Modeling, Analysis, and Design of 5G Networks using
Stochastic Geometry

Dissertation by
Konpal Shaukat Ali

In Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy

King Abdullah University of Science and Technology
Thuwal, Kingdom of Saudi Arabia

October, 2018
The dissertation of Konpal Shaukat Ali is approved by the examination committee.

Committee Chairperson: Mohamed-Slim Alouini
Committee Members: Tareq Al-Naffouri, Meriem T. Laleg, Hikmet Sari (External)
ABSTRACT

Modeling, Analysis, and Design of 5G Networks using Stochastic Geometry

Konpal Shaukat Ali

Improving spectral-utilization is a core focus to cater the ever-increasing demand in data rate and system capacity required for the development of 5G. This dissertation focuses on three spectrum-reuse technologies that are envisioned to play an important role in 5G networks: device-to-device (D2D), full-duplex (FD), and non-orthogonal multiple access (NOMA). D2D allows proximal user-equipments (UEs) to bypass the cellular base-station and communicate with their intended receiver directly. In underlay D2D, the D2D UEs utilize the same spectral resources as the cellular UEs. FD communication allows a transmit-receive pair to transmit simultaneously on the same frequency channel. Due to the overwhelming self-interference encountered, FD was not possible until very recently courtesy of advances in transceiver design. NOMA allows multiple receivers (transmitters) to communicate with one transmitter (receiver) in one time-frequency resource-block by multiplexing in the power domain. Successive-interference cancellation is used for NOMA decoding. Each of these techniques significantly improves spectral efficiency and consequently data rate and throughput; however, the price paid is increased interference. Since each of these technologies allow multiple transmissions within a cell on a time-frequency resource-block, they result in interference within the cell (i.e., intracell interference). Additionally, due to the increased communication, they increase network interference from outside the cell under consideration as well (i.e., increased intercell interference).

Real networks are becoming very dense; as a result, the impact of intercell interference coming from the entire network is significant. As such, using models that
consider a single-cell/few-cell scenarios result in misleading conclusions. Hence, accurate modeling requires considering a large network. In this context, stochastic geometry is a powerful tool for analyzing random patterns of points such as those found in wireless networks. In this dissertation, stochastic geometry is used to model and analyze the different technologies that are to be deployed in 5G networks. This gives us insight into the network performance, showing us the impacts of deploying a certain technology into real 5G networks. Additionally, it allows us to propose schemes for integrating such technologies, mode-selection, parameter-selection, and resource-allocation that enhance the parameters of interest in the network such as data rate, coverage, and secure communication.
ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Mohamed-Slim Alouini for his guidance, mentoring and support throughout the course of my degree. I want to thank my committee members, Dr. Tareq Al-Naffouri, Dr. Meriem T. Laleg, and Dr. Hikmet Sari for their insightful feedback as well as for the time they spent in going through this thesis. Dr. Hesham ElSawy introduced me to stochastic geometry, and for this I will always be grateful to him. I would also like to thank him for his guidance, mentorship and advice on various situations. Dr. Martin Haenggi honed my stochastic geometry as well as a number of other skills, and for this I am very grateful. I would also like to thank him for always making the time for discussions despite being on the other side of the planet and for his astute feedback. I also want to thank Dr. Anas Chaaban for his guidance on the information theoretic aspect of my work.

I want to thank a number of my friends who have, in the words of a wise man, added pizzazz to my life. In particular, I want to give a shout out (in alphabetical order) to: Annie, Anum, Emaad, Farhan, Fatma, Habeeba, Haris, Huda, Idris, Ikram, Iqra, Jenny, Nada, Naimal, ND-MSA, Rana, Sayyari, Tsiky, Warda, Wardah, and Yasmeen.

I want to thank my family – Mama, Papa, Owais, Taha, and Sanayia – for being my biggest support system, always; I would be nowhere without you guys. Thank you for always having my back and for all the love.
TABLE OF CONTENTS

Examination Committee Page 2
Copyright 3
Abstract 4
Acknowledgements 6
List of Abbreviations 12
List of Figures 14
List of Tables 18

1 Introduction 19

1.1 Background ................................................................. 19
1.2 Motivation behind Stochastic Geometry for Large Wireless Networks 21
1.3 Technologies to be incorporated into 5G and Related Work on Them 22
1.3.1 Device-to-Device (D2D) ........................................ 23
1.3.2 Full Duplex (FD) .................................................. 24
1.3.3 Non-Orthogonal Multiple Access (NOMA) .................. 25
1.4 Integrating Efficient Spectrum Reuse Technologies ............... 28
1.4.1 FD-D2D ............................................................ 29
1.4.2 FD-NOMA ....................................................... 31
1.4.3 D2D-NOMA ..................................................... 33
1.5 On Physical Layer Security ......................................... 34
1.6 On the Meta Distribution ............................................. 35
1.6.1 The Concept and Significance of the Meta Distribution .... 35
1.6.2 Calculating the Meta Distribution .............................. 37
1.6.3 Literature on the Meta Distribution ........................... 40
1.7 Objectives and Contributions ....................................... 42
1.7.1 Full Duplex D2D Communication ............................ 42
2 Stochastic Geometry for Wireless Networks

2.1 Why use Stochastic Geometry?

2.2 The Concept of Point Processes (PPs)

2.2.1 Poisson Point Process

2.2.2 Repulsive Processes/Hard-core Processes

2.2.3 Cluster Processes

2.3 Analyzing Metrics

2.3.1 Characterizing the Statistics of the Interference

2.3.2 Characterizing the Statistics of the SINR/SIR - Laplace Transform of Interference Approach

2.3.3 On the Laplace Transform of Interference

2.3.4 Characterizing the Statistics of the SINR/SIR - Relative Distance Process Approach

3 Modeling Cellular Networks with Full Duplex D2D Communication: A Stochastic Geometry Approach

3.1 Introduction

3.2 System Model

3.2.1 Network Model

3.2.2 Mode Selection

3.2.3 Methodology of Analysis

3.3 On Link Distances and UE Classification

3.3.1 Link Distance Analysis

3.3.2 UE Classification

3.4 Transmit Power Analysis

3.5 SINR Analysis

3.6 Results and Analysis

3.6.1 Parameter Selection

3.6.2 Model Validation

3.6.3 Effects of the Interference Protection Condition
3.6.4 Effects of the Distance Cut-off for the FD Mode . . . . . . . . 91
3.6.5 Analyzing Truncation and SINR Outage . . . . . . . . . . . . 92
3.6.6 Effects of Imperfect SIC . . . . . . . . . . . . . . . . . . . . . 93
3.6.7 Effects of the D2D link distance distribution . . . . . . . . . . 94
3.7 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 95

4 The Effect of Spatial Interference Correlation and Jamming on Secrecy in Cellular Networks 97
4.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 97
4.2 System Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 97
  4.2.1 Network Model . . . . . . . . . . . . . . . . . . . . . . . . . . . 97
  4.2.2 Methodology of Analysis . . . . . . . . . . . . . . . . . . . . . . 98
4.3 SIR Analysis . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 99
  4.3.1 Framework . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 99
  4.3.2 Marginal Probabilities in the Analytical Network of Interest . 101
4.4 The Effect of Spatial Correlation . . . . . . . . . . . . . . . . . . . . 102
4.5 Numerical Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 104
4.6 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 106

5 Non-Orthogonal Multiple Access for Large-Scale 5G Networks: Interference Aware Design 107
5.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 107
5.2 Interference-aware Design for NOMA Networks . . . . . . . . . . . 108
  5.2.1 Utility Function . . . . . . . . . . . . . . . . . . . . . . . . . . . . 108
  5.2.2 Optimization Objective . . . . . . . . . . . . . . . . . . . . . . . 110
  5.2.3 UE Clustering . . . . . . . . . . . . . . . . . . . . . . . . . . . . 111
  5.2.4 UE Sorting . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 112
  5.2.5 Power/Rate Allocation . . . . . . . . . . . . . . . . . . . . . . . 115
5.3 Case Study . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 116
  5.3.1 Impact of Intercell Interference on Rate Utility . . . . . . . . . 116
  5.3.2 Impact of Intercell Interference on Power Allocation and Fairness 117
5.4 Intercell Interference Management . . . . . . . . . . . . . . . . . . . . 121
  5.4.1 Offline Interference Management . . . . . . . . . . . . . . . . . 121
  5.4.2 Online Interference Management . . . . . . . . . . . . . . . . . 122
5.5 Integration of NOMA with Other 5G Technologies . . . . . . . . . . . 124
  5.5.1 Full-Duplex NOMA . . . . . . . . . . . . . . . . . . . . . . . . . 126
  5.5.2 D2D NOMA . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 128
# 6 Downlink Non-Orthogonal Multiple Access (NOMA) in Poisson Networks

## 6.1 Introduction

## 6.2 System Model

### 6.2.1 NOMA System Model

### 6.2.2 OMA System Model

## 6.3 SINR Analysis

### 6.3.1 SINR in NOMA Network

### 6.3.2 Laplace Transform of the Intercell Interference

### 6.3.3 UE Ordering

### 6.3.4 Coverage in NOMA Network

## 6.4 NOMA Optimization

### 6.4.1 Problem Formulation

### 6.4.2 Case Study: $N = 2$

### 6.4.3 Algorithm for Solving $P1$

### 6.4.4 Algorithm for Solving $P2$

## 6.5 Results

### 6.5.1 Performance

### 6.5.2 Complexity

## 6.6 Conclusion

# 7 Meta Distribution of Downlink Non-Orthogonal Multiple Access (NOMA) in Poisson Networks

## 7.1 Introduction

## 7.2 System Model

## 7.3 SIR Analysis

## 7.4 Results

## 7.5 Conclusion

# 8 Concluding Remarks

## 8.1 Summary

## 8.2 Future Research Work

# References
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>BS</td>
<td>Base-Station</td>
</tr>
<tr>
<td>BDMA</td>
<td>Bit Division Multiple Access</td>
</tr>
<tr>
<td>CCDF</td>
<td>Complementary Cumulative Density Function</td>
</tr>
<tr>
<td>CCP</td>
<td>Conditional Coverage Probability</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Density Function</td>
</tr>
<tr>
<td>C-NOMA</td>
<td>Cell-center NOMA</td>
</tr>
<tr>
<td>CoMP</td>
<td>Coordinated Multi-Point</td>
</tr>
<tr>
<td>CSI</td>
<td>Channel State Information</td>
</tr>
<tr>
<td>D2D</td>
<td>Device-to-device</td>
</tr>
<tr>
<td>E-NOMA</td>
<td>Everywhere NOMA</td>
</tr>
<tr>
<td>FD</td>
<td>Full Duplex</td>
</tr>
<tr>
<td>FDJ</td>
<td>Full Duplex Jamming</td>
</tr>
<tr>
<td>FFR</td>
<td>Fractional Frequency Reuse</td>
</tr>
<tr>
<td>HARQ</td>
<td>Hybrid Automatic Repeat Request</td>
</tr>
<tr>
<td>HD</td>
<td>Half Duplex</td>
</tr>
<tr>
<td>HetNets</td>
<td>Heterogeneous Networks</td>
</tr>
<tr>
<td>IDMA</td>
<td>Interleave-Division Multiple Access</td>
</tr>
<tr>
<td>i.i.d.</td>
<td>Independent and Identically Distributed</td>
</tr>
<tr>
<td>INR</td>
<td>Interference to Noise Ratio</td>
</tr>
<tr>
<td>IP</td>
<td>Interference Protection</td>
</tr>
<tr>
<td>ISINR</td>
<td>Instantaneous Signal-to-Intercell-interference-and-Noise-Ratio</td>
</tr>
<tr>
<td>ISP</td>
<td>Instantaneous Signal Power</td>
</tr>
<tr>
<td>LT</td>
<td>Laplace Transform</td>
</tr>
<tr>
<td>MFSINR</td>
<td>Mean-Fading Signal-to-Intercell-interference-and-Noise Ratio</td>
</tr>
<tr>
<td>MGF</td>
<td>Moment Generating Function</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple Input Multiple Output</td>
</tr>
<tr>
<td>mmWave</td>
<td>Millimeter Wave</td>
</tr>
<tr>
<td>MSP</td>
<td>Mean Signal Power</td>
</tr>
<tr>
<td>MU-MIMO</td>
<td>multi-user MIMO</td>
</tr>
<tr>
<td>NOMA</td>
<td>Non-Orthogonal Multiple Access</td>
</tr>
<tr>
<td>Symbol</td>
<td>Meaning</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
</tr>
<tr>
<td>OMA</td>
<td>Orthogonal Multiple Access</td>
</tr>
<tr>
<td>OSSA</td>
<td>Opportunistic Secure Spectrum Access</td>
</tr>
<tr>
<td>PDF</td>
<td>Probability Density Function</td>
</tr>
<tr>
<td>PDMA</td>
<td>Pattern Division Multiple Access</td>
</tr>
<tr>
<td>PGFL</td>
<td>Probability Fenerating Functional</td>
</tr>
<tr>
<td>PMF</td>
<td>Probability Mass Function</td>
</tr>
<tr>
<td>PP</td>
<td>Point Process</td>
</tr>
<tr>
<td>PPP</td>
<td>Poisson Point Process</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>RDP</td>
<td>Relative Distance Process</td>
</tr>
<tr>
<td>RI</td>
<td>Residual Intracell interference</td>
</tr>
<tr>
<td>RSI</td>
<td>Residual Self-Interference</td>
</tr>
<tr>
<td>RSS</td>
<td>Radio Signal Strength</td>
</tr>
<tr>
<td>RV</td>
<td>Random Variable</td>
</tr>
<tr>
<td>SCP</td>
<td>Spatially averaged Coverage Probability</td>
</tr>
<tr>
<td>SI</td>
<td>Self-Interference</td>
</tr>
<tr>
<td>SIC</td>
<td>Self-Interference Cancellation</td>
</tr>
<tr>
<td>SuI</td>
<td>Successive Interference</td>
</tr>
<tr>
<td>SuIC</td>
<td>Successive Interference Cancellation</td>
</tr>
<tr>
<td>SINR</td>
<td>Signal to Interference and Noise Ratio</td>
</tr>
<tr>
<td>SIR</td>
<td>Signal to Interference Ratio</td>
</tr>
<tr>
<td>TDMA</td>
<td>Time Division Multiple Access</td>
</tr>
<tr>
<td>TECR</td>
<td>Total Effective Cell Rate</td>
</tr>
<tr>
<td>TMT</td>
<td>Threshold Minimum Throughput</td>
</tr>
<tr>
<td>tEV</td>
<td>Typical Eavesdropper</td>
</tr>
<tr>
<td>tUE</td>
<td>Typical UE</td>
</tr>
<tr>
<td>UE</td>
<td>User Equipment</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

1.1 Downlink guaranteed-rates per UE for different levels of intercell interference in a NOMA setup with $N = 2$. Here, the guaranteed-rate is the minimum effective-rate of a UE in the setup, where the effective-rate of a UE is the target-rate times the probability that no outage occurs. 45

3.1 Transmission Links and Modes. 68
3.2 The shaded crescent (with area $A$) which represents the region where a closer BS to the r-D2D may exist. 73
3.3 CDF of the distance between the r-D2D UE and its closest BS, $r_e$, for $\omega = 1$ and different BS intensities. 75
3.4 SINR Outage vs. $\theta$ for $\rho_c = -80$dBm, $r_1 = 1$ and $T_d = 0.2$ and different $r_2$. 88
3.5 SINR Outage vs. $\theta$ for $\rho_c = -80$dBm, $r_2 = 1$, and different $r_1$ and $T_d$. 88
3.6 $T_n$ gain w.r.t the traditional cellular network vs. $T_d$ with $r_1 = 0.2$ and $r_2 = 0.2$. 90
3.7 $P_{avg}$ vs. $T_d$ with $r_1 = 0.2$ and $r_2 = 0.2$. 90
3.8 SINR Outage vs. $T_d$ for $\theta = 1$ with $r_1 = 0.2$ and $r_2 = 0.2$. 90
3.9 $\mathcal{O}_{net}$ vs. $T_d$ for $\theta = 1$ with $r_1 = 1$ and $r_2 = 2$. 90
3.10 $T_{avg}$ vs. $r_2$ with $r_1 = 0.01$ and $T_d = 1$. 92
3.11 Truncation and SINR Outages vs. $T_d$ with $r_1 = 0.2$ and $r_2 = 0.2$. 92
3.12 SINR Outage vs. $\theta$ with $T_d = 0.2$, $r_1 = 0.2$, $r_2 = 0.2$, and different $\zeta$. 94
3.13 $T_n$ vs. $T_d$ with $r_1 = 1$, $r_2 = 2$, and different $\zeta$. 94
3.14 $T_n$ vs. $T_d$ with $r_1 = r_2 = 0.2$ and different $\zeta$. 94
3.15 $T_n$ vs. $T_d$ with $r_1 = r_2 = 0.2$ and different $\omega$. 94
4.1 Probabilities vs. $\tau_1$ (and $\tau_2$) at $v = \frac{1}{2\sqrt{\lambda}}$ for the network with jamming. Solid lines show the analysis when tUE employs FDJ ($\nu = -90$ dB) and dash-dotted lines show when it does not. Markers show the Monte Carlo simulations. 103
4.2 Probabilities vs. $v$ without FDJ at the tUE. Ignoring correlation results in $\mathbb{P}(S \mid F) = \mathbb{P}(S)$. 103
4.3 Probabilities vs. $\nu$ with $\nu = -90$ dB. Ignoring correlation results in $P(S \mid F) = P(S)$. Solid lines represent the tUE employing FDJ and dash-dotted lines represent the tUE without FDJ.

5.1 Rate utility of NOMA for different scenarios from Table 5.1. UEs are labeled from the strongest channel (i.e., UE1) to the weakest (i.e., UE6).

5.2 A cell of interest in a downlink NOMA setup with $N = 2$ and two UE clusters operating on different frequency channels. The intended signal and interferences experienced at the UEs of the cluster operating on frequency channel 1, and intended signals at the UEs of the cluster on frequency channel 2 are shown.

5.3 UE order in a two-UE NOMA cellular network, where squares, green circles, and red circles denote the BSs, stronger UEs, and weaker UEs. The Voronoi tessellation represents the coverage region of each BS and the dotted links represents association. The BSs affected by interference-aware sorting are highlighted in blue.

5.4 Power allocation for UEs in a downlink NOMA setup with $N = 2$ in the intercell interference aware and agnostic cases.

5.5 Total and individual UE effective-rates vs. target-rate for uplink-NOMA with $N = 2$. Dashed lines represent full-power transmission, solid lines represent power allocation based on attempted fairness when intercell interference is overestimated.

5.6 Potentials of integrating NOMA with 5G technologies for $N = 2$ case.

5.7 TECR vs. target-rate for FD-NOMA with different interference suppression when $N = 2$.

5.8 TECR vs. target-rate for forward-D2D NOMA with $N = 2$ and different interference suppression.

6.1 A realization of the network with $N = 3$ for Models 1 and 2. The UEs, in-disk (dashed circle) and sector (shaded) for the cell at $\mathbf{o}$ are shown.

6.2 Rate region for NOMA and TDMA with MSP-based UE ordering for Model 1 using different $\beta$ and $N = 2$.

6.3 Rate region for NOMA and TDMA with MSP and ISINR-based UE ordering for Model 1 with $\beta = 0$ and $N = 2$.

6.4 Optimum cell sum rate and individual UE throughputs vs. $P_1$ for NOMA with MSP and ISINR-based ordering for Model 1 with $\beta = 0$ and $N = 2$. 
6.5  SINR coverage vs. $\theta$ (identical transmission rate for all UEs) for Model 1 with $N = 3$ employing the fixed power allocation $P_1 = 1/6$, $P_2 = 1/3$ and $P_3 = 1/2$. Solid (dashed) lines show the analysis for ISINR-based (MSP-based) UE ordering. Markers show the Monte Carlo simulations.

6.6  $R_{\text{tot}}$ vs. $N$ for Model 1 using OMA and NOMA with $\mathcal{T} = 0.3$ and different $\beta$ values. Black lines are for MSP-based UE ordering, while red lines are for ISINR-based ordering. The curves end at the largest $N$ that can be supported given the TMT constraint and $\beta$.

6.7  Comparison of individual NOMA UE power and throughputs between ISINR- and MSP-based UE ordering for $N = 3$ and $N = 6$ using Model 1 with $\beta = 0$ and $\mathcal{T} = 0.3$.

6.8  Individual UE throughput and cell sum rate vs. $N$ using Model 1 with $\beta = 0$ for: $\mathcal{P}1$ with $\mathcal{T} = 0.3$ and $\mathcal{P}2$. Black lines are for MSP-based UE ordering, while red lines are for ISINR-based. The blue line is the TMT achieved by UE$_2$, ..., UE$_N$ in $\mathcal{P}1$.

6.9  $R_{\text{tot}}$ vs. $\beta$ for Model 1 with different values of $N$ and $\mathcal{T}$ using MSP-based ordering. Curves represent NOMA and horizontal lines (independent of $\beta$) represent OMA.

6.10  Rate region for NOMA with MSP-based UE ordering, $\beta = 0$ and $N = 2$ for Models 1, 2, and 3.

6.11  $R_{\text{tot}}$ vs. $N$ for $\mathcal{T} = 0.3$ using NOMA and $\beta = 0$ for MSP-based ordering. The curves end at the largest $N$ that can be supported given $\mathcal{T}$ and $\beta$.

6.12  Complexity vs. $N$ for NOMA UEs selected according to Model 1.

7.1  Meta distribution vs. $\alpha$ with $\theta_1 = 1$ and $\theta_2 = 0.5$ for the E-NOMA and C-NOMA scheme. Solid lines represent $P_1 = 0.5$ and dashed lines represent $P_1 = 0.1$. Markers show the Monte Carlo simulations.

7.2  SCP and variance of the meta distribution vs. $\theta$ (identical target rate for all UEs) with $P_1 = 1/3$ for the C-NOMA scheme using the exact and approximate moments of coverage probability.

7.3  SCP and variance of the meta distribution vs. $\theta$ (identical target rate for all UEs) with $P_1 = 1/3$ for the E-NOMA and C-NOMA scheme. Solid lines are for $\beta = 0$, dashed are for $\beta = 0.1$, and dash-dot are for $\beta = 0.2$. Note: the weakest NOMA UE is unaffected by $\beta$. 

7.4 SCP and variance of the meta distribution vs. $\theta_1$. For C-NOMA:

- TMT=0.1 (black) uses $P_2 = 0.18$ and $\theta_2 = -9$ dB, TMT=0.4 (blue) uses $P_2 = 0.54$ and $\theta_2 = -0.7$ dB.

For E-NOMA: TMT=0.1 (red) uses $P_2 = 0.47$ and $\theta_2 = -7$ dB.
LIST OF TABLES

3.1 TABLE I: List of Symbols ............................................. 68
3.2 Mode of Operation Parameters ...................................... 84
3.3 Parameter Values ....................................................... 85
3.4 Effect of increasing $r_1$, $r_2$, and $T_d$. .......................... 85

5.1 Summary of different utilities and objectives. We denote the vector of powers allocated to transmissions by $\mathbf{P}$, and SINRs associated with target-rates by $\theta$, with $\theta = (\theta_1, \ldots, \theta_N)$. ................................. 109
Chapter 1

Introduction

1.1 Background

Due to the ever-growing number of mobile devices, increased trend in the usage of applications, and more data hungry applications, there is a constant need for increased data rates. It is predicted that there will be a thousand-fold increase in mobile traffic by the next decade. According to a number of predictions, such as [1][2], the number of connected devices was forecasted to exceed 50 billion by 2020. These sources have more recently revised their predictions to more conservative ones of 28 and 30 billion, respectively. Irrespective of the exact figures, it is clear that the data traffic and demands of connected devices is growing very rapidly. To meet the required data rates, improving the currently used technologies to make them more efficient will just not be enough as the currently employed 4G has almost reached the theoretical limit on data rate [3]. New technologies need to be employed to meet the demands of the future and hence researchers have started investigating new 5G techniques. The 5G network ought to achieve 1000 times the system capacity, 10 times the spectral efficiency, energy efficiency and data rate, 25 times the average cell throughput, 10 to 100 times higher number of connected devices, and 5 times reduced end-to-end latency [3][4]. The goal is to connect anyone/anything and achieve seamless communication anywhere, at any time.

The amount of spectrum available to utilize for cellular communication, however, is still limited. Due to the scarcity of this resource, technologies employed to im-
prove communication efficiency often focus on reusing the available spectrum more efficiently. Consequently, 5G requires new techniques that are able to reuse the spectrum as well. It ought to be mentioned that another direction that 5G techniques are taking is the deployment of millimeter wave (mmWave) communication. Currently cellular communication takes place in the microwave spectrum (300 MHz-3 GHz). On the other hand, mmWave communication focuses on communicating in the underutilized higher frequency range, namely 30-300 GHz. It should be noted, however, that mmWave communication comes with its own set of challenges that researchers are working on addressing; we do not delve into these as the focus of this dissertation is on microwave communication. In addition to this, 5G will also make use of techniques such as massive multiple-input multiple-output (MIMO), mobile femtocells, and the deployment of heterogeneous networks (HetNets) to achieve its ambitious targets.

The focus of our dissertation is on the new techniques of 5G that reuse the available microwave spectrum. We study these techniques in the context of performance analysis in terms of improving spectral efficiency and data rates. We also study them in terms of designing the network and resource allocation. In addition to this, we address the deployment of such techniques in the context of improving physical layer secrecy.

A technique to be deployed in 5G that focuses on aggressive spectrum reuse is D2D communication where users lying close by are able to bypass the base-station (BS) and communicate directly. This when done on the same spectrum as the cellular links, referred to as underlay D2D, increases spectrum reuse. Another technique which has only recently become possible is FD communication where the transmitter and receiver transmit using the same frequency resources at the same time. NOMA is also a recently emerging technique to be deployed in 5G where users share resources (i.e., multiple access) in the power domain. D2D, FD, and NOMA techniques will be discussed in greater detail in Section 1.3.
With the increased usage of wireless communication in various applications such as banking services, health, etc., the need for secure information transmission is also increasing. The broadcast nature of wireless communication makes security a major concern due to its susceptibility to eavesdropping. Traditionally security measures in the network layer such as cryptography were used for protecting the transmitted information. However, the trend is now shifting to additionally providing physical layer security due to the large sizes and dynamic nature of modern wireless networks. This leads to the notion of examining the physical properties of the wireless channel to enhance security, and studying the impact of network properties such as interference and spatial distribution of nodes on secure communication. Currently, physical layer secrecy is studied assuming independent interference at the legitimate receiver and eavesdropper. Part of this dissertation will focus on the impact that this interference correlation has on the achievable secure communication in a large wireless network.

1.2 Motivation behind Stochastic Geometry for Large Wireless Networks

In order to deploy new technologies, there is a need to first study the impact that they will have. Previously, analysis for any new technique was done at a small scale such as a single-cell and a few users, for e.g. Although this kind of analysis led to accurate expressions for the setup, it is unrealistic. This is because the new technologies to be incorporated into 5G, particularly, impact the entire network more significantly rather than just a few links. Additionally, with the increased network traffic and deployment of small cells, studying just a few links may not necessarily give meaningful results on the effect of incorporating a new technique and, in fact, may even be misleading. Thus, we are interested in the impact that a technology/technique has on the performance of large scale networks.

Studying the effects on large scale networks is obviously difficult due to the com-
plexity of the analysis involved. In the past, two techniques have been used to analyze large networks; namely, the hexagonal grid model approach and by using system level simulations. The hexagonal grid model approach assumes the cellular network is composed of a grid of hexagons representing each cell with BSs placed at the center of each cell. This is an unrealistic model [5,6] since real networks have variations in the coverage area of cells due to variations in capacity demand at different locations. Additionally, it may not be feasible to deploy BSs at certain locations in the real world (e.g. in a river, on top of a building, etc.). Hence, the idealized grid model does not hold due to randomness in the locations of the BSs and gives inaccurate results. The hexagonal grid model does not provide tractable results for intercell interference either, thereby making it an unappealing model for studying performance metrics of the network. On the other hand, using system level simulations is very expensive as extremely large and complex simulations are required to model the network [7].

Stochastic geometry is a very powerful mathematical tool for modeling random patterns of points. Relatively recently, stochastic geometry techniques have been applied to model cellular networks as they not only capture the topological randomness in the network geometry but also because they lead to tractable and accurate analysis [5,8,13]. Additionally, the results obtained are independent of the topology of the network and are general in terms of network parameters. Hence, we use stochastic geometry tools to model, analyze, and design large cellular/ad-hoc wireless networks with random topologies.

1.3 Technologies to be incorporated into 5G and Related Work on Them

In this section we present an overview on three efficient spectrum reuse technologies to be incorporated into 5G that are studied in this dissertation.
1.3.1 Device-to-Device (D2D)

Recently, the concept of D2D communication has emerged to cater the ever-increasing demand of data rate and system capacity. D2D enables UEs lying in close proximity, that intend to exchange information, to bypass the cellular BS and communicate in a peer-to-peer fashion. It is envisioned that D2D communication between UEs can improve the network performance in terms of spatial frequency reuse, latency, and energy consumption [14]. In other words, D2D communication replaces the conventional two-hop cellular link by a short-range, low-power, and direct D2D link which will improve the latency and power consumption. Further, while the conventional cellular association dictates a single link per channel per cell (thereby prohibiting intracell interference), D2D links aggressively reuse the same channel over the spatial domain with no restriction over the cell boundaries, which increases the spatial frequency reuse. Consequently, D2D communication introduces a new type of interference between the cellular mode and D2D mode UEs, which is denoted as cross-mode interference. The envisioned D2D gains come at the expense of more complicated network management. Choosing the mode of operation (i.e., cellular mode or D2D mode) and neighbor discovery are two new network functions that arise with D2D communication. Further, interference management between cellular and D2D links also needs to be taken into account. Despite the increased interference level imposed by D2D communication, non-trivial gains can be harvested if efficient interference coordination between D2D and cellular links is adopted [15] [20]. In addition to improving the spatial spectral utilization, D2D can potentially bring other performance gains for cellular networks, namely, lower power consumption, higher network capacity, and lower communication latency [21] [25].

Several efforts have been invested in the literature to model D2D communication

\(^1\)Cross-mode interference is used to denote both the cellular-to-D2D interference and the D2D-to-cellular interference
in cellular networks. The operation of HD-D2D communication is studied in [15–27] and promising performance gains are reported. The authors of [17] ensure signal-to-interference-and-noise ratio (SINR) violation for cellular users due to D2D interferers is kept below a threshold, using a simple power control method. The analysis, however, is limited to one user, one cell, and one D2D link. The works in [15, 16, 18, 20, 26, 27] use stochastic geometry, while the remaining are surveys or study single-cell setups. In [18], given an interference threshold to the cellular network, the maximum intensity of D2D devices possible in an uplink cellular network is found. The optimal intensity and transmit power, that maximize capacity for a D2D enabled two-tier uplink network with outage probability constraints is found in [19]. Mode selection and power control are not considered in [18] and [19], and the D2D link distances are assumed to be fixed. D2D distance-based mode selection criteria has been employed in [26] and [15]. A biased distance-based mode selection scheme for underlay is considered in [27], which gives the flexibility to control the amount of traffic offloaded from the cellular to the D2D mode.

1.3.2 Full Duplex (FD)

FD communication involves a transceiver communicating such that it is simultaneously transmitting and receiving on the same frequency channel. In the literature this has been, rightfully, compared to “trying to hear a whisper while shouting at the top of your lungs” [28]. This is due to the overwhelming self-interference (SI) dominating the signal at the receive antenna from the message being transmitted by the transceiver itself. Consequently, FD communication was thought to not be possible for the longest time [29]. However, with recent advances in SI cancellation techniques, that mitigate the overwhelming SI, simultaneous transmission and reception on the same time-frequency resource block has become possible; as a result, FD communication has become a reality [30,31].
FD communication at a transceiver can be realized in two ways: 1) Bi-directional FD transmissions, and 2) the FD 3-node topology \cite{32-34}. As the name suggests, bidirectional FD involves a pair of nodes transmitting to and receiving from each other in the same time-frequency resource block. On the other hand, in the FD 3-node topology, the transceiver that communicates via FD transmits to one HD node and receives from another HD node simultaneously on the same frequency channel.

As FD communication enables simultaneous transmission and reception on a link, it was envisioned that the technology has the potential to double the throughput. However, this is far from true as FD communication also dramatically increases the network interference. Accordingly, several studies for FD communication in large networks have been conducted via stochastic geometry \cite{35-39} to include the impact of the increased network interference. In \cite{35-38}, FD communication is employed at the cellular link. When FD is employed at the cellular link, high rate improvement is observed for the downlink \cite{35,36,39}. However, the authors in \cite{37,38} show that the downlink rate improvement may come at the expense of high degradation in the uplink rate due to the high disparity between the uplink and downlink transmit powers. An adhoc network is considered in \cite{39} and the authors report that FD communication offers rate improvement in both the forward and the reverse links when both have equivalent transmit powers.

### 1.3.3 Non-Orthogonal Multiple Access (NOMA)

NOMA is promoted as one of the promising technologies to be adopted in 5G cellular networks. In contrast to conventional orthogonal multiple access (OMA), NOMA is foreseen to increase the network capacity by improving the spatial utilization of the scarce spectrum. Conventionally, temporal, spectral, and/or spatial\textsuperscript{2} orthogonalization are adopted to avoid intracell interference (i.e., interference among UEs in

\textsuperscript{2}Spatial separation of UEs with MIMO can be used with either OMA or NOMA.
the same cell) in OMA. This permits only one UE to be served per time-frequency resource-block per BS in OMA. In NOMA cellular networks, on the other hand, multiple UEs can be clustered and simultaneously served over the same time-frequency resource-block. Particularly, NOMA UEs share the same time-frequency resources by either having their messages superposed in the power domain or in the code domain. NOMA is therefore a special case of superposition coding. The mutual intracell interference among the clustered UEs is managed by proper resource allocation (e.g., transmission power and rate) in conjunction with successive interference cancellation (SuIC). Decoding techniques using SuIC for multiple-access channels have been studied from an information-theoretic perspective for several decades, and they were implemented on a software radio platform. Consequently, UE clustering and resource allocation are crucial for NOMA operation in order to minimize the effect of intracell interference.

Motivated by its potential to increase the capacity of cellular networks, the design of NOMA transmission via UE clustering and resource allocation has received much attention from the research community. For instance, the importance of power allocation for achieving transmission rate gains in a single-cell NOMA downlink is emphasized. The effect of power allocation on transmission rate fairness among NOMA UEs is manifested. UE clustering is considered where selection strategies are proposed based on the distance between the UEs and the BS. In, it is shown that having NOMA UEs with more distinct channels enhances NOMA gains in a single-cell downlink scenario. The coexistence of a cognitive secondary NOMA system with a primary OMA network is studied. However, the study is limited to a single secondary NOMA BS.

Promising results for NOMA as an efficient spectrum reuse technique have been shown. In power allocation schemes are investigated for universal

---

3There are other NOMA schemes such as pattern division multiple access (PDMA), bit division multiple access (BDMA), and interleave-division multiple access (IDMA).
fairness by achieving identical rates for NOMA UEs. The idea of cooperative NOMA is investigated in [46,49]. Most NOMA works order UEs based on either their distance from the transmitting BS [49,53,55] or on the quality of the transmission channel [46,48,51,56]. A number of works in the literature focus on resource allocation [48,53,56,58]. Resource allocation schemes for maximizing rates with constraints often focus on small NOMA clusters such as the two-user case [53,57,58], though works such as [48,56] consider a general number of UEs in a NOMA cluster.

The works in [46–49, 52, 53, 56–58] consider NOMA in a single-cell and therefore do not account for intercell interference which has a drastic negative impact on the NOMA performance as shown in [59]. Using stochastic geometry-based modeling, a large uplink NOMA network is studied in [55,60], a large downlink NOMA network in [60,61], and a qualitative study on NOMA in large networks is carried out in [59]. However, [61] does not take into account the SuIC chain in the SINR analysis, which overestimates coverage. In [60], two-user NOMA with fixed power allocation is studied. In the downlink, comparisons are made between random UE selection and selection such that the weaker UE’s channel quality is below a threshold and the stronger UE’s is greater than a second threshold. It ought to be mentioned that fixed resource allocation does not allow the system to meet a defined cluster objective and makes a comparison with other schemes such as OMA unfair. The UEs in [55] are modeled using a cluster-process, a fixed number of UEs is assumed in a cluster and constant transmission power is considered.

In this dissertation we focus on power domain NOMA, i.e., where messages are superposed in the power domain. This form of NOMA allows multiple UEs to transmit/receive messages in the same time-frequency resource block by transmitting (receiving) them at different power levels in the downlink (uplink). SuIC techniques are then used for decoding.

With NOMA come a number of challenges, including:
1. Determining the size of the UE cluster, i.e., the number of UEs to be served by a BS.

2. Determining which UEs to include in a cluster, referred to as UE clustering.

3. Ordering UEs within a cluster according to some measure of link quality.

4. The objective of the cluster — prioritizing individual UE performance, total cluster performance, or a middle ground between the two.

5. Resource allocation for the UEs in a cluster.

1.4 Integrating Efficient Spectrum Reuse Technologies

In addition to the standalone D2D, FD, or NOMA networks, the integration of these technologies with other 5G candidate technologies, including but not limited to each other, are also studied in the literature. Such studies are important to realize the 5G network, which is foreseen to adopt several new transmission schemes.

The performance gains offered by D2D communication alone for instance are not sufficient to achieve the ambitious performance requirement defined for 5G networks \[62\]. Therefore, it is believed that the foreseen 5G performance will be fulfilled by integrating several new technologies to the state of the art cellular system \[63\]. This implicitly requires more aggressive usage for the spectrum and more sophisticated interference management techniques. For instance, it is shown in \[64\] that integrating massive MIMO with D2D communication increases the cellular transmissions per cell and mitigates the D2D-to-cellular interference, at the expense of increasing the cellular-to-D2D interference. This improves the network spectral efficiency while providing sufficient protection to the conventional cellular users. However, the channel state information between all D2D transmitters and all antennas is required, which imposes high control burden to the system.
In this dissertation we will focus explicitly on the combinations of D2D and FD referred to as FD-D2D, D2D and NOMA referred to as D2D-NOMA, and FD and NOMA referred to as FD-NOMA.

### 1.4.1 FD-D2D

To avoid any ambiguities, we first define FD-D2D as the simultaneous transmission and reception to and from a D2D UE on one frequency channel. Consequently, FD-D2D scenarios can be as follows: 1) Transmission and reception on the same time-frequency resource block between two D2D UEs; in this case the D2D pair employs FD-D2D. 2) A D2D UE transmits to a second D2D UE and receives from a BS or a third D2D UE at the same time on the same frequency channel; in this case only one D2D UE employs FD-D2D. 3) A D2D UE receives from a second D2D UE and transmits to a BS or a third D2D UE at the same time on the same frequency channel; in this case only one D2D UE employs FD-D2D. It should be noted that the first scenario is the bidirectional FD mentioned in Section 1.3.2 while the second and third follow the FD 3-node topology. Our work in Chapter 3 employs FD-D2D on D2D pairs along the lines of the first scenario.

In-band FD communication is an appealing technology to integrate with D2D communication to further improve the spectral efficiency and network throughput for several reasons. One motivation behind this is the aggressive spectrum reuse that such an integration will allow. Additionally, D2D links operating on FD will have comparable transmit powers; this is in contrast to employing FD communication on the cellular uplink and downlink where there is a significant imbalance between the transmit powers. Moreover, the model in [39] establishes the superiority of FD communication in a large adhoc network; since the distribution of D2D users is often modeled as a collection of adhoc nodes, the work in [39] motivates implementing FD communication to D2D links.
Recent studies on a single-cell and single D2D link scenario have shown that FD D2D communication provides significant improvement in the spectral efficiency (up to 100%) over conventional half-duplex (HD) D2D if sufficient self-interference cancellation (SIC) is achieved. These studies also emphasize the importance of cross-mode interference coordination in order to harvest the FD-D2D gains. However, the conclusions in these studies cannot be directly generalized to large scale networks due to the induced interference by the FD link into the network. To validate the FD-D2D benefits, explicit studies for its operation in realistic large-scale setups is required.

Since each transceiver can simultaneously transmit and receive on the same channel, FD-D2D communication activates two transmitters per D2D-link. Therefore, from a large-scale perspective where channels are reused over the spatial domain, FD-D2D communication can significantly increase the interference associated with D2D communication when compared to its HD counterpart. Note that cross-mode interference is already a performance limiting parameter for HD-D2D communication in cellular networks. Hence, it is hard to predict whether FD communication would improve or diminish the D2D gains due to the imposed interference. Therefore, explicit studies for the FD-D2D effect on the aggregate interference in large cellular networks are required. In this context, stochastic geometry provides a powerful mathematical tool that can be exploited to characterize the impact of interference associated with FD-D2D communication.

At the time of our work on FD-D2D communication in Chapter 3, studies on large-scale setups in this context did not exist. However, there are a number of works available now that utilize stochastic geometry to characterize the impact of interference associated with FD-D2D communication in a large network. In an analytical framework to investigate the performance of FD-D2D is studied using a cluster process to model devices that cache popular content. The work in considers a large cellular network overlaid with FD-D2D; different from other
related works the authors study the network so that a D2D UE pairs with a D2D transmitter that is its \( n \)th nearest neighboring UE. The performance of FD-D2D is studied in a two-tier network in [71] and it is shown that by appropriately allocating different proportions of users into different transmission modes performance gains can be achieved. A study on D2D UEs with FD capability that coexist with LTE small cells in unlicensed bands is conducted in [72].

1.4.2 FD-NOMA

FD-NOMA involves the integration of NOMA and FD which can be of the following forms, having the downlink and uplink on the same channel simultaneously (i.e., FD): 1) NOMA in the downlink and OMA in the uplink, 2) OMA in the downlink and NOMA in the uplink, 3) NOMA in both the uplink and downlink. Additionally, it may involve different users/clusters (depending on whether OMA/NOMA are being used, respectively) for the uplink and downlink, i.e., in a FD 3-node topology fashion, or the same users/clusters in both the uplink and downlink, i.e., in a bidirectional FD fashion. Further, a lot of work on FD-NOMA has focused on cooperative NOMA mainly where a near user or another node employ FD to serve as a cooperative relay to the far user.

The works in [73–76] study FD-NOMA in single-cell setups where FD relaying is used to assist a far user in a cooperative manner. In [73] a two-user scenario is considered where the near user acts as a relay to forward the message of the far user using either FD or HD. The impact of having and not having a direct link from the BS to the far user is studied. In [74] a setup is considered where no direct link exists between the BS and far user. The BS transmits both messages superimposed according to NOMA and the near user decodes them. The near user behaves as a FD relay and transmits the relevant message to the weaker user. The authors obtain the optimum NOMA power allocation required to minimize outage.
integrated NOMA and FD transmission scheme is proposed in [75] to reduce relaying delay and improve the end to end throughput. The BS broadcasts the superimposed message for the strong and weak users which is decoded by the strong user and a FD relay. No direct link between the BS and the weak user exists; the weak user’s message is decoded by the relay and forwarded to it in a future time slot using FD. The authors in [76] exploit FD relaying to improve the reliability of NOMA transmissions, in which the NOMA UE with the stronger channel simultaneously receives its own message and relays an older message designated to the UE with the weaker channel. An adaptive multiple access scheme is also studied that switches between the proposed cooperative NOMA, conventional NOMA, and OMA according to the link quality and level of residual self interference.

Some other works on FD-NOMA that do not involve cooperative relaying include [58, 77]. In [77] the authors use FD to realize simultaneous uplink NOMA and downlink OMA in a single-cell setup. A NOMA framework between FD BSs and half-duplex UE clusters in a single-cell setup is proposed in [58], which extends the well known FD 3-node topology to NOMA clusters. The authors are interested in determining the power and subcarrier allocation policy that maximizes the weighted sum throughput of the system. Since obtaining the optimum resource allocation has very high computational complexity, a suboptimum algorithm is proposed to strike a balance between optimality and computational complexity. The work in [78] does not employ FD and NOMA simultaneously; hence, it is not FD-NOMA per se. Instead, the authors propose a selection criterion between NOMA and FD communication that is based on traffic conditions, network density, and self-interference cancellation capabilities.

To the best of our knowledge, no explicit study on FD-NOMA in large networks exists.
1.4.3 D2D-NOMA

D2D-NOMA involves the employment of NOMA between sets of D2D UEs. In this context, all D2D UEs but one form the NOMA cluster, and the remaining non-cluster D2D UE has the role of the BS in cellular NOMA. We refer to the non-cluster D2D UE as the head-UE. D2D-NOMA transmissions can be classified into two: 1) Forward-D2D NOMA, 2) Reverse-D2D NOMA. In forward-D2D NOMA, the head-UE transmits to the other D2D UEs in one time-frequency resource block by superimposing their messages in the power domain in the usual NOMA fashion; this is similar to downlink NOMA. In reverse-D2D NOMA, the head-UE receives from the other D2D UEs in one time-frequency resource block; this is similar to uplink NOMA.

In [79] a single-cell setup is studied where the cellular uplink is overlaid with D2D-NOMA. $M$ cellular UEs transmit to the BS on separate channels, and $N$ D2D groups are considered where a D2D transmitter employs NOMA to transmit to two D2D receivers, i.e., forward-D2D NOMA. A suboptimum algorithm is proposed for subchannel assignment and a non-convex problem for power allocation is solved so that the system sum rate is maximized while satisfying the interference constraints of cellular users. In [80] D2D-NOMA is studied in a large network where a D2D transmitter transmits via NOMA to multiple D2D receivers that are ordered according to their quality of service (QoS) requirements. The authors consider cellular overlaid D2D so that the D2D do not suffer from cellular interference. Two power allocation policies are compared and it is shown that the policy based on the QoS requirements of users outperforms the fixed power allocation policy. A large network with overlay D2D is considered in [81] so that the D2D has its own dedicated spectrum. A D2D device serves two D2D receivers via cooperative hybrid automatic repeat request (HARQ) assisted NOMA. The authors take into account both the spatial interference correlation at the NOMA receivers as well as the temporal in-
interference correlation across HARQ transmissions and show that overlooking these correlations highly overestimates performance resulting in misleading insights.

A single-cell setup is considered in [82] where a downlink cellular NOMA network is underlaid with D2D. Multiple subchannels are considered and a subchannel is underlaid with at most one D2D pair. The goal is to maximize the sum D2D throughput subject to a minimum performance threshold for cellular communication. Accordingly, optimal power control and channel assignment for the D2D pairs is determined. The coexistence between NOMA-based cellular multi-user MIMO (MU-MIMO) and OMA D2D network is studied in [83] in a single-cell setup. A suboptimum algorithm to maximize the system sum throughput subject to a minimum rate constraint on the NOMA users is proposed. Note that although the works in [82, 83] involve NOMA and D2D, they do not employ NOMA at the D2D users; hence, they are not D2D-NOMA by our definition.

1.5 On Physical Layer Security

As mentioned before, with the wide range of applications of wireless communication, secure information transmission has become very important. The focus now is on physical layer security. Various works have been done that focus on exploiting the nature of the wireless network to enhance security [84–94]. In our work on physical layer security, we are particularly interested in the impact that interference has on achievable secure communication.

Until recently, network interference has been considered an undesirable performance limiting parameter for wireless communication. This conception is changing with the emerging requirement for physical layer security where interference may help in increasing transmission secrecy [91, 94]. Secure information transmission opportunities on the legitimate link arise from the random fluctuation of interference power at the receiver and eavesdropper. Particularly, exploiting the time intervals where the
receiver (eavesdropper) experience low (high) interference powers, the transmitter can send information at a rate that is higher than the capacity of the eavesdropper link. This leads to the event of opportunistic secure spectrum access (OSSA) in which the eavesdropper cannot decode the message while the legitimate receiver can; transmission secrecy is therefore attained on the legitimate link.

The effect of interference on transmission secrecy is mainly studied and quantified assuming independent interference at the legitimate-receiver and eavesdropper [91–94]. However, since the interference experienced at the receiver and eavesdropper arises from the same set of transmitters, the network interference is spatially correlated [95, 96]. Studying the impact of correlated interference is important as it may degrade OSSA performance more than that anticipated for independent interference, in which the degradation becomes more prominent as the eavesdropper gets closer to the receiver. Hence, the performance of OSSA should be characterized in terms of the distance between the receiver and eavesdropper. To the best of our knowledge, this problem has not been investigated in the literature. Numerous works have considered the impact of correlated fading between the eavesdropper and receiver in small networks on the achievable secrecy [97,98]; however, our focus is on independent fading but correlated interference in a large network. This case is more practically relevant since the spacing between legitimate users and eavesdroppers certainly exceeds the coherence length of the fading, while the interference is correlated over much larger distances.

1.6 On the Meta Distribution

1.6.1 The Concept and Significance of the Meta Distribution

Until very recently, stochastic geometry based studies have focused on the spatial averages of performance metrics only. The most frequently used metric used in performance analysis is the spatially averaged coverage probability (SCP), which averages
the coverage probability over all fading, activity, and network realizations; the SCP is generally referred to simply as the coverage probability. In particular, the SCP can be understood as the performance of the average/typical link in the network. However, the actual distribution of the performance of majority of the links may not necessarily be very close to the SCP; for instance, a fraction of the links could be much worse than the SCP and another fraction could be much better. As an example, an SCP of 0.7 could be attained by having: 1) each link in the network with a coverage probability of 0.7, 2) 50% of the links with a coverage probability of 0.65 and 50% of the links with a coverage probability of 0.75, 3) 25% of the links with a coverage probability of 0.1 and 75% with a coverage probability of 0.9, which are three very different network scenarios with the same SCP. Incontestably, spatial averages do not reveal the percentile information of performance metrics that are used as an indication of the quality of service that the network provides.

Network operators and vendors, however, are generally interested in the percentile performance of UEs, where the fading and activity change while the network realization is kept constant. This need for more fine-grained information on the performance leads to the notion of studying the distribution of the conditional coverage probability (CCP) which is the coverage probability given a fixed network realization [99]. For a fixed, yet arbitrary, realization of the network we can write the CCP at a receiver as the probability that the signal to interference ratio (SIR) exceeds some threshold $\theta$ as follows:

$$P_c(\theta) = \mathbb{P}(\text{SIR} > \theta | \Phi).$$

Here, the CCP is taken over the fading and activity (if any) of the interferers. In this context, [99] defined the meta distribution as the complementary cumulative density
function (CCDF) of the CCP, i.e.,

\[ \bar{F}_{P_c}(\alpha) \triangleq \mathbb{P}(P_c(\theta) > \alpha), \quad 0 \leq \alpha \leq 1. \]

The \( b^{th} \) moment of the CCP, denoted by \( \mathcal{M}_b \), can be obtained using the meta distribution as follows:

\[ \mathcal{M}_b = \mathbb{E}[P_c(\theta)^b] = \int_0^1 b\alpha^{b-1}\bar{F}_{P_c}(\alpha)d\alpha. \]

It should be noted that by definition, the SCP is the first moment of the CCP. Hence, it can also be calculated using the meta distribution as

\[ \mathbb{P}(\text{SIR} > \theta) = \mathcal{M}_1 = \mathbb{E}[P_c(\theta)] = \int_0^1 \bar{F}_{P_c}(\alpha)d\alpha. \]

### 1.6.2 Calculating the Meta Distribution

Clearly, the meta distribution is a very powerful metric as it reveals the percentile performance of links across an arbitrary network realization. However, calculating the meta distribution in a closed-form is difficult.

Let \( i^2 = -1 \), then the characteristic function of a random variable (RV) \( X \) is defined as \( \phi_X(t) = \mathbb{E}[e^{itX}] \). The characteristic function of the RV \( \ln P_c(\theta) \) can thus be written as

\[ \phi_{\ln P_c(\theta)}(t) = \mathbb{E}[e^{it\ln P_c(\theta)}] = \mathbb{E}[P_c(\theta)^it] = \mathcal{M}_{it}. \]

The standard method to calculate the meta distribution was presented in [99] and uses the Gil-Pelaez theorem in [100]. According to the theorem, the CCDF of the RV
\( \ln \mathcal{P}_c(\theta) \) is given as

\[
\bar{F}_{\ln \mathcal{P}_c(\theta)}(\alpha) = \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \frac{\mathcal{I}(e^{-it\alpha} \mathcal{M}_{it})}{t} dt,
\]

where \( \mathcal{I}(z) \) is the imaginary part of the complex number \( z \). Since \( \mathbb{P}(\mathcal{P}_c(\theta) > \alpha) = \mathbb{P}(\ln \mathcal{P}_c(\theta) > \ln \alpha) \), the meta distribution can be written as

\[
\bar{F}_{\mathcal{P}_c(\theta)}(\alpha) = \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \frac{\mathcal{I}(e^{-it\ln \alpha} \mathcal{M}_{it})}{t} dt.
\]

This expression for the meta distribution is exact; however, the approach requires the integration of imaginary variables and a careful selection of the range of the numerical integration. Additionally, there is no simple way to bound the error of the resulting approximation \([101]\). Consequently, we turn to the moments of the CCP, \( \mathcal{M}_b \), in order to approximate the meta distribution. To the best of our knowledge, only the following three approaches to approximate the meta distribution using the moments of the CCP exist in the literature currently:

- The beta distribution approach: In \([99]\), using moment matching, the beta distribution was proposed as a very accurate approximation for the meta distribution. This approach requires the first two moments of the CCP only and is currently the most popular approach, though it should be mentioned that the literature on the meta distribution is currently limited.

- The truncated Fourier-Jacobi expansion approach: In \([102]\), a general methodology to reconstruct the meta distribution from the moments of the CCP is proposed using the truncated Fourier-Jacobi expansion. The authors motivate this approach as it provides better accuracy than the beta distribution approximation and can be applied to more general networks such as uplink networks and cellular networks underlaid with D2D.
• The binomial mixtures approach: Another approach to approximate the meta
distribution using binomial mixtures is proposed in [101]. It is an efficient
approach as it involves simple linear transforms of the moments of the CCP
and the only parameter required is selecting the number of moments to be
used.

In our work on the meta distribution in Chapter 7, we use the first approach
proposed in [99] which requires only the first two moments of the CCP. Our main
motivation behind this is that closed-form expressions for the moments of the CCP
cannot be obtained for one of the network models we study. The other two approaches,
although more accurate, require a larger number of moments (10 are used in [102] and
25 to 400 moments have been demonstrated in [101]), and are therefore inefficient in
the context of our work on the meta distribution.

As the RV $P_c(\theta)$ is a probability and therefore lies in the interval [0,1], the beta
distribution is a natural choice to approximate the distribution of the CCP [99].
Consequently, the CCDF of the CCP, i.e., the meta distribution, can be approximated
using the CCDF of the beta distribution and moment matching. The probability
density function (PDF) of a beta distributed RV $Y$ is given by

$$f_Y(y \mid a, b) = \frac{y^{a-1}(1-y)^{b-1}}{B(a, b)},$$

where $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ is the beta function, and $a, b > 0$ are the shape parameters.
The mean and variance of $Y$ are given by

$$\mathbb{E}[Y] = \frac{a}{a + b} \text{ and } \sigma_Y^2 = \frac{ab}{(a + b)^2(a + b + 1)}.$$
The CCDF of $Y$ is given by

$$
\bar{F}_Y(y \mid a, b) = 1 - \mathcal{I}_y(a, b),
$$

where $\mathcal{I}_y(a, b) = \int_0^y l^{a-1}(1 - l)^{b-1}dl$. In the context of the meta distribution approximation via the beta distribution, $\mathcal{M}_1 = \mathbb{E}[Y]$ and $\mathcal{M}_2 - \mathcal{M}_1^2 = \sigma_Y^2$. Hence, we can approximate the meta distribution using the first two moments of the CCP as follows:

$$
\bar{F}_{P_c}(\alpha) \approx 1 - \mathcal{I}_\alpha \left( \frac{\tilde{\beta}\mathcal{M}_1}{1 - \mathcal{M}_1}, \tilde{\beta} \right),
$$

where $\tilde{\beta} = \frac{(\mathcal{M}_1 - \mathcal{M}_2) (1 - \mathcal{M}_1)}{\mathcal{M}_2 - \mathcal{M}_1^2}$ and $\mathcal{I}_\alpha(a, b) = \int_0^\alpha l^{a-1}(1 - l)^{b-1}dl$.

1.6.3 Literature on the Meta Distribution

As mentioned, the meta distribution was first formally defined in [99] and the $b^{th}$ moment of the CCP was derived for Poisson networks and downlink cellular networks. The beta distribution as an approximation for the meta distribution was also proposed. The work in [103] considered D2D underlaid with a downlink cellular network. The meta distribution and the mean local delay were studied for both the D2D UE and the cellular UE. In [104] a Poisson bipolar network is considered where the receivers have interference cancellation capability. The moments of the CCP, the meta distribution, and bounds are derived. In [105] the meta distribution of the CCP was studied in uplink cellular networks with with fractional path loss inversion. The authors show that decreasing the activity factor and/or increasing the path loss compensation factor decreases the variation in the SCP, i.e., decreases the discrepancies in the performance of the links. The authors in [106] analyzed the meta distribution in both uplink and downlink cellular networks employing fractional power control. In [107] the meta distribution, spatial outage capacity, and mean delay were studied in millimeter-wave D2D networks. The authors used the SINR for defining the CCP,
instead of the usual SIR used in most meta distribution analysis. Additionally, due to the nature of millimeter-wave communication being different from traditional microwave communication, the beta distribution approximation based on the first two moments of the CCP does not work well; to deal with this, the authors provide a modified approximation using higher moments of the CCP which fits better. In [108] the meta distribution was studied in downlink heterogeneous cellular networks where different BS cooperation techniques were employed. The authors consider the general user as well the worst case user. The meta distribution for uplink and downlink NOMA networks is studied in [109]. While the aforementioned works focus on the meta distribution of the CCP, the authors in [110] study the meta distribution of the secrecy rate. The network considered is comprised of one legitimate link and PPP distributed eavesdroppers; two scenarios are considered: with and without eavesdropper collusion. It ought to be mentioned that the works in [103–110] use the beta distribution to approximate the meta distribution, highlighting its popularity due to its simplicity as well as its accuracy.

Using stochastic geometry, the superiority of NOMA has been established for large-scale interference prone networks [55, 59, 111, 112]. Such studies usually focus on the SCP (which averages the coverage probability over all fading, activity, and network realizations). However, since network operators and vendors are usually more interested in the percentile performance of UEs, where the fading and activity change while the network realization is kept constant, we are interested in the meta distribution of the coverage probability in NOMA networks. The meta distribution for uplink and downlink NOMA networks is studied in [109]. However, the work in [109] focuses on the scheme where NOMA is utilized for UEs located everywhere in the network and does not take into account the joint decoding associated with SuIC.
1.7 Objectives and Contributions

1.7.1 Full Duplex D2D Communication

In Chapter 3, we develop a tractable analytical framework, based on stochastic geometry, for a single tier cellular network underlaid with D2D devices that share the cellular uplink resources and have FD communication enabled. The developed model accounts for a flexible D2D link distance distribution that captures different social interactions between the D2D devices. The UEs have limited transmit powers, employ truncated channel inversion power control, and follow a flexible D2D and FD/HD mode selection criterion. Based on the developed model, the FD-D2D enabled cellular network performance is assessed under perfect and imperfect self-interference assumptions. While imperfect SIC represents a practical operation scenario, perfect SIC has theoretical significance because it shows the explicit contribution of the FD-D2D communication to the aggregate interference level and reveals the subsequent effects on network performance. Different from [15, 18–20, 27, 68] where the cellular network was overlaid with HD-D2D, a cellular network overlaid with FD-D2D is considered in Chapter 3. Additionally, different from [39] where a FD ad-hoc network is considered, the work in Chapter 3 considers FD enabled in an ad-hoc setting (D2D) underlaid with the cellular network. The contributions and findings of Chapter 3 can be summarized as follows:

- The tradeoff, imposed by FD-D2D communication, between increasing the aggregate network interference and improving the spatial frequency reuse is mathematically modeled in terms of outage probability, defined as the probability that the SINR falls below a predefined threshold $\theta$, and the ergodic rate, defined by the seminal Shannon capacity formula.

- We propose a flexible D2D and FD/HD mode selection and power control mechanism to balance the outage probability and spatial spectral efficiency tradeoff
imposed by FD-D2D communication. The proposed mode selection and power control mechanisms are tailored to enable D2D communication, either in FD or HD modes, as long as a certain extent of interference protection (IP) is enforced for cellular users.

- The results show that enforcing FD-D2D communication may highly deteriorate the network performance due to the increased aggregate interference level in the network. On the other hand, non-trivial gains can be harvested from the underlay FD-D2D communication with the proper design of the power control mechanism and D2D FD/HD mode selection criteria. For instance, the results show 64% and 254% spatial spectral efficiency gains harvested by the proposed FD-D2D communication when compared to the HD-D2D communication and conventional (i.e., D2D disabled) cellular network counterparts, respectively.

- The gains that can be obtained by FD-D2D communication in terms of aggregate network throughput, per user throughput, and transmit power reduction are quantified. We also show that there exist optimal values for the design parameters that maximize each of these gains.

- From a mathematical perspective, an accurate approximation for the distance between the D2D-receiver and its closest BS is proposed, which is necessary for modeling and designing the FD-D2D operation.

### 1.7.2 Physical Layer Security

In Chapter 4, we study and quantify the effect of interference correlation on fixed transmission rate OSSA in a large downlink cellular network. We condition on a UE located at the origin $o$, which, under expectation over the point process (PP), becomes the typical UE, denoted as tUE. The eavesdropper closest to the tUE’s serving BS is denoted by tEV. Each UE in the network employs full-duplex jamming (FDJ),
independently of other UEs, by transmitting a jamming signal based on the location of the eavesdropper closest to its serving BS. We let $S$ denote the event that the tUE successfully receives its serving BS’s message and $F$ denote the event that tEV fails to extract information from that message. The OSSA probability is given by $\mathbb{P}(S \cap F)$.

To this end, we explore the effect of FDJ by the tUE on OSSA. The results show that there exists a separation distance between the tUE and tEV below which employing FDJ for the tUE improves OSSA, otherwise FDJ at the tUE does not improve OSSA but increases network interference and power consumption. We also observe that the difference between the OSSA probability when considering and ignoring interference correlation is significant, particularly when the distance between the tUE and tEV is small. This highlights the impact on the achievable secrecy when tEV lies near the tUE; not considering interference correlation would give overly optimistic estimates of the OSSA.

### 1.7.3 NOMA for Large Networks

A number of the aforementioned NOMA studies are myopic in the sense that each BS independently utilizes the local information (e.g., channel state information (CSI), UE locations, target-rates, power constraints, etc.) to perform UE clustering and resource allocation. Such myopic schemes ignore intercell interference (i.e., interference from other cells), which is a fundamental performance limiting parameter in cellular networks. In Chapter 5 we study NOMA for large networks where intercell interference is accounted for. To illustrate the effect of intercell interference on NOMA rate, we show Fig. 1.1 which plots guaranteed versus target UE rates for a two-UE fixed-rate NOMA transmission in a large-scale downlink cellular network. The simulation environment considered to plot Fig. 1.1 is detailed in Appendix I. As shown in the figure, the guaranteed UE rate is always less than the target-rate due to transmission outages, which occur when the fading, noise, and interference lead to a channel
capacity less than the fixed-transmission rate. Also, we observe the existence of an optimal target-rate that balances the tradeoff between outage probability and channel utilization. The figure shows that the guaranteed-rate becomes significantly less than the target-rate for intercell interference agnostic NOMA design. However, intercell interference awareness and management can significantly reduce outages and improve the UE guaranteed-rate, which highlights the drastic effect of intercell interference.

Having highlighted the importance of interference-aware NOMA design for 5G cellular networks, it remains to note that there are multiple challenges associated with such design such as:

- **Interference’s stochastic nature:** This is due to several sources of uncertainty such as channel gains, network geometry, transmission powers, traffic requirements, etc., which make interference hard to estimate.

- **Location dependence:** Interference is highly dependent on the location of UEs within a BS’s service area: UEs near the cell boundary experience more intercell interference.
• **Dense multi-tier topologies:** Cellular networks are evolving towards an ultra-dense multi-tier topology with irregular cell structure. Hence, the distance between a UE and its serving BS cannot be used to infer its location with respect to the cell boundary.

• **Dominant interferer:** Clustered NOMA UEs may have different dominant interfering sources (i.e., other transmitting UEs or BSs) due to their different spatial locations. This complicates the interference management process.

To the best of our knowledge, such challenges are not addressed in the literature in the context of NOMA cellular networks.

Motivated by this, the work in Chapter 5 focuses on the design of *interference-aware NOMA schemes* in *large-scale 5G cellular networks*, where both intracell and intercell interference are jointly considered. We first present the different design objectives for NOMA cellular networks and discuss their advantages and disadvantages. Then, NOMA design for downlink and uplink scenarios are explicitly presented, in which the drastic impact of intercell interference on power allocation and UE sorting and clustering is discussed. In particular, we show that ignoring intercell interference devastates the performance, while interference suppression (which requires interference awareness) can result in tremendous gains. Since intercell interference has such a large impact on performance, we discuss intercell interference management techniques for NOMA cellular networks. Furthermore, as 5G will integrate several new technologies to achieve the envisioned gains in terms of rate, we highlight the potential of integrating NOMA with other 5G candidate technologies such as FD and D2D communication. We discuss the different kinds of interferences that arise due to each integration, their impact on the network, and possible techniques to handle them.
1.7.4 Downlink NOMA for Poisson Networks

In Chapter 6, we analytically study a large multi-cell downlink NOMA system that takes into account intercell interference and intracell interference, error propagation in the SuIC chain, and the effects of imperfect SuIC for a general number of UEs served by each BS (i.e., a general cluster size). We discuss all of the NOMA challenges enumerated in Section 1.3.3. Our goal is to analyze the performance of such a large network setup using stochastic geometry. We introduce and study three different models to show the impact of location-based selection of NOMA UEs in a cluster, i.e., UE clustering, on performance. We analyze and compare the network performance using two ordering techniques, namely mean signal power- (MSP-) based ordering, which is equivalent to distance-based ordering, and instantaneous signal-to-intercell-interference-and-noise-ratio- (ISINR-) based ordering. In this context, the rate region for the two-user case is studied for both ordering techniques. To the best of our knowledge, an analytical work that compares both ordering techniques does not exist.

We consider two main objectives: 1) maximizing the cell sum rate defined as the sum of the throughput of all the UEs in a NOMA cluster of the cell, subject to a threshold minimum throughput (TMT) constraint of $T$ on the individual UEs 2) maximizing the cell sum rate when all UEs in a cluster have identical throughput, i.e., maximizing the symmetric throughput. Accordingly, we formulate optimization problems and propose algorithms for intercell interference-aware resource allocation for both objectives. We show a significant reduction in the complexity of our proposed algorithms when compared to an exhaustive search. OMA is used to benchmark the gains attained by NOMA.

The contributions of Chapter 6 can be summarized as follows:

- We propose three models for the clustering of UEs (i.e., UE selection), which are governed by two important principles: First, a UE should be served by its closest (or strongest) BS; conversely, a BS chooses its NOMA UE from among
the UEs in its Voronoi cell. Second, only UEs with good channel conditions (on average) should be served using NOMA (i.e., sharing resource blocks). In contrast, using a standard Matern cluster process such as in [55] would lead to the unrealistic situation where UEs from another Voronoi cell may be part of a NOMA cluster.

- From the rate region for the two-user case we show that contrary to the expected result UE ordering based on ISINR, which takes into account information about not only the path loss but also fading, intercell interference and noise, is not always superior to MSP-based ordering. We discuss how resource allocation and intracell interference impact this finding.

- We show that there exists an optimum NOMA cluster size that maximizes the constrained cell sum rate given the residual intracell interference (RI) factor $\beta$.

- We show the existence of a critical level of $\text{SuIC} \ 1 - \beta$ that is necessary for NOMA to outperform OMA.

1.7.5 Meta-Distribution of Coverage Probability for NOMA Users

As mentioned earlier, the superiority of NOMA is well established for large networks in terms of the SCP, however, network operators and vendors are more interested in the percentile performance of UEs, where the fading and activity change while the network realization is kept constant. Thus, we are interested in studying the meta distribution of the coverage probability in NOMA networks.

In Chapter 7, we characterize the meta distribution in downlink cellular networks for two NOMA schemes, namely, everywhere NOMA (E-NOMA) and cell-center NOMA (C-NOMA). The E-NOMA scheme utilizes NOMA for UEs located everywhere in the network, while the C-NOMA scheme restricts NOMA to cell-center UEs
only. Closed-form expressions are derived for the moments of the meta distribution in the E-NOMA scheme. Integral expressions are obtained for the moments in the C-NOMA scheme; consequently, we propose accurate approximate moments to simplify the integral calculation. The meta distribution is then approximated using the beta distribution via moment matching to characterize the UEs percentile performance. The meta distribution of NOMA UEs is also studied in [109]. However, the work in [109] does not consider joint decoding for all SuIC phases, i.e., SuIC based coverage is not considered to be the joint event of decoding all the messages for weaker UEs in the SuIC chain; instead, it is the single event of having SINR of the message of interest above the threshold. Additionally, only the the E-NOMA scheme is studied in [109]. By studying and comparing the two schemes, we are able to show that C-NOMA is a more superior scheme as it not only provides higher spatially averaged coverage probability, but also reduces the variance of the coverage probability across the UEs in the network when compared to the E-NOMA scheme. Our results also highlight the importance of careful resource allocation in NOMA as it drastically impacts not only the SCP but also higher moments of the meta distribution.

1.8 Dissertation Outline

The rest of this dissertation is organized as follows:

1. In Chapter 2, a background on stochastic geometry is presented and details of the tools that are made use of in this dissertation are explained.

2. In Chapter 3, an uplink cellular network overlaid with underlay-D2D communication is considered with FD enabled for the D2D users. A framework with a flexible mode selection scheme is developed for this model and the impact of FD-D2D on a cellular network’s overall performance is studied. The cases of perfect and imperfect SIC for the FD-D2D are considered and the gains that
can be harvested by tuning network parameters appropriately are explored.

3. In Chapter 4, the achievable physical layer security at the typical receiver in the presence of an eavesdropper is studied in a downlink cellular network where receivers have FDJ enabled to enhance their OSSA. Different from current work, correlated interference at the receiver and eavesdropper is considered. The OSSA with correlated interference is characterized in terms of the distance between the receiver and eavesdropper and is compared with the case of independent interference. The impact of employing FDJ when the distance between the typical receiver and its eavesdropper grows is also studied.

4. In Chapter 5, NOMA is studied in the context of large networks; the emphasis of the work is the significant impact of intercell interference. Interference-aware NOMA designs are proposed for 5G networks and their benefits and shortcomings are discussed. Designs for downlink and uplink NOMA are presented and the impact of interference-aware power allocation, UE sorting, and clustering is discussed. Intercell interference management in the context of NOMA networks is discussed. The potential gains that can be reaped by integrating NOMA with other 5G technologies such as D2D and FD are highlighted, the challenges that arise with such integrations and possible solutions are discussed.

5. In Chapter 6, a downlink NOMA network for a general number of NOMA UEs is studied. The developed framework considers NOMA decoding as a joint event of decoding all of the messages in the SuIC chain and imperfections in SuIC are also considered. Three models for UE clustering are proposed and two UE ordering techniques are analyzed and compared in each case. Two problems are considered that maximize the cell sum rate subject to different constraints. Two algorithms for efficient resource allocation that solve the problems are proposed.

6. In Chapter 7, the meta distribution in a downlink cellular network employing
NOMA is characterized. Two NOMA schemes based on the location of UEs are studied in this context. This work is used to highlight the superiority of employing NOMA for cell center users in terms of not only the spatially averaged coverage probability but also higher moments of the coverage probability. The importance of efficient resource allocation is also emphasized by its impact on the moments of the coverage probability.

7. In Chapter 8, a brief summary is provided and the dissertation is concluded. Directions for future work are also discussed.

1.9 Notations

In the rest of this dissertation, unless specified otherwise, we use the following notations:

- The PDF, CDF, and CCDF of the RV $X$ are denoted by $f_X()$, $F_X()$, and $ar{F}_X()$.
- The mean of the RV $X$ is denoted by $\mathbb{E}[X]$.
- The probability of event $A$ is given by $\mathbb{P}(A)$.
- The ordinary hypergeometric function is denoted by $_2F_1(.,.;.;.)$.
- The following standard functions are used: $\gamma(m,n) = \int_0^n x^{m-1}e^{-x}dx$, $\Gamma(m,n) = \int_n^\infty x^{m-1}e^{-x}dx$, $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}dt$, and $\text{erfc}(x) = 1 - \text{erf}(x)$.
- Vectors are denoted using bold text and lower case letters.
- The Euclidean norm of the vector $\mathbf{z}$ is denoted by $\|\mathbf{z}\|$.
- A ball/disk centered at $\mathbf{z}$ with radius $r$ is denoted by $b(\mathbf{z},r)$.
- A sector of the disk centered at $\mathbf{x}$ with radius $r$ and angle $\phi$ is denoted by $s(\mathbf{x},r,\phi)$. When $\phi = \pi$, we denote the half-disk by $s(\mathbf{x},r)$.
• The Laplace transform (LT) of the PDF of the RV \( X \) is denoted by \( \mathcal{L}_X(s) = \mathbb{E}[e^{-sX}] \).
Chapter 2

Stochastic Geometry for Wireless Networks

2.1 Why use Stochastic Geometry?

To analyze the performance of network elements such as BSs and UEs, we are generally interested in the SINR or SIR experienced at the network element. In particular, the interference experienced at any network element is a function of the network geometry as it is influenced by not only the locations of the different types of interferers in the network, but also their intensities, and any interference protection boundaries that the receiver of interest may have from any particular class of interferers. Hence, the network geometry plays a significant role on the interference experienced at receivers.

The signal component of the SINR/SIR is a random variable owing to the instantaneous random channel power gain. It may also have additional randomness if a constant transmit power was employed and the legitimate link distance is random. Additionally, due to the random number and locations of the interferers, random instantaneous channel power gains (from the interferers to the receiver), and possibly transmit power, the interference experienced at a receiver is also a random variable. The thermal noise experienced is always random. Consequently, the SINR/SIR experienced at a receiver is a random variable. In order to characterize any performance metric of interest such as coverage, spectral efficiency, and capacity we require the statistics of the random SINR/SIR. It ought to be noted that the interference is the largest source of randomness in the SINR/SIR. Since the locations, number, and
instantaneous channel gains of the interferers are a consequence of the network geometry, the SINR/SIR and therefore its statistics have a significant dependence on the network geometry. Additionally, they vary from one time instant to another.

Stochastic geometry is a very powerful mathematical tool used to provide spatial averages of random patterns of points, taken over a large number of nodes at various locations or over many network realizations. Equivalently, stochastic geometry averages over all network topologies weighted by their probability of occurrence \cite{9,10}. These averages are done for different quantities of interest, particularly the interference, in our dissertation.

2.2 The Concept of Point Processes (PPs)

The notion of analyzing a random configuration of points in an $n$-dimensional space leads to the concept of PPs. A PP is, loosely speaking, a random collection of points that reside in some space \cite{10}. In this dissertation it is used to statistically describe random patterns of points in an $n$-dimensional space, $\mathbb{R}^n$. PP models permit statements about entire classes of networks rather than just one specific configuration of the network \cite{10}.

We denote by $N(B)$ the number of points in a compact set $B \subset \mathbb{R}^n$ that the PP has. A PP can thus be interpreted as a random counting measure. Conversely, it can also be viewed as a random set; this interpretation, however, is restrictive because in principle a PP can have multiple points at one location. A random set, on the other hand, can not capture this. From here on, we focus on the interpretation of a PP as a random counting measure.

Different PPs have various characteristics and properties. We use $\Lambda(B)$ to denote the average number of points of the PP in a compact set $B \subset \mathbb{R}^n$. If the particular
PP admits a density, then the density (also referred to as the intensity) of the PP is denoted by \( \lambda(x) \), where \( x \) is a location in the \( n \)-dimensional space \( \mathbb{R}^n \) and \( \Lambda(B) = \int_B \lambda(x) \, dx \).

Since we model cellular networks in this dissertation, PPs are used to model the spatial locations of different network elements such as BSs, users, D2D transmitters and receivers, in the 2-dimensional space.

### 2.2.1 Poisson Point Process

The Poisson point process (PPP) is the most commonly used PP. The points in a PPP occur independently of one another. The PPP in \( \mathbb{R}^n \) is such that for a compact set \( B \subset \mathbb{R}^n \), the number of points in \( B \) has a Poisson distribution with mean \( \Lambda(B) = \int_B \lambda(x) \, dx \). Additionally, if \( B_1, B_2, \ldots B_m \) are disjoint bounded sets then the number of points in each of the disjoint sets is an independent random variable.

We are, in particular, interested in homogeneous PPs, i.e., PPs with intensity function that is constant over the space of interest, i.e., \( \lambda(x) = \lambda \). In this case the number of points in a compact set \( B \subset \mathbb{R}^n \) simplifies to a Poisson random variable with mean \( \Lambda(B) = \lambda \int_B \, dx = \lambda |B| \), where \( |B| \) denotes the Lebesgue measure\(^1\) of the set \( B \). Thus, in this case, the number of points in the compact set \( B \), \( N(B) \), has the following probability mass function (PMF),

\[
\mathbb{P}(N(B) = k) = \frac{(\lambda |B|)^k e^{-\lambda |B|}}{k!}.
\]

Using this we can derive the probability of a subset being empty, i.e., the void probability of the PPP, to be \( \exp(-\Lambda(B)) \). In the case of the homogeneous PPP this simplifies to \( \exp(-\lambda |B|) \). The simplicity of this expression helps in various analysis as will be seen in further chapters.

\(^1\)The Lebesgue measure is a standard way of computing the size of a subset of an \( n \)-dimensional Euclidean space in measure theory. In particular, for \( n = 1 \) it corresponds to the length of an interval, for \( n = 2 \) it is the area of the subset, and for \( n = 3 \) it is the volume of the subset.
The independence between the points of a PPP greatly simplifies the analysis. In particular, removing a point or adding a point to a realization, thereby conditioning the PPP, does not change the statistics of the PPP. The reduced Palm distribution\(^2\) of a PPP is therefore the distribution of the PPP itself (Slivnyak’s Theorem) \([10, 113]\). Extrapolating from this, the PPP seen from an arbitrary location is the same as conditioning a point to exist at that location. This simplifies analysis greatly for the PPP as analyzing metrics of interest at an arbitrary location is equivalent to analyzing the metric at the origin.

Since the points in a PPP occur independently of one another, it is in the middle of the spectrum of PPs where one extreme is clustered PPs and the other is repulsive PPs also referred to as hard-core PPs. The PPP is more regular (less clustered) than a clustered process but more clustered (less regular) than a repulsive process. We give an overview on these in the following subsections.

### 2.2.2 Repulsive Processes/Hard-core Processes

As the name suggests, a repulsive PP has some sort of repulsion between the points. Often this is achieved by maintaining a certain minimum (hard-core) distance between the points. Such PPs can be formed by starting with an unrestricted PP and then removing points that do not adhere to the minimum distance requirement; this results in a thinning of the original PP. Of course, there are multiple ways to apply these rules.

\(^2\)The Palm distribution of a PP is the probability of an event given that the PP contains a point at a particular location. We are often interested in not considering the point that the PP is conditioned on, such as in the context of wireless networks where for instance we would not want the intended transmitter to be included in the set of interferers. This leads to the notion of the reduced Palm distribution, where the point that the PP is conditioned on, is not considered in the distribution of the PP. \([10]\)
Matérn Hard-core Process

The Matérn hard-core process starts with a homogeneous PPP and then thins it until each remaining point satisfies a minimum distance requirement from every other remaining point. There are two models for this:

- The Matérn hard-core process of type I: here all of the points that have a neighbor closer than a certain minimum distance $d$ are flagged and then deleted.

- The Matérn hard-core process of type II: here all of the points that have a neighbor closer than a certain minimum distance $d$ are also flagged. However, each point in the original process is given an independent mark which is a RV distributed uniformly in $[0,1]$. Flagged points that have a neighbor within $d$ that has a smaller mark are then deleted. This way, type II deletes fewer unnecessary points than type I at the expense of more complicated analysis.

2.2.3 Cluster Processes

As implied by the name, cluster processes have some sort of attraction between the points. Cluster processes are formed by taking a parent PP and one daughter PP per parent point. The daughter processes are then translated to the position of their parent. The cluster process is then the union of all the translated daughter PPs; note that the cluster process does not include the parent points which are only for translating the daughter PPs to their respective locations. Some common examples of the parent process are PPP (as described below) or a lattice.

It should be noted that when the daughter processes are independent of one another and the parent process, but have identical distributions, and the parent process is stationary, then the cluster process has homogeneous independent clustering. In such a situation, only the statistics of a representative cluster need to be specified, and are equivalent to the statistics of all other clusters in the process.
Poisson Cluster Process

The Poisson Cluster Process (PCP) is a cluster process where the parent process is a PPP. Consequently, the daughter PPs are translated to the locations of the points of the PPP.

A PCP where the daughter points in a cluster are independent of each other, identically distributed, and random in number is called a Neyman-Scott cluster process.

Two popularly used PCPs, that are Neyman-Scott processes, are:

- Matérn Cluster Process: here the daughter clusters are formed by having points uniformly distributed a disk. The same fixed radius is used for each daughter disk. The Matérn Cluster Process is formed by translating the center of each daughter disk to its parent’s location.

- Thomas Cluster Process: the process is formed by having a cluster of points scattered using a normal distribution around each parent point.

A process used for the clustering of NOMA UEs in Chapters 6 and 7 of this dissertation is based on the Stienen Model [114]. It is formed by having a fixed number of NOMA UEs located uniformly in disks centered at the BSs. The radius of each disk of UEs is half of the distance between the serving BS and its nearest interfering BS. The notion behind using this disk radius is to ensure that the NOMA UEs lie in the largest disk centered at the BS that can fit inside the Voronoi cell. This ensures that NOMA UEs are not cell-edge UEs which is important as these UEs are sharing resource blocks and should therefore not have poor channel conditions to be able to afford such sharing.

In terms of clustered PPs, such a process of NOMA UEs is similar to a Matérn cluster process in that it is formed by having daughter processes (i.e., the NOMA UEs in each cell) uniformly distributed in disks around the points of the parent process (i.e., the BSs) which is a PPP. However, different from the Matérn cluster process,
the disks do not all have the same radius and the number of points in our disks are fixed. Since each disk radius depends on the distance between the parent point and its nearest neighbor, the disk radius, and consequently the daughter PPs are not independent of the parent PPP; such a process, therefore, does not have homogeneous independent clustering. It is, however, a PCP, as the parent process is a PPP. It is also not a Neyman-Scott process as the number of points in a cluster (disk) is fixed.

It is important to emphasize the superiority of such a clustered process for modeling clustered NOMA UEs over using a Matérn cluster process: such a clustered process does not allow an overlap of disks as each disk is limited to its own Voronoi cell. Consequently, we can guarantee that each UE is served by its closest BS, which cannot be guaranteed in a Matérn cluster process.

2.3 Analyzing Metrics

2.3.1 Characterizing the Statistics of the Interference

As mentioned before, in the context of wireless networks, interference is the main parameter to be characterized using stochastic geometry. The aggregate interference from a certain class of nodes is a function of the PP of those nodes. Hence, in general the interference experienced at a node located at $x$ from a set of interfering nodes modeled by a PP $\Phi$, is given by,

$$I = \sum_{y_i \in \Phi} P_i h(x, y_i) ||x - y_i||^{-\eta}. \quad (2.1)$$

Here, $y_i$ denotes the location of the $i^{th}$ interferer in $\Phi$, $P_i$ is the transmit power of the $i^{th}$ interferer, $h(x, y_i)$ is the random instantaneous channel power gain between the interferer and receiver at $x$. Additionally, the signal power decays with distance at a rate $r^{-\eta}$, where $\eta$ is the path loss exponent. Hence the interference from a certain class of nodes is a stochastic process that depends on the time varying locations ($y_i$)
of the interferers, which is captured by the PP $\Phi$, and the time varying channel power gains $h(x, y_i)$.

Any statistic can be completely characterized using its PDF or equivalently using the cumulative density function (CDF). In general, it is not possible to calculate the PDF of the interference in a large wireless network. Stochastic geometry gives us statistics of the averaged interference from a certain class of nodes at a generic time in the form of the LT of the PDF of the interference. Since interference is a strictly positive random variable, its LT always exists, and is given by $\mathcal{L}_I(s) = \mathbb{E}[e^{-s I}]$. Stochastic geometry provides a systematic method to obtain the LT of the interference from a certain class of nodes which are modeled using a particular PP. It should be noted, however, that the LT can not be inverted in general to retrieve the PDF of the statistic. It can however be used to generate the existing moments of the statistic. In particular, the LT of the interference from a particular class of nodes can be inverted to obtain the moments as follows:

$$\mathbb{E}[I^n] = (-1)^n \frac{d^n \mathcal{L}_I(s)}{ds^n} \bigg|_{s=0}.$$ 

2.3.2 Characterizing the Statistics of the SINR/SIR - Laplace Transform of Interference Approach

Unlike the case of interference from a certain set of nodes modeled using a PP, deriving the statistics of metrics such as coverage (and equivalently SINR/SIR), capacity, and data rate exactly using the LT is not possible. In our dissertation, we resort to other techniques to evaluate the statistics of the SINR/SIR.

Our network models assume Rayleigh fading and hence the channel power gains follow an exponential distribution, $\exp(\mu)$. This is a common assumption made in the literature due to its analytical tractability as well as it being realistic in most scenarios of interest. Using this assumption for the legitimate link’s channel gain allows us to
characterize the CDF of the SINR as follows: denote by $r$ the legitimate link distance, the legitimate links channel power gain by $h_0$, the legitimate transmitter’s power by $P_0$, the noise power by $\sigma^2$, and let the interference $I$ be given by (2.1). Additionally, $\eta$ is the path loss exponent and like the interferers, the legitimate signal power decays at the rate $r^{-\eta}$. Then the CDF of the SINR is given as follows,

$$F_{\text{SINR}}(\theta) = \mathbb{P}(\text{SINR} \leq \theta)$$

$$= \mathbb{P} \left( \frac{P_0 h_0 r^{-\eta}}{I + \sigma^2} \leq \theta \right)$$

$$= \mathbb{P} \left( h_0 \leq \frac{\theta r^{-\eta}}{P_0} (I + \sigma^2) \right)$$

$$= \int_0^\infty F_{h_0} \left( \frac{\theta r^{-\eta}}{P_0} (y + \sigma^2) \right) f_I(y) dy$$

$$= \int_0^\infty \left( 1 - \exp \left( - \frac{\theta r^{-\eta}}{P_0} (y + \sigma^2) \right) \right) f_I(y) dy$$

$$= 1 - \exp \left( - \frac{\theta r^{-\eta}}{P_0} \sigma^2 \right) \int_0^\infty \exp \left( - \frac{\theta r^{-\eta}}{P_0} y \right) f_I(y) dy$$

$$= 1 - \exp \left( - \frac{\theta r^{-\eta}}{P_0} \sigma^2 \right) \mathcal{L}_I \left( \frac{\theta r^{-\eta}}{P_0} \right).$$

(2.2)

It should be noted that we are able to derive the exact CDF of the SINR in this case because of the Rayleigh fading assumption for the legitimate link’s channel power gain. Works that do not have this assumption use other techniques to either approximate the statistics or complex numerical inversions to obtain the statistics of interest. These techniques are not discussed in our dissertation.

To obtain the CDF (and therefore the statistics) of the SIR, we use the same technique described above but set $\sigma^2$, the noise power, to zero.

### 2.3.3 On the Laplace Transform of Interference

In using LTs of interference in stochastic geometry, a strong emphasis is placed on the accuracy of the intensity function of the interferers and the boundary from where
the interferers exist. This is because these are the two characters of the interference that are required for modeling. The importance of this will become clear as we derive the LT of the interference.

In general, let the interference be of the form (2.1). In our dissertation we assume channel power gains come from fading and do not depend on the locations of the transmitter and receiver, hence $h(x, y_i)$ in (2.1) can be replaced by $h_i$. Then the LT of this interference follows as,

$$L_I(s) = E_I\left[\exp(-s I)\right]$$

(2.3)

$$= E_{\Phi, P, h}\left[\exp\left(-s \sum_{y_i \in \Phi} P_i h_i ||x - y_i||^{-\eta}\right)\right]$$

(2.4)

$$\overset{(a)}{=} E_{\Phi}\left[E_{P, h}\left[\exp\left(-s \sum_{y_i \in \Phi} P_i h_i ||x - y_i||^{-\eta}\right)\right]\right]$$

(2.5)

$$= E_{\Phi}\left[\prod_{y_i \in \Phi} E_{P, h}\left[\exp\left(-s P_i h_i ||x - y_i||^{-\eta}\right)\right]\right].$$

(2.6)

Here (a) follows due to the independence of fading.

When the interferers are modeled to be a PPP, we can further simplify the LT of the interference. From Slivnyak’s Theorem, it is known that the interference seen at a generic location is equivalent to the interference observed at the origin so the path loss term $||x - y_i||^{-\eta}$ can be replaced by $||y_i||^{-\eta}$. Consequently, in the case of PPP interferers, the LT is given as

$$L_I(s) \overset{(b)}{=} \exp\left(- \int_{0}^{2\pi} \int_{c_0}^{\infty} \lambda(r) E_{P, h} \left[1 - \exp\left(-s Phr^{-\eta}\right)\right] rdrd\theta\right),$$

(2.7)

where (b) follows from the probability generating functional (PGFL)\(^3\) of the PPP and a transformation to polar coordinates. We use $\lambda(r)$ to denote the intensity

\(^3\)The PGFL of a PP is the $E[\prod_{x \in \Phi} v(x)]$, where $\Phi$ is the PP on $\mathbb{R}^n$ and $v(x)$ is a measurable function $v : \mathbb{R}^n \to [0, 1]$. For the PPP, the PGFL is $\exp\left(\int_{\mathbb{R}^n} (v(x) - 1)\lambda(x)dx\right)$. 
function of the PPP and \(c_0\) denotes the minimum allowed distance of the nearest interferer. This will be referred to as the interference-protection boundary further in this dissertation. It takes on different values depending on the nature of the set of the interferers; for instance, in the case of adhoc interferers where there is no constraint on the nearest interferers \(c_0 = 0\), while interference coming from BSs in a single-tier downlink network would be restricted so that \(c_0\) is the length of the legitimate link. We will show other network setups in Chapter 6 that allow larger interference protection boundaries than the legitimate link length.

If we assume Rayleigh fading, the channel power gains follow \(h \sim \exp(\mu)\). This further simplifies the LT of the interference as follows

\[
\mathcal{L}_{I}(s) \overset{(c)}{=} \exp \left( - \int_{0}^{2\pi} \int_{c_0}^{\infty} \lambda(r) E_{P} \left[ 1 - \frac{1}{1 + \frac{s}{\mu P R^{-\eta}}} \right] r dr d\theta \right). \tag{2.8}
\]

Here \((c)\) follows due to the moment generating function (MGF) of \(h \sim \exp(\mu)\).

If we further consider the PPP to be homogeneous, and hence \(\lambda(r) = \lambda\), the LT simplifies to,

\[
\mathcal{L}_{I}(s) = \exp \left( - 2\pi \int_{c_0}^{\infty} E_{P} \left[ 1 - \frac{1}{1 + \frac{s}{\mu P R^{-\eta}}} \right] r dr \right). \tag{2.9}
\]

Let's assume constant transmit powers for all interferers; without loss of generality, we can set \(P_i\) in \((2.1)\) equal to 1. Then for the homogeneous PPP, under the Rayleigh
fading assumption, the LT of interference can further be simplified as follows

\[
\mathcal{L}_I(s) = \exp \left( -2\pi \lambda \int_{c_0}^{\infty} \left( 1 - \frac{1}{1 + \frac{s}{\mu} r^{-\eta}} \right) r \, dr \right)
\]

\[
\overset{(d)}{=} \exp \left( -2\pi \lambda c_0^2 \int_{1}^{\infty} \left( 1 - \frac{1}{1 + \frac{s}{\mu} c_0^{-\eta} t^{-\eta}} \right) t \, dt \right)
\]

\[
\overset{(e)}{=} \exp \left( -2\pi \lambda c_0^2 \int_{0}^{1} \left( 1 - \frac{1}{1 + \frac{s}{\mu} c_0^{-\eta} g^{-\eta}} \right) g^{-3} \, dg \right)
\]

\[
\overset{(f)}{=} \exp \left( -\pi \lambda c_0^2 \left( \frac{\Gamma(1, -\frac{2}{\eta}, 1 - \frac{s}{\mu} c_0^{-\eta})}{\Gamma(1 - \frac{s}{\mu} c_0^{-\eta})} - 1 \right) \right) \quad (2.10)
\]

where \((d)\) follows by using the transformation \(r/c_0 \rightarrow t\), \((e)\) follows from the transformation \(t^{-1} \rightarrow g\), and \((f)\) follows from how the hypergeometric function is defined.

### 2.3.4 Characterizing the Statistics of the SINR/SIR - Relative Distance Process Approach

Another approach to characterize the CCDF of the SINR is described in this subsection and does not rely on the LT of the intercell interference. For simplicity we will assume all interferers transmit with unit power, i.e., \(P_i = 1\). Using the independent and identically distributed (i.i.d.) Rayleigh fading assumption as before, i.e., the channel power gains are independent of one another and follow \(h \sim \exp(\mu)\), we can rewrite the CCDF of the SINR as follows:

In terms of the CCP, the coverage probability (i.e., the CCDF of the SINR) can
be written as the first moment of the CCP as follows:

\[
\mathbb{P}(\text{SINR} \geq \theta) = \mathbb{E}_\Phi [\mathbb{P}(\text{SINR} \geq \theta | \Phi)]
\]

\[
= \mathbb{E}_\Phi \left[ \mathbb{P} \left( \frac{P_0 h_0 r^{-\eta}}{\sum_{y_i \in \Phi} h_i ||y_i||^{-\eta} + \sigma^2} \geq \theta | \Phi \right) \right]
\]

\[
= \mathbb{E}_\Phi \left[ \mathbb{P} \left( h_0 \geq \frac{\theta P_0}{\sigma} \left( \sum_{y_i \in \Phi} h_i ||y_i||^{-\eta} \right) \right) \right]
\]

\[
\overset{(a)}{=} \mathbb{E}_{\Phi, h_i} \left[ \exp \left( -r_0^\eta \frac{\mu \theta}{P_0} \sum_{y_i \in \Phi} h_i ||y_i||^{-\eta} \right) \right] \exp \left( -r_0^\eta \frac{\mu \theta}{P_0} \sigma^2 \right)
\]

\[
\overset{(b)}{=} \mathbb{E}_\Phi \left[ \prod_{y_i \in \Phi} \left( 1 + \left( \frac{r}{||y_i||} \right)^\eta \frac{\theta P_0}{\sigma^2} \right)^{-1} \right] \exp \left( -r_0^\eta \frac{\mu \theta}{P_0} \sigma^2 \right), \quad (2.11)
\]

where (a) follows using the CDF of \( h_0 \sim \exp(\mu) \) and (b) follows from the MGF of the independent RVs \( h_i \sim \exp(\mu) \).

From [115], the relative distance process (RDP) is given by

\[
\mathcal{R} = \{ y_i \in \Phi : r/||y_i|| \}. \quad (2.12)
\]

The PGFL of the RDP \( \mathcal{R} \) is

\[
\mathcal{G}_\mathcal{R}[f] \triangleq \mathbb{E} \left[ \prod_{y_i \in \mathcal{R}} f(x) \right] = \mathbb{E} \left[ \prod_{y_i \in \Phi} f \left( \frac{r}{||y_i||} \right) \right]
\]

\[
\overset{(a)}{=} \exp \left( -2\pi \lambda \int_{c_0}^{\infty} \left( 1 - f \left( \frac{r}{a} \right) \right) a \, da \right),
\]

\[
\overset{(b)}{=} \exp \left( -2\pi \lambda c_0^2 \int_{1}^{\infty} t \left( 1 - f \left( \frac{r}{c_0 t} \right) \right) \, dt \right), \quad (2.13)
\]

where we obtain (a) using the PGFL of the PPP and a change to polar coordinates, and (b) by the substitution \( a/c_0 \to t \).

Using the RDP, the expectation in the first term of (2.11) can be simplified. Conse-
quently, the CCDF of the SINR can be rewritten as

\[ P(SINR \geq \theta) = E \left[ \prod_{z \in \mathcal{R}} \left( 1 + z^{\eta} \frac{\theta}{P_0} \right)^{-1} \right] \exp \left( -r^{\eta} \frac{\mu \theta}{P_0} \sigma^2 \right). \] (2.14)

When the PGFL of the RDP is in the form of that given in (2.13), the CCDF of the
SINR becomes

\[ P(SINR \geq \theta) \overset{(a)}{=} \exp \left( -2\pi \lambda c_0^2 \int_1^{\infty} t \left( 1 - \left( 1 + \left( \frac{r}{c_0 t} \right)^{\eta} \frac{\theta}{P_0} \right)^{-1} \right) dt \right) \exp \left( -r^{\eta} \frac{\mu \theta}{P_0} \sigma^2 \right) \]

\[ \overset{(b)}{=} \exp \left( -2\pi \lambda c_0^2 \int_0^{1} g^{-3} \left( 1 - \left( 1 + g^{\eta} \frac{\theta r^{\eta}}{P_0 c_0^\eta} \right)^{-1} \right) dg \right) \exp \left( -r^{\eta} \frac{\mu \theta}{P_0} \sigma^2 \right) \]

\[ \overset{(c)}{=} \exp \left( -\pi \lambda c_0^2 \left( \frac{2F_1 \left( 1, \frac{-2}{\eta}, 1 - \frac{2}{\eta}, -\frac{\theta r^{\eta}}{P_0 c_0^\eta} \right)}{1} \right) \right) \exp \left( -r^{\eta} \frac{\mu \theta}{P_0} \sigma^2 \right). \] (2.15)

In (2.15), (a) follows from applying the PGFL of the RDP; note that this requires
using \( f \left( \frac{r}{c_0 t} \right) \). Using the substitution \( t^{-1} \rightarrow g \), we obtain (b), and (c) follows from
the definition of the hypergeometric function. We observe the first term is equal to
the LT in (2.10), and (2.15) is in the form of the CCDF corresponding to (2.2).

To obtain the CDF (and therefore the statistics) of the SIR, we use the same
technique described above but set \( \sigma^2 \), the noise power, to zero.

It should be noted that we have analyzed the statistics of the SINR/SIR in terms
of the CCDF of the SINR/SIR, which by definition is the first moment of the CCP.
In general, other moments of the CCP can also be derived similarly using the PGFL
of the RDP. This notion is used when we are deriving the meta distribution using the
moments of the CCP as will be seen in Chapter 7.

Based on the system models in the next and coming chapters, we similarly build
on the LTs of interferences, the RDP, and related metrics.
Chapter 3

Modeling Cellular Networks with Full Duplex D2D Communication: A Stochastic Geometry Approach

3.1 Introduction

FD communication is optimistically promoted to double the spectral efficiency if sufficient SIC is achieved. However, this is not true when deploying FD-communication in a large-scale setup due to the induced mutual interference. Therefore, a large-scale study is necessary to draw legitimate conclusions about gains associated with FD-communication. In this context, the stochastic geometry tools mentioned in the last chapter for modeling and analyzing large networks are used. This chapter studies the FD operation for underlay D2D communication sharing the uplink resources in cellular networks. We propose a disjoint fine-tuned selection criterion for the D2D and FD modes of operation. Then, we develop a tractable analytical paradigm, based on stochastic geometry, to calculate the outage probability and rate for cellular and D2D users. The results reveal that even in the case of perfect SIC, due to the increased interference injected to the network by FD-D2D communication, having all proximity UEs transmit in FD-D2D is not beneficial for the network. However, if the system parameters are carefully tuned, non-trivial network spectral-efficiency gains (64% shown) can be harvested. We also investigate the effects of imperfect SIC and D2D-link distance distribution on the harvested FD gains.
Table 3.1: TABLE I: List of Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
<th>Symbol</th>
<th>Definition</th>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Psi$</td>
<td>PPP to constitute cellular BSs</td>
<td>$\omega$</td>
<td>Control factor for D2D link distance distribution</td>
<td>$P_u$</td>
<td>Maximum transmit power of a UE</td>
</tr>
<tr>
<td>$\Phi$</td>
<td>PPP to constitute D2D UEs</td>
<td>$\rho$</td>
<td>Power control cutoff threshold at cellular UE</td>
<td>$P_c$</td>
<td>Transmit power of generic cellular UE</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>Fraction of residual SI</td>
<td>$\rho$</td>
<td>Power control cutoff threshold at D2D receiver</td>
<td>$P_d$</td>
<td>Transmit power of generic D2D UE</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Intensity of cellular BSs</td>
<td>$\rho$</td>
<td>Power control cutoff threshold at D2D receiver</td>
<td>$P_e$</td>
<td>Transmit power of generic r-D2D UE</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Intensity of cellular UEs</td>
<td>$\xi$</td>
<td>Power control cutoff threshold at generic $x$-D2D UE</td>
<td>$P_{\text{avg}}$</td>
<td>Network throughput</td>
</tr>
<tr>
<td>$\lambda_d$</td>
<td>Intensity of D2D UEs</td>
<td>$\rho$</td>
<td>Power control cutoff threshold at generic $x$-D2D UE</td>
<td>$\tau$</td>
<td>Cellular truncation outage probability</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>Reciever sensitivity</td>
<td>$\gamma$</td>
<td>Distance from generic $x$-D2D UE to its nearest BS</td>
<td>$T_n$</td>
<td>Celluar truncation outage probability</td>
</tr>
<tr>
<td>$\eta$</td>
<td>Small scale fading channel gain</td>
<td>$I_{\kappa\chi}$</td>
<td>Interference from a transmitter in mode $\kappa$ to a receiver in mode $\chi$</td>
<td>$\sigma^2$</td>
<td>Noise power</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Required SNR threshold</td>
<td>$R_{\chi}$</td>
<td>Link spectrum efficiency for a UE in mode $\chi$</td>
<td>$\eta_d$</td>
<td>Path-loss of D2D link</td>
</tr>
</tbody>
</table>

Figure 3.1: Transmission Links and Modes

3.2 System Model

3.2.1 Network Model

We model a single-tier D2D-enabled cellular network, in which the D2D links are allowed to share the uplink cellular spectrum. The D2D UEs are equipped with FD transceivers and are allowed to operate in FD mode. Imperfect SIC is assumed such that $0 \leq \zeta \leq 1$ fraction of the transmit power leaks back into the receiver chain of the FD transceiver. We define the cellular link as the link from a UE to a BS, the forward-D2D (f-D2D) link as that from the D2D-transmitter to the D2D-receiver, and the reverse-D2D (r-D2D) link as that from the D2D-receiver to the D2D-transmitter. Hence, we refer to the D2D-transmitter that can transmit in f-D2D mode as the f-D2D UE, and the D2D-receiver that can transmit in the r-D2D mode as the r-D2D UE. Note, we use the term D2D UEs to refer to both f-D2D and r-D2D UEs. Consequently, the FD-D2D mode is active only if both the f-D2D link and r-D2D link are established between a D2D transmit-receive pair as shown in Fig. 3.1. As will be
discussed later, we enable a flexible and disjoint mode selection scheme for the f-D2D and r-D2D links.

Independent Poisson Point Processes (PPPs) $\Psi$ and $\Phi_c$ are used to model the cellular BSs and the cellular UEs with intensities $\lambda$ and $\lambda_c$, respectively. We assume $\lambda_c \gg \lambda$ so that each BS always has a UE to serve. Cellular UEs associate to BSs based on the average radio signal strength (RSS), which reduces to the nearest BS association in single-tier networks. When multiple UEs associate to the same BS, they equally share its resources. The cellular network is overlaid by potential D2D transmitters modeled via an independent PPP $\Phi_d$ with intensity $\lambda_d$. Each D2D-transmitter (f-D2D UE) has a D2D-receiver (r-D2D UE) located within the D2D-proximity and can therefore bypass the BS and communicate in the D2D mode. The D2D-proximity is defined as the region where the D2D-transmitter is able to invert path-loss and achieve at least a power of $\rho_{\text{min}}$ at its receiver while satisfying a maximum power constraint, where $\rho_{\text{min}}$ is the receiver-sensitivity. Note, a D2D UE does not necessarily transmit in the D2D mode; it transmits in the D2D mode only if it satisfies the criteria required for D2D communication explained in the next subsection. D2D UEs that do not select the D2D mode are offloaded to out of band frequencies\footnote{The offloading effect of D2D users to cellular mode and vice versa is studied in \cite{27,68}. However, the offloading effect is not considered in this chapter to avoid unnecessary complications to the analysis without providing additional insights.}. The performance of such D2D nodes is out of the scope of the work in this chapter as they do not affect either the interference or the spectral efficiency within the band of interest.

From the PPP assumption, the cellular link distance distribution, denoted by $r_c$, is given by $f_{r_c}(x) = 2\pi \lambda xe^{-\pi \lambda x^2}, \ x \geq 0$. There is no common agreement on the D2D link distance distribution in the literature as it may depend on the underlying application as well as the social interactions between the D2D UEs \cite{15,16}. Therefore, we adopt
the flexible distribution suggested in [116], which is given by

\[ f_{r_d}(x) = \frac{(2 - \omega)x^{1 - \omega}}{\bar{R}^{2 - \omega}}, \quad 0 \leq x \leq \bar{R}, \quad (3.1) \]

where \( r_d \) is a RV denoting the D2D link distance, \( \bar{R} = \left( \frac{P_u}{p_{\text{min}}} \right)^{\frac{1}{\eta_d}} \) is the maximum transmission range of the D2D UE, and \( 0 \leq \omega < 2 \) is a control factor for the distance distribution. Substituting \( \omega = 0 \) in \( f_{r_d}(\cdot) \) gives the no social interaction case where the D2D receiver is uniformly located in a circle with radius \( \bar{R} \) around the D2D transmitter as in [27]. Also, \( \omega = 1 \) represents the case of equiprobable distances in the range of \([0, \bar{R}]\) as in [68], and \( 1 < \omega < 2 \) represents the case with high social interactions which gives higher weights to shorter D2D link distances. It is worth noting that the distance from an f-D2D UE to its nearest BS is identical in distribution to \( r_c \) and so we denote it by \( r_c \) as well. However, the distance from the r-D2D UE to its nearest BS is denoted by \( r_e \) and follows the distribution proposed in the next section.

A distance dependent power-law path-loss model is considered in which the signal power decays at the rate \( r^{-\eta} \) with the distance \( r \), where \( \eta > 2 \) is the path-loss exponent. Since the D2D and cellular links may experience different propagation conditions, we discriminate between the path-loss exponents of the f-D2D link (\( \eta_d \)) and the cellular link (\( \eta_c \)). Assuming channel reciprocity, the r-D2D link has the same path-loss exponent as the f-D2D link. In addition to path-loss attenuation, transmitted signals experience Rayleigh fading with unit-mean exponential channel power gains. It is assumed that the channel gains are independent from the locations of the transmitters, receivers, and independent from one another.

It is assumed that all UEs have a unified maximum transmit power constraint of \( P_u \). Due to the limited transmit power of the UEs, a truncated channel inversion power control is employed. Hence, only UEs that can compensate for the path-
loss and maintain a predefined average power level at their receivers are allowed to transmit. The cutoff threshold for the power control of each of the communication modes is different; for link establishment we require the transmitters to maintain an average power of $\rho_\chi$ at their respective receivers, where $\chi \in \{c,d,e\}$ corresponds to \{cellular, f-D2D, r-D2D\} modes. Such decoupled power control thresholds offer flexible network design and lead to an enhanced network performance. For the sake of simple presentation, we define $r_1 = \frac{\rho_c}{\rho_d}$ and $r_2 = \frac{\rho_d}{\rho_e}$. A cellular (f-D2D, r-D2D) connection can therefore be established if the power required to achieve $\rho_c$ ($\rho_d$, $\rho_e$) at the BS (r-D2D UE, f-D2D UE) does not exceed $P_u$, otherwise the transmitting UE goes into truncation outage. Due to the PPP assumption, the cellular-truncation outage probability can be expressed as $O_p = e^{-\pi \lambda (\frac{P_u}{\rho_c})^2 \eta_c}$.

Universal frequency reuse is assumed across the entire network with no intra-cell interference between cellular users. D2D links reuse the same uplink frequency with no restrictions on cell boundaries, but subject to the mode selection criterion described in the sequel. Without loss in generality, we analyze the system for one uplink channel.

### 3.2.2 Mode Selection

We consider a flexible mode selection criterion based on the bias factor $T_d$ to impose a tunable IP for the BSs. The IP is enforced via the following mode-selection inequalities $r'^{fd}_d \rho_d \leq T_d \rho'^{ec}_c \rho_c$ and $r'^{rd}_d \rho_e \leq T_d \rho'^{ec}_e \rho_c$, for the f-D2D and r-D2D UEs, respectively. In particular, the f-D2D does not operate in the D2D mode unless $r'^{fd}_d \rho_d \leq T_d \rho'^{ec}_e \rho_c$ is satisfied and the r-D2D does not operate in the D2D mode unless $r'^{rd}_d \rho_e \leq T_d \rho'^{ec}_e \rho_c$ is satisfied. These inequalities, denoted by IP conditions, ensure that a D2D link is not established unless the average interference power from the transmitting D2D UE (i.e., f-D2D or r-D2D) to its nearest BS is strictly less than $T_d \rho_c$, in which $T_d$ is a tunable design parameter to control the D2D contribution to the aggregate interference level.
Consequently, $T_d$ controls the extent to which D2D is enabled in the network. Setting $T_d = 0$ turns off D2D communication (both f-D2D and r-D2D) altogether and nullifies the D2D interference, while $T_d = \infty$ enforces D2D communication with no constraint on the D2D interference. Note that the D2D power control cutoff thresholds $\rho_d$ and $\rho_e$ can be also manipulated to encourage/discourage f-D2D and r-D2D link establishment, respectively, for a given $T_d$ without affecting the cellular IP (i.e., $T_d \rho_e$). The employed mode selection scheme is summarized as follows:

- f-D2D UEs transmit in the f-D2D mode if they satisfy the IP-condition and maximum transmit power constraint. Otherwise, they go into truncation outage.\(^2\)

- r-D2D UEs transmit in the r-D2D mode if they satisfy the IP-condition and maximum transmit power constraint. Otherwise, they go into truncation outage.

### 3.2.3 Methodology of Analysis

We begin by analyzing the PDFs of the link distances. This is followed by calculating the probabilities of transmitting in the f-D2D and r-D2D modes, and the probability of a D2D pair being in FD. The PDFs of the transmission powers in each mode of operation is then evaluated and the moments of the transmission powers are found. We characterize the SINR by its CDF, which requires calculation of the LTs of the interferences PDFs. We use the CDF of the SINR to infer link outage probability and spectral efficiency. To this end, we evaluate the FD-D2D enabled cellular network performance in terms of coverage, network spectral efficiency and power consumption.

For the sake of brevity, we define the \textit{network of interest} in this chapter as a cellular network overlaid with FD-D2D that has BS intensity $\lambda$, D2D-UE intensity $\lambda_d$, D2D

\(^2\)Cellular truncation outage occurs due to unsatisfied power control cutoff threshold only. However, D2D truncation outage occurs due to either unsatisfied power control cutoff threshold and/or unsatisfied IP at the nearest BS.
3.3 On Link Distances and UE Classification

3.3.1 Link Distance Analysis

Based on the network realization and the relative positions of the f-D2D and r-D2D UEs, an f-D2D–r-D2D pair may or may not share the same nearest BS. Conditioning on the relative positions of the f-D2D UE, the r-D2D UE, and the nearest BS to the f-D2D UE, Fig. 3.2 shows two different instances of the shaded crescent formed by the two disks, namely, the red disk centered at the r-D2D UE with radius \( r_c \), and the blue disk centered at the f-D2D UE with radius \( r_c \). If a BS exists in the shaded crescent area, then the f-D2D and r-D2D UEs will not share a common nearest BS. Note that, by definition, a BS may only exist outside the circle of radius \( r_c \) around the f-D2D UE, as the BS in Fig. 3.2 is the nearest BS to the f-D2D UE. Consequently, the region where the r-D2D UE may have a nearer BS is limited by the area of the shaded crescent shown in Fig. 3.2. The area of the shaded crescent depends on the relative values of the D2D link distance \( r_d \), the distance between the f-D2D UE and its nearest BS \( r_c \), and the distance between the r-D2D UE and the BS nearest to the f-D2D UE \( r_{c2} = \sqrt{r_c^2 + r_d^2 - 2r_cr_d \cos \theta} \). Let \( r_e \) be the distance between the r-D2D
and its nearest BS. Then, \( r_e = r_{c_2} \) if the shaded crescent contains no BS. Otherwise, \( r_e < r_{c_2} \). Finding the distribution of \( r_{c_2} \) is by itself a difficult problem because it is a function of three random variables, let alone the distribution of \( r_e \) which is a function of \( r_{c_2} \). Therefore, we propose a Rayleigh PDF approximation for the PDF of \( r_e \) and verify its accuracy by simulations.

The intuition behind our approximation is to use a Rayleigh distribution for \( r_e \), which stems from the fact that the closest point from a 2-D PPP to any point in \( \mathbb{R}^2 \) follows the Rayleigh distribution. Exploiting the moment matching method, we only need the mean of \( r_e \) for the Rayleigh distribution fitting. The following proposition formalizes the approximation of the distribution of \( r_e \).

**Proposition 3.1:** The distance between the r-D2D UE and its nearest BS (which may not be the same BS closest to its f-D2D UE) in the network of interest is accurately approximated by the following Rayleigh distribution:

\[
f_{r_e}(x) = \frac{\pi}{2\mu_{r_e}^2} x \exp\left(-\frac{\pi}{4\mu_{r_e}^2} x^2\right), \ x \geq 0, \tag{3.2}\]

where \( \mu_{r_e} \) is an approximation of \( \mathbb{E}[r_e] \) and is given by

\[
\mu_{r_e} = (1 - P_{r_e \neq r_{c_2}}) \mu_{r_{c_2}} + P_{r_e \neq r_{c_2}} \left( \mu_{r_{c_2} | r_{c_2} > r_d} \frac{\mathbb{E}[r_c | r_c > r_d] - \mathbb{E}[r_d | r_c > r_d]}{2} + (1 - P_{r_c > r_d}) \frac{\mathbb{E}[r_c | r_c < r_d] - \mathbb{E}[r_c | r_c < r_d]}{4} \right). \tag{3.3}\]

In (3.3), \( P_{r_e \neq r_{c_2}} \) denotes the probability that the BS closest to the f-D2D UE is not the closest to the r-D2D UE and is given by \( P_{r_e \neq r_{c_2}} = 1 - e^{-\lambda A} \), where

\[
A = \mathbb{E}\left[r_{c_2}^2 (\phi - \frac{\sin(2\phi)}{2}) - r_e^2 (\bar{\theta} - \frac{\sin(2\bar{\theta})}{2})\right]. \tag{3.4}\]
is the average area of the shaded crescents shown in Fig. 3.2 and \( f_{\theta}(\theta) = \frac{1}{\pi}, 0 \leq \theta \leq \pi \). The probability that the f-D2D UE lies closer to the r-D2D UE than its nearest BS is \( P_{r_{c2}} = \frac{2-\omega}{2R^2-\omega} (\pi\lambda) \frac{\omega^2}{\omega^2 + 2\pi \lambda R^2} \]. The mean of \( r_{c2} \) is approximated by \( \mu_{r_{c2}} = \sqrt{\frac{1}{2\lambda} + \left(\frac{2-\omega}{2\omega} \bar{R}\right)^2} \), and the conditional expectations are given by \( \mu_{r_{c2}} | r_{c} > r_{d} = \sqrt{\langle r_{c} | r_{c} > r_{d} \rangle^2 + \langle r_{d} | r_{c} > r_{d} \rangle^2} \) and \( \mu_{r_{c2}} | r_{c} < r_{d} = \sqrt{\langle r_{c} | r_{c} < r_{d} \rangle^2 + \langle r_{d} | r_{c} < r_{d} \rangle^2} \), where:

\[
\begin{align*}
\mathbb{E}[r_{d} | r_{c} > r_{d}] &= \frac{1}{\sqrt{\lambda}} \frac{\Gamma\left(\frac{3-\omega}{2}\right) - \Gamma\left(\frac{3-\omega}{2}, \pi\lambda R^2\right)}{\Gamma\left(\frac{1}{2}\right)}, \\
\mathbb{E}[r_{c} | r_{c} > r_{d}] &= \sqrt{\frac{4(\pi\lambda)^{\frac{1}{2}}}{\Gamma\left(\frac{1}{2}\right)} \left[\frac{\Gamma\left(\frac{5-\omega}{2}\right) - \Gamma\left(\frac{5-\omega}{2}, \pi\lambda R^2\right)}{2(\pi\lambda)^{\frac{1}{2}}} + \frac{\bar{R}^2}{\frac{4\pi\lambda}{2(\pi\lambda)}} \right] \left[\frac{\Gamma\left(\frac{3-\omega}{2}\right) - \Gamma\left(\frac{3-\omega}{2}, \pi\lambda R^2\right)}{2(\pi\lambda)^{\frac{1}{2}}} \right]} \\
\mathbb{E}[r_{d} | r_{c} < r_{d}] &= \frac{2-\omega}{3-\omega} \left(\frac{\bar{R}}{3-\omega}\right)^{\frac{1}{2}} \left[\frac{\Gamma\left(\frac{3-\omega}{2}\right) - \Gamma\left(\frac{3-\omega}{2}, \pi\lambda R^2\right)}{2(\pi\lambda)^{\frac{1}{2}}} \right], \\
\mathbb{E}[r_{c} | r_{c} < r_{d}] &= \frac{\text{erfc}(\frac{\sqrt{\pi} R}{\bar{R}}) - \text{erfc}(\frac{\sqrt{\pi} \bar{R}}{\bar{R}})}{\left(1 - P_{r_{c} > r_{d}}\right)}.
\end{align*}
\]

**Proof:** See Appendix A.

Fig. 3.3 verifies the distribution of \( r_{c} \) in Proposition 3.1 by plotting the CDF obtained from (3.2) for \( \omega = 1 \), against simulations. Similar results are obtained for other values of \( \omega \); which are not plotted for brevity. Hereafter, we will use the notation \( b = \frac{\pi}{4\mu_{r_{c}}} \) for the sake of simple exposition.
3.3.2 UE Classification

The probabilities that D2D transmitters and receivers select their respective modes of operation are given by the following:

**Lemma 3.1:** The probability that an r-D2D UE in the network of interest transmits in the r-D2D mode is given by

\[
P_e = \frac{(2 - \omega) \eta_c}{2 \eta_d R^{2 - \omega}} \left( \frac{T_d \rho_c}{\rho_c b} \right)^{\frac{2 - \omega}{\eta_d}} \gamma \left( \frac{(2 - \omega) \eta_c}{2 \eta_d}, b \left( \frac{P_u}{T_d \rho_c} \right)^{\frac{2}{\eta_c}} \right).
\]

The intensity of the r-D2D links (i.e., intensity of transmitting r-D2D UEs) is given by \( \mathcal{U}_e = \lambda_d P_e \).

*Proof:* See Appendix B.

**Lemma 3.2:** The probability that an f-D2D UE in the network of interest transmits in the f-D2D mode is given by

\[
P_d = \frac{(2 - \omega) \eta_c}{2 \eta_d R^{2 - \omega}} \left( \frac{T_d \rho_c}{(\pi \lambda)^{\frac{\eta_c}{\eta_d}}} \right)^{\frac{2 - \omega}{\eta_d}} \gamma \left( \frac{(2 - \omega) \eta_c}{2 \eta_d}, \pi \lambda \left( \frac{P_u}{\rho_c T_d} \right)^{\frac{2}{\eta_c}} \right).
\]

The intensity of the f-D2D links (i.e., intensity of transmitting f-D2D UEs) is given by \( \mathcal{U}_d = \lambda_d P_d \).

*Proof:* See Appendix B.

**Lemma 3.3:** The probability that a D2D pair is in FD, i.e., both f-D2D and
r-D2D UE are transmitting, in the network of interest is given by
\[
P_{FD} = \int_0^\infty \frac{f_r(g)}{1 + \hat{q}} \left( \gamma \left( \frac{2 - \omega}{2\eta_d}, \frac{\min(P_u, T_d g \eta_c \rho_c)}{T_d \rho_c \rho_d} \right) \right) \left( 2 - \omega \right) \eta_c \left( \frac{T_d \rho_c}{\rho_d (\pi \lambda)^{\frac{2}{2}} \gamma} \right)^{\frac{2 - \omega}{\eta_d}} \] 
\[ \left( \min(P_u, T_d g \eta_c \rho_c) \right)^{\frac{2 - \omega}{\eta_d}} \hat{q} \right) dg, \]
where \( \hat{q} = e^{-\pi \lambda (\frac{P_u}{\rho_c T_d})^{\frac{2}{2}}} \).

Proof: See Appendix C.

3.4 Transmit Power Analysis

Due to the random network topology along with the employed truncated channel inversion power control, the transmit powers of the cellular, f-D2D, and r-D2D communication modes are all random variables. In this section, we characterize the PDF of the transmit powers of each mode as well as their moments.

Forward-D2D Mode

An f-D2D UE selects the f-D2D mode of operation if 1) it satisfies the maximum transmit power constraint, i.e., \( r_d^{\eta_d} \rho_d < P_u \) 2) it satisfies IP to the cellular mode \( r_d^{\eta_d} \rho_d < T_d r_c^{\eta_c} \rho_c \). The transmit power of a UE operating in the f-D2D mode can therefore be written as \( P_d = r_d^{\eta_d} \rho_d \), with PDF given by the following Lemma.

Lemma 3.4: In the network of interest the PDF of the transmit-power of a UE operating in the f-D2D mode is given by,
\[
f_{P_d}(x) = \frac{2x^{\frac{2 - \omega}{\eta_d} - 1} e^{-\pi \lambda (\frac{P_u}{\rho_c T_d})^{\frac{2}{2}} \left( \pi \lambda \left( \frac{2 - \omega}{2\eta_d} \right) \right)^{\frac{2 - \omega}{\eta_d}}} \left( \pi \lambda \left( \frac{P_u}{\rho_c T_d} \right)^{\frac{2}{2}} \right)}{\eta_c (\frac{2 - \omega}{\eta_d})^{\frac{2 - \omega}{2\eta_d}} \gamma \left( \frac{(2 - \omega) \eta_c}{2\eta_d}, \pi \lambda \left( \frac{P_u}{\rho_c T_d} \right)^{\frac{2}{2}} \right)}, \]
for $0 \leq x \leq P_u$. The $\alpha^{th}$ moment of $P_d$ is given by,

$$
\mathbb{E}[P_d^\alpha] = \frac{(T_d \rho_c)^\alpha \gamma \left( \frac{\alpha \eta_c}{2} + \frac{(2-\omega)\eta_c}{2 \eta_d}, \pi \lambda \left( \frac{P_u}{\rho_c T_d} \right) \right) }{(\pi \lambda)^{\frac{\alpha \eta_c}{2}} \gamma \left( \frac{(2-\omega)\eta_c}{2 \eta_d}, \pi \lambda \left( \frac{P_u}{\rho_c T_d} \right) \right) }.
$$

Proof: See Appendix D.

Reverse-D2D Mode

An r-D2D UE selects the r-D2D mode of operation if 1) it satisfies the maximum transmit power constraint, i.e., $r_d \rho_e < P_u$, 2) it satisfies IP to the cellular mode of operation, i.e., $r_d \rho_e < T_d r_e \rho_c$. The transmit power of an r-D2D UE operating in the r-D2D mode can therefore be written as $P_e = r_d \rho_e$, with PDF given by the following Lemma.

**Lemma 3.5:** In the network of interest, the PDF of the transmit-power of a UE operating in the r-D2D mode is given by,

$$
f_{P_e}(x) = \frac{2^b x \eta_c^{(2-\omega)\eta_c} e^{b(x \eta_c r_d \rho_e)^{\frac{2}{\eta_c}}}}{\eta_c (T_d \rho_c)^{(2-\omega)\eta_c} \gamma \left( \frac{(2-\omega)\eta_c}{2 \eta_d}, b \left( \frac{P_u}{T_d \rho_c} \right)^{\frac{2}{\eta_c}} \right) },
$$

for $0 \leq x \leq P_u$, and the $\alpha^{th}$ moment of $P_e$ is,

$$
\mathbb{E}[P_e^\alpha] = \frac{(T_d \rho_c)^\alpha \gamma \left( \frac{(2-\omega)\eta_c}{2 \eta_d} + \frac{\alpha \eta_c}{2}, b \left( \frac{P_u}{T_d \rho_c} \right)^{\frac{2}{\eta_c}} \right) }{b^\alpha \eta_c \gamma \left( \frac{(2-\omega)\eta_c}{2 \eta_d}, b \left( \frac{P_u}{T_d \rho_c} \right)^{\frac{2}{\eta_c}} \right) }.
$$

Proof: See Appendix E.
Cellular Mode

A cellular UE selects the cellular mode of operation when it is not in cellular truncation outage, i.e., $r_{c}^{h_{c}} \rho_{c} < P_{u}$. The transmit power of the UEs operating in the cellular mode is written as $P_{c} = r_{c}^{h_{c}} \rho_{c}$, and the PDF is given by the following Lemma.

**Lemma 3.6:** In the network of interest, the PDF of the transmit-power of a UE operating in the cellular mode is,

$$f_{P_{c}}(x) = \frac{2\pi \lambda x e^{-\frac{x}{2} - \frac{\lambda (P_{u}^{2})^{\frac{1}{2}}}{\rho_{c}^{2}}}}{\rho_{c}^{2} \left( 1 - e^{-\pi \lambda (P_{u}^{2})^{\frac{1}{2}} / \rho_{c}^{2}} \right)}, \quad 0 \leq x \leq P_{u},$$

and the $\alpha^{th}$ moment of the transmit power is given by,

$$\mathbb{E}[P_{c}^{\alpha}] = \frac{\rho_{c}^{\alpha \gamma} \left( \frac{\alpha \rho_{c}}{2} + 1, \pi \lambda \left( \frac{P_{u}}{\rho_{c}} \right) \frac{\gamma}{\rho_{c}} \right)}{(\pi \lambda)^{\frac{\alpha \rho_{c}}{2}} \left( 1 - e^{-\pi \lambda (P_{u}^{2})^{\frac{1}{2}} / \rho_{c}^{2}} \right)}.$$

**Proof:** See Appendix F.

The intensity of the cellular UEs that are not in truncation is given by $(1 - O_{p}) \lambda_{c}$. Since only one UE is allowed to transmit per BS at a time on a given channel, the number of simultaneously transmitting cellular UEs on the same channel is limited by the number of BSs. Hence, the intensity of simultaneously active cellular UEs is limited by $\lambda$.

### 3.5 SINR Analysis

Let the PPs $\tilde{\Phi}_{c} \subset \Phi_{c}$ and $\tilde{\Phi}_{d} \subset \Phi_{d}$ denote the set of interfering cellular UEs and the set of interfering f-D2D UEs, respectively. Also, we define $\tilde{\Phi}_{e}$ as the set of interfering r-D2D UEs. Although we have assumed $\Phi_{c}$ and $\Phi_{d}$ to be independent PPPs, neither $\tilde{\Phi}_{e}$ nor $\tilde{\Phi}_{d}$ is a PPP and both are mutually correlated due to their interactions (i.e., by scheduling and mode selection) with $\Psi$. Furthermore, $\tilde{\Phi}_{e}$ is mutually correlated.
with $\tilde{\Phi}_d$, and hence, is not a PPP. For tractability, we ignore the mutual correlations between $\tilde{\Phi}_c$, $\tilde{\Phi}_d$, and $\tilde{\Phi}_e$, and assume that each of them constitutes an independent PPP. We formally state these approximations as follows:

**Approximation 1:** The set of interfering cellular UEs ($\tilde{\Phi}_c$) constitutes a PPP with intensity $\lambda$, in which the transmit powers of the UEs are independent.

**Approximation 2:** The set of interfering f-D2D UEs ($\tilde{\Phi}_d$) constitutes a PPP with intensity $U_d$, in which the transmit powers of the UEs are independent.

**Approximation 3:** The set of interfering r-D2D UEs ($\tilde{\Phi}_e$) constitutes a PPP with intensity $U_e$, in which the transmit powers of the UEs are independent.

**Approximation 4:** The sets $\tilde{\Phi}_c$, $\tilde{\Phi}_d$, and $\tilde{\Phi}_e$ are independent of one another.

**Remark:** It is worth mentioning that Approximations 1, 2, 3, and 4 only ignore the mutual correlations between interfering UEs. However, the correlation between the interfering UEs and the test-receiver is captured through the proper calculation of the IP boundaries. Similar approximations are done in [15, 27, 117, 118] for tractability, and are shown to be accurate. Such approximations maintain the model tractability and lead to simple yet accurate expressions for the distribution of the SINRs for each mode of operation. The accuracy of the aforementioned approximations and the distribution of $r_e$ in Proposition 3.1 are validated in Section 3.6.

We characterize the SINR by its CDF. For notational convenience we have defined the set $\chi \in \{c,d,e\}$ where $c$, $d$, and $e$ denote the cellular, the f-D2D, and r-D2D modes of operation, respectively. Hence, we can define a unified SINR expression for all modes of operation as

$$\text{SINR}_\chi = \frac{\rho_\chi h_0}{\sigma^2 + I_{c\chi} + I_{d\chi} + I_{e\chi} + \zeta P_\chi \mathbb{1}_{FD}},$$

where the noise power is denoted by $\sigma^2$, $I_{\kappa\chi}$ is the interference from UEs transmitting in mode $\kappa \in \{c,d,e\}$ to the receiver of the UE transmitting in mode $\chi$, and $\mathbb{1}_{FD}$
is the event that both the f-D2D and r-D2D UEs are active, i.e., the FD-D2D mode is active. For $\chi = c$, $\mathbb{1}_{FD} = 0$; for $\chi \in \{d, e\}$, $\mathbb{1}_{FD}$ is 1 with probability $P_{FD}$ and is 0 otherwise. The interference $I_{\kappa \chi} = \sum_{u_i \in \hat{\Phi}_\kappa} P_{\kappa_i} h_i ||y - u_i||^{-n_\kappa}$, where $y$ and $u_i$ denote the positions of the test receiver and the $i^{th}$ interferer, respectively, $P_{\kappa_i}$ denotes the transmit power of the $i^{th}$ interferer, and $h_i$ denotes the channel between the $i^{th}$ interferer and receiver. The SINR outage is evaluated as:

$$
\mathbb{P}(\text{SINR}_\chi \leq \theta) = \mathbb{P}(h_0 \leq \frac{\theta}{\rho_\chi}(\sigma^2 + I_{c\chi} + I_{d\chi} + I_{e\chi} + \zeta P_{\chi} \mathbb{1}_{FD}))
= 1 - e^{-\frac{\theta}{\rho_\chi}(\sigma^2 + I_{c\chi} + I_{d\chi} + I_{e\chi} + \zeta P_{\chi} \mathbb{1}_{FD})}
= 1 - e^{-\frac{\theta}{\rho_\chi} \sigma^2 \mathcal{L}_{P_\chi} \left( \frac{\theta \mathbb{1}_{FD}}{\rho_\chi} \right) \prod_{\kappa \in \{c,d,e\}} \mathcal{L}_{I_{\kappa \chi}} \left( \frac{\theta}{\rho_\chi} \right).}
$$

(3.5)

where the second equality follows from the exponential distribution of $h_0$, and $\mathcal{L}_X(s)$ denotes the LT of the PDF of the RV $X$ evaluated at $s$. It is worth noting that at the event $\mathbb{1}_{FD} = 0$, the LT $\mathcal{L}_{P_\chi}(0) = 1$. In particular, when the imperfect SIC scenario is considered, the SINR outage for the f-D2D and r-D2D UEs ($\chi \in \{d, e\}$) is calculated as $\frac{P_{FD}}{P_{\chi}} \mathbb{P}(\text{SINR}_\chi \leq \theta|\mathbb{1}_{FD} = 1) + (1 - \frac{P_{FD}}{P_{\chi}}) \mathbb{P}(\text{SINR}_\chi \leq \theta|\mathbb{1}_{FD} = 0)$. The weights account for the fraction of the D2D UEs transmitting in FD and HD, respectively. For the perfect SIC scenario, (3.5) can be used directly (with $\mathcal{L}_{P_\chi}(0) = 1$). The LTs for the aggregate interferences are given by the following lemma.
Lemma 3.7: For the network of interest, the LTs of the interferences PDFs are:

\[
\mathcal{L}_{I_{ec}}(s) = \exp \left( -s \mathbb{E} \left[ \frac{\sqrt{P_{e}}}{\eta_{c}} \right] \Gamma \left( 1 + \frac{2}{\eta_{c}} \right) \Gamma \left( 1 - \frac{2}{\eta_{c}} \right) \right)
\]

\[
\mathcal{L}_{I_{dc}}(s) = \exp \left( -s \mathbb{E} \left[ \frac{\sqrt{P_{d}}}{\eta_{c}} \right] \Gamma \left( 1 + \frac{2}{\eta_{c}} \right) \Gamma \left( 1 - \frac{2}{\eta_{c}} \right) \right)
\]

\[
\mathcal{L}_{I_{cc}}(s) = \exp \left( -s \mathbb{E} \left[ \frac{\sqrt{P_{c}}}{\eta_{c}} \right] \Gamma \left( 1 + \frac{2}{\eta_{c}} \right) \Gamma \left( 1 - \frac{2}{\eta_{c}} \right) \right)
\]

\[
\mathcal{L}_{I_{ed}}(s) = \exp \left( -\pi \sqrt{s} \mathbb{E} \left[ \sqrt{P_{e}} \right] \Gamma \left( 1 + \frac{2}{\eta_{d}} \right) \Gamma \left( 1 - \frac{2}{\eta_{d}} \right) \right)
\]

\[
\mathcal{L}_{I_{dd}}(s) = \exp \left( -\pi \sqrt{s} \mathbb{E} \left[ \sqrt{P_{d}} \right] \Gamma \left( 1 + \frac{2}{\eta_{d}} \right) \Gamma \left( 1 - \frac{2}{\eta_{d}} \right) \right)
\]

\[
\mathcal{L}_{I_{cd}}(s) = \exp \left( -\pi \sqrt{s} \mathbb{E} \left[ \sqrt{P_{c}} \right] \Gamma \left( 1 + \frac{2}{\eta_{d}} \right) \Gamma \left( 1 - \frac{2}{\eta_{d}} \right) \right)
\]

An important scenario of interest is the case of \( \eta_{c} = 4 \), which does not only simplify the analysis but also represents a practical value for outdoor cellular communications in urban environments \[5,6,13,119,121\].

Corollary 3.1: For the network of interest, at path-loss exponent \( \eta_{c} = 4 \), the LTs of the interferences the cellular UEs experience from each communication mode reduce to:

\[
\mathcal{L}_{I_{cc}}(s)_{\eta_{c}=4} = \exp \left( -\pi \mathbb{E} \left[ \sqrt{P_{c}} \right] \arctan \left( \sqrt{s} \rho_{c} T_{d} \right) \right)
\]

\[
\mathcal{L}_{I_{dc}}(s)_{\eta_{c}=4} = \exp \left( -\pi \mathbb{E} \left[ \sqrt{P_{d}} \right] \arctan \left( \sqrt{s} \rho_{c} T_{d} \right) \right)
\]

\[
\mathcal{L}_{I_{cd}}(s)_{\eta_{c}=4} = \exp \left( -\pi \sqrt{s} \mathbb{E} \left[ \sqrt{P_{c}} \right] \arctan \left( \sqrt{s} \rho_{c} \right) \right)
\]

Proof: See Appendix G. \( \square \)
\[ P(SINR_c \geq \theta) = \exp \left( -\frac{\theta}{\rho_c} \sigma^2 - 2\pi \left( \frac{\theta}{\rho_c} \right) \frac{2F_1(1, \frac{\eta_c-2}{\eta_c}; \frac{2\eta_c-2}{\eta_c}; -\theta)}{(\lambda E[P_c^{\frac{2}{\eta_c}}])^{-1}((\eta_c - 2)\theta)^{\frac{2}{\eta_c} - 1} + \frac{2F_1(1, \frac{\eta_c-2}{\eta_c}; \frac{2\eta_c-2}{\eta_c}; -\theta)\eta_c}{(U_d E[P_d^{\frac{2}{\eta_d}}])^{-1}((\eta_c - 2)(\theta T_d)^{\frac{2}{\eta_c} - 1})} \right) \] (3.6)

\[ P(SINR_d \leq \theta) = E[e^{-\zeta P_d^{\eta_d} F_d}] \exp \left( -\frac{\theta}{\rho_d} \sigma^2 - \pi \left( \frac{\theta}{\rho_d} \right)^{\frac{2}{\eta_d}} \left( \lambda E[P_d^{\frac{2}{\eta_d}}] \Gamma(1 + \frac{2}{\eta_d}) \Gamma(1 - \frac{2}{\eta_d}) + \frac{2}{\eta_d} \Gamma(1 - \frac{2}{\eta_d}) \right) \right) \] (3.7)

\[ P(SINR_e \leq \theta) = E[e^{-\frac{\zeta P_e^{\eta_e} F_e}{\rho_e^{\eta_e}}}] \exp \left( -\frac{\theta}{\rho_e} \sigma^2 - \pi \left( \frac{\theta}{\rho_e} \right)^{\frac{2}{\eta_e}} \left( \lambda E[P_e^{\frac{2}{\eta_e}}] \Gamma(1 + \frac{2}{\eta_e}) \Gamma(1 - \frac{2}{\eta_e}) + \frac{2}{\eta_e} \Gamma(1 - \frac{2}{\eta_e}) \right) \right) \] (3.8)

\[ P(SINR_f \leq \theta) = E[e^{-\zeta P_f^{\eta_f} F_f}] \exp \left( -\frac{\theta}{\rho_f} \sigma^2 - \pi \left( \frac{\theta}{\rho_f} \right)^{\frac{2}{\eta_f}} \left( \lambda E[P_f^{\frac{2}{\eta_f}}] \Gamma(1 + \frac{2}{\eta_f}) \Gamma(1 - \frac{2}{\eta_f}) + \frac{2}{\eta_f} \Gamma(1 - \frac{2}{\eta_f}) \right) \right) \] (3.9)

Proof: Simplify the expressions in Lemma 3.6 at \( \eta_c = 4 \). ☐

Using the LTs of the interference PDFs, the outage probabilities of the D2D and cellular links are given in the following theorem:

**Theorem 3.1:** For the network of interest with residual SI fraction \( \zeta \), the success probability for a UE operating in the cellular mode is given by (3.6) in general and by (3.7) when \( \eta_c = 4 \), the success probability for a UE operating in the f-D2D mode is given by (3.8), and the success probability for a UE operating in the r-D2D mode is given by (3.9).

Proof: Using (3.5) and the LTs of the interferences found in Lemma 3.6, we obtain the SINR outage probability expressions for each mode of operation. ☐

Let \( \xi(t) = \frac{(t^2-1)}{\rho_x} \); the link spectrum efficiency for a UE operating in mode \( \chi \) is
In this section, we validate the developed mathematical model and benchmark the 
3.6 Results and Analysis

In this section, we validate the developed mathematical model and benchmark the 
FD-D2D operation against the D2D enabled cellular network with HD UEs, denoted
as the HD-network, and the traditional cellular network where D2D is disabled. The cellular network overlaid with FD-D2D being considered in this chapter will be referred to as the FD-network. Let $A_\chi$ be the probability of the joint event that a randomly selected user is operating in mode $\chi$ and is not in truncation outage, $\text{PUR}(\chi)$ and $\text{PCR}(\chi)$ be the average per-user rate and average per-cell rate, respectively, of mode $\chi$, $\Lambda(\chi)$ be the intensity of users operating in mode $\chi$, and $\text{Tx}(\chi)$ be the average transmit power of users operating in mode $\chi$, where the value of each of these parameters in each network scenario is given in Table 3.2. Note that $\text{PUR}(\chi)$ in Table 3.2 has a multiplication factor of $\frac{1}{2}\beta$ in which the factor $\frac{1}{2}$ reflects the two-hop nature (i.e., uplink then downlink) of the cellular links and the factor $\beta = \frac{\text{BS intensity}}{\text{intensity of UEs in cellular mode}} = \frac{\lambda}{(1-O_p)\lambda_c}$ reflects the share each user get from the uplink spectrum when equal sharing among the users is assumed. On the other hand, $\text{PCR}(c)$ does not incorporate the two hops or the spectrum sharing factors because we look at the total uplink rate from the BS side. Assuming a round robin scheduling for the cellular UEs, $\text{Tx}(c)$ is also multiplied with the factor $\beta$ to reflect the activity of the UEs.

Table 3.3: Parameter Values

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>10 BS/km$^2$</td>
<td>$\rho_d$</td>
<td>$\frac{2}{\pi}$</td>
</tr>
<tr>
<td>$\lambda_c$</td>
<td>100 UE/km$^2$</td>
<td>$\rho_e$</td>
<td>$\frac{2}{\pi} - \frac{\lambda}{\rho_c}$</td>
</tr>
<tr>
<td>$\lambda_r$</td>
<td>100 UE/km$^2$</td>
<td>$\rho_d$</td>
<td>4</td>
</tr>
<tr>
<td>$P_u$</td>
<td>200 mW</td>
<td>$\eta$</td>
<td>4</td>
</tr>
<tr>
<td>$\rho_{\text{ms}}$</td>
<td>-90 dBm</td>
<td>$\omega$</td>
<td>1</td>
</tr>
<tr>
<td>$\rho_c$</td>
<td>-80 dBm</td>
<td>$\sigma^2$</td>
<td>-90 dBm</td>
</tr>
</tbody>
</table>

Table 3.4: Effect of increasing $r_1$, $r_2$, and $T_d$.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>t-D2D</th>
<th>v-D2D</th>
<th>Cellular</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U_d$</td>
<td>$\rho_d$</td>
<td>$\rho_e$</td>
<td>IP</td>
</tr>
<tr>
<td>Increasing $r_1$</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
</tr>
<tr>
<td>Increasing $r_2$</td>
<td>–</td>
<td>–</td>
<td>↑</td>
</tr>
<tr>
<td>Increasing $T_d$</td>
<td>↑</td>
<td>–</td>
<td>↑</td>
</tr>
</tbody>
</table>
As shown in Table 3.2, cellular and f-D2D links share the spectrum in the HD-network, while the spectrum is explicitly used by cellular links in the traditional cellular case. Therefore, the rates in Table 3.2 are explicitly defined for each network scenario to reflect their different interference environments, where $\mathcal{R}_\chi^{(\text{Conv})} > \mathcal{R}_\chi^{(\text{HD})} > \mathcal{R}_\chi^{(\text{FD})}$. Note that $\mathcal{R}_\chi^{(\text{FD})}$ is given in (3.10). The rates $\mathcal{R}_\chi^{(\text{HD})}$ and $\mathcal{R}_\chi^{(\text{Conv})}$ are evaluated via (3.10) by eliminating the LT of the SI along with $\mathcal{L}_{I_e}(\cdot)$ and $\mathcal{L}_{I_d}(\cdot)\mathcal{L}_{\mathcal{I}_e}(\cdot)$, respectively.

From the user side, we define two performance metrics to assess the per-user gain in the FD-network when compared to the HD-network and traditional network. The first metric is the per-user rate, defined as $T_{\text{avg}} = \sum \chi \mathcal{A}_\chi \text{PUR}(\chi)$. The second metric is the average transmit power, defined as $P_{\text{avg}} = \sum \chi \mathcal{A}_\chi \text{Tx}(\chi)$. The performance gain from the network side is evaluated by the network throughput, which is defined as $T_n = \lambda \sum \chi \text{PCR}(\chi)$.

### 3.6.1 Parameter Selection

In Sections 4.5-B-4.5-E, we focus on the FD-network performance assuming perfect SIC (i.e., $\zeta = 0$) to study the explicit contribution of the FD communication to the aggregate interference and the subsequent effect on outage and rate. Once FD-D2D gains over the HD and traditional networks are highlighted, the effect of imperfect SIC is studied (i.e., $\zeta > 0$) in Section 4.5-F. Section 4.5-G focuses on the effect of the link distance distribution and so assumes perfect SIC. For the simulation scenario, unless stated otherwise, the parameter values in Table 3.3 are used. Proposition 3.1 is used for the distribution of $f_{r_e}(x)$ in the analysis.

For a fixed $\rho_c$, we control our network using three main parameters: $r_1$, $r_2$, and $T_d$. In particular, $r_1$ controls the power required at both the f-D2D and r-D2D UEs; decreasing $r_1$ implies higher $\rho_d$ and $\rho_e$ at the receiver UEs. Using $r_2$ we control the power required at the r-D2D link’s receiver only; decreasing $r_2$ implies higher $\rho_e$ at
the f-D2D UEs. The amount of IP provided to BSs is controlled by $T_d$; increasing $T_d$ loosens the IP conditions by increasing the maximum allowed interference (i.e., $T_d\rho_c$), which allows more f-D2D and r-D2D UEs to satisfy the mode selection inequalities and transmit. Table 3.4 summarizes the effects of varying $r_1$, $r_2$, and $T_d$ on the network, where (↑), (↓), and (−) denote increase, decrease, and no change, respectively.

### 3.6.2 Model Validation

We first validate Approximations 1, 2, 3, and 4, as well as Proposition 3.1 for an FD-network by showing that the analysis is a good estimate of the simulations. Note that the simulation does not enforce any of these approximations. In each simulation run, a PPP cellular network with BS intensity $\lambda$ is simulated in a 1000 km$^2$ area. We then generate cellular and f-D2D UEs with intensities $\lambda_c$ and $\lambda_d$, respectively. For each f-D2D UE, an r-D2D UE is generated within a radius of $\bar{R}$ according to the PDF in (3.1). The cellular UEs are scheduled to transmit if they are not in truncation and if there is no other UE scheduled in the same Voronoi cell. The f-D2D and r-D2D UEs are scheduled to transmit if they satisfy the maximum transmit power constraint and IP. All UEs employ channel inversion power control.

Fig. 3.4 is a plot of SINR outage against $\theta$ with $r_1$ fixed to 1 (i.e., $\rho_d = \rho_c$). The figure shows that increasing $r_2$ (i.e., decreasing $\rho_c$) worsens SINR for all three communication modes. This occurs due to the increase in the number of transmitting r-D2D UEs which increases network interference. The r-D2D mode, however, is impacted significantly more than the cellular and f-D2D modes as increasing $r_2$ not only increases interference but also worsens the received signal power ($\rho_e$) of the r-D2D mode.

Fig. 3.5 shows the SINR outage of all three communication modes increases when $r_1$ is increased (for fixed $r_2$ and $T_d$) and when $T_d$ is increased (for fixed $r_1$ and $r_2$). Increasing $r_1$ and $T_d$ each increases outage due to the increased interference
that results. However, $r_1$ impacts the f-D2D and r-D2D modes more significantly than the cellular mode as it additionally affects their received signal powers. Since varying $T_d$ only affects the network interference, the impact on the SINR of all three communication modes is similar. This difference in sensitivity to the parameters $r_1$ and $r_2$ gives the network operator the ability to alter the respective parameter if the performance of one of these modes needs to be altered without affecting the other modes too much, thereby giving flexibility to control the performance of a mode without affecting the other modes significantly.

### 3.6.3 Effects of the Interference Protection Condition

Increasing $T_d$ decreases the amount of IP for the cellular mode and thereby allows a larger number of D2D UEs to transmit. This increases spatial frequency reuse but also increases network interference. Hence, $T_d$ imposes a tradeoff between the number of simultaneously active links and the transmission rate per link.

We first show the network throughput vs $T_d$ for the FD-network and HD-networks normalized w.r.t. the traditional cellular network in Fig. 3.6. Note that $T_n$ of the traditional cellular network does not change with $T_d$ as D2D communication is prohibited. The figure shows the existence of an optimal $T_d$ that maximizes the network
throughput. An optimal $T_d$ exists because increasing $T_d$, at first, has a larger positive impact on the overall performance by increasing the number of transmitting D2D UEs, thereby increasing spatial frequency reuse. Beyond the optimal $T_d$, the negative impact of the D2D interference dominates the network performance and increasing $T_d$ deteriorates the network throughput. An important observation from the figure is that the FD-network offers non-trivial throughput gains when compared to the HD-network and traditional cellular network, 64% and 245%, respectively. These high gains are observed because we look at the performance from the network perspective in which the FD-network allows an additional $U_e$ and additional $(U_d + U_e)$ links per unit area to efficiently reuse the spectrum when compared to the HD-network and traditional network, respectively. It is also worth mentioning that the HD-network allows an additional $U_d$ links per unit area to reuse the spectrum compared to the traditional network, which gives 110% increase in the throughput.

The D2D bias factor $T_d$ also effects the average transmit power, $P_{avg}$, of the D2D devices as shown in Fig. 3.7. The figure shows that D2D communication generally reduces the average transmission power when compared to the traditional cellular network for low values of $T_d$. This occurs because low $T_d$ only allows D2D UEs that have lower path-loss attenuation to transmit which reduces the transmission power due to the employed channel inversion power control. Also, for lower $T_d$, the FD-network offers a lower power consumption than the HD-network because it allows a larger number of devices to exploit good channel conditions and communicate in the D2D mode. However, for larger $T_d$, UEs with higher transmit-powers are allowed to transmit; since, the FD-network allows a larger number of these than the HD-network, its average power consumption exceeds the HD-network’s. Interestingly, the $T_d$ that optimizes spectral efficiency (cf. Fig. 3.6) falls in the region that offers high transmit power reduction w.r.t. the traditional network, implying that using the correct value of $T_d$ enables the network to simultaneously consume less power per transmitting UE
Figure 3.6: $T_n$ gain w.r.t the traditional cellular network vs. $T_d$ with $r_1 = 0.2$ and $r_2 = 0.2$.

Figure 3.7: $P_{avg}$ vs. $T_d$ with $r_1 = 0.2$ and $r_2 = 0.2$.

Figure 3.8: SINR Outage vs. $T_d$ for $\theta = 1$ with $r_1 = 0.2$ and $r_2 = 0.2$.

Figure 3.9: $O_{net}$ vs. $T_d$ for $\theta = 1$ with $r_1 = 1$ and $r_2 = 2$.

and gain maximum throughput.

Fig. 3.8 shows the negative impact of the increased D2D communication on the explicit SINR outage of each mode of operation for all values of $T_d$. In particular, for each mode of operation (cellular, f-D2D, and r-D2D), inducing more D2D communication deteriorates the SINR outage. Hence the SINR outage of common modes of operation is highest for the FD-network, followed by the HD-network, and finally the traditional cellular network. Also, the figure manifests the crucial role of IP on the cellular network outage probability and the drastic rate of outage increase with increasing $T_d$.

Fig. 3.9 is a plot of the average SINR network-outage against $T_d$, which is defined as $O_{net} = \sum \frac{A(\lambda)}{\sum \Lambda(\kappa)} O_{\lambda}$. Note that we differentiate between the outage probabili-
ties in each network scenario, namely, $O^{(FD)}$, $O^{(HD)}$, and $O^{(Conv)}$, according to the interference environment. Hence, $O^{(FD)}$ is given in (3.5) and the outages $O^{(HD)}$ and $O^{(Conv)}$ are evaluated via (3.5) by eliminating the LT of the SI along with $L_{e_{i_{e}}} (\cdot)$ and $L_{e_{d}} (\cdot) L_{e_{c_{x}}} (\cdot)$, respectively. As shown in the figure, the FD-network has the highest network-outage, followed by the HD-network, and finally the traditional cellular network. The figure manifests the importance of $T_{d}$ and shows that the FD-network requires a much more stringent IP-condition to maintain the same outage performance as the HD-network.

Figs. 3.6 and 3.9 clearly show the tradeoff between spectral efficiency and outage probability. It could be concluded that despite the increased outage probability, the overall network capacity increases due to the improved spatial frequency reuse. It is worth mentioning that the high numerical values for outage probabilities in Figs. 3.8 and 3.9 are common in stochastic geometry based analysis due to the simplified system model and the employed simplistic interference management scheme to maintain tractability. Nevertheless, despite the increased outage, FD-D2D communication provides potential gains to the per-user as well as the aggregate network throughputs in cellular networks. In practice, the ignored effect of shadowing and propagation along with employing more sophisticated interference management schemes are expected to reduce SINR outage and increase the harvested FD-D2D gains.

3.6.4 Effects of the Distance Cut-off for the FD Mode

Increasing $r_{2}$ (i.e., decreasing $\rho_{e}$) decreases the power required by r-D2D UEs to invert their channel and thereby increases the number of transmitting r-D2D UEs. This increases spatial frequency reuse but also increases network interference. Additionally, the received intended signal power of the r-D2D links decreases.

Fig. 3.10 shows the existence of an optimal $r_{2}$ that maximizes $T_{\text{avg}}$. An optimal exists because increasing $r_{2}$, is beneficial at first as it increases the number of
transmitting r-D2D UEs that make a useful contribution to the network. Increasing $r_2$ beyond this causes deterioration to the overall performance due to the increased interference, as well as due to the decreased power of the intended signal of the r-D2D UEs (i.e., $\rho_e$). Fig. 3.10 shows that at the optimal $r_2$, the FD-network outperforms the HD-network by 18%. Both the HD-network and the FD-network outperform the traditional cellular network significantly. The high performance gain offered by the D2D communication (both FD-D2D and HD-D2D) w.r.t. the traditional cellular network can be attributed to the explicit utilization by each D2D link for the available uplink channel when compared to the share $\beta$ that UEs get when scheduled in the cellular mode. Furthermore, D2D communication establishes a direct (i.e., one-hop) link between two UEs compared to the two-hop (i.e., uplink then downlink) communication via the BS.

### 3.6.5 Analyzing Truncation and SINR Outage

Fig. 3.11 is a plot of the truncation and SINR outages of the individual transmission modes with increasing $T_d$ for an FD-network. Note, the r-D2D links and f-D2D links have higher truncation outage than the cellular links due to the small values of $r_1$ and $r_2$ (and therefore high values of $\rho_d$ and $\rho_e$) being used. Our goal is to observe the effect of increasing $T_d$ on the truncation and SINR outages. Since increasing $T_d$ decreases
IP, the truncation outages of the f-D2D and r-D2D transmission modes decrease with $T_d$ until they settle to a constant. This occurs when the inability to invert the channel to the receiver becomes the bottleneck of truncation outage and not the inability to comply with the IP-condition. At the same time we see that increasing $T_d$, which allows more f-D2D and r-D2D links, increases SINR outage. Increasing $T_d$ allows more D2D transmissions that cause more interference to the BSs; this occurs either when the links have high power and/or when the transmitting UE is closer to the BS.

### 3.6.6 Effects of Imperfect SIC

In this set of results, we investigate the effect of imperfect SIC on the FD-D2D network performance. First, we look at SINR outage probability for different values of the residual SI fraction $\zeta$ in Fig. 3.12. As expected, increasing $\zeta$ deteriorates the outage probability for f-D2D and r-D2D UEs due to the increased residual SI. It ought to be highlighted that only a fraction $P_{FD}$ of the f-D2D and r-D2D links operate in FD and experience residual SI. Additionally, in Fig. 3.12 we note that the r-D2D mode is impacted more severely by the residual SI than the f-D2D mode. This occurs because of the $r_2$ being used, which increases $\rho_e$ and therefore r-D2D transmission powers, which in turn leads to more SI for the r-D2D UEs.

We also look at the effect of imperfect SIC on the total network throughput in Figs. 3.13 and 3.14 for different values of $r_1$ and $r_2$. In both cases, the figures show that increasing $\zeta$ deteriorates the network throughput due to the imposed SI on the FD links. The HD-network is included in the figures to benchmark the FD-D2D with imperfect SIC. Figs. 3.13 and 3.14 show a $\zeta$ dependent threshold where the HD-network outperforms the FD-network when $\zeta$ is high. In particular, Fig. 3.14 shows that if the network parameters are properly tuned, higher $\zeta$ values become more tolerable allowing the FD-network to outperform the HD-network. This highlights the importance of properly tuning the network parameters allowing, theoretically, the
3.6.7 Effects of the D2D link distance distribution

Finally, we inspect the effect of the link distance distribution parameter $\omega$ on the network throughput in Fig. 3.15. The figure shows that increasing $\omega$ increases $T_n$ for a given $T_d$. This can be explained by the fact that larger $\omega$ values give higher weights to shorter distances, which results in less D2D transmission power due to the employed power control, and hence, less network interference and improved network throughput.
3.7 Conclusion

This chapter presents a tractable framework for large-scale cellular networks overlaid with FD-D2D UEs that have imperfect SIC capabilities and a tunable D2D link distance distribution. We first propose a flexible network design where the flexibility comes from imposing tunable design variables that control the extent to which D2D communication is enabled in the network along with the interference protection provided for cellular users. We also propose a disjoint mode selection for the forward (f-D2D) and reverse (r-D2D) links, which depends on their relative positions from the nearest BS. To carry out our analysis, we propose an accurate approximation for the PDF of the distance between the r-D2D UE and its nearest BS. We then characterize the aggregate interference and derive the outage probability and ergodic rate. The results show that enforcing all potential D2D links to operate in D2D can severely degrade the network performance due to the imposed interference. Hence, the extent to which the D2D is enabled in the network has to be carefully tuned to balance the tradeoff between spatial frequency reuse and aggregate interference level. Due to the imposed aggregate interference, the FD-D2D communication does not double the network rate when compared to the HD-D2D operation even with perfect SIC at the optimal design variables. Nevertheless, FD-communication offers non-trivial gains compared to its HD counterpart, if the design parameters are carefully selected (64% in Fig. 3.6). In the case of imperfect SIC, a minimum level of SIC is required to achieve gains from employing FD-D2D compared to the HD-network. However, if the network parameters are tuned carefully, this minimum level of SIC can be decreased. Finally, we investigate the effect of the link distance distribution on the FD-network performance and show its prominent effect. While this chapter shows potentials for FD-D2D communication, it also highlights the importance of sophisticated interference management to maintain an acceptable outage probability and boost the harvested FD gains.
Interference at each receiving node in this chapter was assumed to be independent. However, as the interference comes from the same set of transmitters and as some of the receiving nodes are close by (such as the D2D pair), the interference experienced at such nodes may be correlated. In the context of the work in this chapter, it was not important to consider this interference correlation as the proximity receivers were not looking for the same message. However, when proximity receivers are interested in the same message, interference correlation becomes more relevant. It can be particularly useful (or detrimental) in the context of cooperative communication or in the context of secure communication. In the next chapter we study the impact of interference correlation, which is generally ignored, in the context of physical layer security.
Chapter 4

The Effect of Spatial Interference Correlation and Jamming on Secrecy in Cellular Networks

4.1 Introduction

Recent studies on secure wireless communication have shed light on a scenario where interference has a desirable impact on network performance. Particularly, assuming independent interference-power fluctuations at the eavesdropper and the receiver, opportunistic secure-information transfer can occur on the legitimate-link. However, interference is spatially correlated due to the common set of interfering sources, which may diminish the OSSA probability. In this chapter, we study and quantify the effect of spatial interference correlation on OSSA in cellular-networks. We also investigate the potential of FDJ solutions. The results highlight the scenarios where FDJ improves OSSA performance.

4.2 System Model

4.2.1 Network Model

We focus on the downlink of a network where the BSs form a homogeneous PPP $\Phi_b$ with intensity $\lambda$. The PP of UEs, $\Phi_u$, is obtained by placing a point uniformly at random in each Voronoi cell of $\Phi_b$. This is the user model of type I in [122]. These UEs are served in the same time-frequency resource block. Eavesdroppers are distributed according to an arbitrary stationary PP independent of $\Phi_b$. A UE employs FDJ if the nearest eavesdropper is closer than a certain distance, and if it does, it transmits
with power $P_u$, contributing to the aggregate interference. UEs are able to cancel a fraction $1 - \nu$ of their own jamming power.

We assume an interference-limited regime where all BSs transmit with the same power $P_b$. A Rayleigh fading environment is considered, hence the fading coefficients follow a unit mean exponential distribution. Also, it is assumed that all channel gains are independent from one another. A power law path loss model is considered where the signal power decays at the rate $r^{-\eta}$ with the distance $r$, where $\eta > 2$ is the path loss exponent and we use $\delta$ to denote $2/\eta$. In this chapter, we refer to the network described in this section as the actual network of interest.

### 4.2.2 Methodology of Analysis

A BS transmits a message at a rate of $R_t = \log(1 + \tau_1)$. According to Wyner’s encoding scheme [123], $R_s = [R_t - R_c]^+$ is the rate of secure communication, where $R_c = \log(1 + \tau_2)$ is the cost of securing the message. Hence, for $R_s > 0$, we require $\tau_1 > \tau_2$. The message can therefore be decoded securely at the tUE if the SIR exceeds $\tau_1$ (i.e., the event $S$ occurs). For the eavesdroppers, let $F$ be the event that the SIR at each eavesdropper is below $\tau_2$. Since $F$ is cumbersome to analyze, we focus on the event $F \supset F$ that the SIR at tEV is below $\tau_2$. The two main performance metrics are the probability of OSSA, i.e., $P(S \cap F)$, and the success probability of the tUE, $P(S)$. To characterize the effect of spatial interference correlation, we also derive the metrics assuming independent interference.

---

1The SIR at tEV is not necessarily the largest among all eavesdroppers due to interference and fading (and the BS does not have information about the CSI of the eavesdroppers). Thus, we have $P(S \cap F) > P(S \cap F)$, but the two probabilities are close thanks to the large-scale path loss.
4.3 SIR Analysis

4.3.1 Framework

By construction, the PP of UEs $\Phi_u$ is stationary but not Poisson since there is only one UE per cell. It is a soft-core process, where the likelihood of having two points very close is much smaller than in a PPP. To enable an analysis of the interference caused by FDJ, we use the approximation in [122], where the PP of UEs as seen from the tUE is a non-homogeneous (but isotropic) PPP with radial intensity function $\lambda(r) = \lambda g(r)$, where $g(r) \approx 1 - e^{-3\sqrt{\lambda}r}$ is the pair correlation function [10, Def. 6.6] of the PP of UEs. This way, the repulsion between the tUE and the interfering UEs is captured.

Since $\Phi_b$ is a PPP, the legitimate link distance $R$ is Rayleigh distributed with mean \(\frac{1}{2\sqrt{\lambda}}\) and $b = \frac{13}{10}$, where the correction factor is due to the fact that the tUE resides in the typical cell, not the Crofton cell. Hence we use $f_R(r) = 2b\pi \lambda r e^{-b\pi \lambda r^2}$, $r \geq 0$ (see [122], (12)). The orientation of a UE w.r.t. its BS, $\theta$, is uniform on $[0, 2\pi]$. In this chapter, we refer to the network where the UEs, as seen from the tUE, form a PPP with radial intensity function $\lambda(r) = \lambda \left(1 - e^{-3\sqrt{\lambda}r}\right)$ and link distance PDF $f_R$ as the analytical network of interest. A UE employs FDJ independently of other UEs with some probability $q$, which is the probability that the eavesdropper nearest to its serving BS lies closer to the UE than some threshold distance. The PP of jamming UEs, $\Phi_j$, is therefore a thinned version of $\Phi_u$ with intensity function $\lambda_j(r) = q\lambda \left(1 - e^{-3\sqrt{\lambda}r}\right)$. To characterize the effect of interference correlation as a function of the distance between tUE and tEV, we condition on tEV to be located at $v = (v, 0)$. It should be noted that in our analysis the tUE does not employ FDJ with probability $q$, instead we study its performance with and without FDJ. The performance of tUE if it uses FDJ with probability $q$ is easily obtained by averaging the performance with and without FDJ.
Let $\text{SIR}_u$ and $\text{SIR}_e$ be the SIRs at the tUE and tEV, respectively. The distance from the tUE’s BS to tEV is $\|\mathbf{R} - \mathbf{v}\|$, where $\mathbf{R} = (R \cos \theta, R \sin \theta)$.

\begin{align*}
\text{SIR}_u &= \frac{P_b h_0 R^{-\eta}}{P_b I_{bu} + P_u I_{uu} + 1 J P_u \nu}, \\
\text{SIR}_e &= \frac{g_0 \|\mathbf{R} - \mathbf{v}\|^{-\eta}}{I_{be} + \frac{P_b}{P_c} I_{ue} + 1 J \frac{P_u}{P_c} \tilde{g}_0 \nu^{-\eta}},
\end{align*}

(4.1) (4.2)

where $I_{bu} = \sum_{x \in \Phi_b} h_x \|x\|^{-\eta}$, $I_{uu} = \sum_{y \in \Phi_j} h_y \|y\|^{-\eta}$, $I_{be} = \sum_{x \in \Phi_b} g_x \|x - \mathbf{v}\|^{-\eta}$, and $I_{ue} = \sum_{y \in \Phi_j} g_y \|y - \mathbf{v}\|^{-\eta}$.

The fading coefficients from the tUE’s BS to itself and tEV are $h_0$ and $g_0$, respectively; while $h_x (h_y)$ and $g_x (g_y)$ are the fading coefficients from the BS located at $x$ (jamming UE located at $y$) to the tUE and tEV, respectively. If jamming at the tUE is active, the indicator function $1_J$ is 1, otherwise it is 0. The fading coefficient from the tUE to the tEV is $\tilde{g}_0$, and $\nu$ is the fraction of residual self-interference (RSI) experienced at the tUE because of imperfect cancellation of the jamming signal. We define the probability of successful reception at the tUE as $P(S) = P(\text{SIR}_u \geq \tau_1)$ and failure to extract information from the received message at tEV as $P(F) = P(\text{SIR}_e < \tau_2)$.

The normalized interference (i.e., assuming unit transmit power) from the BSs to the tUE (tEV) is denoted by $I_{bu}$ ($I_{be}$), and from the jamming UEs to the tUE (tEV) by $I_{uu}$ ($I_{ue}$). Additionally, we use the notations $\mathbf{x} = (x \cos \phi, x \sin \phi)$ and $\mathbf{r} = (r \cos \theta, r \sin \theta)$.

**Lemma 4.1:** In the analytical network of interest, conditioned on the link distance $R$, the LTs of the interferences are

$$
\mathcal{L}_{I_X}(s) = \exp \left( - \int_{\mathbb{R}^2} \frac{s}{s + \|z_X\|^\eta} \lambda_X(\mathbf{x}) d\mathbf{x} \right),
$$

(4.3)
for \( \chi \in \{\text{bu, be, uu, ue}\} \). The intensity functions \( \lambda_\chi(x) \) are:

\[
\lambda_{\text{bu}}(x) = \lambda_{\text{be}}(x) = \lambda (1 - 1\{x \in b(o, R)\})
\]

\[
\lambda_{\text{uu}}(x) = \lambda_{\text{ue}}(x) = q\lambda \left(1 - e^{-3\sqrt{q\lambda}}\right).
\]

For \( \chi \in \{\text{bu, uu}\} \), \( z_\chi = x \) and for \( \chi \in \{\text{be, ue}\} \), \( z_\chi = x - v \).

For the case of \( \mathcal{I}_{\text{bu}} \), the LT simplifies to

\[
\mathcal{L}_{\mathcal{I}_{\text{bu}}}(s) = \exp \left( -\frac{2\pi \lambda s}{(\eta - 2)R^\eta - 2} F_1 \left(1, 1 - \delta; 2 - \delta; -\frac{s}{R^\eta} \right) \right)
\]

\[
\eta = 4 = e^{-\pi \lambda \sqrt{s} \tan^{-1} (\sqrt{s})}.
\]

**Proof:** The LTs follow from the PGFL of the PPP.

### 4.3.2 Marginal Probabilities in the Analytical Network of Interest

The probabilities of the events \( S \) and \( F \) are

\[
\mathbb{P}(S) = \mathbb{P} \left( h_0 \geq \tau_1 R^\eta \left( \mathcal{I}_{\text{bu}} + \frac{P_u}{P_b} \mathcal{I}_{\text{uu}} + \mathbb{1}_J \frac{P_u}{P_b} \nu \right) \right)
\]

\[
= \int_0^\infty \mathcal{L}_{\mathcal{I}_{\text{bu}}} (\tau_1 R^\eta) L_{\mathcal{I}_{\text{uu}}} \left( \frac{P_u}{P_b} \tau_1 R^\eta \right) e^{-1 \frac{P_u}{P_b} \nu \tau_1 R^\eta} f_R(r) dr.
\]

\[
\mathbb{P}(F) = \mathbb{P} \left( g_0 < \tau_2 \|R - v\|^\eta \left( \mathcal{I}_{\text{be}} + \frac{P_u}{P_b} \mathcal{I}_{\text{ue}} + \mathbb{1}_J \frac{P_u}{P_b} \tilde{g}_0 v^{-\eta} \left\| \right\| \right) \right)
\]

\[
= 1 - \int_{\mathbb{R}^2} \mathbb{E}_{\tilde{g}_0} \left[ e^{-\frac{1}{\tau_2 \nu} \|R - v\|^\eta} \right] L_{\mathcal{I}_{\text{ue}}} \left( \frac{\|r - v\|^\eta}{\tau_2 \nu} \right) L_{\mathcal{I}_{\text{be}}} (\tau_2 \|r - v\|^\eta) f_R(\|r\|) \frac{2\pi}{2\pi} dr
\]

\[
= 1 - \int_{\mathbb{R}^2} L_{\mathcal{I}_{\text{ue}}} \left( \frac{\|r - v\|^\eta}{\tau_2 \nu} \right) L_{\mathcal{I}_{\text{be}}} (\tau_2 \|r - v\|^\eta) f_R(\|r\|) \frac{2\pi}{2\pi} dr.
\]
4.4 The Effect of Spatial Correlation

We are interested in the OSSA event $S \cap F$ and the event $S \mid F$ that the tUE has successful reception given that tEV is in SIR outage. The effect of interference correlation on OSSA is demonstrated by comparing the probabilities of these events when the spatial correlation is considered between tEV and tUE, and when it’s not.

**Theorem 4.1:** In the analytical network of interest, the OSSA probability is

$$P(S \cap F) = P(S) - \int_0^{2\pi} \int_0^\infty \frac{e^{-1} \frac{P_u}{P_b} \frac{\tau_1 r^\eta}{\nu} e^{-\lambda(B(r, \theta)+F(r, \theta))} f_R(r)}{2\pi} dr d\theta$$

(4.8)

where $P(S)$ is from (4.6) and

$$B(r, \theta) = \int_0^{2\pi} \int_r^\infty \left( 1 - \frac{(1 + \tau_1 r^\eta x^{-\eta})^{-1}}{1 + \tau_2 \|r-v\|^\eta \|x-v\|^\eta} \right) x dx d\phi$$

(4.9)

$$F(r, \theta) = \int_0^{2\pi} \int_0^\infty q \left( 1 - e^{-3\sqrt{\lambda} \|x\|} \right) \left( 1 - \frac{(1 + \frac{P_u}{P_b} \frac{\tau_1 r^\eta x^{-\eta}}{\nu})^{-1}}{1 + \frac{P_u}{P_b} \tau_2 \|r-v\|^\eta \|x-v\|^{-\eta}} \right) x dx d\phi.$$  

(4.10)

**Proof:** See Appendix [H] \[\square\]

**Corollary 4.1:** In the analytical network of interest, ignoring the fact that interference is correlated at the tUE and tEV, $P(S \cap F) = P(S)P(F)$ and $P(S \mid F) = P(S)$.

**Proof:** Assuming independent interference makes $S$ and $F$ independent. \[\square\]
Figure 4.1: Probabilities vs. $\tau_1$ (and $\tau_2$) at $v = \frac{1}{2\sqrt{\lambda}}$ for the network with jamming. Solid lines show the analysis when tUE employs FDJ ($\nu = -90$ dB) and dash-dotted lines show when it does not. Markers show the Monte Carlo simulations.

Figure 4.2: Probabilities vs. $v$ without FDJ at the tUE. Ignoring correlation results in $P(S \mid F) = P(S)$. 
Figure 4.3: Probabilities vs. $v$ with $\nu = -90$ dB. Ignoring correlation results in $P(S \mid F) = P(S)$. Solid lines represent the tUE employing FDJ and dash-dotted lines represent the tUE without FDJ.

4.5 Numerical Results

The intensity of the BSs and UEs used in this section is $\lambda = 10$ /km$^2$. We take the transmit powers to be $P_b = 1$ and $P_u = 0.2$, $\eta = 4$ and $\nu = -90$ dB. $\tau_2 = \tau_1 - 0.1$ and $q = 0.5$ for all of the figures, and in Figs. 4.2 and 4.3 we set $\tau_1 = 0$ dB.

Fig. 4.1 is a plot of the different probabilities for the analytical network of interest against $\tau_1$ (and $\tau_2$) when FDJ is employed at the tUE (solid lines) and when it is not (dash-dotted lines). Our assumptions are verified as the simulation for the actual network of interest matches the analysis well.

Fig. 4.2 is a plot of $P(S \cap F)$ and $P(S \mid F)$ without FDJ at the tUE as a function of $v$. For comparison, the probabilities obtained when interference correlation is ignored are also shown. We observe at smaller values of $v$ the difference between $P(S \cap F)$ when considering and ignoring interference correlation is significant. This highlights the impact on achievable secrecy when tEV lies near the tUE; not considering interference correlation would give overly optimistic estimates of the OSSA. As $v$ increases, the interference experienced at tEV and tUE decorrelates, and the two curves eventually meet. Further, both curves meet the curve of $P(S)$ at high $v$, as larger values of $v$
imply tEV is far and always in outage, thereby simplifying $\mathbb{P}(S \cap F)$ to $\mathbb{P}(S)$.

Fig. 4.3 is a plot of the probabilities of interest as a function of $v$ both when the tUE employs FDJ (solid lines) and when it does not (dash-dotted lines). It is observed that the curve for $\mathbb{P}(F)$ when the tUE employs FDJ has a dip. This occurs because at smaller values of $v$, FDJ from the tUE dominates the interference at tEV causing it to be in outage. As $v$ increases, the effect of jamming decreases when compared to the interference until the situation is reversed. When the tUE employs FDJ, $\mathbb{P}(S | F)$ (when interference correlation is considered) and $\mathbb{P}(S \cap F)$ reflect this trend. Additionally, due to RSI, for the case of FDJ at the tUE, all the probabilities that are impacted by $S$ are reduced.

In Fig. 4.3 we observe that at lower $v$ employing FDJ at the tUE significantly enhances secure communication. However, after a certain $v$, the values of $\mathbb{P}(S \cap F)$ and $\mathbb{P}(S | F)$ with FDJ at the tUE do not outperform the case without FDJ. This shows that FDJ at the tUE is useful for enhancing OSSA when the distance between the tUE and tEV is small. However, at larger values of $v$, the achievable secrecy with FDJ at the tUE is not larger than that without FDJ. This occurs because the FDJ signal from the tUE does not dominate the deterioration of the performance of tEV anymore at larger $v$. Employing FDJ after this distance is not useful as it increases network interference and is a waste of power; additionally, when $\nu$ is large enough to deteriorate $\mathbb{P}(S)$, it reduces the maximum achievable OSSA from the case without FDJ at the tUE.

Due to space constraints we do not include results for other values of $\nu$. With FDJ at the tUE, increasing $\nu$ reduces the probabilities of events that include $S$; the trends for the curves, however, remain the same. It ought to be mentioned that if $\nu$ is too high, the OSSA probability would be too low due to very large RSI, making the employment of FDJ at the tUE a hindrance to secure communication even at small $\nu$.
4.6 Conclusion

We analyze the probability of secure communication in the downlink of a Poisson cellular network when an eavesdropper lies near the UE under consideration. The UEs are assumed to be equipped with FDJ capability. The interference powers at the UE and eavesdropper are correlated since they lie nearby. To highlight the impact of the correlated interference, we compare the results with those obtained if interference was assumed to be independent. We consider two scenarios: with jamming and without jamming by the UE under consideration. We find that in both cases ignoring interference correlation overestimates the probability of secure communication, particularly at smaller distances between the eavesdropper and UE of interest. The numerical results show that there exists a critical distance between the legitimate receiver and eavesdropper after which jamming by the UE under consideration does not enhance the achievable secrecy. We conclude that jamming is an effective solution to combat the negative effect of interference correlation on secrecy when the eavesdropper lies close to the receiver and RSI is tolerable.

In the next chapter we introduce NOMA for large networks and emphasize the negative impact of intercell interference in such networks. We propose interference-aware NOMA designs in both the downlink and uplink for 5G networks.
Chapter 5

Non-Orthogonal Multiple Access for Large-Scale 5G Networks: Interference Aware Design

5.1 Introduction

NOMA is promoted as a key component of 5G cellular networks. As the name implies, NOMA operation introduces intracell interference (i.e., interference arising within the cell) to the cellular operation. The intracell interference is managed by careful NOMA design (e.g., user clustering and resource allocation) along with SuIC. However, most of the proposed NOMA designs are agnostic to intercell interference (i.e., interference from outside the cell), which is a major performance limiting parameter in 5G networks. This chapter sheds light on the drastic negative-impact of intercell interference on the NOMA performance and advocates interference-aware NOMA design that jointly accounts for both intracell and intercell interference. To this end, a case study for fair NOMA operation is presented and intercell interference mitigation techniques for NOMA networks are discussed. This chapter also investigates the potential of integrating NOMA with two important 5G transmission schemes, namely, full duplex and device-to-device communication. This is important since the ambitious performance defined by the 3rd Generation Partnership Project (3GPP) for 5G is foreseen to be realized via seamless integration of several new technologies and transmission techniques.
5.2 Interference-aware Design for NOMA Networks

There are different types of performance criteria (utility) that can be used to assess the performance of a network. The different utilities impose a trade-off between performance, optimization complexity, and signaling overhead. There are also different optimization frameworks that impose a trade-off between the value of the overall network utility and the fairness among UEs, as shown Table 5.1. As such, the first step in NOMA design is to define the utility function and the optimization objective, based on the operator key performance indicator.

5.2.1 Utility Function

The utility function is usually related to the transmission rate, which is a function of the SINR. Consider $N$ NOMA UEs per cell. The SINR associated with the $i^{th}$ strongest UE is determined as

\[ \text{SINR}_i = \frac{P_i F_i}{\sigma_i^2 + I_i + C_i} \]  

(5.1)

where $P_i$ is the transmit power related to the $i^{th}$ NOMA UE, $F_i$ is the fading (distance dependent large-scale and multi-path small-scale) power gain, $\sigma_i^2$ is the noise power, $I_i$ is the intercell interference power, and $C_i$ is the intracell interference power.

If interference is treated as noise, Shannon’s capacity specifies the maximum rate that can be reliably transmitted is

\[ T_i = \log(1 + \text{SINR}_i) \]

\[ = \log(1 + \frac{P_i F_i}{\sigma_i^2 + I_i + C_i}) \]  

(5.2)

Since the channel gains and interference may randomly change over time, the UE clustering, power allocation, and transmission rates should be continuously adapted.
to the instantaneous values of $F_i$, $I_i$, and $C_i$ for all $i \in \{1, \ldots, N\}$. In this case, the utility is defined in terms of the ergodic rate $T_{\text{erg}}^i = \mathbb{E}[T_i]$, which necessitates high overhead to feedback the channel gains and the interference related information for each UE. Since such overhead is not affordable in large-scale networks, the utility should be defined according to some fixed transmission rates that are vulnerable to outage. A per-UE fixed transmission rate can be selected as $\log(1 + \theta_i)$, which is subject to decoding outages if $\text{SINR}_i < \theta_i$. Fixed rate transmissions reduce the required signaling overhead at the expense of an effective-rate of

$$T_{\text{eff}}^i = \mathbb{P}\{\text{SINR}_i > \theta_i\} \log(1 + \theta_i), \quad (5.3)$$

which is less than the target-rate. Hence, UE clustering, power allocation, and target-rates should be carefully selected, according to interference and channel gain statistics, to maximize the effective-rates (cf. Fig. 1.1). These statistics can be obtained offline via stochastic geometry analysis [12]. A simpler approach is to define a global rate utility for all UEs (i.e., $\log(1+\theta)$) and focus on the UE clustering and power allocation only, which simplifies the network design. The aforementioned utility definitions, which are summarized in Table 5.1, lead to different utility formulations that impose a tradeoff between performance and optimization complexity.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>UE Utility function</th>
<th>Max. Total Utility</th>
<th>Max. Total Utility s.t. QoS</th>
<th>Max. Fairness</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static Channel</td>
<td>$T_i = \log(1 + \text{SINR}_i)$</td>
<td>$\max_P \sum_i T_i$</td>
<td>$\max_P \sum_i T_i$, s.t. $T_i &gt; \text{QoS}$</td>
<td>$\max_P \min_i T_i$</td>
<td>simple to study</td>
<td>impractical</td>
</tr>
<tr>
<td>Fading Channel</td>
<td>$T_{\text{erg}}^i = \mathbb{E}[T_i]$</td>
<td>$\max_P \sum_i T_i^\text{erg}$</td>
<td>$\max_P \sum_i T_i^\text{erg}$, s.t. $T_i^\text{erg} &gt; \text{QoS}$</td>
<td>$\max_P \min_i T_i^\text{erg}$</td>
<td>1) practical</td>
<td>high overhead (instantaneous CSI)</td>
</tr>
<tr>
<td>Adaptive Rate</td>
<td>$T_{\text{eff}}^i = \mathbb{P}{\text{SINR}_i &gt; \theta_i}$</td>
<td>$\max_P \sum_i T_i^\text{eff}$</td>
<td>$\max_P \sum_i T_i^\text{eff}$, s.t. $T_i^\text{eff} &gt; \text{QoS}$</td>
<td>$\max_P \min_i T_i^\text{eff}$</td>
<td>1) low overhead</td>
<td>1) outage prone</td>
</tr>
<tr>
<td>Fading Channel</td>
<td>$T_{\text{eff}}^i = \mathbb{P}{\text{SINR}_i &gt; \theta_i} \times \log(1 + \theta)$</td>
<td>$\max_P \sum_i T_i^\text{eff}$</td>
<td>$\max_P \sum_i T_i^\text{eff}$, s.t. $T_i^\text{eff} &gt; \text{QoS}$</td>
<td>$\max_P \min_i T_i^\text{eff}$</td>
<td>2) requires CSI stats</td>
<td>2) low utility</td>
</tr>
<tr>
<td>Fading Channel</td>
<td>$\tilde{T}_{\text{eff}}^i = \mathbb{P}{\text{SINR}_i &gt; \theta_i}$</td>
<td>$\max_P \sum_i \tilde{T}_i^\text{eff}$</td>
<td>$\max_P \sum_i \tilde{T}_i^\text{eff}$, s.t. $\tilde{T}_i^\text{eff} &gt; \text{QoS}$</td>
<td>$\max_P \min_i \tilde{T}_i^\text{eff}$</td>
<td>1) low overhead</td>
<td>1) outage prone</td>
</tr>
<tr>
<td>Global target-rate</td>
<td>$\max_P \sum_i \tilde{T}_i^\text{eff}$</td>
<td>$\max_P \sum_i \tilde{T}_i^\text{eff}$, s.t. $\tilde{T}_i^\text{eff} &gt; \text{QoS}$</td>
<td>$\max_P \min_i \tilde{T}_i^\text{eff}$</td>
<td>$\max_P \min_i \tilde{T}_i^\text{eff}$</td>
<td>2) requires CSI stats</td>
<td>2) low utility</td>
</tr>
</tbody>
</table>

Table 5.1: Summary of different utilities and objectives. We denote the vector of powers allocated to transmissions by $P$, and SINRs associated with target-rates by $\theta$, with $\theta = (\theta_1, \ldots, \theta_N)$.  


5.2.2 Optimization Objective

Having specified the utility function, the next step is to determine the optimization framework. In this chapter we consider three types of optimization frameworks:

**Total Utility Maximization**

The total utility maximization represents an extreme scenario for the optimization formulation, which does not prioritize individual per-UE rate but only the total cell rate. Such formulation allocates all the resources to the UE with the best channel, which maximizes SINR, and hence, the effective-rate. In this extreme case, all UEs other than the UE with the strongest channel have zero rates and both NOMA and OMA would have the same total utility.

**Total Utility Maximization with quality of service (QoS) constraint.**

A QoS constraint is added to the total utility maximization formulation to guarantee a minimum effective-rate for all UEs. In this case, the optimization framework will operate all but the UE with the best channel at the QoS constraint and allocate all of the remaining resources (e.g., power) to the best UE.

**Maximum fairness.**

Universal fairness among the UEs can be achieved via a max-min optimization framework which will achieve the maximum possible symmetric effective-rate for all UEs.

The different optimization frameworks are summarized in Table 5.1. Fig. 5.1 depicts a realization of each of these optimization frameworks in terms of the rate utility of each UE. As can be seen in the figure, the total utility maximization framework results in all resources being given to the strongest UE, i.e., UE1. Consequently it
Figure 5.1: Rate utility of NOMA for different scenarios from Table 5.1. UEs are labeled from the strongest channel (i.e., UE1) to the weakest (i.e., UE6).

has the highest rate utility while all other UEs have zero rate utility. The total utility maximization with QoS constraint framework on the other hand ensures all UEs meet the QoS constraint and then dedicates the remaining resources to UE1 to maximize its utility. In the maximum fairness framework all UEs are guaranteed identical (symmetric) effective-rate and the resources are distributed so that the maximum symmetric effective-rate is achieved. It ought to be noted that the total utility after applying the QoS constraints is lower than that of the unconstrained utility maximization which allocates all resources to UE1. Enforcing the symmetric effective-rate (i.e., maximum fairness) further reduces the total utility than both of the aforementioned schemes. After specifying the design objective, several design problems must be tackled based on the selected objective.

5.2.3 UE Clustering

One design problem concerns forming the NOMA UE clusters and determining their sizes. The goal is that UEs within the same cluster operate in NOMA within the same
time-frequency resource-block. Different clusters within the same BS are assigned orthogonal resources, and hence, intracell interference management is only required among UEs within the same cluster. However, clusters across different cells may share the same time-frequency resource block, which induces mutual intercell interference. This is illustrated in Fig. 5.2 which shows two different two-UE clusters operating on different frequency channels in the cell of interest. In particular, we show all the signals (intended and interfering) that the UEs on channel 1 receive and only the intended signal for the UEs on channel 2. Also shown are some nearby interfering BSs. As shown in the figure, the NOMA cluster in the cell of interest on frequency channel 1 does not receive interference from the signal of the serving BS to the cluster on channel 2. This cluster does, however, receive intercell interference from the signals of all other BSs on frequency channel 1. Also shown is the intracell interference from the signal intended for the other UEs in the NOMA cluster as the serving BS superposes the messages intended for all UEs in a cluster.

UE clustering is commonly based on sorting UEs in terms of their channel qualities \[46,48\] or link distance \[49,54\] (discussed in the next subsection). However, clustering should not only incorporate the UE’s local information (distance from BS, channel gain, target-rate, etc.), but also the intercell interference. Accounting for intercell interference can change the UEs’ effective channel qualities, and consequently their sorting. Intercell interference also affects clustering from a different perspective. Namely, clustering UEs that share the same dominant intercell interferer facilitates further interference management, which again highlights the importance of interference-aware design.

5.2.4 UE Sorting

SuIC is an important ingredient in NOMA for mitigating the intracell interference. The intracell interference, which arises within the cell, is comprised of the interference
Figure 5.2: A cell of interest in a downlink NOMA setup with $N = 2$ and two UE clusters operating on different frequency channels. The intended signal and interferences experienced at the UEs of the cluster operating on frequency channel 1, and intended signals at the UEs of the cluster on frequency channel 2 are shown.

Figure 5.3: UE order in a two-UE NOMA cellular network, where squares, green circles, and red circles denote the BSs, stronger UEs, and weaker UEs. The Voronoi tessellation represents the coverage region of each BS and the dotted links represents association. The BSs affected by interference-aware sorting are highlighted in blue.
from messages that can not be canceled in the SuIC chain as well as imperfections in SuIC leading to residual successive interference (SuI) from the canceled messages. The decoding order for SuIC in uplink and downlink NOMA is not the same. In particular, a UE in the downlink decodes the messages of all weaker UEs (i.e., UEs after itself in the ordering) before decoding its own message. In the uplink on the other hand, a message is decoded after all messages from stronger UEs (i.e., UEs before itself in the ordering) have been decoded. Hence intracell interference can be written in the downlink as

$$C_i = F_i \left( \sum_{j=1}^{i-1} P_j + \beta \sum_{j=i+1}^{N} P_j \right)$$ (5.4)

and in the uplink as

$$C_i = \sum_{j=i+1}^{N} P_j F_j + \beta \sum_{j=1}^{i-1} P_j F_j.$$ (5.5)

The factor $\beta$ reflects SuIC efficiency; perfect SuIC implies that $\beta = 0$. As mentioned in the last subsection, UEs are often sorted in terms of either their channel gain or link distance. However, to obtain good results, UEs must be sorted in descending order of a better channel quality indicator. One possible such indicator is the ratio

$$G_i = \frac{F_i}{I_i + \sigma_i^2},$$

which accounts for the UE’s channel taking into account the intercell interference.

Intercell interference $I_i$ can drastically change the UEs sorting as shown in the example in Fig. 5.3 where a network with two-UE NOMA clusters is considered. The UEs with better channel quality indicator in each cell are marked in green and the weaker UEs in red. In Fig. 5.3a UEs are sorted according to channel gain and hence do not take into account intercell interference. However, in Fig. 5.3b the UEs are ordered according to intercell-interference based channel quality. We observe here that the cells with BSs marked in blue have the order of their UEs reversed from that in Fig. 5.3a. This is attributed to the irregular cell structure which can cause UEs...
with the best channel gain to the serving BS to have the highest interference-to-noise ratio. Thus, intercell interference has a detrimental effect in UE sorting and must be taken into account.

5.2.5 Power/Rate Allocation

Due to UE-ordering based SuIC and corresponding intracell interference of the NOMA scheme (as can be seen from (5.4) and (5.5)), one should allocate transmission powers and rates such that:

- In the downlink, UE \( i \) can decode and cancel messages designated to UEs \( i + 1, \ldots, N \) and tolerate interference from UEs \( 1, \ldots, i - 1 \).

- In the uplink, the BS can decode the messages with the allocated rates successively from the message of UE 1 to that of UE \( N \), every time canceling the decoded message from its received signal.

This results in allocation of higher power and/or lower rate for weaker UEs in the downlink and stronger UEs in the uplink.

As a consequence of the use of SuIC in NOMA, the condition of successful decoding of a message is transformed into a joint condition wherein successful decoding of a message depends on successful decoding of the preceding messages in the SuIC chain. In particular, the coverage expression for the \( i^{th} \) UE can be written as follows

\[
\mathcal{P}_i = \mathbb{P} (\text{SINR}_{i,j} \geq \theta_j \ \forall j \in \mathcal{D}),
\]

where SINR\(_{i,j}\) is the SINR of the \( j^{th} \) message at the \( i^{th} \) decoder\(^1\), \( \theta_j \) is the target-SINR corresponding to the target-rate of the \( j^{th} \) message, and \( \mathcal{D} = \{1, \ldots, i\} \) for the

\(^1\)In the case of uplink there is only one receiver, namely the BS. This notation can be interpreted as the scenario where the BS is attempting to decode the \( i^{th} \) message and for this to be succesful all messages for UEs \( 1 < j < i \) also need to be decodable.
uplink and $\mathcal{D} = \{i, \ldots, N\}$ for the downlink. It ought to be noted that (5.6) is different from the coverage used in (5.3) as SuIC-based coverage results in a joint event. Additionally, the optimal power and rate allocation have to be done while taking intercell interference into account in order where resource allocation done based on an intercell-interference agnostic network results in very poor network performance.

5.3 Case Study

For the sake of exposition, we focus this section on the Max. Fairness utility (cf. Table 5.1) i.e., symmetric effective-rate for all UEs in the downlink and corresponding power allocation. In the uplink, we first show why employing power allocation that leads to symmetric effective-rates is challenging in a large network (detailed in Section 5.3.2). Full power transmissions, which do not lead to symmetric effective-rates, are then employed for the uplink. The impact of the interference on the rate utility and power allocation for the downlink and uplink scenarios are detailed. The operation scenario used in this case study and the rest of the chapter is detailed in Appendix I.

It ought to be highlighted that the operation scenario (in Appendix I) uses a global target-rate.

5.3.1 Impact of Intercell Interference on Rate Utility

Due to the SuIC, successful decoding of a message at the $i^{th}$ UE in the downlink is a joint event of decoding (for cancellations purpose) all messages for weaker UEs in addition to its own message. In the uplink, the BS is interested in all messages, and hence, successful decoding incorporates all messages starting with the strongest UE’s message. Consequently, the effective-rate associated with the $i^{th}$ NOMA UE is

$$\tilde{T}_i^{\text{eff}} = \log(1 + \theta)P_i$$

(5.7)
It should be mentioned that since the optimization framework considered in Appendix I is based on global target-rate, we have $\theta$ in (5.7), otherwise we would employ $\theta_i$ for calculating $T_{\text{eff}, i}$. Similarly, due to the global target-rate, $\theta_j = \theta$, $\forall j$ in (5.6). The intercell interference is usually overlooked in the literature, as only small-scale (i.e., single-cell) NOMA scenarios are considered [46–48]. However, since NOMA is foreseen to be deployed in 5G networks, which are intrinsically ultra-dense and interference limited, neglecting intercell interference is not a justifiable practice as highlighted in Fig. 1.1.

As specified before, Fig. 1.1 is a plot of the guaranteed-rate (i.e., minimum effective-rate according to (5.7) for all UEs) against the target-rate in the downlink with $N = 2$. It ought to be mentioned that since maximum fairness (i.e., symmetric effective-rate) is the objective of our case study, the effective-rate is identical for all UEs except when intercell interference is not accounted for in power allocation (detailed in the next subsection). The figure depicts the achievable gains in guaranteed-rate for a UE as intercell interference mitigation is improved. The upper-limit for this is the plot of full interference suppression, which, although impractical in reality, corresponds to no frequency reuse among BSs. Equivalently, it can be viewed as a single-cell setup. Fig. 1.1 thereby also emphasizes how studies on small-scale setups significantly overestimate actual guaranteed-rate and consequently network performance, stressing the importance of studying NOMA in the context of a real large-scale network.

### 5.3.2 Impact of Intercell Interference on Power Allocation and Fairness

Since signals are multiplexed in the power domain, NOMA requires power allocation for each message given a power constraint.
Downlink

In the downlink, a particular BS needs to transmit \( N \) messages using power \( P \). Interestingly, when all messages are sent at the same target-rate, unlike conventional power allocation schemes such as water filling strategies, downlink NOMA-UEs with poor channel conditions are allocated larger power than those with stronger channels \([45]\). This is done to ensure that a UE can treat the messages intended for UEs with stronger channels than itself as noise. Assuming the same target-rate for all messages, to achieve maximum fairness (i.e., symmetric effective-rate) we require a power allocation scheme that equalizes SINR for each message at its respective receiver. This will ensure that coverage will be identical and due to a global target-rate, the effective-rate will be symmetric.

Having a power allocation scheme that aims to achieve symmetric effective-rate without taking into account intercell interference in a large-scale network is far from optimal. In fact, a symmetric effective-rate is never achieved in this case because excluding intercell interference from the power allocation makes the SINR of stronger channels appear much larger than that of weaker channels. This results in a power allocation scheme that gives a lot more power to the weaker channels than it ought to. This is demonstrated in Fig. 5.4 where a two-UE downlink NOMA setup is considered. Intercell interference aware and agnostic power allocation are compared. We observe that intercell interference agnostic power allocation results in a 92% decrease in the allocated power of the stronger UE compared to its intercell interference aware counterpart, while the weaker UE has a 19% increase. In a large-scale downlink setup, the intercell interference agnostic approach deteriorates the performance of the UEs with stronger channels significantly as the impact of intercell interference dominates the impact of intracell interference and noise. Additionally, in a real network, such an intercell interference agnostic power allocation scheme is not able to provide fairness, i.e., a symmetric effective-rate, either. This in turn deteriorates the total effective
Figure 5.4: Power allocation for UEs in a downlink NOMA setup with $N = 2$ in the intercell interference aware and agnostic cases.

cell rate (TECR), which is the sum of the individual effective rates of all the UEs in the NOMA cluster of a cell. We use the TECR as a measure of network performance. Incorporating intercell interference in the power allocation results in a less dramatic difference between powers allocated to the stronger and weaker channels, and better network performance.

In Fig. 1.1, we observe that the intercell interference agnostic power allocation guarantees a much lower effective-rate than its interference aware counterpart. This is due to the stronger UE getting significantly lower power than its requirement causing it to become the bottleneck of the guaranteed effective-rate in the agnostic case. This highlights that the large impact of intercell interference on the SINRs and therefore correct power allocation. This intercell interference can be taken into account by large-system analysis using stochastic geometry.

**Uplink**

Although a more involved power allocation strategy can be used for uplink NOMA, it is difficult to equalize SINRs associated with each message (to achieve symmetric
effective-rate) at the BS in this case. This is because intercell interference in the uplink comes from transmitting UEs from interfering cells, which like the UEs in the cell of interest require a power allocation strategy of their own. Game-theoretic approaches to find solutions to this may be employed. We do not delve into such approaches in this chapter and propose some simpler alternatives. One solution would be to overestimate the intercell interference by assuming that the interfering UEs transmit using their full power and then do power allocation that attempts to equalize SINR based on the overestimated intercell interference. Another approach is to enforce all uplink NOMA-UEs to transmit with their maximum power as done in [55].

Power allocation based on overestimated intercell interference results in much worse network performance than full-power transmissions as shown in Fig. 5.5 where a two-UE uplink NOMA setup is considered. Dashed lines represent full-power transmissions and solid lines are for power allocation based on overestimated intercell interference. We observe that attempting to equalize SINR using overestimated intercell interference does not result in fairness, i.e., in symmetric effective-rates for the UEs. In fact, the weaker UE performs only a little better than the full-power transmission case, while the stronger UE does much worse. This deteriorates the overall performance without achieving symmetric effective-rates in the case of overestimated intercell interference. It can be explained by the fact that the overestimated intercell interference, although impacts both UEs, makes the power allocation tilt in favor of the weaker UE which is not able to improve performance as much even with the additional resource. This too highlights the significance of intercell interference and how its inaccurate estimation deteriorates network performance. Additionally, it emphasizes that accounting for intercell interference inaccurately is worse than inefficient resource allocation.
Figure 5.5: Total and individual UE effective-rates vs. target-rate for uplink-NOMA with $N = 2$. Dashed lines represent full-power transmission, solid lines represent power allocation based on attempted fairness when intercell interference is overestimated.

5.4 Intercell Interference Management

As highlighted in the previous sections and shown in Fig. 1.1, there is significant room for improving the NOMA performance via intercell interference management. Roughly, interference management techniques can be divided into two main categories: offline and online. Both techniques are briefly discussed in the sequel.

5.4.1 Offline Interference Management

The main advantage of the offline interference management techniques is that they require no operational overhead. However, the price paid is a rigid design that may lead to spectrum underutilization. Common examples of offline interference management techniques are

- **Frequency reuse**: the entire spectrum is partitioned and allocated to BSs such that neighboring BSs do not share the same set of channels [124]. Despite its fundamental role in previous generations of cellular networks, frequency reuse cannot be employed in 5G networks for the following reasons; i) it underutilizes
the spectrum in less loaded cells; and ii) there is no efficient frequency reuse scheme that avoids interference among neighboring BSs due to the irregular structure of cells, especially due to dense multi-tier topologies.

- **Fractional frequency reuse (FFR):** cells are partitioned into cell-center and cell-edge regions. The spectrum is divided into two main chunks, the cell-center chunk and the cell-edge chunk. The cell-center chunk is universally reused over all cell-center regions. The cell-edge chunk is allocated to cell-edge regions such that neighboring cells do not share common cell-edge channels. FFR leads to a better frequency utilization than conventional frequency reuse. However, the notion of cell-edge and cell-center should be based on the interference to noise ratio (INR), which depends on intercell interference, rather than the cell geometry.

- **Sectoring:** the service-area of each BS is divided into several sectors using directional antennas. FFR among the sectors can be applied such that neighboring BSs’ sectors do not share the same set of cell-edge channels. Due to irregular cell structure, it is hard to equalize the region covered by each sector while ensuring no cell-edge frequencies overlap among adjacent cells.

**NOMA specific challenges:** The offline interference management techniques assign frequencies to geographical areas which restricts the NOMA clustering process and deteriorates the multi-user diversity. Furthermore, there could be regions over (under) populated with UEs more (less) than the affordable cluster size, which may lead to UE blocking (resource underutilization).

### 5.4.2 Online Interference Management

Online interference management is conducted by means of BSs cooperation and/or coordination, which may involve high signaling overhead and increases the resource
allocation complexity. Some examples of online interference management techniques are:

- **Silencing**: this is a simple form of online interference management, where the serving BS sends a silencing request, on a designated frequency band, to dominant interfering BSs and/or UEs [129]. However, this leads to spectrum underutilization in the cells of the silenced interferers.

- **Cognitive spectrum access**: BSs are divided into spectrum owners and cognitive BSs. Cognitive BSs can reuse the spectrum via underlay, overlay, or interweave techniques [130]. In the underlay scheme, a cognitive BS should know the CSI at the primary receiver and operate subject to an interference constraint. In the overlay, the cognitive BS allocates some of its power to aid the primary transmission and the rest of the power to its own transmission. The interweave technique is an opportunistic spectrum access technique where the cognitive BS uses the idle channels of the primary BSs.

- **Cooperative beamforming**: BSs employ multiple antennas and are divided into clusters. BSs with the same cluster perform joint precoding and power allocation to align the mutual interference at the scheduled UEs [125,131–133]. Such a technique necessitates high signaling overhead for CSI feedback and imposes high resource allocation complexity.

- **Coordinated multi-point transmission (CoMP)**: UEs at the cell boundary between multiple BSs can be jointly served by these BSs in a network-MIMO fashion [127]. Such a technique significantly enhances the cell-edge UEs, however, at the expense of i) high signaling overhead for CSI; and ii) high backhaul utilization for sharing the UEs messages among the cooperating BSs.

**NOMA specific challenges**: The main challenge imposed by NOMA for online interference management is that the NOMA UEs may have different dominant inter-
cell interferers, which complicates the intercell interference management. Clustering the NOMA UEs that share common dominant intercell interference source may not be efficient from the intracell interference perspective. Balancing the tradeoff between intracell and intercell interference in NOMA clustering is a fundamental open problem. Determining which interferer to silence, for instance, can be challenging as decisions such as whether to silence the interferer that most deteriorates one receiver or silencing the interferer that is most detrimental to all receivers in a cluster in the downlink needs to be made. Similarly, in the uplink one needs to determine how many interferers from a cluster to silence and if this number can vary in different clusters. In other words, one needs to know which NOMA receiver/transmitter needs to be prioritized most. In the case of beamforming and CoMP interference management, the complexity of the precoding and resource allocation, along with the CSI signaling overhead, increases with the cluster size. In the cognitive spectrum access scenario, there are multiple primary UEs that should be simultaneously considered by cognitive NOMA UEs. Consequently, the spectrum sensing complexity increases with NOMA cluster size.

5.5 Integration of NOMA with Other 5G Technologies

5G defines a performance leap of 1000 times capacity when compared to current 4G networks. Such ambitious performance gain is foreseen to be fulfilled via integrating several key technologies \cite{134,135} including NOMA, FD, and D2D communication. This section highlights the opportunities and challenges for the integration of NOMA with the aforementioned technologies. Before delving into these details we specify the differences between our integration and what exists in the literature.

Often works that integrate the aforementioned technologies with NOMA simply employ them on different devices or at different times. They do not necessarily integrate the technologies in the sense of being deployed simultaneously by the same
Figure 5.6: Potentials of integrating NOMA with 5G technologies for $N = 2$ case.

devices. In particular, in [78] either NOMA or FD are employed in a time frame and the two can not be employed simultaneously. In [75,76] the BS transmits using NOMA and one of the receivers relays an older message in a FD but non-NOMA fashion. Similarly, in [83] D2D and NOMA are not integrated by the same devices since the cellular MU-MIMO employs NOMA while overlaid with simple D2D communication.

We are interested in studying the impact of deploying these technologies with NOMA simultaneously at the same devices. Studies that integrate the technologies in such a way include [58]. Here the BS deploys NOMA and FD simultaneously by communicating with uplink and downlink NOMA clusters. The uplink and downlink clusters, however, do not operate in FD. It should be noted that [58] focuses on a single-cell scenario thereby not studying the impact of integrating the two technologies in a large network. Similarly, NOMA is employed by D2D devices in [136], thereby integrating the two technologies at the same devices. Again, however, the study is on a single-cell setup, thereby not accounting for the impact of the integration in a large network. This large network aspect is what we discuss next.
Figure 5.7: TECR vs. target-rate for FD-NOMA with different interference suppression when $N = 2$.

5.5.1 Full-Duplex NOMA

FD communication allows a transmit-receive pair to communicate simultaneously on the same channel, i.e., in the same time-frequency resource block. Conventionally, FD communication was rendered infeasible due to the overwhelming self-interference. Thanks to the recent advances in digital and analog circuit design, sufficient self-interference cancellation to operate in FD mode is now viable.

Integrating NOMA with FD communication enables simultaneous uplink and downlink communication between the BS and all UEs within the NOMA cluster, which further improves the spectral utilization when compared to the standalone NOMA or FD scenarios. However, FD-NOMA communication increases the aggregate interference dramatically. Since the uplink and downlink operate on the same frequency band, FD-NOMA experiences the following sources of interference:

- **Intra-mode intercell interference**: downlink-to-downlink and uplink-to-uplink interference from other cells.

- **Inter-mode intercell interference**: downlink-to-uplink and uplink-to-downlink interference from other cells.
• **Intra-mode intracell interference:** downlink-to-downlink and uplink-to-uplink NOMA interference from transmitters within the same cluster in the same cell.

• **Inter-mode intracell interference:** uplink-to-downlink only, which arises from other uplink transmissions of UES within the same cluster.

• **Residual self-interference:** due to the imperfection of self-interference cancellation.

Fig. 5.6a plots TECR against target-rate for a two-UE NOMA setup. Traditional downlink NOMA (without FD) is compared against FD-NOMA. The figure highlights the potential of harvesting rate gains via FD-NOMA without any interference management. However, the above discussion highlights that the integration between FD and NOMA requires sophisticated interference management techniques to alleviate the aforementioned types of interference. Note that alleviating one source of interference may aggravate another. For instance, inter-mode intracell interference management may be facilitated by clustering UEs that are sufficiently separated across the cell. However, sufficiently separated UEs may have different dominant intercell interferers, which complicates the intercell interference management.

Fig. 5.7 is a plot of TECR against target-rate for the two-UE FD-NOMA setup with different interference suppression. We observe from the figure that suppressing intercell interference has the most significant impact on FD-NOMA performance as only a 30% intercell interference suppression results in significant gains in TECR. Techniques to mitigate intercell interference may require NOMA UEs to be clustered in order to share common strong interferers for instance. However, at the UEs, the inter-mode intracell interference impacts the SINR experienced significantly more than any other interference source. As observed in Fig. 5.7, suppressing intra-mode intracell interference (without any intercell interference suppression), which only im-
pacts the downlink, also results in approximately the same TECR improvement as 30% intercell interference suppression. This can be explained by the fact that the source of the inter-mode intracell interference, i.e., the UE cluster, occurs inside the cell thereby having the most detrimental impact on downlink performance. Suppressing this interference, by imposing a certain minimum distance between the UEs in a cluster for instance, is therefore of utmost importance for downlink performance. This sheds light on the need to prioritize dealing with interference sources; suitable inter-cell interference mitigation techniques can then be employed to improve the network performance further.

Despite the imposed interference challenges, we observe the potential of FD-NOMA to improve the rate from Figs. 5.6a and 5.7. Additionally, from Fig. 5.7 we can extrapolate that intercell interference is overall most detrimental to FD-NOMA network performance.

\section*{5.5.2 D2D NOMA}

D2D communication allows proximate UEs to bypass BSs and communicate directly in a peer-to-peer fashion. Such short range direct D2D communication is foreseen to relieve BS congestion, improve spatial spectrum utilization, reduce power consumption, and reduce latency \cite{21,68}. Integrating NOMA with D2D enables one-to-many (denoted as forward-D2D) and many-to-one (denoted as reverse-D2D) communication, which can further improve the D2D gains. To be specific, forward-D2D involves NOMA transmission from one D2D transmitter to a cluster of D2D receivers, while reverse-D2D involves NOMA transmission from a cluster of D2D transmitters to one D2D receiver. Note that D2D-UEs operate under a limited power budget when compared to BSs, and hence, the induced interference from D2D-NOMA communication can be affordable.

The D2D-NOMA communication can either share the uplink or downlink resources
of the cellular communicants. For the sake of exposition, we consider a D2D-NOMA framework sharing a downlink OMA cellular transmission. Additionally, we assume only one D2D-NOMA cluster in each cell on the resource-block being considered.

**Forward-D2D NOMA**

The forward-D2D NOMA experiences the following interference sources:

- **Cellular interference** from all cellular interferers, which can be divided into intracell and intercell according to the D2D cluster position.

- **Inter-cluster D2D interference** from transmitters of other D2D-NOMA clusters.

- **NOMA interference** from the messages designated to other D2D UEs within the same cluster.

**Reverse-D2D NOMA**

The reverse-D2D NOMA experiences the following interference sources:

- **Cellular interference** which can be divided into intracell and intercell according to the D2D cluster position.

- **Inter-cluster D2D interference** from transmitters of other D2D-NOMA clusters.

- **NOMA interference** from other D2D UEs within the same cluster.

Fig. 5.6b is a plot of TECR against target-rate. We plot both forward and reverse D2D-NOMA with cluster size $N = 2$. These are compared against forward and reverse D2D time division multiple access (TDMA), i.e., OMA, where the two-UE cluster shares time resources. For completeness, we also plot traditional downlink
cellular NOMA with cluster size $N = 2$. Despite the increased interference, Fig. 5.6b depicts the potential of harvesting rate gains from D2D-NOMA. We observe that the reverse D2D-NOMA always outperforms D2D-TDMA. However, forward D2D-NOMA outperforms the D2D-TDMA prior to the 5 dB, after which the case is reversed. This is because in the forward D2D-NOMA the transmitting D2D-UE splits its limited power budget among its receivers. In contrast, the reverse D2D-NOMA UEs use their full power budget for transmission. Hence, it is recommended to use TDMA in the forward D2D link if high target-rate is required.

Fig. 5.8 plots TECR against target-rate for the two-UE forward-D2D NOMA setup with different interference suppression. Suppressing cellular intracell interference (i.e., from the BS) at the D2D receivers is of utmost importance as it has the most significant impact on their performance. This is due to the high transmit power of BSs and small distance from the serving BS inside the cell. From the figure we observe that suppressing cellular intracell interference offers significant gains from the case without any interference suppression. Additionally, the severity of this interference is reflected in the fact that suppressing it leads to similar gains as an intercell interference suppression of 65% (compared to only a 30% intercell interference suppression in the FD-NOMA case), emphasizing the importance of managing it. Of course intercell interference still has the largest impact on TECR as its suppression leads to the largest gains. Suppressing cellular intracell interference may require techniques such as scheduling D2D receivers at a certain distance from the BS in the cell. Unlike the case of FD-NOMA, this complies with techniques to enhance intercell interference management by clustering the D2D UEs far from the BS for instance. Since intercell interference impacts the performance significantly as well, intercell interference management techniques should be employed in addition to cellular intracell interference management techniques.
Figure 5.8: TECR vs. target-rate for forward-D2D NOMA with $N = 2$ and different interference suppression.

Similarly for the case of reverse-D2D NOMA, the cellular intracell interference (i.e., from the BS inside the cell) to the D2D receiver has the most significant impact as the BS transmits at high power and is inside the cell. Additionally, although low power, the multiple transmitting D2D UEs cause significant interference at the cellular receiver as they may be large in number depending on the cluster size and lie inside the cell. Like its forward-D2D counterpart, interference between the intracell transmitters and receivers needs to be handled in addition to intercell interference. The results for this are similar to the forward-D2D NOMA case and are not included for the sake of exposition.

5.6 Summary

In this chapter we demonstrate the negative impact of intercell interference on NOMA performance in 5G networks. We also present an interference aware NOMA design that jointly accounts for intercell and intracell interference. Particularly, we discuss interference aware UE sorting along with the different interference aware design objectives and optimization frameworks. To this end, a case study for symmetric
NOMA transmission rate utility to maximize fairness in the downlink and full-power transmission in the uplink, is presented and interference mitigation techniques in the context of NOMA are highlighted. Last but not least, we consider the potentials and challenges of integrating NOMA with other 5G candidate technologies, namely, full-duplex and device-to-device communication, in a large network. The different kinds of interferences that arise, their impact, and how they can be handled are discussed.

The main recommendations we make in this chapter can be summarized as follows:

- Intercell interference must be accounted for to avoid significantly overestimating network performance (single-cell).

- Downlink NOMA power allocation should account for intercell interference to avoid drastically deteriorating performance.

- Uplink NOMA performs better using full-power transmissions rather than power allocation based on overestimated intercell interference. This highlights that accounting for intercell interference inaccurately damages performance significantly more than suboptimal resource allocation.

- Intercell interference awareness can change the UE sorting order to a more efficient one.

- Accounting for intercell interference can improve UE clustering which plays a role in improving the efficiency of interference management techniques.

- Intercell interference by far has the most detrimental impact on network performance. However, when NOMA is integrated with other technologies, certain receivers may be more severely impacted by another interference source and need to be handled accordingly.

  - In FD-NOMA the UEs are drastically impacted by inter-mode intracell interference. Techniques to manage this may contradict managing intercell
interference. However, handling this should be prioritized first followed by other methods to handle intercell interference.

– In D2D-NOMA the D2D receiver(s) are most impacted by the cellular intracell interference. Techniques to manage this also help mitigate intercell interference and should be used alongside other intercell interference management techniques.

All the above points represent interesting problems for future investigation, which can lead to practical design guidelines for NOMA systems.

This chapter focuses on emphasizing the importance of interference-aware NOMA design, in both the downlink and uplink. The next chapter, on the other hand, explicitly studies a large downlink NOMA network with a general NOMA cluster size. An analytical framework is developed in this context for three different UE clustering techniques, each with two different UE ordering techniques. Resource allocation algorithms are developed for two different network objectives.
Chapter 6

Downlink Non-Orthogonal Multiple Access (NOMA) in Poisson Networks

6.1 Introduction

A network model is considered where Poisson distributed BSs transmit to $N$ power-domain NOMA UEs each that employ SuIC for decoding. We propose three models for the clustering of NOMA UEs and consider two different ordering techniques for the NOMA UEs: mean signal power-based and instantaneous signal-to-intercell-interference-and-noise-ratio-based. For each technique, we present a signal-to-interference-and-noise ratio analysis for the coverage of the typical UE. We plot the rate region for the two-user case and show that neither ordering technique is consistently superior to the other. We propose two efficient algorithms for finding a feasible resource allocation that maximize the cell sum rate $R_{\text{tot}}$, for general $N$, constrained to: 1) a minimum throughput $T$ for each UE, 2) identical throughput for all UEs. We show the existence of: 1) an optimum $N$ that maximizes the constrained $R_{\text{tot}}$ given a set of network parameters, 2) a critical SuIC level necessary for NOMA to outperform orthogonal multiple access. The results highlight the importance in choosing the network parameters $N$, the constraints, and the ordering technique to balance the $R_{\text{tot}}$ and fairness requirements. We also show that interference-aware UE clustering can significantly improve performance. Additionally, we show that our algorithms are significantly faster than an exhaustive search, thereby, motivating their use.
6.2 System Model

(a) Model 1: sector corresponds to in-disk.  (b) Model 2: sector is half of the in-disk.

Figure 6.1: A realization of the network with $N = 3$ for Models 1 and 2. The UEs, in-disk (dashed circle) and sector (shaded) for the cell at $o$ are shown.

6.2.1 NOMA System Model

We consider a downlink cellular network where BSs transmit with a total power budget of $P = 1$. Each BS serves $N$ UEs in one time-frequency resource block by multiplexing the signals for each UE with different power levels; here $N$ denotes the cluster size. The BSs use fixed-rate transmissions, where the rate can be different for each UE, referred to as the transmission rate of the UE. Such transmissions lead to effective rates that are lower than the transmission rate; we refer to the effective rate of a UE as the throughput of the UE. The BSs are distributed according to a homogeneous PPP $\Phi$ with intensity $\lambda$. To the network we add an extra BS at the origin $o$, which, under expectation over the PPP $\Phi$, becomes the typical BS serving UEs in the typical cell. In this chapter we study the performance of the typical cell. Note that since $\Phi$ does not include the BS at $o$, $\Phi$ is the set of the interfering BSs for the UEs in the typical cell. Denote by $\rho$ the distance between the BS at $o$ and its
nearest neighbor. Since $\Phi$ is a PPP, the distance $\rho$ follows the distribution

$$f_\rho(x) = 2\pi \lambda xe^{-\pi \lambda x^2}, \quad x \geq 0. \quad (6.1)$$

Consider a disk centered at the $o$ with radius $\rho/2$. We refer to this as the in-disk as shown in Fig. 6.1. The in-disk is the largest disk centered at the serving BS that fits inside the Voronoi cell. UEs outside of this disk are relatively far from their BS, have weaker channels and thus are better served in their own resource block (without sharing) or even using CoMP transmission if they are near the cell edge [139, 140]. These UEs will not be discussed further in this chapter. We focus on UEs inside the in-disk since they have good channel conditions, yet enough disparity among themselves, and thus can effectively be served using NOMA.

We consider three models for the clustering of UEs. Each model results in a Poisson cluster process with $N$ points distributed uniformly and independently in each cluster. Let $x$ be the parent point, i.e., the BS, and $\rho_x$ the distance to its nearest neighbor $y_x$; for brevity, $\rho_o$ is denoted by $\rho$. The points in the cluster are:

- **Model 1**: distributed on the disk $b(x, \rho_x/2)$

- **Model 2**: distributed on the half-disk $s(x, \rho_x/2)$ such that all points in $s$ have distance at least $\rho_x$ from $y_x$.

- **Model 3**: distributed on the line segment $s(x, \rho_x/2) \cap \ell(x, y_x)$, where $\ell(x, y_x)$ is the line through $x$ and $y_x$.

More compactly, let $s(x, \rho_x/2, \phi) \subseteq b(x, \rho/2)$ be the (closed) disk sector of angle $\phi$ whose curved boundary has midpoint $z_x = (3x - y_x)/2$. Then for Model 1, $\phi = 2\pi$, for Model 2, $\phi = \pi$, and for Model 3, $\phi = 0$. A realization of the cell at $o$, its in-disk, and the surrounding cells are shown in Fig. 6.1; the sectors $s(x, \rho_x/2, \phi)$ are shown shaded for Models 1 and 2.
For Model 1, the union of all the disks \( \bigcup_{x \in \Phi} b(x, \rho_x/2) \) is the so-called Stienen model \([114]\). The area fraction covered by the Stienen model is 1/4. This means that if all users form a stationary point process, 1/4 of them are served using NOMA in Model 1 and 1/8 in Model 2 (and 0 in Model 3). More generally, for arbitrary \( \phi \), the area fraction is \( \phi/(8\pi) \). Note that the NOMA UEs form a Poisson cluster process where a fixed number of daughter points are placed uniformly at random on disks of random radii. The radii are correlated since the in-disks of two cells whose BSs are mutual nearest neighbors have the same radius and all disks are disjoint, but given the radii, the \( N \) daughter points are placed independently. Hence, there are three important differences to (advantages over) Matern cluster processes: the number of daughter points is fixed, the disk radius is random, and the disks do not overlap.

Focusing on the typical cell, the link distance \( R \) between a UE uniformly distributed in the sector of the in-disk \( s(o, \rho/2, \phi) \) with \( \phi > 0 \) and the BS at \( o \), conditioned on \( \rho \), follows

\[
f_{R|\rho}(r | \rho) = \frac{8r}{\rho^2}, \quad 0 \leq r \leq \frac{\rho}{2}.
\]

(6.2)

Since \( \phi > 0 \) for Models 1 and 2, the statistics of their link distances are according to (6.2). For Model 3, however, the sector becomes a line segment as \( \phi \to 0 \). Consequently, \( R \), conditioned on \( \rho \), in Model 3 follows the distribution

\[
f_{R|\rho}(r | \rho) = \frac{2}{\rho}, \quad 0 \leq r \leq \frac{\rho}{2}.
\]

(6.3)

**Remark 1:** Given \( \rho \), the exact distance between the UE and the interferer nearest to \( o \) in Model 3 is \( z = R + \rho \).

**Remark 2:** As there is no interfering BS inside \( b(o, \rho) \), a UE located at \( u \) in \( s(o, \rho/2, \phi) \), for any \( \phi \), is \( \rho - R \) away from the boundary of this disk. Hence, all three models guarantee that there is no interfering BS in \( b(u, \rho - R) \).
It makes sense to employ NOMA for UEs that have good channel conditions and thus can afford to share resource blocks with other UEs rather than those that cannot. Accordingly, any user close to a cell edge is worse off than the cell center users, on average. As $\phi$ decreases, users are located in the in-disk farther from any cell edge, particularly the edge closest to $o$, and consequently have better intercell interference conditions. In this context, Model 2 can be used as a technique to improve the performance by selecting UEs for NOMA operation, i.e., UE clustering, more efficiently based on their locations, and Model 3 can be viewed as a method to upper bound the achievable performance via UE clustering.

A Rayleigh fading environment is assumed such that the fading coefficients are i.i.d. with a unit mean exponential distribution. A power law path loss model is considered where the signal power decays at the rate $r^{-\eta}$ with distance $r$, where $\eta > 2$ denotes the path loss exponent and $\delta = \frac{2}{\eta}$. In this chapter we use $I^o$ to denote the intercell interference and $I^\circ$ to denote the intracell interference.

SuIC is employed for decoding NOMA UEs. According to the NOMA scheme, the power allocation and transmission rate are designed such that the $i^{th}$ strongest UE is able to decode the messages intended for all those UEs weaker than itself. This requires ordering of UEs based on the quality of the transmission link. We order UEs in such a way that the $i^{th}$ UE, referred to as UE$_i$, has the $i^{th}$ strongest transmission link. There are various ways to define what comprises the link quality. The link quality should include the effect of path loss (and therefore link distance), fading and intercell interference. The impact of the large-scale path loss is more stable and dominant than the fading effect which varies on a much shorter time scale. Additionally, accounting for intercell interference and fading necessitates very high feedback overhead. Since for small values of $N$ the path loss dominates the channel relative to fading, considering the quality of a channel to be based on the distance between a UE and its BS is often assumed to be a reasonable approximation
The link quality can be determined by ordering the UEs of the typical cell from strongest to weakest according to descending

- Mean signal power (MSP)\(^1\): this ignores fading and therefore orders UEs based on descending \( R^{-\eta} \). Equivalently it can be viewed as ordering based on ascending link distance \( R \).

- Instantaneous signal power (ISP): this includes fading and therefore orders UEs based on descending \( hR^{-\eta} \).

- Mean-fading signal-to-intercell-interference-and-noise ratio (MFS\( \text{INR} \)): this assumes channels with the mean fading gain of 1 in both the transmission from the serving BS and in the intercell interference and therefore orders UEs based on descending \( \frac{R^{-\eta}}{\sum_{x \in \Phi} \frac{||x-u||^{-\eta} + \sigma^2}{||x||}} \) where \( ||x|| \) and \( ||u|| \) are the locations of the interfering BSs and UE, respectively and \( \sigma^2 \) is the noise power.

- Instantaneous signal-to-intercell-interference-and-noise ratio (IS\( \text{INR} \)): this includes fading and therefore orders UEs based on descending \( \frac{hR^{-\eta}}{I^{-\eta} + \sigma^2} \).

Analyzing the SINR for ordering based on ISP and MFS\( \text{INR} \) is out of the scope of the work in this chapter. Hence, we analyze and compare the following two schemes:

- MSP-based: the UEs of the typical cell are indexed according to their ascending ordered distance \( R_i \); the \( i^{th} \) closest UE from \( o \) is referred to as UE\( _i \), for \( 1 \leq i \leq N \).

- IS\( \text{INR} \)-based: UEs of the typical cell are indexed with respect to their descending ordered IS\( \text{INR} \) \( Z_i^{\text{ISINR}} \); hence, UE\( _i \) has the \( i^{th} \) largest IS\( \text{INR} \), for \( 1 \leq i \leq N \).

\(^1\)It should be noted that this ordering is based on the total unit power transmission received at the UE.

\(^2\)Note that \( Z_i \) is equivalent to SINR\( _i^{\text{OMA}} \) in (6.28). We use the notation \( Z_i \) for brevity and to differentiate between the context it is being used in.
The power for the signal intended for UE\textsubscript{i} is denoted by \( P_i \), hence \( P = \sum_i P_i \).

To successfully decode its own message, UE\textsubscript{i} must therefore be able to decode the messages intended for all UEs weaker than itself, i.e., UE\textsubscript{i+1}, \ldots, UE\textsubscript{N}. This is achieved by allocating higher powers and/or lower transmission rates to the data streams of the weaker UEs. Correspondingly, UE\textsubscript{i} is not able to decode any of the streams sent to UEs stronger than itself, i.e., UE\textsubscript{1}, \ldots, UE\textsubscript{i−1} due to their smaller powers and/or higher transmission rates. Assuming perfect SuIC, the intracell interference experienced at UE\textsubscript{i} when decoding its own message, \( I_{i}^{c} \), is comprised of the powers from the messages intended for UE\textsubscript{1}, \ldots, UE\textsubscript{i−1}. Since in practice SuIC is not perfect, our mathematical model additionally considers a fraction \( 0 \leq \beta \leq 1 \) of RI from the UEs weaker than UE\textsubscript{i} in \( I_{i}^{c} \). When perfect SuIC is assumed, \( \beta = 0 \), while \( \beta = 1 \) corresponds to no SuIC at all. Additionally, UE\textsubscript{i} suffers from intercell interference, \( I_{i}^{\text{o}} \), arising from the power received from all the other BSs in the network, and noise power \( \sigma^2 \). For the NOMA network, \( 2N - 1 \) parameters are to be selected, namely \( N \) transmission rates and \( N - 1 \) powers. Note that MSP-based ordering of UEs is agnostic to intercell interference and fading; however, our resource allocation (choice of the \( 2N - 1 \) parameters) is not. For the case of IS\text{INR}-based ordering, both ordering and resource allocation are intercell interference- and fading-aware.

### 6.2.2 OMA System Model

We compare our NOMA model with a traditional OMA network where only one UE is served by each BS in a single time-frequency resource block. We focus on TDMA. For a fair comparison with the NOMA system, the BS serves \( N \) UEs distributed uniformly at random in (part of) the in-disk as in the NOMA setup according to the model being employed. Each TDMA message is transmitted using full power \( P = 1 \) for a duration \( T_i \). Without loss of generality, a unit time duration is assumed for a NOMA transmission and therefore \( \sum_i T_i = 1 \). Consequently, \( 2N - 1 \) parameters are
to be selected for the OMA network, namely $N$ transmission rates and $N - 1$ fractions of the time slot. We compare both MSP-based UE ordering and ISINR-based ordering for the OMA model, too.

6.3 SINR Analysis

6.3.1 SINR in NOMA Network

In the NOMA network, the SINR at UE$_i$ of the message intended for UE$_j$ in the typical cell for $i \leq j \leq N$ is

$$\text{SINR}^i_j = \frac{h_iR_i^{-\eta}P_j}{h_iR_i^{-\eta}\left(\sum_{m=1}^{j-1} P_m + \beta \sum_{k=j+1}^{N} P_k\right) + \sum_{x \in \Phi} g_{y_i} ||y_i||^{-\eta} + \sigma^2},$$

where $y_i = x - u_i$, $u_i$ is the location of UE$_i$, $h_i$ ($g_{y_i}$) is the fading coefficient from the serving BS (interfering BS) located at $o$ ($x$) to UE$_i$. The intracell interference experienced when UE$_i$ decodes the message for UE$_j$ is denoted by $I_{j,i}^\circ$. We use $I_i^\circ$ to denote $I_{i,i}^\circ$.

6.3.2 Laplace Transform of the Intercell Interference

We analyze the LT of the intercell interference at both the unordered UE and the UE ordered based on MSP. Upon taking the expectation over the BS PPP and the unordered UE’s (ordered UE’s) location, the UEs in the cell with the BS at $o$ (UE$_1$ to UE$_N$) become the typical unordered UEs (typical ordered UEs, from UE$_1$ to UE$_N$.).

**Lemma 6.1:** The LT of $I_i^\circ$ ($I_i^\circ$) at the typical unordered UE (ordered UE$_i$) conditioned on $R_i$ and $\rho$, where $u = \rho - R_i$ ($u_i = \rho - R_i$), in Model 1 is approximated
as

$$
\mathcal{L}_{I^o|R_{s},\rho}(s) \approx \exp \left( -\frac{2\pi \lambda s}{(\eta - 2) u^\eta - 2} 2F_1 \left( 1, 1 - \delta; 2 - \delta; \frac{-s}{u^\eta} \right) \right) \frac{1}{1 + s\rho^{-\eta}} \tag{6.4}
$$

\begin{align*}
\eta = & \frac{4}{e^{-\pi \lambda \sqrt{s} \tan^{-1} \left( \frac{\sqrt{s}}{u_i} \right)}} \frac{1}{1 + s\rho^{-\eta}} \tag{6.5}

\mathcal{L}_{I^o|R_{s},\rho}(s) \approx \exp \left( -\frac{2\pi \lambda s}{(\eta - 2) u_i^\eta - 2} 2F_1 \left( 1, 1 - \delta; 2 - \delta; \frac{-s}{u_i^\eta} \right) \right) \frac{1}{1 + s\rho^{-\eta}} \tag{6.6}

\eta = & \frac{4}{e^{-\pi \lambda \sqrt{s} \tan^{-1} \left( \frac{\sqrt{s}}{u_i} \right)}} \frac{1}{1 + s\rho^{-\eta}}. \tag{6.7}
\end{align*}

**Proof:** Let \( y = x - u \), where \( \|x\| \) and \( \|u\| \) are the locations of the interfering BSs and the UE, respectively. The fading coefficient from the interfering BS at \( x \) to the UE is \( g_y \). The intercell interference experienced at the UE is

$$
I^o = \sum_{x \in \Phi, \|x\| > \rho} g_y \|y\|^{-\eta} + \sum_{x \in \Phi, \|x\| = \rho} g_y \|y\|^{-\eta}. \tag{6.8}
$$

The first term of the LT accounts for the first term in (6.8) corresponding to the non-nearest interferers from \( o \) lying at a distance at least \( u \) (\( u_i \)) from the unordered UE (ordered UE\(_i\)). It is obtained from employing Slivnyak’s theorem, the PGFL of the PPP, and MGF of \( g_y \sim \exp(1) \). However, this does not include the BS at distance \( \rho \) from \( o \), which is accounted for by the second term in (6.8) using the MGF of \( g_y \). Denote by \( z \) the distance between this interferer and the typical UE. Then using the MGF of \( g_y \), the exact expression for the LT of the second term in (6.8) is

$$
\mathbb{E}_{z|\rho} \left[ (1 + sz^{-\eta})^{-1} \right].
$$

For simplicity we approximate it using the approximate mean of this distance. Since the average position of the typical UE distributed uniformly in the in-disk is \( o \), \( \mathbb{E}[z | \rho] \approx \rho \). \( \square \)

**Note:** The first term of the LT of \( I^o \) \((I^o) \) is pessimistic since the interference guard zone in our model \( u \) (\( u_i \)) is smaller than the actual one. For the second term, an exact
evaluation (by simulation) shows that the difference between $\mathbb{E}[z \mid \rho]$ and $\rho$ is less than 3.2%.

In the case of Model 2 the distance between the UEs and the interferer closest to $o$ is larger than in the case of Model 1. This corresponds to a change in the impact of the second term of (6.8). The LT of intercell interference changes accordingly.

**Lemma 6.2:** The LT of $I^o (I^o_i)$ at the typical unordered UE (ordered UE $i$) conditioned on $R (R_i)$ and $\rho$, where $u = \rho - R (u_i = \rho - R_i)$, in Model 2 is approximated as

$$
\mathcal{L}_{I^o | R, \rho}(s) \approx \exp \left(-\frac{2\pi \lambda s}{(\eta - 2)u^\eta} 2F_1 \left(1, 1-\delta; 2-\delta; \frac{-s}{u^\eta} \right) \right) \frac{1}{1 + s(1.25\rho)^{-\eta}} \quad (6.9)
$$

$$
\eta = 4 e^{-\pi \lambda \sqrt{\pi} \tan^{-1} \left( \frac{\sqrt{\pi}}{s} \right)} \frac{1}{1 + s(1.25\rho)^{-4}} \quad (6.10)
$$

**Proof:** The proof follows according to that in Lemma 6.1. However, in the second term, $\mathbb{E}[z \mid \rho] \approx 1.25\rho$. We use this approximation because a UE located in Model 2, i.e., in the half-disk away from the interferer nearest to $o$, has $\rho \leq \mathbb{E}[z \mid \rho] \leq 1.5\rho$; consequently, we approximate the average position of a UE in this model and $z$ accordingly. An exact evaluation (by simulation) for Model 2 shows that the difference between $\mathbb{E}[z \mid \rho]$ and 1.25$\rho$ is less than 0.92%.

In the case of Model 3 the distance between the UEs and the interferer closest to $o$ is exactly $z = R + \rho$. This too corresponds to a change in the impact of the second term of (6.8). The LT of intercell interference changes accordingly.

**Lemma 6.3:** The LT of $I^o (I^o_i)$ at the typical unordered UE (ordered UE $i$) distributed according to Model 3, conditioned on $R (R_i)$ and $\rho$, where $u = \rho - R$...
(u_i = \rho - R_i), a_1 = \frac{(1.5\rho)^\eta}{s}$ and $a_2 = \frac{\eta^\eta}{s}$, is approximated as

$$
L_{I|R,\rho}(s) \approx \exp \left( -\frac{2\pi \lambda s}{(\eta - 2)u_i^{\eta - 2}2F_1 \left( 1, 1 - \delta; 2 - \delta; \frac{-s}{u_i^\eta} \right)} \right) \left( 1 - 3 \left( \frac{1}{\eta}; \frac{\eta + 1}{\eta}; -a_1 \right) + 2 \left( \frac{1}{\eta}; \frac{\eta + 1}{\eta}; -a_2 \right) \right)
$$

(6.13)

$$
\eta = e^{-\pi \sqrt{s} \tan^{-1} \left( \frac{\eta}{\sqrt{s}} \right)} \times \left( 1 - \frac{-\tan^{-1} \left( \frac{\eta}{\sqrt{s}} \right) + \tanh^{-1} \left( \frac{\eta}{\sqrt{s}} \right)}{\frac{2 \pi a_1^2}{\eta a_1^2}} \right) + \frac{-\tan^{-1} \left( \frac{\eta}{\sqrt{s}} \right) + \tanh^{-1} \left( \frac{\eta}{\sqrt{s}} \right)}{\frac{2 \pi a_1^2}{\eta a_1^2}}.
$$

(6.14)

Proof: The first term of (6.13) and (6.15) follows from the first term of the LTs in Lemma 6.1. The exact second term is $E_{z |\rho} \left[ (1 + sz^{-\eta})^{-1} \right]$. Since $z = R + \rho$, using (6.3), $f_{z |\rho}(y | \rho) = 2/\rho$, $\rho \leq y \leq 3\rho/2,$

$$
E_{z |\rho} \left[ (1 + sz^{-\eta})^{-1} \right] = \int_{\rho}^{1.5\rho} \frac{1}{1 + sy^{-\eta}} f_{z |\rho}(y | \rho) dy
$$

$$
= 1 - 3 \left( \frac{1}{\eta}; \frac{\eta + 1}{\eta}; -a_1 \right) + 2 \left( \frac{1}{\eta}; \frac{\eta + 1}{\eta}; -a_2 \right).
$$

\[\square\]
6.3.3 UE Ordering

Since the order of the UEs is known at the BS, we use order statistics for the PDFs of the link quality. These are derived using the distribution of the unordered link quality statistics and the theory of order statistics [143].

**MSP-Based Ordering**

UEs are ordered based on the ascending ordered link distance $R_i$. Hence, $R_i$ is the distance between the $i^{th}$ nearest UE, i.e., UE$_i$ to its serving BS, given $\rho$, for $1 \leq i \leq N$.

Using the distribution of the unordered link distance $R$ conditioned on $\rho$ in (6.2) for Models 1 and 2 we have

$$f_{R_i | \rho}(r | \rho) = \binom{N - 1}{i - 1} \frac{8rN}{\rho^2} \left( \frac{4r^2}{\rho^2} \right)^{i-1} \left( 1 - \frac{4r^2}{\rho^2} \right)^{N-i} \tag{6.17}$$

for $0 \leq r \leq \rho/2$, where $\binom{c}{d} = \frac{c!}{d!(c-d)!}$.

Similarly, using the distribution of the unordered link distance $R$ conditioned on $\rho$ in (6.3) for Model 3 we have

$$f_{R_i | \rho}(r | \rho) = \binom{N - 1}{i - 1} \frac{N^2}{\rho} \left( \frac{2r}{\rho} \right)^{i-1} \left( 1 - \frac{2r}{\rho} \right)^{N-i} \tag{6.18}$$

for $0 \leq r \leq \rho/2$.

Note that MSP-based ordering guarantees that the nearest interfering BS from UE$_i$ is farther than $\rho - R_i$.

**ISÎNKR-Based Ordering**

UEs are ordered based on descending ordered ISÎNKR, $Z_i$. The unordered ISÎNKR is

$$Z = \frac{hR^{-\eta}}{I^0 + \sigma^2}.$$

**Lemma 6.4:** The CDF of the unordered ISÎNKR $Z$ conditioned on $\rho$ is approxi-
mated as

\[
F_{Z|\rho}(x) \approx 1 - \int_0^{\rho/2} \mathcal{L}_{I^o|R,\rho}(xr^{\eta}) \exp(-xr^{\eta}\sigma^2) f_{R|\rho}(r) dr,
\]  

(6.19)

where \( \mathcal{L}_{I^o|R,\rho}(s) \) is approximated in Lemmas 6.1, 6.2, and 6.3 for Models 1, 2, and 3, respectively, and \( f_{R|\rho}(r) \) is given in (6.2) for Models 1 and 2 and in (6.3) for Model 3.

**Proof:** By definition of \( Z \),

\[
F_{Z|\rho}(x) = \mathbb{E}_{R,I^o} \left[ \mathbb{P} \left( h \leq xR^o(I^o + \sigma^2) \mid R, I^o \right) \right] 
\]

\[
= \mathbb{E}_{R,I^o} \left[ 1 - \exp(-xR^oI^o) \exp(-xR^o\sigma^2) \right] 
\]

\[
\approx 1 - \int_0^{\rho/2} \mathcal{L}_{I^o|R,\rho}(xr^{\eta}) \exp(-xr^{\eta}\sigma^2) f_{R|\rho}(r) dr.
\]

Here (a) follows from \( h \sim \text{exp}(1) \) and (b) using the definition of the LT of \( I^o \) conditioned on \( R \) and \( \rho \).

**Corollary 6.1:** The CDF of the ordered IS\( \text{INR} \) \( Z_i \) conditioned on \( \rho \) is approximated as

\[
F_{Z_i|\rho}(x) \approx \sum_{k=N+1-i}^{N} \binom{N}{k} \left( F_{Z|\rho}(x) \right)^k \left( 1 - F_{Z|\rho}(x) \right)^{N-k},
\]  

(6.20)

where \( F_{Z_i|\rho}(x) \) is given in (6.19).

### 6.3.4 Coverage in NOMA Network

In order to decode its intended message, UE\(_i\) needs to decode the messages intended for all UEs weaker than itself. We use \( \theta_j \) to denote the SINR threshold corresponding to the transmission rate associated with the message for UE\(_j\). Coverage at UE\(_i\) is
defined as the event

\[ C_i = \bigcap_{j=i}^{N} \left\{ \text{SINR}_i^j > \theta_j \right\} = \bigcap_{j=i}^{N} \left\{ h_i > R_i^j (I_i^j + \sigma^2) \frac{\theta_j}{P_j} \right\}, \tag{6.21} \]

where \( \bar{P}_j = P_j - \theta_j \left( \sum_{m=1}^{j-1} P_m + \beta \sum_{k=j+1}^{N} P_k \right). \)

**Remark 3:** We observe that the impact of the intracell interference is that of a reduction in the effective transmit power to \( \bar{P}_j; \) without intracell interference, \( \bar{P}_j \) in (7.5) would be replaced by \( P_j. \) This reduction and thus \( \bar{P}_j \) is dependent on the transmission rate of the message to be decoded.

We introduce the notion of **NOMA necessary condition** for coverage, which is coverage when only intracell interference, arising from NOMA UEs within a cell, is considered. By definition we can write the signal-to-intracell-interference ratio (SIR) of the message for UE \( j \) at UE \( i \) as

\[ \text{SIR}_i^j = \frac{h_i R_i^j \bar{P}_j}{h_i \left( \sum_{m=1}^{j-1} P_m + \beta \sum_{k=j+1}^{N} P_k \right)} = \frac{P_j}{\sum_{m=1}^{j-1} P_m + \beta \sum_{k=j+1}^{N} P_k}. \tag{6.22} \]

From (6.22), the SIR of the message for UE \( j \) is independent of the UE (i.e., UE \( i \)) it is being decoded at; hence, it can be rewritten as \( \hat{\text{SIR}}_j \). In order for the message of UE \( j \) to satisfy the NOMA necessary condition for coverage, we require

\[ \hat{\text{SIR}}_j > \theta_j \iff \bar{P}_j > 0. \tag{6.23} \]

The above condition constrains the transmission rate for the message of UE \( j \) to be less than a certain value that is a function of the power distribution among the NOMA UEs. If this condition is not satisfied, the message for UE \( j \) cannot be decoded since \( \hat{\text{SIR}}_j \) is an upper bound on \( \text{SINR}_i^j, \ j \geq i \). Consequently UE \( i \) will be in outage as \( \bar{P}_j \) will not be positive. Note that for the particular case of UE \( 1 \) with perfect SuIC
(i.e., \( \beta = 0 \)), there is no intracell interference and \( \bar{\text{SIR}}_1 = \infty \) implying UE_1 always satisfies the NOMA necessary condition for coverage when SuIC is perfect; equivalently, when \( \beta = 0, \bar{P}_1 = P_1 \). Hence, failing to satisfy the NOMA necessary condition for coverage guarantees outage for all UEs that need to decode that message simply because the transmission rate is too high for the given power allocation. This shows the importance of resource allocation in terms of power allocation and transmission rate choice.

Using \( M_i = \max \frac{\theta_i}{P_j} \), \( C_i \) in (7.5) can be rewritten compactly as

\[
C_i = \{ h_i > \frac{R_i^0(I_i^0 + \sigma^2)M_i}{I_{\theta i}|R_i, \rho} \}.
\] (6.24)

Next, we derive the coverage probabilities for UEs using each ordering technique.

**Coverage for UEs Ordered Based on MSP**

**Theorem 6.1:** If \( \bar{P}_j > 0 \), the coverage probability of the typical UE_1 ordered based on MSP, is approximated as

\[
\mathbb{P}(C_i) \approx \int_0^{\infty} \int_0^{x/2} e^{-r^\theta \sigma^2M_i} \mathcal{L}_{R_i|R_i, \rho}(r^\theta M_i) f_{R_i|R_i, \rho}(r | x) \, dr \, f_\rho(x) \, dx,
\] (6.25)

where \( f_\rho(x) \) is given in (6.1), \( f_{R_i|R_i, \rho}(r | x) \) in (7.4) for Models 1 and 2 and in (6.18) for Model 3, and \( \mathcal{L}_{R_i|R_i, \rho} \) is approximated in Lemmas 6.1, 6.2, and 6.3 for Models 1, 2, and 3, respectively.

**Proof:**

\[
\mathbb{P}(C_i) = \mathbb{E}_\rho \left[ \mathbb{E}_{R_i} \left[ e^{-R_i^0 \sigma^2M_i \mathbb{E} \left[ e^{-R_i^0 M_i} | R_i, \rho \right]} \right] \right],
\]

as \( h_i \sim \exp(1) \). The inner expectation is the conditional LT of \( I_i^0 \) (given \( R_i \) and \( \rho \)).
From this we obtain (6.25).

Coverage for UEs Ordered Based on IS\(^8\)InR

**Theorem 6.2:** If \(\tilde{P}_j > 0\), the coverage probability of the typical UE\(_i\) ordered based on IS\(^8\)InR, is approximated as

\[
P(C_i) \approx \int_0^\infty (1 - F_{Z_i|\rho}(M_i \mid x)) f_{\rho}(x)dx,
\]

where \(f_{\rho}(x)\) is given in (6.1).

**Proof:** (6.26) follows using \(P(C_i) = P(Z_i > M_i)\).

For a given SINR threshold \(\theta_i\), corresponding to a transmission (normalized) rate of \(\log(1 + \theta_i)\), the throughput of the typical UE\(_i\) is

\[
R_i = P(C_i) \log(1 + \theta_i).
\]

The cell sum rate is \(R_{\text{tot}} = \sum_{i=1}^N R_i\).

**6.3.5 OMA Network**

The SINR for UE\(_i\) of the typical cell is

\[
\text{SINR}_{\text{OMA}}^i = \frac{h_i R_i^{-\eta}}{\sum_{x \in \Phi} g_{y_i} \|x - u_i\|^{-\eta} + \sigma^2},
\]

where \(u_i\) is the location of UE\(_i\), \(h_i\) \((g_{y_i})\) is the fading coefficient from the serving BS (interfering BS) located at \(o\) \((x)\) to UE\(_i\). Coverage at UE\(_i\) is defined as \(\tilde{C}_i = \{\text{SINR}_{\text{OMA}}^i > \theta_i\}\).

**Lemma 6.5:** In the OMA network, the coverage probability of the typical UE\(_i\)
ordered based on MSP is approximated as

\[ P(\tilde{C}_i) \approx \int_0^{\infty} \int_0^{x/2} e^{-\theta_i r \sigma^2} \mathcal{L}_{I_i | R_i, \rho}(r | x) f_{R_i | \rho}(r | x) dx f_{\rho}(x) dx, \]  \hspace{1cm} (6.29)

where \( f_{\rho}(x) \) is given in (6.1), \( f_{R_i | \rho}(r | x) \) in (6.4) for Models 1 and 2 and (6.18) for Model 3, and \( \mathcal{L}_{I_i | R_i, \rho} \) is approximated in Lemmas 6.1, 6.2, and 6.3 for Models 1, 2, and 3, respectively.

**Proof:** Using the exponential distribution of \( h_i \) and the LT of \( I_i^o \) conditioned on \( R_i \) and \( \rho \) we obtain (6.29).

**Lemma 6.6:** In the OMA network, the coverage probability of the typical UE \( i \) ordered based on IS\( ^{\text{INR}} \) is approximated as

\[ P(\tilde{C}_i) \approx \int_0^\infty \left( 1 - F_{Z_i | \rho}(\theta_i | x) \right) f_{\rho}(x) dx, \]  \hspace{1cm} (6.30)

where \( F_{Z_i | \rho}(y | \rho) \) is given in (6.19) and \( f_{\rho}(x) \) in (6.1).

**Proof:** (6.30) follows from \( P(\tilde{C}_i) = P(Z_i > \theta_i) \).

Denote by \( T_i \) the fraction of the time slot allotted to \( U_E_i \). For a given SINR threshold \( \theta_i \) and corresponding transmission (normalized) rate \( \log(1 + \theta_i) \), the throughput of the typical UE \( i \) is

\[ \mathcal{R}_i = T_i P(\tilde{C}_i) \log(1 + \theta_i). \]  \hspace{1cm} (6.31)

### 6.4 NOMA Optimization

#### 6.4.1 Problem Formulation

Determining the optimization objective of the NOMA framework can be complicated. The objective of NOMA is to provide coverage to multiple UEs in the same time-frequency resource block. Naturally we are interested in maximizing the cell sum rate.
It is well known that the cell sum rate is maximized by allocating all resources (power in the NOMA network) to the best UE [144]. However, this comes at the price of a complete loss of fairness among NOMA UEs, which is one of the main motivations behind serving multiple UEs in a NOMA fashion. Hence, we constrain the objective of maximizing cell sum rate to ensure multiple UEs are served. In addition to the power constraint we consider two kinds of constraints: 1) constraining resources so that each of the typical UEs achieves at least the TMT, 2) constraining resources so that the typical UEs achieve symmetric (identical) throughput. Formally stated, these objectives are:

- $P_1$ - Maximum cell sum rate subject to the TMT $T$:

$$\max_{(P_i, \theta_i)_{i=1,...,N}} \mathcal{R}_{tot} \quad \text{subject to} \quad \sum_{i=1}^{N} P_i = 1 \text{ and } \mathcal{R}_i \geq T.$$  

Because this problem is non-convex, an optimum solution, i.e., choice of $P_i$ and $\theta_i$ that result in the maximum constrained $\mathcal{R}_{tot}$, cannot be found using standard optimization methods. However, from the rate region for static channels we know that a resource allocation that results in all UEs achieving the TMT $T$, while all of the remaining power being allocated to the nearest UE, i.e., UE$_1$, to maximize its throughput is the optimum solution for that problem. An example of this for the two-user case is presented in [57].

- $P_2$ - Maximum symmetric throughput:

$$\max_{(P_i, \theta_i)_{i=1,...,N}} \mathcal{R}_{tot} \quad \text{subject to} \quad \sum_{i=1}^{N} P_i = 1 \text{ and } \mathcal{R}_1 = \ldots = \mathcal{R}_N.$$  

This is equivalent to maximizing the smallest UE throughput. Solving this results in a resource allocation that achieves the largest symmetric throughput (universal fairness), i.e., $\mathcal{R}_1 = \ldots = \mathcal{R}_N$. Since this problem is also non-convex,
an optimum solution cannot be found using standard optimization methods.

**Remark 4:** The maximum symmetric throughput is the largest TMT that can be supported.

**Remark 5:** Due to outage, the typical UEs that achieve the same throughput ($R_i$) do not necessarily have the same individual transmission rates (and corresponding $\theta_i$’s).

The same objectives hold for OMA networks. The constrained resource allocated to the UEs, however, is time for TDMA instead of power for NOMA, i.e., $\sum_i T_i = 1$. The OMA UEs enjoy full power in their transmissions. Optimization over transmission rate is done similarly to NOMA.

### 6.4.2 Case Study: $N = 2$

In this subsection we focus on the two-user case for which we can plot the maximum throughput for each UE subject to any power distribution for NOMA. This gives us the rate regions for the $N = 2$ scenario as shown in Figs. 6.2 and 6.3. We use Model 1, $\lambda = 10$, $\sigma^2 = -90$ dBm, $\eta = 4$ in this subsection. The rate regions in Fig. 6.2 are using different $\beta$ values and MSP-based ordering, while Fig. 6.3 uses both MSP- and ISINR-based UE ordering with perfect SuIC (i.e., $\beta = 0$). Since the OMA scheme employed is TDMA, the resource allocation in this case is not in terms of power but of time. We use the optimal $\theta_i$ for a given power (time) distribution between the two NOMA (TDMA) UEs.

In the rate regions in Figs. 6.2 and 6.3 each point on the curve is obtained from optimal transmission rate allocation that maximizes throughput given a power (time) distribution for the two NOMA (TDMA) UEs. A zero throughput of UE$_1$ (the intersection of the curves with the y-axis) corresponds to all the power being allocated to UE$_2$ in the case of NOMA and all the time being allotted to UE$_2$ in the case of TDMA and vice versa for zero throughput of UE$_2$ (the intersection of the curves with
Figure 6.2: Rate region for NOMA and TDMA with MSP-based UE ordering for Model 1 using different $\beta$ and $N = 2$.

Figure 6.3: Rate region for NOMA and TDMA with MSP and ISINR-based UE ordering for Model 1 with $\beta = 0$ and $N = 2$.

Figure 6.4: Optimum cell sum rate and individual UE throughputs vs. $P_1$ for NOMA with MSP and ISINR-based ordering for Model 1 with $\beta = 0$ and $N = 2$.

The rest of the points in each NOMA curve (TDMA curve) are made of all possible power (time) distributions between the two UEs. Each curve is the boundary of the corresponding rate region. Optimal resource allocation allows us to operate on the boundary of the rate region. This sort of graph also reveals what areas of throughput operation result in higher cell sum rate given a TMT constraint on the UEs. Additionally, if a symmetric throughput is required, the rate region shows us the maximum throughput possible. Obtaining the rate region for larger $N$, however, is impractical as it requires exhaustively going through the $2N - 1$ parameters for resource allocation.
With perfect SuIC (i.e., $\beta = 0$), if resource allocation is optimum, i.e., if we operate at the boundary of the rate region, NOMA outperforms TDMA for both the symmetric-throughput objective ($P_2$) and given any TMT ($P_1$) as shown in Figs. 6.2 and 6.3. In Fig. 6.2 we observe that increasing $\beta$ deteriorates performance by pushing the boundary of the rate region inward. Also, if $\beta$ is too high, with optimum resource allocation, TDMA always outperforms NOMA. Additionally, the rate region graphs shed light on the importance of resource allocation; suboptimum resource allocation can result in significant deterioration in performance as one could be operating inside the rate region far from the boundary. Thus, appropriate resource allocation is very important to fully exploit the potential of NOMA.

In Fig. 6.4 we plot the optimum cell sum rate and individual UE throughput for $N = 2$ against increasing $P_1$ (decreasing $P_2$) for NOMA with the two UE ordering techniques. Intuitively, UE ordering that incorporates more information about the channel is more accurate and should result in superior performance given any power distribution. Accordingly, one may anticipate that ISI\(\text{INR}\)-based ordering, which takes into account path loss, fading, intercell interference and noise, to always be superior in terms of $R_{\text{tot}}$ to MSP-based ordering, which only accounts for path loss. Contrary to this expectation, we observe that ISI\(\text{INR}\)-based ordering is not always superior. In particular, below a certain $P_1$, MSP-based ordering outperforms ISI\(\text{INR}\)-based ordering in terms of cell sum rate. ISI\(\text{INR}\)-based ordering exceeds in performance when $P_1$ is increased beyond this. In Fig. 6.3 we observe that this holds for TDMA, too. This occurs because:

1. ISI\(\text{INR}\)-based ordering does, in fact, incorporate more information about the channel; the weakest (strongest) UE in this case on average is weaker (stronger) than the weakest (strongest) UE in MSP-based ordering. This can be seen in Fig. 6.4 for $N = 2$ where the weak (strong) UE of ISI\(\text{INR}\)-based ordering consistently underperforms (outperforms) its MSP counterpart. This applies to
both NOMA and OMA as it depends on the UE ordering.

2. Additionally for NOMA, which employs SuIC, UE\_N is unable to cancel SuI for the messages of any other UE and therefore suffers the largest intracell interference. In the case of IS\_INR-based ordering, unlike its MSP counterpart, UE\_N may not necessarily be the farthest UE from the BS making the impact of intracell interference larger; this further deteriorates the SINR and therefore the throughput of the IS\_INR-based UE\_N.

Hence, when P\_1 is small in Fig. 6.4, the impact on R\_tot of the larger R\_2 in MSP-based ordering is more significant than the impact on R\_tot of the larger R\_1 in IS\_INR-based ordering. When P\_1 increases the impact of the significance is reversed.

From Fig. 6.3 we observe that for higher TMT values (including the symmetric throughput), MSP-based ordering outperforms IS\_INR-based ordering in terms of R\_tot. This will become obvious in the results section as well.

### 6.4.3 Algorithm for Solving P1

Since standard optimization techniques cannot be employed for any N, the optimum solution to P\_1 can only be found exhaustively by searching over all combinations of power and transmission rate for each of the N NOMA UEs. This, however, is an extremely tedious approach, particularly as N increases. In this subsection we propose an efficient algorithm that, while not guaranteeing an optimum solution, finds a feasible solution, i.e., a solution that satisfies the constraints (but there is no guarantee that the cell sum rate is close to the global maximum).

Given a certain power, UE\_1 is able to achieve a larger throughput from this resource than any other UE. It therefore makes sense to solve P\_1 by first ensuring that all UEs other than the strongest achieve TMT with the smallest powers possible. This will leave the largest P\_1 for UE\_1. UE\_1 can then maximize the cell sum rate
by maximizing $R_1$ with this power by finding the appropriate transmission rate. In other words, our algorithm for $P_1$ solves

$$\max_{(P_i, \theta_i)_{i=1,\ldots,N}} R_1 \quad \text{subject to} \quad \sum_{i=1}^{N} P_i = 1 \text{ and } R_j = T, \quad 2 \leq j \leq N.$$  

We tackle this problem by decoupling the choice of power and transmission rate; our algorithm finds the minimum possible power and corresponding smallest transmission rate$^3$ that achieve $T$ for UE$_2$ to UE$_N$ and allocates the remaining power to UE$_1$. UE$_1$ then optimizes its transmission rate (and therefore $\theta_1$) with the remaining power to maximize its throughput. If a UE cannot attain $T$, the available power is insufficient and the algorithm is unable to find a feasible solution as the cluster size $N$ is too large to attain this TMT for all UEs. This can be remedied by either decreasing $T$ or $N$. Formally, we state the working of the algorithm in Algorithm $\text{[1]}$.

Since SuIC requires knowledge of resource allocation for the weaker UEs in the decoding chain, we start with UE$_N$ in line 1. Note that the range for transmission rate is $\theta_i \geq 0$; to make our search finite, our algorithm searches in the range $\theta_{LB} \leq \theta_i \leq \theta_{UB}$. In lines 3 to 16, starting with zero power, we search for the smallest corresponding transmission rate, starting with $\theta_{LB}$ and increasing in steps of $\Delta \theta$, that can meet the TMT constraint. For a given $P$, if no $\theta$ is found until $\theta_{UB}$ that achieves TMT, the power is increased in steps of $\Delta P$. Once the minimum power that can meet the TMT is found, this power and the corresponding minimum transmission rate is saved. We then move on to the next weakest UE, using the stored power and transmission rate for the stronger UEs. This process is repeated for the $N - 1^{th}$ strongest UE to UE$_2$. If the TMT cannot be met for UE$_i$, $i \in \{2, \ldots, N\}$, even when

$^3$For an $i \in \{1, \ldots, N\}$, the function $R_i(\theta_i)$ is monotonically increasing from zero and then monotonically decreasing to zero, with a unique maximum at a finite $\theta_i > 0$. This is because for small $\theta_i$, $P(C_i)$ is close to 1, hence $R_i$ increases linearly with $\log(1 + \theta_i)$, while for large $\theta_i$, $P(C_i)$ goes to zero more quickly than $\log(1 + \theta_i)$ grows. Hence, each $R_i$ (except the maximum) can be satisfied by two $\theta_i$ values. We select the smaller value since it increases the coverage probability for all UEs that are to decode the $i^{th}$ message.
Algorithm 1 Resource allocation of a feasible solution to $P_1$

1: Begin with $\text{UE}_N$, $i = N$, $P = [\ ]$, $\theta = [\ ]$, $\mathcal{R} = [\ ]$

2: while $i > 0$ do

3: if $i > 1$ then

4: for $P_i = 0 : \Delta_P : 1 - \sum_{k=i+1}^{N} P_k$ do

5: for $\theta_i = \theta_{LB} : \Delta_\theta : \theta_{UB}$ do

6: Calculate $\mathcal{R}_i$ using (6.27) with (6.25) for MSP-based (with (6.26) for ISINR-based) UE ordering

7: if $\mathcal{R}_i \geq \mathcal{T}$ then

8: Update: $P = [P_i; P]$; $\theta = [\theta_i; \theta]$; $\mathcal{R} = [\mathcal{R}_i; \mathcal{R}]$; $i = i - 1$

9: Go to 2

10: end if

11: end for

12: if $P_i = 1 - \sum_{k=i+1}^{N} P_k$ then

13: TMT cannot be met for all UEs; exit

14: end if

15: else

16: end if

17: $P_1 = 1 - \sum_{k=2}^{N} P_k$

18: if $P_1 > 0$ then

19: $\mathcal{R}_{1}^{\text{vec}} = [\ ]$

20: for $\theta_1 > 0$ do

21: Calculate $\mathcal{R}_1$ using (6.27) with (6.25) for MSP-based (with (6.26) for ISINR-based) UE ordering

22: Update $\mathcal{R}_{1}^{\text{vec}} = [\mathcal{R}_1^{\text{vec}}; \mathcal{R}_1]$

23: end for

24: Update: $\mathcal{R}_1 = \max(\mathcal{R}_1^{\text{vec}})$ and corresponding $\theta_1$

25: if $\mathcal{R}_1 \geq \mathcal{T}$ then

26: Update: $P = [P_1; P]$; $\theta = [\theta_1; \theta]$; $\mathcal{R} = [\mathcal{R}_1; \mathcal{R}]$; $i = 0$

27: Go to 2

28: else

29: TMT cannot be met for all UEs; exit

30: end if

31: else

32: TMT cannot be met for all UEs; exit

33: end if

34: end if

35: end while

all of the remaining power $1 - \sum_{k=i+1}^{N} P_k$ is allocated to it, the power budget is not sufficient for the current TMT and we exit the algorithm in line 13. Otherwise, the throughput achieved in line 7 is $\mathcal{R}_i = \mathcal{T}$. After the minimum required powers to achieve the TMT have been allocated to $\text{UE}_2, \ldots, \text{UE}_N$, we use the remaining power in lines 17 to 34 to maximize $\mathcal{R}_1$ by finding the appropriate $\theta_1$. If the remaining power is 0 or if $\mathcal{R}_1 < \mathcal{T}$ in lines 32 and 29, respectively, the TMT cannot be met for all UEs and we exit the algorithm.

The same algorithm is employed for OMA, except that throughputs are calculated using using (6.31) with (6.29) for MSP-based (with (6.30) for ISINR-based) UE or-
dering, and the contending resource is \( T \) instead of \( P \). Note that since our problem includes intercell interference, our resource allocation is intercell interference-aware.

We compared the solutions of Algorithm 1 with those found using an exhaustive search for the case of \( N = 2 \) and different values of \( T \). It turns out that for \( N = 2 \) our solutions are, in fact, optimum. It is of course not possible to compare the results of Algorithm 1 with those from an exhaustive search for larger \( N \) as it is computationally too expensive.

6.4.4 Algorithm for Solving \( \mathcal{P}_2 \)

Since \( \mathcal{P}_2 \), like \( \mathcal{P}_1 \), is non-convex, the optimum solution cannot be found using standard optimization techniques. As mentioned in the previous subsection, doing an exhaustive search over all combinations of power and transmission rate for each of the \( N \) NOMA UEs is extremely tedious. We propose an algorithm which, while not guaranteeing an optimum solution, finds a feasible solution. Denote by \( \mu \) the threshold throughput that each UE must achieve. Then our algorithm for \( \mathcal{P}_2 \) solves the following:

\[
\max_{(P_i, \theta_i) = 1, \ldots, N} \mu \quad \text{subject to} \quad \sum_{i=1}^{N} P_i = 1 \quad \text{and} \quad R_i = \mu.
\]

As with Algorithm 1, starting with UE\(_N\), we aim to find the smallest power that can achieve \( \mu \). Unlike Algorithm 1 which does this up to UE\(_2\) only, in this case we do it until UE\(_1\), i.e., for all the UEs so that the UEs have symmetric throughput \( \mu \). If the total power consumed is less than the power budget, we increase \( \mu \). However, if each UE cannot achieve \( \mu \), the threshold throughput is too high and needs to be reduced. In this way we update \( \mu \) until the highest \( \mu \) that can be achieved by all UEs while consuming the full power budget is found. Formally, the algorithm is stated in Algorithm 2.
Algorithm 2 Resource allocation of a feasible solution to $\mathcal{P}_2$

1: Begin with $\mu = 0.3$, $\mu_H = \infty$, $\mu_L = 0$, $\zeta = 0$, $a = 1.3$, $n = 1$  
2: while $n$ do  
3:  if $\zeta = 0$ then  
4:     if $\mu_H = \infty$ then  
5:         $\mu = a\mu$  
6:     else  
7:         $\mu = \frac{\mu_H + \mu}{2}$  
8:     end if  
9: else  
10:     $\mu = \frac{\mu_L + \mu}{2}$  
11: end if  
12: Begin with UE $N$, $i = N$, $P = [ \ ]$, $\theta = [ \ ]$  
13: while $i > 0$ do  
14:     for $P_i = 0 : \Delta P : 1 - \sum_{k=i+1}^{N} P_k$ do  
15:         for $\theta_i = \theta_{LB} : \Delta \theta : \theta_{UB}$ do  
16:             Calculate $\mathcal{R}_i$ using (6.27) with (6.25) for MSP-based (with (6.26) for IS\textsuperscript{INR}-based) UE ordering  
17:         end if  
18:     end for  
19: end while  
20: if $\zeta = 1$ then  
21:     $\mu_H = \mu$  
22: else  
23:     $\mu_L = \mu$  
24: end if  
25: if $\mu_H - \mu_L < 0.01\mu_H$ then  
26:     Algorithm has converged, update: $n = 0$  
27: end if  
28: Go to (28)  
29: end if  
30: end while  
31: if $\mu_H - \mu_L < 0.01\mu_H$ then  
32:     Algorithm has converged, update: $n = 0$  
33: end if  
34: end if  
35: end if  
36: end while

In Algorithm 2 all UEs must achieve the threshold throughput of $\mu$, which is executed in lines 12 to 27. This is done by starting with UE$_N$ to find the smallest power and its corresponding smallest transmission rate that can attain $\mu$; once found, these are stored. We then move on to the next weakest UE, using the stored power and transmission rate for the stronger UEs. This process is repeated until UE$_1$. If there isn’t sufficient power for a UE to attain $\mu$, the flag $\zeta$ in line 23 is updated from 0 to 1 denoting that the current threshold throughput $\mu$ is too high and we exit the while loop to update $\mu$; otherwise, the flag $\zeta = 0$. We begin the algorithm assuming the last $\mu = 0.3$ and $\zeta = 0$. The upper bound on the threshold throughput (which not all of the UEs can attain at once), $\mu_H$, is initially set to $\infty$ and the lower bound on the threshold throughput (which all of the UEs can attain), $\mu_L$, is set to 0. We
update $\mu_H$ in line 29 when a smaller value of $\mu$ is found which all of the UEs fail to attain, i.e., when $\zeta = 1$. Similarly, $\mu_L$ is updated in line 31 when a larger value of $\mu$ is found which all of the UEs can attain, i.e., when $\zeta = 0$. This way we iteratively update $\mu$ to be the average of the most updated upper and lower bounds in lines 7 and 10. When the difference between $\mu_H$ and $\mu_L$ is smaller than a certain value, such as 1% in line 37, we assume the algorithm has converged. This way we are able to find the largest symmetric throughput. It should be noted that we use the coefficient $a$ such that $a > 1$; this allows us to update $\mu$ when we do not have available a finite $\mu_H$ in line 5. Also, note that although the algorithm begins with $\mu = 0.3$ and $\zeta = 0$, the choice of these parameters is arbitrary; even if $\mu = 0.3$ is not achievable by all of the UEs, since $\mu_H$ will be updated in the next iteration, the algorithm will not function incorrectly.

The same algorithm is employed for OMA, except that throughputs are calculated using using (6.31) with (6.29) for MSP-based (with (6.30) for ISINR-based) UE ordering, and the contending resource is $T$ instead of $P$.

6.5 Results

In this section we consider BS intensity $\lambda = 10$, noise power $\sigma^2 = -90$ dB and $\eta = 4$.

6.5.1 Performance

We first show using simulations that the approximations in Theorems 6.1 and 6.2 are tight. The results are shown in Fig. 6.5 which considers a system with $N = 3$ employing Model 1, a fixed power allocation scheme where $P_1 = 1/6$, $P_2 = 1/3$ and $P_3 = 1/2$, and both MSP-based and ISINR-based UE ordering. For clarity of presentation we choose the same transmission rate for all three UEs in both cases and plot coverage of each UE against the corresponding SINR threshold. The figure verifies the accuracy of our SINR analysis as the coverage expressions for both types of
Figure 6.5: SINR coverage vs. $\theta$ (identical transmission rate for all UEs) for Model 1 with $N = 3$ employing the fixed power allocation $P_1 = 1/6$, $P_2 = 1/3$ and $P_3 = 1/2$. Solid (dashed) lines show the analysis for ISÎãîR-based (MSP-based) UE ordering. Markers show the Monte Carlo simulations.

UE ordering match the simulation closely. We observe that ISÎãîR-based UE ordering is superior for all UEs other than UE$_N$. As explained previously, this is because UE$_N$ for ISÎãîR-based ordering is weaker than its MSP counterpart.

Resource allocation for the remaining figures is done according to Algorithms 1 and 2 depending on whether the objective is constrained to a TMT (i.e., $\mathcal{P}1$) or symmetric throughput (i.e., $\mathcal{P}2$), respectively. Unless specified otherwise, Model 1 is employed.

Fig. 6.6 is a plot of the cell sum rate against the number of UEs in a NOMA cluster, $N$, for both MSP-based and ISÎãîR-based UE ordering given a TMT constraint. We have included $N = 1$ in these plots which has the same $R_{\text{tot}}$ for all the curves since it only has one UE in a resource block (.: independent of $\beta$) which maximizes its throughput (.: independent of $\mathcal{T}$). Given $\mathcal{T}$ and $\beta$, there exists an optimum $N$ that maximizes $R_{\text{tot}}$. When $\beta$ is large we observe that using NOMA may not necessarily be more beneficial compared to OMA in terms of $R_{\text{tot}}$. Otherwise, for small $N$, increasing $N$ enhances $R_{\text{tot}}$ because interference cancellation is efficient in this regime, and more UEs are covered. Also, increasing $N$ results in a stronger UE$_1$ as it decreases $R_1$ on average in the case of MSP-based ordering and improves $Z_1$ on average for
ISI\textsubscript{NR}-based ordering; this enhances $R_1$ given a $P_1$. This in turn enhances $R_{tot}$ which in the TMT constraint problem ($P1$) receives the largest contribution from $R_1$. However, increasing $N$ beyond the optimum leaves too little power for UE\textsubscript{1} to boost $R_1$ with. For a given $T$ and $\beta$, the resources are only sufficient to support a maximum cluster size; after this $N$ (discontinuation of the plots), not all of the UEs are able to achieve $T$. Increasing $\beta$ results in a decrease in the maximum cluster size that can be supported. Similarly, increasing $T$ results in a decrease in the maximum cluster size that can be supported \cite{112}; this has not been shown for brevity. NOMA outperforms OMA significantly if $\beta$ is small and can support the same number of UEs or more.

In Fig. 6.6 we observe that for a given $\beta$ ISI\textsubscript{NR}-based UE ordering outperforms MSP-based ordering when $N$ is not larger than a certain value. After this, MSP performs better. In Fig. 6.7 we compare the individual NOMA UE powers and throughputs when $\beta = 0$ and $T = 0.3$ for $N = 6$ (where MSP-based ordering outperforms) with $N = 3$ (where ISI\textsubscript{NR}-based ordering outperforms). Unlike the other
Table 6.8: Individual UE throughput and cell sum rate vs. $N$ using Model 1 with $\beta = 0$ for $P_1$ with $T = 0.3$ and $P_2$. Black lines are for MSP-based UE ordering, while red lines are for IS\ñINR-based. The blue line is the TMT achieved by UE$_2$, ..., UE$_N$ in $P_1$.

<table>
<thead>
<tr>
<th>$N$</th>
<th>Throughput</th>
<th>Cell Sum Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1.5</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>2.5</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>3.5</td>
</tr>
</tbody>
</table>

Figure 6.9: $R_{\text{tot}}$ vs. $\beta$ for Model 1 with different values of $N$ and $T$ using MSP-based ordering. Curves represent NOMA and horizontal lines (independent of $\beta$) represent OMA.

For smaller $N$ ($N = 3$ in Fig. 6.7) when IS\ñINR-based ordering is employed, despite the increased power requirement by UE$_N$, there is still enough power left for UE$_1$ to maximize its throughput with so that $R_{\text{tot}}$ exceeds the MSP case. Although UE$_1$ in the IS\ñINR-based case is stronger, when $N$ is larger, the $P_1$ left is too little and $R_{\text{tot}}$ is lower than its MSP-based counterpart. The figures highlight that when $N$ is large, using the simpler MSP-based ordering scheme results in better performance than the complex IS\ñINR-based ordering.

Fig. 6.8 is a plot of both the $R_{\text{tot}}$ and individual UE throughput against the
number of UEs in a NOMA cluster for MSP and IS\textsuperscript{INR}-based UE ordering. We compare the maximum symmetric throughput objective in \( P2 \) (dashed lines) with the TMT constrained objective where \( T = 0.3 \) in \( P1 \) (solid lines). The curves for \( P1 \) end at the largest \( N \) that can be supported given the TMT constraint and \( \beta \). The symmetric throughput objective does not have a limit on the largest \( N \) that can be supported but for comparison with \( P1 \), we plot them up to the same value of \( N \). The TMT constraint of \( T = 0.3 \) always outperforms the symmetric throughput objective in terms of \( R_{\text{tot}} \) for the values of \( N \) that it can support; this is in accordance with what is anticipated from the rate region in Fig. 6.3. Additionally, for \( P2 \), MSP-based ordering always has a superior \( R_{\text{tot}} \) compared to its IS\textsuperscript{INR}-based counterpart, which is in line with our conclusions from Fig. 6.3.

In Fig. 6.8 the symmetric throughput objective of \( P2 \), like the case of \( P1 \), has an optimum \( N \) that maximizes \( R_{\text{tot}} \). For the problem in \( P2 \), the largest symmetric throughput is limited by the weakest UE, UE\(_N\), which requires the largest power. As \( N \) grows, the weakest UE becomes worse and the total power needs to be shared among a larger number of UEs. This causes the individual UE’s (symmetric) throughput to decrease with \( N \) as shown. However, increasing \( N \) at first enhances \( R_{\text{tot}} \) because SuIC is efficient in this regime and more UEs are covered, so the gains from the larger number of UEs are more significant. For larger \( N \), the individual UE throughput becomes too small. Consequently \( R_{\text{tot}} \) starts decreasing after the optimum \( N \). Additionally, as long as \( N \) is not too large, the individual UEs perform better for \( P2 \) compared to UE\(_2\), \ldots, UE\(_N\) in \( P1 \) which achieve \( T \). Also, \( R_1 \) in \( P1 \) outperforms the individual UE throughput in \( P2 \) as anticipated. More interestingly, \( R_1 \) has an optimum \( N \) for which it is maximized. This is due to the improving strength of UE\(_1\) as \( N \) grows, followed by a decrease in \( R_1 \) because of lower available \( P_1 \) when \( N \) is too large.

Fig. 6.9 plots the cell sum rate against \( \beta \) for different \( N \) and \( T \) using MSP-
based ordering. Since OMA does not use SuIC, the corresponding $R_{\text{tot}}$ plots are horizontal lines independent of $\beta$. The figure shows the existence of a maximum $\beta$ value until which a NOMA system with a particular $N$ and $T$ is able to outperform the corresponding OMA system. This highlights that there is a critical minimum level of SuIC required for NOMA to outperform OMA. We also observe that the decrease in $R_{\text{tot}}$ as a function of $\beta$ is steeper for larger $N$ and $T$ highlighting their increased susceptibility to RI.

The results highlight the importance and impact of choosing network parameters such as $N$ and the UE ordering technique depending on the network objective and $\beta$. As an example, if complete user fairness is required, i.e., the objective is $P_2$, MSP-based ordering would result in higher $R_{\text{tot}}$, while $N$ would be chosen according to $\beta$. Similarly, if the network requires a certain TMT, the objective is $P_1$. To enhance $R_{\text{tot}}$, ISINR-based ordering may be chosen if $T$ is not too high with a smaller $N$; otherwise MSP-based ordering would be a better option. The value of $N$ would also depend on $\beta$.

![Figure 6.10: Rate region for NOMA with MSP-based UE ordering, $\beta = 0$ and $N = 2$ for Models 1, 2, and 3.](image1)

![Figure 6.11: $R_{\text{tot}}$ vs. $N$ for $T = 0.3$ using NOMA and $\beta = 0$ for MSP-based ordering. The curves end at the largest $N$ that can be supported given $T$ and $\beta$.](image2)

Fig. 6.10 is a plot of the rate regions for the $N = 2$ case using MSP-based UE ordering and with $\beta = 0$ for Models 1, 2, and 3. We observe that the optimum
performance (rate region boundary) of Model 2 can be viewed as an upper bound on that of Model 1. This can be explained by the fact that Model 2 selects UEs that are located farther from the closest cell edge than the UEs in Model 1 resulting in better average interference conditions and consequently, performance. Similarly Model 3, which selects UEs that are located even farther from the closest cell edge than both Models 1 and 2 and thereby have the best average interference conditions, upper bounds the other two models in terms of performance.

Fig. 6.11 is a plot of the cell sum rate with increasing $N$ for all three models with MSP-based ordering, $\beta = 0$ and $T = 0.3$. In general, the smaller the sector angle $\phi$, the more clustered NOMA UEs are towards the center of the cell and the better their performance is on average. Accordingly, we observe that Model 2 outperforms Model 1 for each value of $N$, and that Model 3 outperforms both Models 1 and 2. This highlights how a superior UE clustering technique that selects UEs with good interference conditions is able to significantly improve the performance. In particular, for the given $\beta$ and $T$ at its optimum $N$, Model 2 outperforms the optimum of Model 1 by 15.5%, and Model 3 outperforms Models 1 and 2 by 129% and 98.4%, respectively. Additionally, we observe that a more superior clustering technique is able to support a larger maximum cluster size given a TMT constraint; for $T = 0.3$, Models 2 and 3 are able to support a largest $N$ of 7 and 9, respectively, compared to the largest $N$ of 6 in the case of Model 1.

It ought to be highlighted that although Model 3 shows a significant improvement in performance, its main purpose is to serve as an upper bound. In a practical setting, a model with a sector angle such as $\phi = \pi/2$ (i.e., “Model 2.5”) may provide a very good trade off between having enough UEs available for NOMA and the interference conditions.
6.5.2 Complexity

In this subsection we discuss the complexity of the proposed algorithms and compare them with an exhaustive search. As mentioned in Section IV, since the range for transmission rate is $\theta_i \geq 0$, our algorithms search in $\theta_{LB} \leq \theta_i \leq \theta_{UB}$ to make the search finite. For a fair comparison we use the same search space for the exhaustive search. We search in $-10 \text{ dB} \leq \theta_i \leq 22 \text{ dB}$ and use step size $\Delta_\theta = 1$ in this chapter. As a result there are $\hat{\Delta}_\theta = 33$ choices of $\theta_i$. Since the range of power allocated to a UE is $0 \leq P_i \leq 1$, based on the step size $\Delta_P$ there are $1/\Delta_P$ choices of $P_i$. We define the complexity as the sum of the number of times throughput is calculated for each UE, i.e., the number of power-transmission throughput combinations the algorithm iterates over. Consequently, for an exhaustive search the complexity is $(\hat{\Delta}_\theta/\Delta_P)^N$.

Fig. 6.12(a) plots the complexity against $N$ for different values of TMT and step size $\Delta_P$ for Algorithm 1. As anticipated, decreasing $\Delta_P$ increases the complexity. Also, decreasing the TMT decreases complexity as the non-strongest UEs find the least power required to achieve the TMT more quickly. We observe that for a given $\mathcal{T}$ and $\Delta_P$, IS\textsuperscript{INR}-based ordering has lower complexity than its MSP-based counterpart until large $N$ where its complexity becomes larger. This is based on similar reasons to Fig. 6.7 where UE\textsubscript{N} in IS\textsuperscript{INR}-based ordering requires larger power than its MSP counterpart, which increases the complexity and becomes the dominant factor at high $N$. The results suggest that the complexity is of the form $c^N$, where $c = \hat{\Delta}_\theta/\Delta_P$ for an exhaustive search. In the case of our algorithms we observe that $c \ll (\hat{\Delta}_\theta/\Delta_P)$. In particular, for Algorithm 1 with $\Delta_P = 0.01$, our $c$ is about 3 and thus 1000 times smaller than $\hat{\Delta}_\theta/\Delta_P = 3300$.

In Fig. 6.12(b) we additionally plot the complexity curves for an exhaustive search and Algorithm 2. Since Algorithm 2 has Algorithm 1 nested in it, it repeats Algorithm 1 multiple times in a way; consequently, the complexity is higher. Increasing the step size has a similar effect as in Algorithm 1, we do not show these for brevity. It
should be noted that the complexity of Algorithm 2 does not increase monotonically with $N$ as in the case of Algorithm 1. This is due to the more heuristic nature of Algorithm 2 because of the choice of the initial parameters $a$ and $\mu$ which result in a varying number of iterations before the largest symmetric throughput is achieved. As a result, complexity can change depending on the choice of these parameters. For fairness, we chose the same parameters for all $N$ corresponding to the values mentioned in Algorithm 2. Most importantly, from Fig. 6.12(b) we observe the significant difference between the complexity of our algorithms and an exhaustive search. Since an exhaustive search is the only optimum way for solving both non-convex problems $P1$ and $P2$, the stark difference in complexity motivates the use of efficient algorithms such as ours for resource allocation.

6.6 Conclusion

In this chapter a large cellular network that employs NOMA in the downlink is studied. As NOMA requires ordering of the UEs based on some measure of link quality, two kinds of UE ordering techniques are analyzed and compared: 1) MSP-based ordering, 2) ISINR-based ordering. An SINR analysis that takes into account the SuIC
chain and RI from imperfections in SuIC is developed for each ordering technique. We show that neither ordering technique is consistently superior to the other and highlight scenarios where each outperforms the other. Two non-convex problems of maximizing the cell sum rate $R_{\text{tot}}$ subject to a constraint are formulated: a TMT constraint $\mathcal{T}$ in $\mathcal{P}_1$ and the symmetric throughput constraint in $\mathcal{P}_2$. Since the optimum solution for resource allocation to solve each problem requires an exhaustive search, two efficient algorithms for general NOMA cluster size $N$ that give feasible solutions for intercell interference-aware power allocation and choice of transmission rate are proposed. We show that the complexity of the proposed algorithms is significantly lower than an exhaustive search. Additionally, the existence of an optimum NOMA cluster size that maximizes $R_{\text{tot}}$ for each problem is shown. It is observed that $\mathcal{P}_1$ provides a higher $R_{\text{tot}}$; however, $\mathcal{P}_2$ guarantees better individual UE performance. Furthermore, it is shown that NOMA outperforms OMA provided $\beta$ is below a certain critical value. The results highlight the importance and impact of choosing network parameters such as $N$ and the UE ordering technique, depending on the network objective and $\beta$. Three models to show the impact of UE clustering in NOMA are introduced. The models demonstrate that efficient UE clustering which selects UEs with good interference conditions can improve performance significantly; in fact, with efficient UE clustering the cell sum rate can be doubled.

In this chapter we have studied the performance of cell-center NOMA users in terms of metrics based on the standard coverage probability, i.e., the spatially averaged coverage. Network operators are often interested in the percentile of users that can attain a certain coverage. In this context, we study the meta distribution of NOMA users in the next chapter. We use this metric to also justify and promote the employment of NOMA for cell center users (i.e., the clustering techniques) in this chapter by comparing the performance of NOMA for cell center users against when NOMA is employed for users everywhere in the cell (i.e., no particular NOMA
clustering).
Chapter 7

Meta Distribution of Downlink Non-Orthogonal Multiple Access (NOMA) in Poisson Networks

7.1 Introduction

In this chapter we study the meta distribution of the coverage probability in downlink NOMA networks. Two schemes are assessed based on the location of the NOMA users: 1) E-NOMA: users are located everywhere in the network; 2) C-NOMA: users are located in the cell-center only. We derive the moments of the meta distribution for both schemes and the meta distribution is approximated via the beta distribution. For the E-NOMA scheme, closed-form moments are derived. Only exact integral expressions can be derived for the moments of the C-NOMA scheme, however. To simplify these integral calculations, we propose approximate moments for the C-NOMA scheme. Our results show that restricting NOMA to cell-center users provides a significantly higher mean, lower variance and better percentile performance for the coverage probability. In addition to this, our results indicate the importance of careful resource allocation and its drastic impact on not only the SCP but also on higher moments of the meta distribution.

7.2 System Model

We consider a downlink cellular network where BSs are distributed according to a homogeneous PPP $\Phi$ with intensity $\lambda$. Each BS serves $N$ UEs in one time-frequency resource block by multiplexing the signals for each UE with different power levels
using a total power budget $P$. Without loss of generality, we assume $P = 1$ in this chapter. A Rayleigh fading environment is assumed such that the fading coefficients are i.i.d. with a unit mean exponential distribution. A power-law path-loss model is considered where the signal power decays at the rate $r^{-\eta}$ with distance $r$, where $\eta > 2$ denotes the path-loss exponent and $\delta = \frac{2}{\eta}$.

SuIC requires ordering the UEs according to some measure of link strength \cite{111}. For $i \in \{1, \ldots, N\}$, the $i^{th}$ strongest UE is referred to as UE$_i$. In this chapter, we order the UEs based on the link distance $R$. The ordered link distance of UE$_i$ is denoted by $R_i$; consequently, UE$_i$ is nearer to the BS and therefore stronger than UE$_j$ for $i < j$ (i.e., $R_i < R_j$). Exploiting SuIC, UE$_i$ decodes and cancels messages intended for all weaker UEs before decoding its own message. On the other hand, messages for stronger UEs are treated as noise and contribute to the intracell interference. We incorporate imperfect SuIC into our analysis by considering a fraction $\beta$ of residual intracell interference from the canceled messages of weaker UEs. Let $P_i$ and $\log(1+\theta_i)$ denote the power allocated and target rate for UE$_i$; the corresponding SIR threshold for the message of UE$_i$ is $\theta_i$. Note that due to the power budget, $\sum_{i=1}^{N} P_i = 1$. For feasible SuIC, proper resource allocation, i.e., power allocation and rate adaptation (e.g., $P_i \leq P_j$ and/or $\theta_i \geq \theta_j$ for $i < j$), for all UEs is required.

**Lemma 7.1:** For any ascending ordered statistic like $R_i$, based on the statistics of their unordered counterpart $R$, the PDF is

$$f_{R_i}(r) = \binom{N-1}{i-1}Nf_R(r)(F_R(r))^{i-1}(1-F_R(r))^{N-i}. \quad (7.1)$$

In terms of components larger than $i$, (7.1) can be rewritten as

$$f_{R_i}(r) = f_{R_i}(r) + \sum_{m=i+1}^{N} \binom{m-1}{i-1}(-1)^{m-i}f_{R_m}(r), \quad (7.2)$$

where $f_{R_j}(r) = \binom{N-1}{j-1}Nf_R(r)(F_R)^{j-1}$ for $i \leq j \leq N$. In terms of components smaller
than \(i\), (7.1) can be rewritten as

\[
f_{R_i}(r) = f_{\tilde{R}_i}(r) + \sum_{m=1}^{i-1} \frac{(N-m)!(-1)^{i-m}}{(m-1)!(i-m)!} f_{\tilde{R}_m}(r),
\]  

(7.3)

where \(f_{\tilde{R}_j}(r) = \binom{N-1}{j-1} N f_R(r) (1 - F_R)^{N-j} \) for \(1 \leq j \leq i\).

We denote the distance between a BS and its nearest neighboring BS by \(\rho\). Since \(\Phi\) is a PPP, the PDF of \(\rho\) is \(f_{\rho}(x) = 2\pi \lambda x e^{-\pi \lambda x^2}, x \geq 0\). Consider a disk around each BS located at \(x\) with radius \(\rho/2\), i.e., \(b(x, \rho/2)\); like Chapter 6, we refer to this as the in-disk. The in-disk is the largest disk centered at a BS that fits inside its Voronoi cell. We study and compare NOMA for the following two schemes.

**Everywhere NOMA (E-NOMA)**

\(N\) UEs are distributed uniformly and independently in each Voronoi cell. Consequently, the distribution of the unordered link distance \(R\) follows \(f_R(r) = 2\pi \lambda r e^{-\pi \lambda r^2}, r \geq 0\). Using this PDF and its CDF \(F_R(r)\), the ordered distance distribution \(f_{R_i}(r)\), \(r \geq 0\), in the E-NOMA scheme follows (7.1).

**Cell-Center NOMA (C-NOMA)**

\(N\) UEs are distributed uniformly and independently in the in-disk \(b(x, \rho/2)\) of each BS at \(x\). Consequently, the link distance \(R\), conditioned on \(\rho\), follows \(f_{R|\rho}(r \mid \rho) = \frac{8r}{\rho^2}, 0 \leq r \leq \frac{\rho}{2}\). Using (7.1) the PDF of \(R_i\), conditioned on \(\rho\), in the C-NOMA scheme follows

\[
f_{R_i|\rho}(r \mid \rho) = \binom{N-1}{i-1} 8r N \left(\frac{4r^2}{\rho^2}\right)^{i-1} \left(1 - \frac{4r^2}{\rho^2}\right)^{N-i}, \quad 0 \leq r \leq \frac{\rho}{2}.
\]  

(7.4)

**Remark:** C-NOMA restricts the link distance to \(\rho/2\); the notion is that NOMA is better suited for UEs that are closer to the serving BS. UEs with relatively larger
7.3 SIR Analysis

SuIC requires a UE to successfully decode all of the messages intended for weaker UEs. Consider a randomly selected BS located at $x_0$ and its associated UEs; the SIR at UE $i$ of the message intended for UE $j$ for $i \leq j \leq N$ is

$$\text{SIR}_i^j = \frac{h_i R_i^{-\eta} P_j}{h_i R_i^{-\eta} \left( \sum_{m=1}^{j-1} P_m + \beta \sum_{k=j+1}^{N} P_k \right) + \sum_{x \in \Phi \setminus x_0} g_y_i \|y_i\|^{-\eta}},$$

where $y_i = x - u_i$, $u_i$ is the location of UE $i$, $\| \cdot \|$ denotes the Euclidean norm, and $h_i$ ($g_y_i$) is the fading power gain from the serving (interfering) BS to UE $i$.

Accordingly, due to SuIC decoding, coverage at UE $i$ is defined via the following joint event

$$C_i = \bigcap_{j=1}^{N} \left\{ \text{SIR}_i^j > \theta_j \right\} = \bigcap_{j=1}^{N} \left\{ h_i > R_i^{-\eta} \bar{P}_j \frac{\theta_j}{\sum_{x \in \Phi \setminus x_0} g_y_i \|y_i\|^{-\eta}} \right\},$$

(7.5)

where $\bar{P}_j = P_j - \theta_j \left( \sum_{m=1}^{j-1} P_m + \beta \sum_{k=j+1}^{N} P_k \right)$. We rewrite (7.5) as $C_i = \left\{ h_i > R_i^{-\eta} M_i \sum_{x \in \Phi} g_y_i \|y_i\|^{-\eta} \right\}$ using $M_i = \max_{i \leq j \leq N} \frac{\theta_j}{\bar{P}_j}$.

For a fixed, yet arbitrary, realization of the network, the CCP of UE $i$ in a randomly selected cell, $\mathcal{P}_{C_i}$, is

$$\mathcal{P}_{C_i} = \mathbb{P}(C_i | \Phi) \overset{(a)}{=} \mathbb{E}_{g_y_i} \left[ \exp \left( - R_i^{-\eta} M_i \sum_{x \in \Phi \setminus x_i} g_y_i \|y_i\|^{-\eta} \right) \right] | \Phi]$$

$$\overset{(b)}{=} \prod_{x \in \Phi \setminus x_0} \frac{1}{1 + R_i^{-\eta} M_i \|y_i\|^{-\eta}}.$$

(7.6)

where (a) follows using the CDF of $h_i \sim \exp(1)$ and (b) follows from the MGF of the

link distances are better served in their own resource block without sharing [111].
independent RVs $g_{x_i} \sim \exp(1)$.

Denote the $b^{th}$ moment of the CCP of UE$_i$ across all links in an arbitrary fixed realization of the network by $M_{i,b}$. Then,

$$M_{i,b} = \mathbb{E} \left[ \prod_{x \in \Phi \setminus x_0} (1 + R_i^b M_i \|y_i\|^{-\eta})^{-b} \right].$$  \hfill (7.7)

**Remark:** If $\tilde{P}_j < 0$, the CCP is zero. Henceforth we assume resource allocation such that $\tilde{P}_j \geq 0$.

**Note:** If $b = 1$ in (7.7), we obtain the SCP of UE$_i$.

Through moment matching, the meta distribution of UE$_i$ is approximated using the beta distribution [99] as follows

$$\tilde{F}_{P_{C_i}}(\alpha) = \mathbb{P}(P_{C_i} > \alpha) \approx 1 - L_\alpha \left( \frac{\tilde{\beta}_i M_{i,1}}{1 - M_{i,1}}, \tilde{\beta}_i \right),$$  \hfill (7.8)

where $\tilde{\beta}_i = \frac{(M_{i,1} - M_{i,2})(1 - M_{i,1})}{M_{i,2} - M_{i,1}}$ and $L_\alpha(a, b) = \int_0^\alpha t^{a-1} (1 - t)^{b-1} dt$. The variance of the meta distribution of UE$_i$ is defined as

$$\sigma_i^2 = M_{i,2} - M_{i,1}^2.$$  \hfill (7.9)

The ordered RDP for UE$_i$, which is the RDP in [115] using ordered link distance, is

$$\mathcal{R}_i = \{ x \in \Phi \setminus \{x_0\} : R_i/\|y_i\| \}. $$  \hfill (7.10)
Using the PGFL of the PPP in (a), the PGFL of $\mathcal{R}_i$ is

$$G_{\mathcal{R}_i}[f] \triangleq \mathbb{E} \left[ \prod_{x \in \mathcal{R}_i} f(x) \right] = \mathbb{E} \left[ \prod_{x \in \Phi \setminus \{x_0\}} f \left( \frac{R_i}{\|y_i\|} \right) \right]$$

$$\overset{(a)}{=} \mathbb{E}_{R_i} \left[ \exp \left( -2\pi \lambda \int_{R_i} \left( 1 - f \left( \frac{R_i}{a} \right) \right) a \, da \right) \right]. \quad (7.11)$$

Using the ordered RDP for UE$_i$, the expectation in (7.7) can also be evaluated as

$$M_{i,b} = \mathbb{E} \left[ \prod_{y \in \mathcal{R}_i} (1 + M_i y^n)^{-b} \right]. \quad (7.12)$$

1) E-NOMA Scheme

We characterize the PGFL of the ordered RDPs and obtain closed for expressions for $M_{i,b}$.

**Lemma 7.2:** The PGFL of $\mathcal{R}_i$ for $1 \leq i \leq N$ in E-NOMA is

$$G_{\mathcal{R}_i}[f] = \tilde{G}_{\mathcal{R}_i}[f] + \sum_{m=1}^{i-1} \frac{(N-m)!}{(m-1)!(i-m)!} \tilde{G}_{\mathcal{R}_m}[f], \quad (7.13)$$

where for $1 \leq j \leq i$

$$\tilde{G}_{\mathcal{R}_j}[f] = \frac{(N-j)N}{(N-j+1) + 2 \int_1^\infty \left( 1 - f \left( \frac{y}{y-1} \right) \right) y \, dy}. \quad (7.14)$$

**Proof:** We obtain (7.13) using (7.3) in (7.11). Also using (7.11),

$$\tilde{G}_{\mathcal{R}_j}[f] = \int_0^\infty f_{\mathcal{R}_j}(x) \exp \left( -2\pi \lambda \int_{R_i} \left( 1 - f \left( \frac{x}{a} \right) \right) a \, da \right) dx$$

$$\overset{(a)}{=} \frac{N}{j-1} \pi \lambda N \int_0^\infty e^{-2\pi \lambda m \int_1^\infty (1-f(y^{-1})) y \, dy} e^{-\pi \lambda (N-j+1)m} dm$$

where (a) is obtained by changing variables and (7.14) is obtained using the MGF of $m \sim \exp(\pi \lambda (N-j+1))$. \qed
**Corollary 7.1:** $\mathcal{M}_{i,b}$ for $1 \leq i \leq N$ in E-NOMA is

$$
\mathcal{M}_{i,b} = \tilde{\mathcal{M}}_{i,b} + \sum_{m=1}^{i-1} \frac{(N-m)!(−1)^{i-m}}{(m-1)!(i-m)!} \tilde{\mathcal{M}}_{m,b},
$$

(7.15)

where for $1 \leq j \leq i$

$$
\tilde{\mathcal{M}}_{j,b} = \binom{N-1}{j-1} \frac{N}{N-j + 2 \, F_1(b, -\delta, 1 - \delta, -\mathcal{M}_{i})}.
$$

(7.16)

**Proof:** (7.15) is obtained using (7.13), where we define using (7.12)

$$
\tilde{\mathcal{M}}_{j,b} = G_{\mathcal{R}_{j}}[(1 + M_{i}y) - b]
$$

(7.14)

$$
\frac{(N-1)}{(j-1)} \frac{N}{N-j + 2 \int_{1}^{\infty} \left(1 - (1 + M_{i}y^{-b})\right) y \, dy}.
$$

We obtain (a) using (7.14), and (7.16) follows by $y \to g^{-1}$.


2) C-NOMA Scheme

We obtain integral expressions for $\mathcal{M}_{i,b}$. We also propose approximate PGFLs of the ordered RDP and use these to evaluate $\mathcal{M}_{i,b}$ in a simpler form.

**Lemma 7.3:** The $b$th moment of the CCP for UE$_{i}$ in the C-NOMA scheme is

$$
\mathcal{M}_{i,b} \approx E_{\rho, R_{i}} \left[ e^{-2\pi\lambda \int_{\rho-R_{i}}^{\infty} \left(1 - \frac{M_{i}R_{i}y}{\rho y}\right)^{-b} r dr} \left(1 + \frac{M_{i}R_{i}y}{\rho y}\right)^{-\delta} \right].
$$

(7.17)

**Proof:** In the C-NOMA model each UE is conditioned to have an interferer $\rho$ away from the serving BS. Hence, using (7.7)

$$
\mathcal{M}_{i,b} = E \left[ \prod_{x \in \Phi \setminus x_{0}, \|x-x_{0}\| > \rho} \left(1 + M_{i} \frac{R_{i}^{\eta}}{\|y_{i}\|^{\eta}}\right)^{-b} \prod_{x \in \Phi \setminus x_{0}, \|x-x_{0\|= \rho} \left(1 + M_{i} \frac{R_{i}^{\eta}}{\|y_{i}\|^{\eta}}\right)^{-b} \right].
$$

We obtain the first term in (7.17) using the PGFL of the PPP and the guard zone
\(b(u_i, \rho - R_i)\) in the C-NOMA scheme. The average location of a UE distributed uniformly in the in-disk is the center of the disk, i.e, \(x_0\). Accordingly, we approximate the average distance between a UE and the BS \(\rho\) away from \(x_0\) as \(\rho\); hence, the second term in (7.17) is obtained. This approximation has been validated to be tight in [111,112].

Consider the following two approximations:

- **A1**: UE \(i\) is guaranteed to have no interfering BS in \(b(u_i, R_i)\), which is not the largest guard zone around the UE.
- **A2**: Deconditioning on the BS \(\rho\) away from the serving BS.

**Remark**: The two approximations have opposing effects; A1 overestimates inter-cell interference while A2 underestimates it.

Calculating \(M_{i,b}\) using Lemma 7.3 requires a triple integral. However, exploiting A1 and A2, we provide an approximation to calculate \(M_{i,b}\) that requires a single integration.

**Lemma 7.4**: Using A1 and A2, the PGFL of \(R_i\) conditioned on \(\rho\) for \(1 \leq i \leq N\) in the C-NOMA scheme is

\[
\mathcal{G}_{R_i|\rho}[f] = \mathcal{G}_{R_i|\rho}[f] + \sum_{m=i+1}^{N} \binom{m-1}{i-1} (-1)^{m-i} \mathcal{G}_{R_{m-1}|\rho}[f], \quad (7.18)
\]

where for \(i \leq j \leq N\)

\[
\mathcal{G}_{R_j|\rho}[f] = \frac{(N-1)\left(\Gamma(j) - \Gamma\left(j, \frac{\pi \rho^2}{2} \int_1^\infty \left(1 - f\left(\frac{1}{y}\right)\right) dy\right)\right)}{\frac{1}{N} \left(\frac{\rho^2}{2} \pi \lambda \int_1^\infty \left(1 - f\left(\frac{1}{y}\right)\right) dy\right)^j}. \quad (7.19)
\]

**Proof**: We obtain (7.18) using (7.2) in (7.11). Also using (7.11),

\[
\mathcal{G}_{R_j|\rho}[f] = \int_0^\infty f_{R_j}(x) \exp\left(-2\pi \lambda \int_{R_i}^\infty \left(1 - f\left(\frac{x}{a}\right)\right) a da\right) dx
\]

\[
\begin{align*}
&\overset{(a)}{=}\binom{N-1}{j-1} \frac{4^j}{\rho^{2j}} \int_0^\infty e^{-2\pi \lambda m \int_{1-f(y^{-1})}^\infty y dy} m^{j-1} dm
\end{align*}
\]
(a) follows by changing variables, and (7.19) by integration.

We approximate $M_{i,b}$ by substituting the approximate PGFL of $R_i$, conditioned on $\rho$, into (7.12) and averaging over $\rho$.

**Corollary 7.2:** Using A1 and A2, $M_{i,b}$ for $1 \leq i \leq N$ in C-NOMA is

$$M_{i,b} = \tilde{M}_{i,b} + \sum_{m=i+1}^{N} \binom{m-1}{i-1} (-1)^{m-i} \tilde{M}_{m,b}, \quad (7.20)$$

where for $i \leq j \leq N$

$$\tilde{M}_{j,b} = \mathbb{E}_\rho \left[ \frac{\Gamma(j) - \Gamma\left( j, \frac{\pi \lambda \rho^2}{4} \frac{2 F_1 \left( b, -\delta, 1 - \delta, -M_i - 1 \right) - 1}{\Gamma(j - 1)} \rho \right)}{\left( \frac{\pi \lambda j}{\pi \rho} \int_{1}^{\infty} \left( 1 - (1 + M_i y^-)^{-b} \right) y dy \right)^j} \right]. \quad (7.21)$$

**Proof:** (7.20) is obtained using (7.18) where we define using (7.12)

$$\tilde{M}_{j,b} = \mathbb{E}_\rho \left[ \frac{\Gamma(j) - \Gamma\left( j, \frac{\pi \lambda \rho^2}{4} \frac{2 F_1 \left( b, -\delta, 1 - \delta, -M_i - 1 \right) - 1}{\Gamma(j - 1)} \rho \right)}{\left( \frac{\pi \lambda j}{\pi \rho} \int_{1}^{\infty} \left( 1 - (1 + M_i y^-)^{-b} \right) y dy \right)^j} \right].$$

We obtain (a) using (7.19), and (7.21) follows by $y \to g^{-1}$. 

**7.4 Results**

In this section, we select the following parameters: $\lambda = 10$, $\eta = 4$, $\beta = 0$ and $N = 2$, unless stated otherwise. Simulations are repeated 50,000 times. Since the power budget is $P = 1$, $P_2 = 1 - P_1$. Unless stated otherwise, Lemma 7.3 is used for the moments of the CCP in the C-NOMA model.

Fig. 7.1 verifies the approximation of the meta distribution in (7.8) using simulations for both NOMA schemes with different values of $P_1$. The approximation is tighter (looser) for C-NOMA (E-NOMA) because of its larger (smaller) interference-exclusion disk with radius $\rho - R_i$ ($R_i$). The figure shows that the fraction of UE, that attain a given coverage probability is always much larger for the C-NOMA scheme.
when compared to the E-NOMA scheme, which highlights the superiority of restricting NOMA to cell-center UEs. When $P_1 = 0.5$, 98.9% (92.1%) of UE1 (UE2) achieve a coverage probability of at least 0.5 in the C-NOMA scheme, while only 61.5% (19.9%) of UE1 (UE2) achieve the same coverage probability in the E-NOMA scheme. Decreasing $P_1$ worsens the performance of UE1 and improves UE2; consequently, decreasing $P_1$ in Fig. 7.1 increases the fraction of UE2 that attains a certain coverage probability at the expense of reducing the fraction of UE1 achieving a given coverage probability.

Fig. 7.2 plots the mean and variance of the meta distribution for the NOMA UEs in the C-NOMA scheme. We compare using the moments obtained with and without the approximations A1 and A2. We observe that the approximation is tight for the SCP and overestimates the variance, particularly for UE2 near the peak.

Fig. 7.3 plots the mean and variance of the meta distribution of the UEs for both schemes using identical resource allocation. We observe that C-NOMA outperforms the E-NOMA scheme in terms of both SCP and variance. Increasing $\beta$ deteriorates performance of the non-weakest UEs, decreasing SCP and increasing variance. For a given $\beta$, the higher SCP of the C-NOMA scheme can be attributed to the fact that the
UEs are closer to the BS on average than the E-NOMA scheme. The lower variance is also due to the limited vicinity leading to lower disparity than the E-NOMA model. Furthermore, $\sigma_i^2$ peaks at high $\theta$ for the C-NOMA scheme (corresponding to low SCP); which is not the case for the E-NOMA scheme. This implies the existence of $\theta$ with high SCP and low $\sigma_i^2$ in C-NOMA, thereby highlighting its superiority with careful resource allocation. The C-NOMA is also a more consistent scheme as both SCP and variance are better for UE$_1$ than UE$_2$; this is not the case for the E-NOMA scheme.

Fig. 7.4 plots the mean and variance of the meta distribution for an optimized power and rate adaptation for UE$_2$ such that the total rate is maximized subject to a TMT constraint. The rate of a UE is defined as the SCP times target rate. Resource allocation is done according to the algorithm in [112] and results in UE$_2$ having rate equal to the TMT. We also plot the rate of UE$_1$ in Fig. 7.4. In the C-NOMA scheme (and the E-NOMA scheme, not shown for brevity), increasing the TMT increases $\sigma_2^2$ while the peak $\sigma_1^2$ occurs at lower $\theta_1$ but does not change in value. When the TMT is 0.1, the SCP of UE$_2$ and $\sigma_2^2$ are worse for the E-NOMA scheme. Although the peak
\[ \sigma_1^2 \text{ is higher for the C-NOMA scheme than the E-NOMA scheme, at the optimum } \theta_1 \text{ that maximizes the rate of UE}_1, \sigma_1^2 \text{ is lower for the C-NOMA scheme. Other than highlighting the superiority of the C-NOMA scheme, this also emphasizes the importance of optimum resource allocation for not just the SCP, but also for higher moments of the meta distribution.} \]

7.5 Conclusion

In this chapter we study the meta distribution of the CCP of NOMA UEs distributed according to two schemes: one that places no location-based restrictions on the NOMA UEs (i.e., the E-NOMA scheme), and the other that restricts NOMA UEs to the cell center (i.e., the C-NOMA scheme). Closed form expressions for the moments of the meta distribution in the E-NOMA scheme are derived. The C-NOMA scheme requires a triple integral so we propose approximate moments that reduce to a single integration. Our results show that employing NOMA for cell-center users is significantly more beneficial in terms of both the SCP as well as higher moments of the CCP (such as the variance of the meta distribution), than using it for all UEs in a cell. These findings motivate the UE clustering techniques proposed in the works of \[111,112\]. Our findings also emphasize the importance of careful resource allocation in NOMA which significantly affects performance not only in terms of the SCP but also higher moments of the CCP.
Chapter 8

Concluding Remarks

8.1 Summary

In this dissertation we first highlight the very demanding requirements of 5G networks, especially in terms of data rates and achievable throughput. In this context, we focus on different spectrum reutilization techniques. In particular, we introduce and discuss three spectrum reuse techniques, namely D2D, FD and NOMA. We motivate the need to study the performance of large scale wireless networks in order to correctly assess and analyze the impact of these new spectrum reuse techniques in real, dense 5G networks. In Chapter 2 we discuss how stochastic geometry tools are used for modeling different wireless networks as point processes based on their characteristics. We also describe different techniques on how such point processes can be used to analyze metrics of interest.

In Chapter 3 we study the impact of FD-enabled D2D communication underlaid with a traditional OMA cellular network. The purpose of such a network is improved spectrum utilization. We develop a tractable analytical model with a flexible mode selection scheme and quantify the gains that can be obtained by FD-D2D communication. A tradeoff exists between increasing aggregate network interference and improving spatial frequency reuse. We find that enforcing FD-D2D communication deteriorates network performance due to the high interference levels that result, whereas tuning the network parameters correctly allows non-trivial gains to be harvested using this scheme. Also, an accurate approximation for the PDF of
the distance between the D2D-receiver and its closest BS is proposed which can be applied to other similar setups as well.

In Chapter 4 the focus is on achievable physical layer security at a receiver in the presence of an eavesdropper. We consider correlation between the interference experienced at the receiver and eavesdropper and study the probability of OSSA. Interesting trends in the probabilities are discovered. More importantly, we find that considering interference correlation, particularly when the eavesdropper lies near to the legitimate receiver, has a significant impact on the probabilities of secure communication compared to the case where independent interference is considered. To this end, we propose a FD jamming solution to improve the OSSA when the eavesdropper lies closer to the legitimate receiver. However, jamming increases the network interference and deteriorates the OSSA when the eavesdropper lies farther from the user. Hence, a tradeoff exists between using FD jamming to improve the OSSA for nearby eavesdroppers and degrading the OSSA if the eavesdropper lies farther away.

In Chapter 5 we emphasize the significant impact of intercell interference on NOMA design. Interference-aware power allocation, UE sorting and clustering are discussed for both downlink and uplink NOMA along with their benefits and shortcomings. Additionally, intercell interference management in the context of NOMA networks is studied. We also highlight the potential gains that can be attained by integrating NOMA with other technologies such as D2D and FD. Such integrations give rise to a number of challenges that are also highlighted. Potential solutions to deal with these challenges are discussed, too.

In Chapter 6 we analytically study a large downlink NOMA network for a general number of NOMA UEs. Three UE clustering schemes are proposed and two UE ordering techniques are considered. An SINR analysis is developed for each clustering scheme using both UE ordering techniques. Our developed framework considers
NOMA decoding as the joint event of decoding the messages for all weaker UEs and takes into account imperfections in the SuIC chain. Two problems that maximize the cell sum rate subject to different constraints are formulated. Two algorithms are developed, one for each problem, that efficiently allocate resources to solve the problems. We show that neither UE ordering technique is consistently superior to the other. Also, we find that NOMA is able to outperform OMA only if a certain quality of SuIC can be guaranteed. Additionally, we show the existence of an optimum cluster size that maximizes cell sum rate given the network objective and residual SuIC factor. Our results indicate the significant improvement in performance that comes with interference-aware UE clustering. These results emphasize the importance of choosing network parameters such as cluster size and ordering technique depending on the objective of the network and network conditions such as the quality of SuIC.

In Chapter 7 we study the meta distribution in a downlink NOMA network. Two UE clustering schemes are considered: one with UEs distributed everywhere in the cell, and one with cell-center NOMA UEs. The moments of the meta distribution are derived for each scheme and the performance of the two schemes is compared. This chapter emphasizes the superiority of NOMA for cell-center UEs over UEs located anywhere. It also highlights the importance of careful resource allocation on not just the SCP but also on higher moments of the CCP.

8.2 Future Research Work

The work presented in this dissertation can be extended in a number of ways. Naturally, studying the impact of other efficient spectrum utilization techniques that are to be incorporated in 5G would be one direction. Analyzing the effects of integrating spectrum reuse techniques studied in this dissertation with other techniques such as cognitive radio would also be fruitful. Additionally, all of the work in this dissertation is done on single-tier networks. To achieve high data rates, 5G networks will be het-
erogeneous with multiple tiers operating simultaneously, some of which will be over the same frequency bands. Analyzing spectrum reuse technologies in this context will also be interesting.

We emphasized the importance of interference correlation when receivers are close by in the context of secrecy. It should be noted that NOMA UEs can be close by as well, particularly when clustering techniques for cell-center UEs are used such as those proposed in this dissertation. Additionally, because of the nature of SuIC decoding, NOMA UEs are looking to decode some of the same messages. Incorporating the impact of interference correlation for decoding messages of proximity NOMA UEs and using this information for retransmissions among NOMA UEs in a cooperative manner may be an interesting direction.

Another possible direction to extend the work of this dissertation is mmWave networks. As mentioned in Chapter 1, the currently available microwave spectrum is scarce while a large bandwidth is available at mmWave frequencies. Additionally, it has been shown that mmWave frequencies are suitable for cellular communication as long as the cell radius is limited [145]. However, mmWave networks are intrinsically different from microwave networks and hence the mathematical frameworks we have available to model currently deployed networks can not be used directly to compare the performance of mmWave networks. Due to their different nature, there is a need to study which of the currently used spectrum reuse technologies can be deployed in mmWave networks, and will actually benefit mmWave communication. The strong directionality requirement along with the vulnerability to blocking requires nontrivial changes to the modeling and analysis of mmWaves compared to microwave networks. Hence, it is important to develop a framework so that the performance mmWave networks can be fairly evaluated.
REFERENCES


ACCEP TED AND UNDER PREPARATION
PUBLICATIONS

Conference Papers


Journal Papers


A Proof of Proposition 3.1

The exact area of the shaded crescent, given \( r_c, r_d \) and \( \bar{\theta} \), can be found by \( A_{r_c, r_d, \bar{\theta}} \), where

\[
A_{r_c, r_d, \bar{\theta}} = r_c^2(\phi - \frac{\sin(2\phi)}{2}) - r_c(\bar{\theta} - \frac{\sin(2\bar{\theta})}{2}).
\]

The angle \( \phi = \pi - \arccos(\frac{r_d^2 + r_c^2 - r_c^2}{2r_c r_d}) \) and the distribution \( f_{\bar{\theta}}(\bar{\theta}) = \frac{1}{\pi}, \ 0 \leq \bar{\theta} \leq \pi \) is used instead of \( f_{\theta}(\theta) \) due to symmetry.

The average area, \( A \), of the shaded crescent is found numerically as,

\[
A = \int_0^\pi \int_0^R \int_0^\infty f_{r_c}(r_c) f_{r_d}(r_d) f_{\bar{\theta}}(\bar{\theta}) A_{r_c, r_d, \bar{\theta}} \ dr_c \ dr_d \ d\bar{\theta}.
\]

Define \( P_{r_c \neq r_c^2} \) as the probability that at least one BS lies in the crescent in Fig. 3.2. Due to the PPP assumption the number of BSs in an area is a Poisson RV and \( P(\text{no BSs lie in crescent}) = e^{-\lambda A} \). Therefore, by definition \( P_{r_c \neq r_c^2} = 1 - e^{-\lambda A} \).

The probability that the f-D2D UE lies closer to the r-D2D UE than its nearest BS, \( P_{r_c > r_d} \), is:

\[
P_{r_c > r_d} = \mathbb{P}(r_c > r_d) = \int_0^R (1 - F_{r_c}(x)) f_{r_d}(x) dx
\]

\[
= \int_0^R \frac{(2 - \omega)x^{1-\omega}}{R^{2-\omega}} e^{-\pi \lambda x^2} dx
\]

\[
= \frac{2 - \omega}{2R^{2-\omega}} \pi \Gamma(\frac{2 - \omega}{2}) \left[ \Gamma(\frac{2 - \omega}{2}) - \Gamma(\frac{2 - \omega}{2}, \pi \lambda R^2) \right].
\]
Using Jensen’s inequality,

\[ E[r_{c_2}] = E[\sqrt{r_c^2 + r_d^2 - 2r_c r_d \cos \theta}] \leq \sqrt{E[r_{c_2}^2]}. \]

Additionally, by Jensen’s \( E[X^2] \geq (E[X])^2 \), and \( E[r_c r_d \cos \theta] = 0 \) due to independence.

We thus approximate the mean of \( r_{c_2} \) as,

\[ \mu_{r_{c_2}} = \sqrt{(E[r_c])^2 + (E[r_d])^2} = \sqrt{\frac{1}{4\lambda}(\frac{2 - \omega}{3 - \omega})^2}. \]

Similarly, \( \mu_{r_{c_2}|r_c > r_d} = \sqrt{(E[r_c|r_c > r_d])^2 + (E[r_d|r_c > r_d])^2} \)

and \( \mu_{r_{c_2}|r_c < r_d} = \sqrt{(E[r_c|r_c < r_d])^2 + (E[r_d|r_c < r_d])^2} \).

The conditional PDFs are evaluated as follows,

\[
\begin{align*}
    f_{r_d|r_c > r_d}(x) &= \frac{\int_{x}^{\infty} f_{r_d|r_c}(x|y)f_{r_c}(y)dy}{P(r_d < r_c)} \\
    &= \frac{2x^{1-\omega}(\pi \lambda)^{\frac{3}{2}-\omega}e^{-\pi \lambda x^2}}{\left[\Gamma(\frac{3-\omega}{2}) - \Gamma(\frac{3-\omega}{2}, \pi \lambda R^2)\right]} \quad 0 \leq x \leq \bar{R} \\
    f_{r_c|r_c > r_d}(x) &= \int_{0}^{\min(x, \bar{R})} f_{r_c|r_d}(x|y)f_{r_d}(y)dy \\
    &= \frac{4(\pi \lambda)^{\frac{4}{2}-\omega}e^{-\pi \lambda x^2}x^{(\min(x, \bar{R}))^{2-\omega}}}{(2 - \omega)\left[\Gamma(\frac{3-\omega}{2}) - \Gamma(\frac{3-\omega}{2}, \pi \lambda R^2)\right]} \quad 0 \leq x \leq \infty \\
    f_{r_d|r_c < r_d}(x) &= \frac{\int_{0}^{x} f_{r_d|r_c}(x|y)f_{r_c}(y)dy}{P(r_d > r_c)} \\
    &= \frac{2 - \omega x^{1-\omega}(1 - e^{-\pi \lambda x^2})}{R^{2-\omega}(1 - P_{r_c > r_d})} \quad 0 \leq x \leq \bar{R} \\
    f_{r_c|r_c < r_d}(x) &= \frac{\int_{x}^{\bar{R}} f_{r_c|r_d}(x|y)f_{r_d}(y)dy}{P(r_d > r_c)} \\
    &= \frac{2\pi \lambda x^{(\bar{R}^{2-\omega} - x^{2-\omega})}}{e^{\pi \lambda x^2}R^{2-\omega}(1 - P_{r_c > r_d})} \quad 0 \leq x \leq \bar{R}.
\end{align*}
\]

By integrating, the conditional expectations follow as,

\[
\begin{align*}
    E[r_d|r_c > r_d] &= \frac{\Gamma(\frac{3-\omega}{2}) - \Gamma(\frac{3-\omega}{2}, \pi \lambda R^2)}{\sqrt{\pi \lambda}(\frac{3-\omega}{2}) - \Gamma(\frac{3-\omega}{2}, \pi \lambda R^2)} \\
    E[r_c|r_c > r_d] &= \frac{4(\pi \lambda)^{\frac{4}{2}-\omega}}{(2 - \omega)\left[\Gamma(\frac{3-\omega}{2}) - \Gamma(\frac{3-\omega}{2}, \pi \lambda R^2)\right]} \left[\frac{\Gamma(\frac{3-\omega}{2}) - \Gamma(\frac{3-\omega}{2}, \pi \lambda R^2)}{2(\pi \lambda)^{\frac{2}{2}-\omega}} + R^{2-\omega}\left(\frac{erfc(\bar{R}\sqrt{\pi \lambda})}{4\pi \lambda^{0.5}} + \frac{e^{-\pi \lambda R^2}}{2\pi \lambda}\right)\right]
\end{align*}
\]
\[\mathbb{E}[r_d | r_c < r_d] = \frac{2^{\frac{\omega}{3-\omega}}}{(1-P_{r_c > r_d})} \left( \frac{R}{3^{\frac{\omega}{3-\omega}}} - \frac{\pi \lambda}{2R^{3-\omega}} [\Gamma(\frac{3-\omega}{2}) - \Gamma(\frac{3-\omega}{2}, \pi \lambda R^2)] \right)\]

\[\mathbb{E}[r_c | r_c < r_d] = \frac{\text{erf}(\sqrt{\pi \lambda R})}{2\sqrt{\pi \lambda}} - \frac{\Gamma(\frac{5-\omega}{2}) - \Gamma(\frac{5-\omega}{2}, \pi \lambda R^2)}{(\pi \lambda)^{\frac{3-\omega}{2}} R^{3-\omega}(1-P_{r_c > r_d})} .\]

When \( r_c > r_d \) the mean of \( r_e (\neq r_{c_2}) \) is approximated by \( \frac{\mu_{r_{c_2}}_{r_c > r_c} + \mathbb{E}[r_c | r_c > r_d] - \mathbb{E}[r_d | r_c > r_d]}{2} .\)

This is because the r-D2D UE lies outside the crescent and we find \( r_e (\neq r_{c_2}) \) by averaging between the inner and outer radius of the crescent due to the distribution of the BSs being homogeneous. When \( r_c < r_d \), the r-D2D UE lies inside the crescent. To approximate the mean of \( r_e \neq r_{c_2} \), the averaging (due to homogeneity) is done assuming the distance from the r-D2D UE to the boundary of the arc varies from 0 to \( r_d - r_c \) half the time and from 0 to \( r_{c_2} \) for the other half; this results in the term \( \frac{\mu_{r_{c_2}} + \mathbb{E}[r_d | r_c < r_d] - \mathbb{E}[r_e | r_e < r_d]}{4} . \) These result in the approximation in (3.3).
B Proof of Lemma 3.1 & Lemma 3.2

Lemma 3.1

An r-D2D UE transmits if it satisfies the maximum transmit power constraint and the IP-condition. Hence,

\[
P_e = \mathbb{P}(r_d \eta d \rho_e < P_u \cap r_d \eta d \rho_e < T_d \eta d \rho_c)
= \mathbb{P}(r_d < \left(\frac{P_u}{\rho_e}\right)^{\frac{1}{n_d}} \cap r_e > \left(\frac{r_d \eta d \rho_e}{T_d \rho_c}\right)^{\frac{1}{n_e}})
= \int_0 \left(\frac{r_u}{\eta d}\right)^{\frac{1}{n_d}} f_{rd}(r)[1 - F_{re}\left(\left(\frac{r_d \eta d \rho_e}{T_d \rho_c}\right)^{\frac{1}{n_e}}\right)]dr.
\]

Lemma 3.1 is obtained by evaluating the above integral with \(f_{rd}(\cdot)\) given in (3.1) and \(F_{re}(\cdot)\) obtained from (3.2).

Lemma 3.2

An f-D2D UE transmits if it satisfies the IP-condition and the maximum transmit power constraint. Hence,

\[
P_d = \mathbb{P}(r_d \eta d \rho_d < T_d \rho_c \cap r_d \eta d \rho_d < P_u)
= \mathbb{P}(r_c > \left(\frac{r_d \eta d \rho_d}{T_d \rho_c}\right)^{\frac{1}{n_c}} \cap r_d < \left(\frac{P_u}{\rho_d}\right)^{\frac{1}{n_d}}).
\]

This can be calculated using similar steps to the proof of the r-D2D case shown in Lemma 3.1.
C Proof of Lemma 3.3

Denote by $Z$ the maximum allowed distance between a D2D pair for the f-D2D UE to be transmitting, i.e., $Z = \min\left(\left(\frac{P_u}{\rho_d}\right)^{\frac{1}{\eta_d}}, \left(\frac{T_d\rho_c\rho_e}{\rho_d}\right)^{\frac{1}{\eta_d}}\right)$ and $f_Z(z) = 2\pi\lambda \frac{\eta_d}{\eta_c} \left(\frac{\rho_d}{T_d\rho_c}\right)^{\frac{2}{\eta_c}} \times \frac{z^{2\eta_d-1} e^{-\pi\lambda \left(\frac{2\eta_d}{T_d\rho_c}\right)^{\frac{2}{\eta_c}}}}{1-\dot{q}}$, $0 \leq z \leq \left(\frac{P_u}{\rho_d}\right)^{\frac{1}{\eta_d}}$, where $\dot{q} = e^{-\pi\lambda \left(\frac{P_u}{\rho_e}\right)^{\frac{2}{\eta_c}}}$. For a D2D pair to be operating in FD we require both the r-D2D and f-D2D UE to be transmitting. Thus,

$$P_{FD} = \mathbb{P}(\text{r-D2D is transmitting} \cap \text{f-D2D is transmitting})$$

$$= \mathbb{P}\left(r_d^{\eta_d} < \frac{P_u}{\rho_c} \cap r_d^{\eta_d} < \frac{T_d\rho_c\rho_e}{\rho_d} \bigg| r_d < Z \right) P_d,$$

where $P_1$ denotes the conditional probability $\mathbb{P}(\text{r-D2D is transmitting}|\text{f-D2D is transmitting})$ and $P_d$ is given in Lemma 3.2. We evaluate $P_1$ as,

$$P_1 = \int_0^\infty f_{r_c}(g) \int_0^{\left(\frac{\min(P_u, T_d g \rho_c)}{\rho_c}\right)^{\frac{1}{\eta_d}}} \left(\frac{P_u}{\rho_d}\right)^{\frac{1}{\eta_d}} f_{r_d|z}(x|z) f_z(z) d\gamma \frac{d\gamma}{d\rho_c}$$

$$= \int_0^\infty f_{r_c}(g) \left(\frac{2-\omega}{2\eta_d}\right) \left(\frac{2-\omega}{\eta_c}\right) \frac{(2-\omega)\eta_c}{2\eta_d} \left(\frac{T_d\rho_c}{\rho_d(\pi\lambda)^{\frac{2-\omega}{\eta_d}}} - \dot{q} \times \left(\frac{\min(P_u, T_d g \rho_c)}{\rho_c}\right)^{\frac{2-\omega}{\eta_d}}\right) dg.$$
D \hspace{1em} \text{Proof of Lemma 3.4}

Denote by $X_c$ and $X_d$ the unconditional transmit powers required to invert the channel to the nearest BS and r-D2D UE, respectively; hence, $X_c = r_c^{\eta_c} \rho_c$ and $X_d = r_d^{\eta_d} \rho_d$. Using the PDFs of $r_c$ and $r_d$ we obtain $f_{X_c}(x) = \frac{2 \pi \lambda x^{\frac{2-\omega}{\eta_c c}} e^{-\pi \lambda (\frac{x}{\rho_c})^2 \frac{2}{\eta_c c}}}{\eta_c \rho_c^{\frac{2}{\eta_c c}}} 0 \leq x \leq \infty$ and $f_{X_d}(x) = \frac{2 - \omega}{R^2} x^{\frac{2-\omega}{\eta_d d} - 1} \frac{2 \pi \lambda x^{\frac{2-\omega}{\eta_d d}} e^{-\pi \lambda (\frac{x}{\rho_d})^2 \frac{2}{\eta_d d}}}{\eta_d \rho_d^{\frac{2}{\eta_d d}}} 0 \leq x \leq \bar{R}^{\eta_d} \rho_d$. As successful communication requires satisfying the IP-condition and the maximum transmit power constraint, the PDF of the transmit power in the f-D2D mode is given by,

$$f_{P_d}(x) = \int_{x}^{\infty} f_{X_d|T_c X_c}(x|y) f_{X_c}(y) dy$$

$$= \frac{2 - \omega}{R^2} x^{\frac{2-\omega}{\eta_d d} - 1} \frac{2 \pi \lambda y^{\frac{2-\omega}{\eta_c c}} e^{-\pi \lambda (\frac{y}{\rho_c})^2 \frac{2}{\eta_c c}}}{\eta_c \rho_c^{\frac{2}{\eta_c c}}} dy$$

$$= \frac{2 x^{\frac{2-\omega}{\eta_d d} - 1} e^{-\pi \lambda (\frac{x}{\rho_c T_d})^2 (\frac{2}{\eta_c c})}}{\eta_c (\rho_c T_d)^{\frac{2-\omega}{\eta_d d}} \gamma \left( \frac{2 - \omega}{\eta_d d}, \pi \lambda \left( \frac{P_u}{\rho_c T_d} \right)^{\frac{2}{\eta_c c}} \right)}, \hspace{1em} 0 \leq x \leq P_u.$$
Denote by $X_i$ and $X_e$ the unconditional transmit powers required to invert the channel to the nearest BS and f-D2D UE, respectively; hence, $X_i = r_i^n \rho_{e}$ and $X_e = r_e^n \rho_{e}$. Using the PDFs of $r_e$ and $r_d$ we obtain

$$f_{X_i}(x) = \frac{2bx^{\frac{\omega-2}{\eta_d} - 1} e^{-b(\frac{x}{\rho_{c}})^\frac{2}{\eta_c}}}{\eta_c \rho_{c}^\frac{2}{\eta_c}} \quad 0 \leq x \leq \infty,$$

and

$$f_{X_e}(x) = \frac{2-\omega}{\omega} x^{\frac{2-\omega}{\eta_d} - 1} \quad 0 \leq x \leq \bar{R}^n \rho_{e}. \quad \text{As successful communication in the r-D2D mode requires satisfying the IP-condition and the maximum transmit power constraint, the PDF of the transmit power in the r-D2D mode is,}$$

$$f_{P_e}(x) = \int_x^\infty f_{X_i|T_d X_i}(x|y) f_{X_i}(y) dy$$

$$= \frac{2-\omega}{\omega} \frac{x^{\frac{2-\omega}{\eta_d} - 1}}{\eta_d \rho_{e}^\frac{2}{\eta_d}} \frac{2b}{\eta_c \rho_{c}^\frac{2}{\eta_c}} \int_x^\infty \frac{2by^{\frac{2-\omega}{\eta_d} - 1} e^{-b(\frac{y}{\rho_{c}})^\frac{2}{\eta_c}}}{\eta_c \rho_{c}^\frac{2}{\eta_c}} dy$$

$$= \frac{2(T_d \rho_{e})^{(2-\omega) \eta_{d} b}}{\eta_{c}^{\frac{2-\omega}{2\eta_{d}}}} \frac{2bx^{(2-\omega) \eta_{d} - 1} e^{-b(\frac{x}{\rho_{c}})^\frac{2}{\eta_c}}}{\eta_{c}^{\frac{2}{2\eta_{d}}}}$$

$$= \frac{2(T_d \rho_{e})^{(2-\omega) \eta_{d} b}}{\eta_{c}^{\frac{2-\omega}{2\eta_{d}}}} \frac{2bx^{(2-\omega) \eta_{d} - 1} e^{-b(\frac{x}{\rho_{c}})^\frac{2}{\eta_c}}}{\eta_{c}^{\frac{2}{2\eta_{d}}}}.$$
F Proof of Lemma 3.6

Due to the PPP assumption, the cellular link distance $r_c$ follows a Rayleigh distribution mentioned in Section 3.2. Denote by $X_c = r_c^\eta_c \rho_c$ the unconditional transmit power required to invert the channel to the nearest BS, where $f_{X_c}(x) = \frac{2 \pi \lambda x^{(\frac{2}{\eta_c} - 1)} e^{-\pi \lambda (\frac{x^2}{\rho_c^2})}}{\eta_c \rho_c^{\frac{2}{\eta_c}}}; \quad 0 \leq x \leq \infty$. Since successful communication requires satisfying the maximum transmit power constraint, we require $P_c < P_u$. Hence, the PDF of the transmit power is obtained as

$$f_{P_c}(x) = \frac{f_{X_c}(x)}{\int_0^{P_u} f_{X_c}(y) dy} = \frac{2 \pi \lambda x^{\frac{2}{\eta_c} - 1} e^{-\pi \lambda (\frac{x^2}{\rho_u^2})}}{\eta_c \rho_c^{\frac{2}{\eta_c}} \left(1 - e^{-\pi \lambda (\frac{P_u}{\rho_c})^{\frac{2}{\eta_c}}}\right)},$$

where, $0 \leq x \leq P_u$. We obtain $E[P_c^\alpha]$ by $\int_0^{P_u} x^\alpha f_{P_c}(x) dx$. 
G Proof of Lemma 3.7

The LTs of the interferences experienced by receivers in mode $\chi$ from UEs in mode $\kappa$ are evaluated as,

$$
\mathcal{L}_{I_{\chi\kappa}}(s) = \mathbb{E}_\Phi_{\kappa} \left[ e^{-s \sum_{u_j \in \Phi_{\kappa}} P_{\kappa} h_j u_j^{-\eta_\chi}} \right]
$$

$$
\overset{(1)}{=} \mathbb{E}_\Phi_{\kappa} \left[ \mathbb{E}_P_{\kappa, h} \left[ e^{-s \sum_{u_j \in \Phi_{\kappa}} P_{\kappa} h_j u_j^{-\eta_\chi}} \right] \right]
$$

$$
= \mathbb{E}_\Phi_{\kappa} \left[ \prod_{u_j \in \Phi_{\kappa}} \mathbb{E}_P_{\kappa, h} \left[ e^{-s P_{\kappa} h_j u_j^{-\eta_\chi}} \right] \right]
$$

$$
\overset{(2)}{=} \exp \left( -2\pi U_\kappa \int_{\text{IP-boundary}_{\kappa\chi}}^{\infty} \mathbb{E}_P_{\kappa, h} \left[ 1 - e^{-sP_{\kappa}hu_j^{-\eta_\chi}} \right] u \, du \right)
$$

$$
\overset{(3)}{=} \exp \left( -2\pi U_\kappa \int_{\text{IP-boundary}_{\kappa\chi}}^{\infty} \mathbb{E}_P_{\kappa} \left[ 1 - \frac{1}{1 + sP_{\kappa}u^{-\eta_\chi}} \right] u \, du \right)
$$

where $\text{IP-boundary}_{\kappa\chi} = 0$ for $\{\chi \in \{d, e\}, \forall \kappa\}$ as the f-D2D and r-D2D modes do not offer any IP to their receivers. For $\kappa \in \{d, e\}$, IP-boundary$_{\kappa\chi} = \left( \frac{P_{\kappa}}{P_{\kappa}T_d} \right)^{\frac{1}{\eta_\chi}}$ as the BSs have IP from the transmitting f-D2D and r-D2D UEs. Additionally, an IP to BSs from cellular interferers exists as only one cellular UE serves the BS on a channel at a time; thus, interfering cellular UEs have IP-boundary$_{\kappa\chi} = \left( \frac{P_{\kappa}}{P_{\kappa}T_c} \right)^{\frac{1}{\eta_\chi}}$. The intensity of the interferers in mode $\kappa$ is denoted by $U_\kappa$, hence $U_\kappa$ for $\kappa \in \{d, e\}$ is $U_d$ and $U_e$, respectively, while $U_c = \lambda$. Due to the independence of fading, we have (1), and (2) follows from using the PGFL of the PPP. Using the MGF of $h \sim \exp(1)$, (3) is obtained. Continuing the integration in (3) and applying a change of variables, we arrive at the LTs of the interferences in Lemma 3.7.
H Proof of Theorem 4.1

By definition,

\[
P(S \cap F) = P(S) - \mathbb{E} \left[ e^{-\tau_1 R^u (I_{su} + \frac{\beta_a}{\beta_b} (I_{su} + 1, u))} - \tau_2 \beta \eta (I_{bu} + \frac{\beta_a}{\beta_b} (I_{bu} + 1, u)) \right].
\]

Since \( \Phi_b \) and \( \Phi_j \) are PPPs, employing the PGFL of the PPP and the MGF of \( \tilde{g}_0 \sim \exp(1) \) we have

\[
A = \mathbb{E}_{R, \theta} \left[ e^{-\lambda R, \theta} e^{-\lambda F(R, \theta)} \right].
\]

Here,

\[
B(r, \theta) = \int_0^{2\pi} \int_r^\infty \mathbb{E}_{h_x, g_x} \left[ 1 - e^{-\tau_1 r h_x ||x||} - \tau_2 \beta \eta g_x \right] xdx d\phi
\]

and

\[
F(r, \theta) = \int_0^{2\pi} \int_0^\infty q \left( 1 - e^{-3 \sqrt{q} ||y||} \right) \times
\]

\[
\mathbb{E}_{h_y, g_y} \left[ 1 - e^{-\tau_1 r h_y ||y||} - \tau_2 \beta \eta g_y \right] ydy d\phi.
\]

Due to the independence of the unit mean exponentially distributed \( h_x \) and \( g_x \) in \( B(r, \theta) \), and \( h_y \) and \( g_y \) in \( F(r, \theta) \) and using their MGFs, we obtain (4.9) and (4.10).
I Operation Scenario in Chapter 5

The operation scenario used in Chapter 5 is detailed below.

- A homogeneous PPP cellular network of intensity $\lambda = 10$ BSs/km$^2$ is employed. The PPP assumption for the locations of BS have several practical and theoretical validations \cite{12}.

- $N$ UEs are dropped uniformly in the Voronoi cell of each BS (cf. Fig. 5.3).

- A Rayleigh fading environment and a power law path-loss model, where the signal decays with distance $r$ as $r^{-\eta}$ and the path-loss exponent $\eta = 4$, are assumed.

- We set the power budgets to $P = 1$ W and $P_u = 0.2$ W for the BS and UEs, respectively.

- In the case of NOMA, the entire time-frequency block is used for transmission. In the case of OMA, the time resource (i.e., TDMA) is split between the $N$ UEs.

- Fixed rate transmission is employed and a global target-rate (i.e., the same target-rate for all UEs in a cluster) is used.

- Fixed rate transmission is vulnerable to outage, correspondingly the effective-rate is defined as coverage probability multiplied by target-rate. For NOMA, this is calculated using (5.7), which employs the coverage probability given in (5.6). Effective-rate for OMA is calculated using the general formula for
effective-rate in (5.3) multiplied by the fraction of the resource being shared (time fraction for TDMA).

- Power allocation and optimization objective for NOMA:
  - Downlink: Optimization objective is maximum fairness (i.e., symmetric effective-rate for all UEs). Power allocation is done accordingly so that effective-rates can be identical. Since global target-rate is used, identical effective-rates are achieved by a power allocation that can equalize SINRs for the UEs.
  - Uplink: In the uplink, superposition is not required by a transmitter (UE). All UEs transmit with full-power and therefore do not achieve symmetric effective-rate. The superiority of this approach and the challenges associated with employing a power allocation that achieves symmetric effective-rate in the uplink are highlighted in Section 5.3.2.