arrays," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 7, pp. 1358–1369, September 2007.

[49] A. Guo and M. Haenggi, "Asymptotic deployment gain: A simple approach to characterize the SINR distribution in general cellular networks," *http://arxiv.org/abs/1404.6556*, April 2014.

[50] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388–404, March 2000.

[51] M. Haenggi and R. Ganti, "Interference in large wireless networks," *Foundations and Trends in Networking*, vol. 3, no. 2, pp. 127–248, 2009.

[52] M. Haenggi, "On distances in uniformly random networks," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3584–3586, October 2005.

[53] ——, "Diversity loss due to interference correlation," *IEEE Communications Letters*, vol. 16, no. 10, pp. 1600–1603, October 2012.

[54] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1029–1046, September 2009.

[55] K. Hamdi, W. Zhang, and K. Letaief, "Opportunistic spectrum sharing in cognitive MIMO wireless networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 8, pp. 4098–4109, August 2009.

[56] K. A. Hamdi, "Capacity of MRC on correlated Rician fading channels," *IEEE Transactions on Communications*, vol. 56, no. 5, pp. 708–711, May 2008.

[57] ——, "Precise interference analysis of OFDMA time-asynchronous wireless ad-hoc networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 1, pp. 134–144, January 2010.

[58] K. A. Hamdi and Y. M. Shobowale, "Interference analysis in downlink OFDM considering imperfect intercell synchronization," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 7, pp. 3283–3291, September 2009.

[59] T. Han and N. Ansari, "Energy efficient wireless multicasting," *IEEE Communications Letters*, vol. 15, no. 6, pp. 620–622, June 2011.

[60] Z. Han, R. Zheng, and H. V. Poor, "Repeated auctions with Bayesian nonparametric learning for spectrum access in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 3, pp. 890–900, March 2011.

[61] R. W. Heath, T. Wu, Y. H. Kwon, and A. C. Soong, "Multiuser MIMO in distributed antenna systems with out-of-cell interference," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4885–4899, October 2011.

[62] R. W. Heath Jr and D. J. Love, "Multimode antenna selection for spatial multiplexing systems with linear receivers," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 3042–3056, August 2005.

[63] F. Hou, L. X. Cai, P.-H. Ho, X. Shen, and J. Zhang, "A cooperative multicast scheduling scheme for multimedia services in IEEE 802.16 networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1508–1519, March 2009.

[64] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, February 2013.

[65] K. Huang, V. Lau, and Y. Chen, "Spectrum sharing between cellular and mobile ad hoc networks: Transmission-capacity trade-off," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1256–1267, 2009.

[66] K. Huang, J. G. Andrews, D. Guo, R. W. Heath, and R. A. Berry, "Spatial interference cancellation for multiantenna mobile ad hoc networks," *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1660–1676, March 2012.

[67] S. Huang, X. Liu, and Z. Ding, "Opportunistic spectrum access in cognitive radio networks," in *Proceedings of IEEE INFOCOM*, 2008, pp. 1427–1435.

[68] H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprashad, "Achieving massive MIMO spectral efficiency with a not-so-large number of antennas," *IEEE Transactions on Wireless Communications*, vol. 11, no. 9, pp. 3226–3239, September 2012.

[69] A. M. Hunter, J. G. Andrews, and S. Weber, "Transmission capacity of ad hoc networks with spatial diversity," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 5058–5071, December 2008.

[70] S. Hur, T. Kim, D. Love, J. Krogmeier, T. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Transactions on Communications*, vol. 61, no. 10, pp. 4391–4403, October 2013.

[71] P. Janis, V. Koivunen, C. B. Ribeiro, K. Doppler, and K. Hugl, "Interference avoiding MIMO schemes for device-to-device radio underlaying cellular networks," in *Proceed-*

ings of IEEE Personal, Indoor and Mobile Radio Communications, 2009, pp. 2385–2389.

[72] L. Jayasinghe, P. Jayasinghe, N. Rajatheva, and M. Latva-aho, "MIMO physical layer network coding based underlay device-to-device communication," in *Proceedings of IEEE International Symposium on Personal Indoor and Mobile Radio Communications*, 2013, pp. 89–94.

[73] M. Ji, G. Caire, and A. F. Molisch, "Fundamental limits of distributed caching in D2D wireless networks," in *Proceedings of IEEE Information Theory Workshop*, 2013, pp. 1–5.

[74] N. Jindal, J. G. Andrews, and S. Weber, "Multi-antenna communication in ad hoc networks: Achieving MIMO gains with SIMO transmission," *IEEE Transactions on Communications*, vol. 59, no. 2, pp. 529–540, February 2011.

[75] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2640–2651, August 2011.

[76] D. Kalathil and R. Jain, "Spectrum sharing through contracts for cognitive radios," *IEEE Transactions on Mobile Computing*, vol. 12, no. 10, pp. 1999–2011, October 2013.

[77] X. Kang, Y.-C. Liang, H. K. Garg, and L. Zhang, "Sensing-based spectrum sharing in cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 8, pp. 4649–4654, October 2009.

[78] B. Kaufman, J. Lilleberg, and B. Aazhang, "Spectrum sharing scheme between cellular users and ad-hoc device-to-device users," *IEEE Transactions on Wireless Communications*, vol. 12, no. 3, pp. 1038–1049, March 2013.

[79] H. P. Keeler, B. Blaszczyszyn, and M. K. Karray, "SINR-based k-coverage probability in cellular networks with arbitrary shadowing," in *Proceedings of IEEE International Symposium on Information Theory*, 2013, pp. 1167–1171.

[80] M. Kountouris and J. G. Andrews, "Transmission capacity scaling of sdma in wireless ad hoc networks," in *Proceedings of Information Theory Workshop*. IEEE, 2009, pp. 534–538.

[81] U. C. Kozat and L. Tassiulas, "Throughput capacity of random ad hoc networks with infrastructure support," in *Proceedings of ACM MobiCom*, 2003, pp. 55–65.

[82] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, February 2014.

[83] C.-H. Lee, C.-Y. Shih, and Y.-S. Chen, "Stochastic geometry based models for modeling cellular networks in urban areas," *Wireless Networks*, vol. 19, no. 6, pp. 1063–1072, August 2013.

[84] J. C. Li, M. Lei, and F. Gao, "Device-to-device (D2D) communication in MU-MIMO cellular networks," in *Proceedings of IEEE Global Communications Conference*, 2012, pp. 3583–3587.

[85] X.-Y. Li, "Multicast capacity of wireless ad hoc networks," *IEEE/ACM Transactions on Networking*, vol. 17, no. 3, pp. 950–961, June 2009.

[86] X. Lin, J. G. Andrews, and A. Ghosh, "Modeling, analysis and design for carrier aggregation in heterogeneous cellular networks," *IEEE Transactions on Communications*, vol. 61, no. 9, pp. 4002–4015, September 2013.

[87] X. Lin and J. G. Andrews, "A general approach to SINR-based performance metrics with application to D2D and carrier aggregation," in *Proceedings of Asilomar Conference on Signals, Systems, and Computers*, November 2013, pp. 1–5.

[88] X. Lin, J. G. Andrews, A. Ghosh, and R. Ratasuk, "An overview of 3GPP device-to-device proximity services," *IEEE Communications Magazine*, vol. 52, no. 4, pp. 40–48, April 2014.

[89] X. Lin, J. G. Andrews, and A. Ghosh, "Spectrum sharing for device-to-device communication in cellular networks," *IEEE Transactions on Wireless Communications*, to appear, 2014.

[90] X. Lin, R. K. Ganti, P. J. Fleming, and J. G. Andrews, "Towards understanding the fundamentals of mobility in cellular networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 4, pp. 1686–1698, April 2013.

[91] X. Lin, R. Ratasuk, A. Ghosh, and J. G. Andrews, "Modeling, analysis and optimization of multicast device-to-device transmissions," *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4346–4359, August 2014.

[92] Y.-D. Lin and Y.-C. Hsu, "Multihop cellular: A new architecture for wireless communications," in *Proceedings of IEEE INFOCOM*, vol. 3, 2000, pp. 1273–1282.

[93] B. Liu, Z. Liu, and D. Towsley, "On the capacity of hybrid wireless networks," in *Proceedings of IEEE INFOCOM*, vol. 2, 2003, pp. 1543–1552.

[94] C.-H. Liu and J. G. Andrews, "Multicast outage probability and transmission capacity of multihop wireless networks," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4344–4358, July 2011.

[95] J. Liu, W. Chen, Z. Cao, and K. B. Letaief, "Dynamic power and sub-carrier allocation for OFDMA-based wireless multicast systems," in *Proceedings of IEEE International Conference on Communications*, 2008, pp. 2607–2611.

[96] R. H. Louie, M. R. McKay, and I. B. Collings, "Open-loop spatial multiplexing and diversity communications in ad hoc networks," *IEEE Transactions on Information Theory*, vol. 57, no. 1, pp. 317–344, January 2011.

[97] P. Madhusudhanan, X. Li, Y. Liu, and T. Brown, "Stochastic geometric modeling and interference analysis for massive MIMO systems," in *Proceedings of International Symposium on Modeling Optimization in Mobile, Ad Hoc Wireless Networks*, 2013, pp. 15–22.

[98] X. F. Mao, X.-Y. Li, and S. J. Tang, "Multicast capacity for hybrid wireless networks," in *Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing*, 2008, pp. 189–198.

[99] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, November 2010.

[100] Y. Medjahdi, M. Terré, D. Le Ruyet, D. Roviras, and A. Dziri, "Performance analysis in the downlink of asynchronous OFDM/FBMC based multi-cellular networks," *IEEE*

Transactions on Wireless Communications, vol. 10, no. 8, pp. 2630–2639, August 2011.

[101] H. Min, J. Lee, S. Park, and D. Hong, "Capacity enhancement using an interference limited area for device-to-device uplink underlaying cellular networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 12, pp. 3995–4000, December 2011.

[102] Y. Mostofi and D. C. Cox, "Mathematical analysis of the impact of timing synchronization errors on the performance of an OFDM system," *IEEE Transactions on Communications*, vol. 54, no. 2, pp. 226–230, February 2006.

[103] H. Q. Ngo, E. Larsson, and T. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436–1449, April 2013.

[104] D. Niyato and E. Hossain, "Competitive spectrum sharing in cognitive radio networks: A dynamic game approach," *IEEE Transactions on Wireless Communications*, vol. 7, no. 7, pp. 2651–2660, July 2008.

[105] T. D. Novlan, H. S. Dhillon, and J. G. Andrews, "Analytical modeling of uplink cellular networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 6, pp. 2669–2679, June 2013.

[106] NSN, "Looking ahead to 5G: Building a virtual zero latency gigabit experience," *white paper*, December 2013.

[107] J. C. Park and S. K. Kasera, "Enhancing cellular multicast performance using ad hoc networks," in *Proceedings of IEEE Wireless Communications and Networking*

Conference, vol. 4, 2005, pp. 2175–2181.

[108] M. Park, K. Ko, H. Yoo, and D. Hong, "Performance analysis of OFDMA uplink systems with symbol timing misalignment," *IEEE Communications Letters*, vol. 7, no. 8, pp. 376–378, August 2003.

[109] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 101–107, June 2011.

[110] T. Pollet, P. Spruyt, and M. Moeneclaey, "The BER performance of OFDM systems using non-synchronized sampling," in *Proceedings of IEEE Global Communications Conference*, 1994, pp. 253–257.

[111] Qualcomm, "Creating a digital 6th sense with LTE Direct," July 2014.

[112] K. Raghunath and A. Chockalingam, "SIR analysis and interference cancellation in uplink OFDMA with large carrier frequency/timing offsets," *IEEE Transactions on Wireless Communications*, vol. 8, no. 5, pp. 2202–2208, May 2009.

[113] R. Ramanathan, J. Redi, C. Santivanez, D. Wiggins, and S. Polit, "Ad hoc networking with directional antennas: a complete system solution," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 3, pp. 496–506, March 2005.

[114] T. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. Wong, J. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013.

[115] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, January 2013.

[116] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Transactions on Communications*, vol. 45, no. 12, pp. 1613–1621, December 1997.

[117] J. Seppala, T. Koskela, T. Chen, and S. Hakola, "Network controlled device-to-device (D2D) and cluster multicast concept for LTE and LTE-A networks," in *Proceedings of IEEE Wireless Communications and Networking Conference*, 2011, pp. 986–991.

[118] S. Shakkottai, X. Liu, and R. Srikant, "The multicast capacity of large multihop wireless networks," *IEEE/ACM Transactions on Networking*, vol. 18, no. 6, pp. 1691–1700, December 2010.

[119] M. Shao, S. Dumitrescu, and X. Wu, "Layered multicast with inter-layer network coding for multimedia streaming," *IEEE Transactions on Multimedia*, vol. 13, no. 2, pp. 353–365, April 2011.

[120] H. Song, J.-P. Hong, and W. Choi, "On the optimal switching probability for a hybrid cognitive radio system," *IEEE Transactions on Wireless Communications*, vol. 12, no. 4, pp. 1594–1605, April 2013.

[121] M. Speth, S. A. Fechtel, G. Fock, and H. Meyr, "Optimum receiver design for wireless broad-band systems using OFDM - Part I," *IEEE Transactions on Communications*, vol. 47, no. 11, pp. 1668–1677, November 1999.

[122] D. Stoyan, W. Kendall, J. Mecke, and L. Ruschendorf, *Stochastic Geometry and Its Applications*. Wiley New York, 1995.

[123] C. Suh, S. Park, and Y. Cho, "Efficient algorithm for proportional fairness scheduling

in multicast OFDM systems," in *Proceedings of IEEE Vehicular Technology Conference*, vol. 3, 2005, pp. 1880–1884.

[124] H. Tang, C. Zhu, and Z. Ding, "Cooperative MIMO precoding for D2D underlay in cellular networks," in *Proceedings of IEEE International Conference on Communications*, 2013, pp. 5517–5521.

[125] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, November-December 1999.

[126] A. M. Tonello, N. Laurenti, and S. Pupolin, "Analysis of the uplink of an asynchronous multi-user DMT OFDMA system impaired by time offsets, frequency offsets, and multi-path fading," in *Proceedings of IEEE Vehicular Technology Conference*, vol. 3, 2000, pp. 1094–1099.

[127] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 641–657, March 1999.

[128] A. M. Tulino and S. Verdú, *Random matrix theory and wireless communications*. Now Publishers Inc., 2004, vol. 1.

[129] E. Tuomaala and H. Wang, "Effective SINR approach of link to system mapping in OFDM/multi-carrier mobile network," in *International Conference on Mobile Technology, Applications and Systems*, 2005, pp. 1–5.

[130] S. Vasudevan, J. Kurose, and D. Towsley, "On neighbor discovery in wireless networks with directional antennas," in *Proceedings of IEEE INFOCOM*, vol. 4, March 2005, pp. 2502–2512.

[131] R. Vaze and R. W. Heath, "Transmission capacity of ad-hoc networks with multiple antennas using transmit stream adaptation and interference cancellation," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 780–792, February 2012.

[132] S. Venkatesan, A. Lozano, and R. Valenzuela, "Network MIMO: Overcoming intercell interference in indoor wireless systems," in *Proceedings of Asilomar Conference on Signals, Systems and Computers*, 2007, pp. 83–87.

[133] S. Verdu, *Multiuser detection*. Cambridge University Press, 1998.

[134] S. Verdú and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 622–640, March 1999.

[135] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2658–2668, October 2003.

[136] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Transactions on Information Theory*, vol. 49, no. 8, pp. 1912–1921, August 2003.

[137] X. Wang, T. T. Tjhung, Y. Wu, and B. Caron, "SER performance evaluation and optimization of OFDM system with residual frequency and timing offsets from imperfect synchronization," *IEEE Transactions on Broadcasting*, vol. 49, no. 2, pp. 170–177, June 2003.

[138] X. Wang, Z. Li, P. Xu, Y. Xu, X. Gao, and H.-H. Chen, "Spectrum sharing in cognitive radio networks - an auction-based approach," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 40, no. 3, pp. 587–596, June 2010.

[139] S. Weber and J. G. Andrews, "Transmission capacity of wireless networks," *Foundations and Trends in Networking*, vol. 5, no. 2-3, pp. 109–281, 2012.

[140] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3936–3964, September 2006.

[141] Wi-Fi Alliance, "Wi-Fi peer-to-peer (P2P) technical specification," *V1.2*, February 2013.

[142] H. Won, H. Cai, K. Guo, A. Netravali, I. Rhee, K. Sabnani *et al.*, "Multicast scheduling in cellular data networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 9, pp. 4540–4549, September 2009.

[143] H. Wu, C. Qiao, S. De, and O. Tonguz, "Integrated cellular and ad hoc relaying systems: iCAR," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 10, pp. 2105–2115, October 2001.

[144] X. Wu, S. Tavildar, S. Shakkottai, T. Richardson, J. Li, R. Laroia, and A. Jovicic, "FlashlinQ: A synchronous distributed scheduler for peer-to-peer ad hoc networks," in *Proceedings of Allerton Conference on Communication, Control, and Computing*, 2010, pp. 514–521.

[145] J. Xu, S.-J. Lee, W.-S. Kang, and J.-S. Seo, "Adaptive resource allocation for MIMO-OFDM based wireless multicast systems," *IEEE Transactions on Broadcasting*, vol. 56, no. 1, pp. 98–102, February 2010.

[146] S. Xu, H. Wang, T. Chen, Q. Huang, and T. Peng, "Effective interference cancellation scheme for device-to-device communication underlaying cellular networks," in

Proceedings of IEEE Vehicular Technology Conference, 2010, pp. 1–5.

[147] C. Yu, K. Doppler, C. Ribeiro, and O. Tirkkonen, "Resource sharing optimization for device-to-device communication underlaying cellular networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2752–2763, August 2011.

[148] A. Zemlianov and G. de Veciana, "Capacity of ad hoc wireless networks with infrastructure support," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 3, pp. 657–667, March 2005.

[149] J. Zhang, R. Chen, J. G. Andrews, A. Ghosh, and R. W. Heath, "Networked MIMO with clustered linear precoding," *IEEE Transactions on Wireless Communications*, vol. 8, no. 4, pp. 1910–1921, April 2009.

[150] R. Zhang and Y.-C. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 88–102, February 2008.

[151] Y. J. Zhang and A. So, "Optimal spectrum sharing in MIMO cognitive radio networks via semidefinite programming," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 2, pp. 362–373, February 2011.

Vita

Xingqin Lin received the B.Eng. degree in Electronic Information Engineering from Tianjin University (China) in 2009 and the M.Phil. degree in Information Engineering from The Chinese University of Hong Kong in 2011. He is currently a Ph.D. candidate in Electrical & Computer Engineering at The University of Texas at Austin. His research interests are mainly in the broad area of wireless communications and networking with current focus on device-to-device communication, massive MIMO, and small cells. He is a Certified Teaching Assistant at UT Austin, and received the Microelectronics and Computer Development (MCD) fellowship from UT Austin in 2012. He was an Exemplary Reviewer for IEEE Wireless Communications Letters in 2013. He held summer internships at Qualcomm CR&D in Bridgewater, NJ in summer 2014, Nokia Siemens Networks in Arlington Heights, IL in summer 2013, and Alcatel-Lucent Bell Labs in Murray Hill, NJ in summer 2012.

Permanent email: linxingqin2008@gmail.com

This dissertation was typeset with LaTeX[†] by the author.

[†]LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.