

ENERGY AWARE MANAGEMENT OF 5G NETWORKS

by

CHANG LIU

B.S., Jilin University, China, 2009

M.S., Jilin University, China, 2012

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Electrical and Computer Engineering
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2016

Abstract

The number of wireless devices is predicted to skyrocket from about 5 billion in 2015 to 25 billion by 2020. Therefore, traffic volume demand is envisioned to explode in the very near future. The proposed fifth generation (5G) of mobile networks is expected to be a mixture of network components with different sizes, transmit powers, back-haul connections and radio access technologies. While there are many interesting problems within the 5G framework, we address the challenges of energy-related management in a heterogeneous 5G networks. Based on the 5G architecture, in this dissertation, we present some fundamental methodologies to analyze and improve the energy efficiency of 5G network components using mathematical tools from optimization, control theory and stochastic geometry.

Specifically, the main contributions of this research include:

- We design power-saving modes in small cells to maximize energy efficiency. We first derive performance metrics for heterogeneous cellular networks with sleep modes based on stochastic geometry. Then we quantify the energy efficiency and maximize it with quality-of-service constraint based on an analytical model. We also develop a simple sleep strategy to further improve the energy efficiency according to traffic conditions.
- We conduct a techno-economic analysis of heterogeneous cellular networks powered by both on-grid electricity and renewable energy. We propose a scheme to minimize the electricity cost based on a real-time pricing model.
- We provide a framework to uncover desirable system design parameters that offer the best gains in terms of ergodic capacity and average achievable throughput for device-to-device underlay cellular networks. We also suggest a two-phase scheme to optimize the ergodic capacity while minimizing the total power consumption.

- We investigate the modeling and analysis of simultaneous information and energy transfer in Internet of things and evaluate both transmission outage probability and power outage probability. Then we try to balance the trade-off between the outage performances by careful design of the power splitting ratio.

This research provides valuable insights related to the trade-offs between energy-conservation and system performance in 5G networks. Theoretical and simulation results help verify the performance of the proposed algorithms.

ENERGY AWARE MANAGEMENT OF 5G NETWORKS

by

Chang Liu

B.S., Jilin University, China, 2009

M.S., Jilin University, China, 2012

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Electrical and Computer Engineering
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2016

Approved by:

Major Professor
Balasubramaniam Natarajan

Copyright

Chang Liu

2016

Abstract

The number of wireless devices is predicted to skyrocket from about 5 billion in 2015 to 25 billion by 2020. Therefore, traffic volume demand is envisioned to explode in the very near future. The proposed fifth generation (5G) of mobile networks is expected to be a mixture of network components with different sizes, transmit powers, back-haul connections and radio access technologies. While there are many interesting problems within the 5G framework, we address the challenges of energy-related management in a heterogeneous 5G networks. Based on the 5G architecture, in this dissertation, we present some fundamental methodologies to analyze and improve the energy efficiency of 5G network components using mathematical tools from optimization, control theory and stochastic geometry.

Specifically, the main contributions of this research include:

- We design power-saving modes in small cells to maximize energy efficiency. We first derive performance metrics for heterogeneous cellular networks with sleep modes based on stochastic geometry. Then we quantify the energy efficiency and maximize it with quality-of-service constraint based on an analytical model. We also develop a simple sleep strategy to further improve the energy efficiency according to traffic conditions.
- We conduct a techno-economic analysis of heterogeneous cellular networks powered by both on-grid electricity and renewable energy. We propose a scheme to minimize the electricity cost based on a real-time pricing model.
- We provide a framework to uncover desirable system design parameters that offer the best gains in terms of ergodic capacity and average achievable throughput for device-to-device underlay cellular networks. We also suggest a two-phase scheme to optimize the ergodic capacity while minimizing the total power consumption.

- We investigate the modeling and analysis of simultaneous information and energy transfer in Internet of things and evaluate both transmission outage probability and power outage probability. Then we try to balance the trade-off between the outage performances by careful design of the power splitting ratio.

This research provides valuable insights related to the trade-offs between energy-conservation and system performance in 5G networks. Theoretical and simulation results help verify the performance of the proposed algorithms.

Table of Contents

List of Figures	xiii
List of Tables	xv
Acknowledgements	xv
Dedication	xvi
Abbreviations	xvii
1 Introduction	1
1.1 What is 5G?	1
1.2 Challenges	2
1.3 Research Questions	4
1.4 Related Work	6
1.4.1 Sleep Modes for Small Cells	6
1.4.2 Energy Harvesting Base Stations	8
1.4.3 D2D Underlay Cellular Networks	11
1.4.4 Energy Considerations in Internet of Things	13
1.5 Contributions	14
1.6 Dissertation Organization	17
2 Background	19
2.1 Stochastic geometry	19

2.1.1	Overview	19
2.1.2	Related Work and Motivation	20
2.2	Point Processes	21
2.2.1	Point Process	21
2.2.2	Poisson Point Process	22
2.2.3	Determinantal Point Processes	23
2.2.4	Ginibre Point Process	23
2.3	Channel Model	24
2.3.1	Channels in a wireless network	24
2.3.2	Propagation Models	24
2.4	Key Performance Metrics	26
2.4.1	Signal-to-interference-plus-noise Ratio	26
2.4.2	Coverage Probability	27
2.4.3	Channel Capacity	27
2.4.4	Ergodic Capacity	28
2.4.5	Average Achievable Throughput	28
2.5	Summary	29
3	Small Cell Base Station Sleep Strategies for Energy Efficiency	30
3.1	System Model	32
3.1.1	Power Consumption Model	32
3.1.2	Sleep Modes for Small Cell BSs	33
3.1.3	Sleep Strategies for Small Cell BSs	34
3.1.4	Heterogeneous Cellular Network Model	34
3.1.5	Energy Efficiency Metric	35
3.2	Random Sleeping	36

3.2.1	Average Achievable Rate for Random Sleeping	36
3.2.2	Coverage Probability Consideration	37
3.2.3	Delay Constraint Consideration	38
3.2.4	Energy Efficiency for Random Sleeping Strategy	38
3.2.5	Problem Formulation	41
3.2.6	Analysis of Optimization Design	42
3.3	Strategic Sleeping	45
3.4	Simulation Results	47
3.4.1	Random Sleeping Strategy	47
3.4.2	Strategic Sleeping Strategy	50
3.5	Summary	52

4 Power Management in Heterogeneous Networks with Energy Harvesting

	Base Stations	53
4.1	System Model	54
4.1.1	Harvested Renewable Energy Model	55
4.1.2	Real-time Electricity Price Model	56
4.1.3	Base Station Management	56
4.1.4	Lattice Model	57
4.2	On-grid Power Cost Minimization Formulation	57
4.3	Stage 1 - Optimization of Transmit Power	61
4.3.1	Simplification of Signal-to-interference Ratio Constraint	61
4.3.2	Convergence of Algorithm 4	64
4.3.3	Simulation Result	65
4.3.4	Problem Transformation	65
4.4	Stage 2 - Controlling Stored Energy Usage - NMPC Framework	67

4.4.1	NMPC Formulation	68
4.4.2	Solution of the NMPC Problem	69
4.5	Simulation Results	72
4.5.1	Case-1	72
4.5.2	Case-2	73
4.5.3	Case-3	74
4.5.4	Case-4: Effect of Time Window	76
4.6	Summary	76

5 Power-aware Performance Analysis and Optimization in D2D Underlay

	Networks	78
5.1	System Model	80
5.1.1	Network Model	80
5.1.2	Success Probabilities	81
5.2	Ergodic Capacity Analysis	82
5.2.1	Ergodic Capacity of CUs	83
5.2.2	Ergodic Capacity of DUs	86
5.2.3	Simulation Results	90
5.3	Optimization of Ergodic Capacity and Power Consumption	91
5.3.1	Stage-1: Ergodic Capacity Optimization	92
5.3.2	Stage-2: Optimization of Ergodic Capacity and Power Consumption .	94
5.4	Average Achievable Throughput Analysis	99
5.4.1	Average Achievable Throughput of CUs	100
5.4.2	Average Achievable Throughput of DUs	104
5.4.3	Simulation Results	108
5.5	Ergodic Capacity with Varying User Distributions	111

5.5.1	Network Model	111
5.5.2	Ergodic Capacity of DUs	111
5.5.3	Simulation Results	118
5.6	Summary	120
6	Modeling and Analysis of Simultaneous Information and Energy Transfer	122
6.1	System Model	123
6.2	Performance Analysis	126
6.2.1	Transmission Outage Probability	126
6.2.2	Power Outage Probability	132
6.3	Convexity of the Maximum Outage Probability Upper Bound	134
6.4	Simulation Results	136
6.5	Summary	140
7	Conclusion	143
7.1	Conclusion	143
7.2	Future Work	145
	Bibliography	148
A	Proofs of Chapter 5	165
A.1	Proof of Lemma 6	165
A.2	Proof of Theorem 4	166

List of Figures

1.1	A view of 5G networks	2
2.1	Channel models	24
3.1	Schematic diagram of d_{ij} , v_i and v_{ij}	48
3.2	q_{on} vs coverage probability	49
3.3	Energy efficiency improvement vs t_{min}	50
3.4	Ratio of EE with strategic sleeping policy to EE with random sleeping policy	51
4.1	Simulation Results of Case-1	73
4.2	Simulation Results of Case-2	74
4.3	Simulation Results of Case-3	75
4.4	Effect of Time Window	77
5.1	C_C vs. P_C	91
5.2	C_C vs. λ_C	91
5.3	C_D vs. P_T	92
5.4	C_D vs. λ_T	92
5.5	γ vs. \hat{C}	93
5.6	DTU density λ_T vs. approximated ergodic capacity \hat{C}	95
5.7	Achievable Area in Case 1	96
5.8	Achievable Area in Case 2	97
5.9	Achievable Area in Case 3	98

5.10	Ergodic capacity and total power consumption for different choices of parameters (P_C, P_T, λ_T). The black pentagram represents the optimal case while green circles represent other cases.	99
5.11	\bar{R}_C vs P_C	109
5.12	\bar{R}_C vs λ_C	110
5.13	\bar{R}_D vs P_T	110
5.14	\bar{R}_D vs λ_T	111
5.15	P_T vs C_D	119
5.16	R_{max} vs C_D	120
6.1	IoT Network Model	124
6.2	IoT Network Model	125
6.3	ϵ vs. P_{outage}^T	137
6.4	P_G vs. P_{outage}^T	138
6.5	ϵ vs. P_{outage}^P	139
6.6	P_G vs. P_{outage}^P	140
6.7	ϵ for Maximum outage probability vs. different θ	141
6.8	ϵ^* for Maximum outage probability vs. different P_G	141
6.9	ϵ for Maximum outage probability vs. different ρ	142

List of Tables

3.1	Wake-up times and power consumption for different sleep modes	34
3.2	Simulation parameters	48
3.3	Simulation results of optimal sleep mode probabilities	49
4.1	Notation Table	58
4.2	Power Optimization Simulation Results	65
4.3	Simulation parameters	72
4.4	Simulation Results	76
5.1	Simulation parameters	90
5.2	Simulation parameters	109
5.3	Simulation parameters	118
6.1	Simulation parameters	137

Acknowledgments

I would like to give my sincere thanks for all people that have helped me in the past four years while I pursued my Ph.D. study and research.

First of all, I would like to express my immense gratitude to my advisor, Dr. Bala Natarajan, for giving me an opportunity to pursue my Ph.D. degree in Kansas State University. His guidance and encouragement have been constantly encouraging and directing me through difficulties not only in my Ph.D. study but in so many difficult and trying life events as well.

Secondly, I am also grateful to my committee for their input, whether in teaching classes from which I benefited, or in the discussions we have had on various academic or applied topics. Prof.s Gruenbacher, Kuhn, Albin and Anderson, thank you.

Thirdly, I am beholden for the chance to be a part of such a unique group of students, faculty and staff in Electrical and Computer Engineering Department at Kansas State University. The members of the wireless communication (WiCom) group in particular have made my experiences memorable. My lab-mates Nick, Siddharth, Sayak, Mohammed, Shafiul, Kan, Wenji, Kumar, Hammad, Dale and Chenyu, have all provided elements of an experience that will shape the rest of my life. Their input and friendship during this time will forever be a part of my model of earnest academic personal development, creative thinking, and sincere conversation.

Also, The professors I have worked under in fulfilling my teaching responsibilities, Prof.s Natarajan, Soldan, Dyer, Mirafzal, Day, Hageman and Chandra, have been most gracious and with whom working has most enjoyable. They have encouraged me to try new teaching tactics and have often provided technical aid when necessary.

Last but not least, to my parents and parents-in-law, who have been giving us a hand with taking care of our little angel Lanyin through all those challenging periods.

Dedication

To Zhenlun Liu and Rongman Xiao, my dear father and mother, Lei Wei and Huixia Wang, my dear father-in-law and mother-in-law. I wish I can always make you proud.

To Hui Wei, my wonderful husband who is too good to be true. You raise me up to more than I can be.

To Emma Lanyin Wei, my precious daughter. You are my little angel who have brought me so much luck and joy.

To Yuye Liu, Yan Zheng, Xiaobin Xu, Jipeng Zhu, Qi Zhao, Fang Cao, Lin Li, Lei Xu, Lezhuang Wang, Chunlai Yang, Qiongqiong Pu, Lei Liu, Tao Luo, Qian Zhao, Guangxing Min, Yu Tian and Lin Wei, my wonderful friends. Thank you all for being there for me as always.

Words cannot express how much I love you!

Abbreviations

5G	the 5 th generation
BS	base station
PA	power amplifier
D2D	device-to-device
IoT	Internet of things
HCN	heterogeneous cellular networks
QoS	quality of service
FCC	Federal Communications Commission
OFDMA	orthogonal frequency division multiple access
FBS	femtocell base station
UE	user equipment
MBS	macrocell base station
PBS	picocell base station
SINR	signal-to-interference-plus-noise ratio
PPP	Poisson point process
LOS	line-of-sight
NLOS	non-line-of-sight
OFDM	orthogonal frequency-division multiplexing
MIMO	multiple-input multiple-output
NMPC	nonlinear model predictive control
GPP	Ginibre point process
DPP	determinantal point process

SIR	signal-to-interference ratio
EE	energy efficiency
RF	radio frequency
FPGA	field-programmable gate array
TCXO	temperature compensated crystal oscillators
CCDF	complementary cumulative distribution function
ROI	region of interest
MPC	model predictive control
CU	cellular user
DU	D2D user
DTU	D2D transmitting user
DRU	D2D receiving user
PDF	probability density function
CDF	cumulative density function

Chapter 1

Introduction

1.1 What is 5G?

When people ask what 5G is, perhaps the only thing we can say with certainty is: it is the 5th generation of mobile networks. This answer may not provide much information, but until around 2020 (the expected time of 5G roll out to meet business and consumer demands [1]), it is too early to define 5G with any certainty. The good news is, we can describe what people are looking for in 5G [2]: to connect the entire world and achieve seamless and ubiquitous communication between people and people (*anybody*), between people and machine, or between machine and machine (*anything*), wherever they are (*anywhere*), whenever they need (*anytime*), by whatever electronic devices/services/networks they wish (*anyhow*). This includes but is not limited to: 1000 times the system capacity, 10 times the spectral efficiency, 10 times the end-user data rate, and 25 times the average cell throughput, all relative to 4G LTE.

To achieve these goals, a general consensus among organizations, governments, academia and industries is that 5G may require integration of several techniques, scenarios and use cases rather than the invention of a new single radio access technology. As is shown in Fig. 1.1, 5G wireless networks are expected to be a mixture of network tiers with different

sizes, transmit power, back-haul connections and radio access technologies. Based on this architectural enhancement, in this dissertation, we wish to explore some promising energy-related technologies that can be deployed to meet 5G requirements.

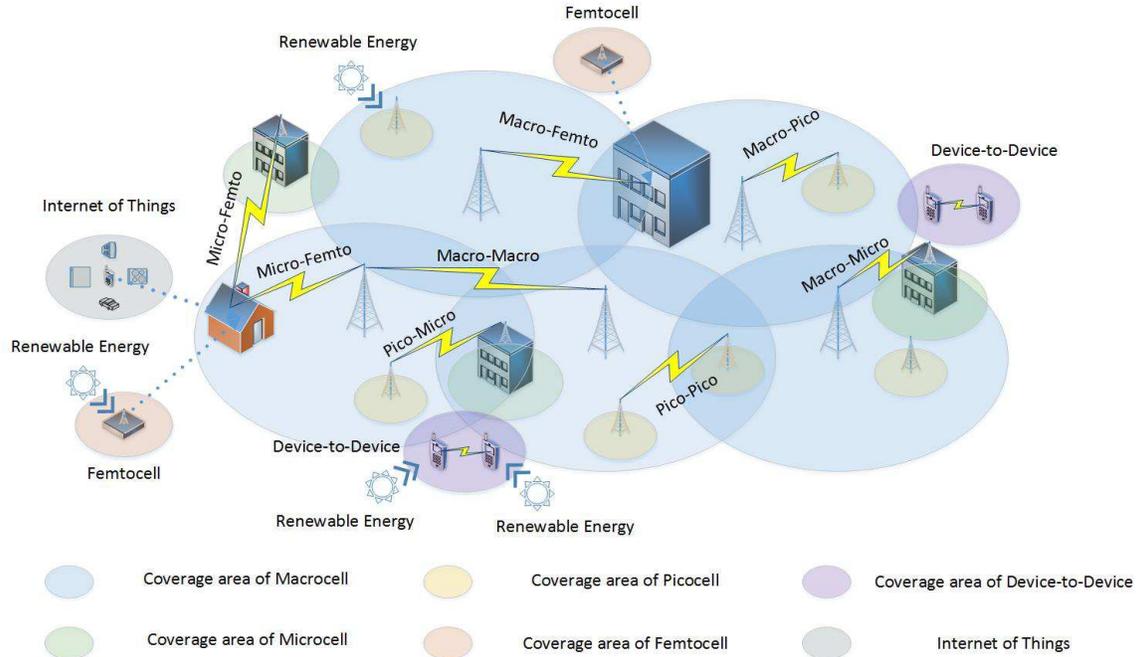


Figure 1.1: *A view of 5G networks*

1.2 Challenges

To support all the requirements outlined in Section 1.1, the densification and expansion of 5G networks will pose new challenges related to energy efficiency. With a predicted wireless data explosion [2], the total energy consumption may easily climb to an unacceptable level. Therefore, the network energy efficiency of the communication - typically measured in either Joules/bit or bits/Joule - will need a several-order-of-magnitude increase to at least maintain the power consumption. In this section, we present some research issues related to power-saving techniques to improve energy efficiency of 5G networks.

- *Base Station Energy Efficiency:* According to the analysis of DOCOMO [3], a Japanese operator, 75 – 80% of the energy consumption is at the level of base station (BS). So the center of attention of cellular network operators is at the BSs. The energy consumption of a typical BS can be reduced by improving the BS hardware design and by including additional software and system features to balance energy consumption and performance. In order to improve hardware design of a BS for energy consumption, we need to address the energy efficiency of the power amplifier (PA) (which dominates the energy consumption of a BS) as given by,

$$\eta_{PA} = \frac{P_{out}^{PA}}{P_{in}^{PA}}. \quad (1.1)$$

Traditional PAs have an efficiency of only about 15% [4]. The excessive energy is transformed into heat. That is, the biggest part of the total energy budget is actually consumed by the hardware (e.g., coolers and circuit energy consumption)

Therefore, the ability to shut down infrastructure nodes (or parts of them) or to adapt the transmission strategy according to the traffic will therefore become an important design aspect of energy-efficient wireless architectures [5, 6]. The modeling, optimization and operation of such systems are highly nontrivial.

- *Renewable Energy Management:* Renewable energy can be obtained from multiple sources, such as wind, solar, hydro-power, geothermal, and biomass. Recently, renewable energy has been introduced as a promising solution to reduce the “non-green” energy consumption of cellular networks. Nevertheless, the use of renewable energy also introduces some technical challenges. The first challenge encountered is their variability. For example, in the case of solar energy, the harvested amount is highly affected by weather and day-night cycle. A second challenge is that the typical power density of renewable power sources is significantly smaller than that of BS loads. As a result, energy harvesting BSs are mostly also powered by on-grid power. The

smart management for cost reduction of cellular networks with hybrid energy supply and variable electricity price has attracted interest from both academia and industry. Also, the feasibility study of low-power wireless nodes (e.g., sensors, actuators) powered solely by renewable energy is an open problem.

- *Energy-Aware Device-to-Device Network Planning:* Device-to-Device (D2D) communication refers to direct communication of cellular users in proximity without going through BSs. 5G networks is likely to incorporate D2D communication as an underlay to improve energy efficiency and meet the growing demand for high data rates. Due to the proximity of devices, gains can be expected in terms of higher data rates at low energy consumption and small latency. However, it is challenging to decide when D2D communication is appropriate and then to design an appropriate protocol. The key to efficient operation in this scenario lies in the coordination of interference and finding an optimal trade-off between cellular and D2D communication.
- *Energy Limitations in Internet of things:* The Internet of Things (IoT), which adds “anything” as an additional dimension to connectivity (in addition to anywhere and anytime), is emerging as a building block in 5G networks. The limited battery power of hardware elements is stretched to its limits while collecting and transmitting data. The lifetime of any resource in IoT depends upon the availability of energy. The loss of energy affects the whole environment under observation. Thus, there is a pressing need to reduce energy consumption for the prolonged lifetime of resources and the effective operation of IoT systems.

1.3 Research Questions

In this dissertation, we ask the following questions related to the energy efficiency challenges:

Question 1. (Power-saving mode design for small cells) *In a heterogeneous cellular network (HCN), can we turn off or partially turn off some small cell BSs? If we turn off some small cells, can we still maintain the quality of service (QoS)? How can we optimize the trade-off between power consumption and network capacity based on an analytical model? How should we design a sleep strategy according to the traffic load to maximize the energy efficiency enhancement?*

Question 2. (Management of energy harvesting base stations) *For BSs powered by both renewable energy and grid power, with time-varying harvested energy and electricity price, how may service providers choose to minimize grid power consumption? Specifically, how should we adjust transmit power of each BS to minimize overall power consumption while guaranteeing proper coverage? Should we save harvested energy for future use when electricity price is higher? If so, how should we manage the stored energy usage?*

Question 3. (Device-to-device performance analysis and optimization) *D2D underlay cellular networks possess advantages but also result in interference between cellular users and D2D terminals. Is there a critical set of system parameters (density of D2D users, cellular BSs, transmit power etc.) that can ensure that the benefits of D2D underlay operation can outweigh its drawbacks? Since D2D terminals are typically in proximity, how should we model D2D underlay cellular networks to capture the line-of-sight (LOS) component in D2D links? How should we evaluate the capacity of D2D underlay cellular networks? Can we jointly optimize the network capacity and the power consumption?*

Question 4. (Simultaneous information and energy transfer in Internet of things) *Is it feasible to integrate simultaneous information and energy transfer in IoT networks to prolong their lifetime? How should we model the IoT network to capture the repulsion among energy sources? How should we analytically evaluate the outage performances? How should we*

manage the power split ratio to balance the trade-off between outage probabilities? How would the network parameters (e.g., density and transmit power of gateways, power splitting ratio and SINR threshold) affect the outage performances?

1.4 Related Work

1.4.1 Sleep Modes for Small Cells

As one of the most effective approaches for small cell energy savings, sleep mode was initially proposed in IEEE 802.11b [7]. The goal was to reduce energy consumption at the link level, especially for non-real time traffic. However, within the sleep mode framework, some key issues must be considered. When BSs are switched off, radio coverage and QoS must still be guaranteed. As BSs are densely deployed, users in sleeping BS coverage can be served by neighboring active BSs by slightly increasing BS transmit power. For sleep mode operation, microcells and picocells can always be managed by operators. According to the Federal Communications Commission (FCC), even femtocells can sometimes be controlled by a wireless BS controller [8].

Recent efforts have been made related to power saving in cellular networks with the introduction of sleep modes [9–14] for BSs. Ref. [10] proposes a protocol for orthogonal frequency division multiple access (OFDMA) macro-femtocell network to allow femtocell base stations (FBSs) to completely switch off its radio communication and associated processing when not involved in an active call. The authors in [11, 12] discuss three different sleep strategies controlled by small cell, core network, and user equipment (UE), respectively. Power consumption is reduced by using sleep mode in low traffic hours in [13]. And [14] presents a simple sleeping scheme named fixed time sleeping as a case study for saving the energy of FBSs.

In all prior efforts, the emphasis was on reducing energy consumption. Also, how to understand the trade-off between energy consumption and QoS has been addressed and

investigated in [15–23]. The work in [15] studies power consumption minimization with QoS constraints, such as detection probability and false alarm probability of cognitive small cell networks. In [16] a trade-off between power consumption and throughput QoS is introduced, and it is concluded that sleep mode operation is effective when the cell size is small and traffic is light. The authors in [17] propose a clustering based power saving algorithm for femtocell networks to keep FBSs in the energy saving mode for the most time while preserving QoS experienced by users. In [18], the authors propose to save power by regularly checking the traffic load level and making sleep decisions based on a certain threshold. Simulation results show that at very high traffic load it is possible to increase end-user performance and decrease energy consumption at the same time. Authors in [19] study the optimal sleep/wake up schemes for FBSs deployed within macrocell base stations (MBSs) to offload part of its traffic. The aim is to minimize the energy consumption of the overall heterogeneous network while preserving the QoS experienced by users. The approach in [20] reduce the energy per information bit by making picocell base stations (PBSs) transmit at full load during good channel conditions and go into sleep otherwise. A maximum packet delay is defined to satisfy required QoS. Authors in [21] investigate and compare three dynamic algorithms to control sleep modes of small cells with lower power nodes to maximize energy saving while minimizing any negative effect on the end user experience. In [22], via adaptation of MBS transmission power and introduction of sleep mode for micro BS, network energy consumption is reduced while requested network capacity is guaranteed. The study in [23] devises a sleep mode mechanism that maximizes energy saving of the network while maintaining both the user throughput and the reliability performance.

Energy efficiency is becoming an important metric in evaluating next generation small cell networks [24]. However, there are very few specific efforts that attempt to address how service providers may choose to optimize the trade-off between power consumption and network throughput based on an analytical model. In [25], optimal sleep modes are determined for MBSs based on power consumption minimization and energy efficiency (EE)

maximization, respectively. However, according to their work, the optimal solution for EE maximization is to switch off all MBSs due to the quickly decreasing rate of network throughput. Simultaneous consideration of other parameters, such as traffic demands which would result in a new trade-off, is still an open issue. In addition, high density of small cell BSs is needed for switching off MBSs without incurring coverage hole.

Instead of traditional on-off mode, several multilevel sleep modes are proposed [26–28]. IEEE 802.16m [26] proposes two power saving mechanisms: sleep mode for UE and low duty operation mode for FBSs. The authors in [27] use active, listening, and sleep strategies for heterogeneous mobile network, respectively. In [28] four sleep modes are further proposed, jointly optimizing power consumption and wake-up time in femtocell network. However, no multilevel sleep modes are applied to HCNs to date.

To summarize, to the best of our knowledge, to study the trade-off between energy consumption and QoS based on an analytical model and optimize EE of HCNs with multilevel sleep modes is still an open problem.

1.4.2 Energy Harvesting Base Stations

As one of most effective approaches for energy savings in HCNs, energy harvesting BSs have been extensively studied recently [29]–[34]. Ericsson [29] has developed a wind-powered tower for cellular network BSs. Nokia Siemens Networks [30] has also developed a green BS which relies on a combination of solar and wind power to avoid using any grid electricity. Authors in [31] study structure, design, and control system of 3 KW wind and solar hybrid power systems for 3G BS. [32] proposes a feasible configuration of a stand-alone PV/Wind Hybrid Energy system with diesel generator as a backup for cellular mobile telephony BS site in isolated areas of central India. [33]–[34] describe a wind turbine generator system for supplying power to a small-scale BS on an island. It is expected that new energy harvesting systems will be integrated into mobile telephony BSs in near future, providing opportunities to reduce energy costs for operators.

There have been many efforts related to resource allocation in energy harvesting networks [35]-[39]. Authors in [35] consider delay minimization for interference networks with renewable energy source, where the transmission power of a node comes from both the conventional AC power and the renewable energy source. [36] considers energy allocation problem for point-to-point wireless communications with energy harvesters to maximize the throughput. [37]-[38] maximize the short-term throughput of an energy harvesting node under a deadline constraint. In [39] an online algorithm jointly manages the energy and makes power allocation decisions for packet transmissions to achieve close-to-optimal utility performance in energy-harvesting networks with finite capacity energy storage devices.

Besides, intelligently operating energy harvesting BSs in HCNs related to power saving has also received considerable interest [40]-[49]. A new tractable model for K-tier HCNs is developed in [40], where each BS is powered solely by a self-contained energy harvesting module. [41] proposes a hand over parameter tuning algorithm and a power control algorithm to guide mobile users to access BSs with green energy supply, thus reducing grid electricity expense and CO_2 emission. [42] deals with several aspects including lower power consumption that must be considered when a radio BS site is driven from a solar power source. Results in [40]-[42] are only valid for systems with a single energy source and not applicable to cellular networks employing hybrid energy BSs. However, as discussed in [43], a BS powered solely by an energy harvester may not be able to maintain stable operation and guarantee quality of service (QoS). Therefore, a hybrid energy harvesting system design is preferable in practice for providing uninterrupted service.

There have been some recent efforts that consider BSs with mixed power supply from both renewable energy sources and power grid. The reliable grid power guarantees that the service requirement is satisfied, while effective renewable energy allocation policy reduces grid power consumption. Authors in [44] proposes a two-stage dynamic programming (DP) algorithm to optimize on-grid power consumption over BSs' on-off states and allocation of BSs' resource blocks. [45] optimizes energy utilization in such networks by maximizing

utilization of green energy, and thus saving on-grid energy. An energy aware cell size adaptation algorithm named ICE is proposed in [46]-[47], which balances the energy consumption among BSs, and enables more users to be served with green energy. Authors in [48] study resource allocation algorithm design for energy-efficient communication in an orthogonal frequency division multiple access (OFDMA) downlink network with hybrid energy harvesting BS. A power optimization problem with average delay constraint on the downlink of a Green Base station is considered in [49]. This Green BS is powered by both renewable energy as well as conventional sources like diesel generators or the power grid. The authors try to minimize energy drawn from conventional energy sources and utilize harvested energy to the maximum extent. The optimal action consists of scheduling the users and allocating the optimal transmission rate for the chosen user.

Alternately, cost reduction for energy harvesting cellular networks has also been investigated [50]-[51]. [50] poses a novel cellular network planning problem to reduce capital and operational expenditure (CAPEX and OPEX, respectively), considering the use of renewable energy sources. A fundamentally new concept of energy balancing, and a novel algorithm to accomplish it is proposed. Authors in [51] present a novel optimization framework enabling cellular network planning to reduce deployment cost that takes into consideration two state-of-the-art technologies: (i) Dynamic Spectrum Access, and (ii) energy harvesting.

In summary, prior efforts on energy harvesting cellular network while extensive, either focus on technical issues or economic issues. It is not only prudent but in fact critical to take a techno-economic perspective of energy harvesting BS management. For example, how should an operator minimize on-grid power expenditure by management of BSs while considering varying real-time electricity price? To the best of our knowledge, this is an open question that requires a more holistic view of both technical and economic aspects of HCNs.

1.4.3 D2D Underlay Cellular Networks

Recently, there has been a surge of interest in interference management of D2D communication in cellular networks [52–54], including mode selection, resource allocation, power control, etc. [23, 55–63]. For example, [55] focuses on the study of neighbor discovery signal design and interference management for D2D discovery under a typical LTE system deployment. Authors in [56] propose a resource scheduling strategy within interference-suppression-area to improve performance of D2D communication underlying cellular networks. [57] exploits clustering of D2D users, frequency reuse over clusters and then uses interference alignment to enhance the sum rate of D2D-enabled cellular networks. [58] proposes an interference management strategy to enhance the overall capacity of cellular networks and D2D systems. Authors in [59] propose a way of applying interference-aware interference mitigation algorithms to D2D communication in cellular networks for system throughput improvement. [60] studies the intelligent resource allocation for D2D pairs. They propose a partial time-frequency resource allocation framework to improve the overall system capacity and stability while meeting the D2D users' QoS requirement. In [61] the authors provide a performance assessment of a power control algorithm with variable target signal-to-interference-plus-noise ratio (SINR) applied to cellular and D2D communication. A new interference management scheme [23] is proposed for D2D multicast to increase the overall capacity and improve resource utilization. [62] proposes a method to properly choose the cellular UE that shares the radio resource with D2D users to mitigate the interference from the cellular link to the D2D receivers. [63] derives an optimal transmission power under an average power constraint to maximize the ergodic capacity, based on a simplified model involving a pair of D2D users, a cellular user, a relay node and eNB.

Stochastic modeling of BS locations has spawned additional efforts in this arena. For example, interference management of D2D communication based on stochastic geometry has been considered in [64–72]. [64] uses tools from stochastic geometry to evaluate the performance of cognitive and energy harvesting-based D2D communication in cellular networks

with general path-loss coefficient. [65] addresses two fundamental issues in D2D enhanced cellular networks, such as how D2D users (DUs) should access spectrum (overlay and underlay) and how DUs should choose between communicating directly or via the BS. Authors in [66] build a tractable model for D2D communication underlay multi-cell cellular networks to evaluate the coverage probability and area spectral efficiency by utilizing stochastic geometry. They adopt exclusion regions around BSs to limit the locations of cellular users and active DUs to mitigate interference in the hybrid system. [67] applies stochastic geometry for D2D network analysis and investigates infinite and finite coverage area and exclusion region. Coverage study is provided in [68] by deriving uplink and downlink SINR distributions for both cellular users and dense DUs based on statistical user distribution and channel information. [69] proposes a two-phase scheme to maximize the ergodic capacity of D2D networks as well as guaranteeing the reliability of D2D communication. The distribution of D2D pairs are modeled by Poisson point processes (PPP). Authors in [70] use results from stochastic geometry to derive the probability of successful content delivery in the presence of D2D interference and noise, and maximize the total probability of content delivery by finding the best caching distribution. [71] proposes a fully-distributed random access protocol for the D2D communication in a cellular network assuming that the locations of devices follow a PPP. In the above literature, all communication links are modeled as Rayleigh fading, which means that all the links are assumed to be dominated by non-line-of-sight (NLOS) components. Unlike prior efforts, in [72] a Rician fading channel is used to characterize LOS components in D2D direct links, which is shown to be a more realistic model for mobile-to-mobile communication systems with dominant LOS [73]. To the best of our knowledge, [72] is the only work which considers Rician fading for D2D links. However, in [72], distances between all D2D pairs are considered as a fixed constant, which reduces the applicability of the derived results.

1.4.4 Energy Considerations in Internet of Things

To improve the energy efficiency of battery-equipped IoT system has emerged as a major research issue. For example, [74] proposes an energy-efficient architecture for IoT consisting of three layers including sensing and control, information processing and presentation. [75] investigates an energy-efficient sleep scheduling with QoS consideration in 3GPP LTE-A networks for IoT. The authors aim to maximize the sleep periods of devices while guaranteeing the QoS especially on the aspects of traffic bit-rate, packet delay and packet loss rate in IoT applications. [76] presents two energy efficient methods for detecting replicas for resource-limited mobile devices in IoT.

Integrating power transfer into wireless communications for supporting simultaneous information and energy transmission (SIET) is a promising technique in energy-constrained wireless networks, including orthogonal frequency-division multiplexing (OFDM) relaying system [77], multiple-input multiple-output (MIMO) wireless broadcast system [78], distributed antenna communication systems [79], full duplex wireless-powered communication network [80], battery-free wireless sensor networks (WSN) [81]. [77] considers SIET in an OFDM relaying system. It first studies transmission protocols for different cases, and then formulates joint resource allocation problems to maximize the system throughput. [78] pursues a unified study on SIET of a MIMO wireless broadcast system. The authors derive the optimal transmission strategy to achieve different trade-offs for maximal information rate versus energy transfer. [79] studies the resource allocation algorithm design for secure information and renewable green energy transfer to mobile receivers in distributed antenna communication systems. [80] designs an efficient protocol to support simultaneous wireless energy transfer in the downlink and wireless information transmission in the uplink for the proposed full duplex wireless-powered communication network. [81] analyzes the performance of a battery-free wireless sensor powered by ambient radio frequency (RF) energy harvesting using a stochastic geometry approach. The authors derive upper bounds of both power outage probability and transmission outage probability. However, to the best of our

knowledge, to study SIET for IoT is still an open problem.

On the other hand, stochastic geometry has been a very powerful mathematical and statistical tool to model and analyze wireless networks with random topologies. PPP has been so far the most popular point process used, mainly due to its analytical tractability. However, with assumption of independent node locations, PPP can not capture repulsion between nodes for most actual wireless networks. The Ginibre point process (GPP), on the other hand, has been proposed as a model for cellular networks recently. For example, [82] introduces β -GPP as a model for wireless networks where the nodes exhibit repulsion. [83] validates that BS locations can be fitted with a β -Ginibre point process. [84] propose two approximation models for uplink cellular networks in which BSs are deployed according to the GPP. [85] considers point-to-point downlink simultaneous wireless information and power transfer (SWIPT) transmission from an access point to a wireless sensor in a network, where ambient RF transmitters are distributed as a Ginibre α -determinantal point process (DPP). To the best of our knowledge, SIET for IoT networks has not been modeled using GPP to date.

1.5 Contributions

- Question 1: Small Cell Base Station Sleep Strategies for Energy Efficiency
 - Using a stochastic geometry-based HCN model, we derive expressions for coverage probability, average achievable rate, and EE in a K -tier HCN with four different sleep modes.
 - We design the sleep mode strategy for EE optimization under random sleeping assumption.
 - We further study a strategic sleeping scheme to manage sleep modes of BSs according to the traffic conditions with low computational complexity.

These contributions are discussed in detail in Chapter 3 and also appear in the following article:

[86]: Liu, C.; Natarajan, B.; Xia, H., “Small Cell Base Station Sleep Strategies for Energy Efficiency,” in Vehicular Technology, IEEE Transactions on , vol.PP, no.99, pp.1-1

- Question 2: Power management in heterogeneous networks with energy harvesting base stations
 - With the introduction of a lattice model, we provide a structured method to obtain coverage probability of HCNs with an irregular deployment of BSs in range of interest (ROI).
 - We conduct a techno-economic analysis of on-grid power consumption in HCNs with hybrid energy supply.
 - Including a coverage probability constraint, we provide an approach to minimize energy consumption of a HCN with hybrid energy supply via controlling transmit power of each BS.
 - We propose a novel treatment of stored energy management problem in the context of nonlinear model predictive control (NMPC) theory.

These contributions are discussed in detail in Chapter 4 and also appear in the following article:

[87]: Chang Liu, Balasubramaniam Natarajan, “Power management in heterogeneous networks with energy harvesting base stations”, Physical Communication, Volume 16, September 2015, Pages 14-24.

- Question 3: Power-Aware Performance Analysis and Optimization in D2D Underlay Cellular Networks

- In a D2D underlay cellular network, we derive upper and lower bounds for ergodic capacity of the cellular network, and recursive closed-form expression and closed-form approximation for ergodic capacity of the D2D communication when the path loss coefficient is equal to 4.
- We identify the D2D user density and transmit power that maximize the global ergodic capacity of the network. Specifically, a two-phase optimization scheme is proposed to optimize ergodic capacity while minimizing overall power consumption.
- We analyze the average achievable throughput of D2D underlay cellular networks.
- We provide closed-form ergodic capacity results (when path loss exponent is 4) for two other cases of the D2D receiver user (DRU) distributions: (1) distance between a DU pair follows a uniform distribution and (2) a DRU is distributed uniformly in the circular area around its serving D2D transmit user (DTU).

These contributions are discussed in detail in Chapter 5 and also appear in the following articles:

[88]: Chang Liu, Balasubramaniam Natarajan, “Maximizing Ergodic Capacity in D2D Underlay Networks”, *IEEE Transactions on Vehicular Technology (under review)*, 2015.

[89]: Chang Liu, Balasubramaniam Natarajan, “Ergodic Capacity in D2D Underlay Networks with Varying User Distribution”, *Consumer Communications and Networking Conference, CCNC 2016 (accepted)*, 2016.

[90]: Chang Liu, Balasubramaniam Natarajan, “Average Achievable Throughput in D2D Underlay Networks”, *Wireless Communications and Networking Conference, WCNC 2016 (under review)*, 2016.

- Question 4: Modeling and Analysis of Simultaneous Information and Energy Transfer in Internet of Things
 - We analyze the network performances using Ginibre point process to capture the geometric characteristics of IoT gateway locations.
 - We present, for the first time, valid closed-form upper bounds for both power outage probability and transmission outage probability in IoT networks assuming a practical case when the path loss coefficient is equal to 4.
 - We provide a tight upper bound for transmission outage probability in computable integral representations assuming a practical case when the path loss coefficient is equal to 4.
 - We prove the convexity of the maximum outage probability upper bounds to identify the optimal power split ratio to balance the trade-off between outage probabilities.

These contributions are discussed in detail in Chapter 6 and also appear in the following article:

[91]: Chang Liu, Balasubramaniam Natarajan, “Modeling and Analysis of Simultaneous Information and Energy Transfer in Internet of Things” (under preparation).

1.6 Dissertation Organization

Chapter 2 provides the background on stochastic geometry, point processes, channel models and network performance metrics. Chapter 3 describes sleep mode design of small cell networks to increase energy efficiency. Chapter 4 proposes a two-stage energy management scheme of cellular BSs with hybrid energy supply to minimize costs from grid power consumption. Chapter 5 presents power-aware D2D performance analysis and optimization in the context of ergodic capacity and average achievable rate. Chapter 6 studies the mod-

eling and analysis of SIET in IoT. Concluding remarks and future research directions are discussed in Chapter 7.

Chapter 2

Background

This chapter discusses the concepts of stochastic geometry, point processes, channel models and network performance metrics, which provides the background for the analysis presented in Chapter 3-6.

2.1 Stochastic geometry

2.1.1 Overview

In telecommunications, wireless network models based on stochastic geometry refer to mathematical models designed to represent aspects of wireless networks using tools from stochastic geometry. The related research efforts include analyzing these models to: (1) better understand wireless communication networks, and (2) predict and improve network performances. The models require using techniques from stochastic geometry and related fields such as point processes, spatial statistics, geometric probability, percolation theory, as well as methods from more general mathematical disciplines such as geometry, probability theory, stochastic processes, information theory, queuing theory, and Fourier analysis.

The primary idea underlying the research of the stochastic geometry models is to assume the node positions or the network structure to be random, due to the unpredictability of

users in wireless networks. Then based on stochastic geometry, closed-form or semi-closed-form expressions for quantities of interest (e.g., signal-to-interference ratio (SIR), coverage probability, channel capacity, etc.) might be derived without resorting to simulation methods or deterministic models (possibly intractable or inaccurate).

To model a heterogeneous wireless network using stochastic geometry, first we assume that it consists of K tiers [92] where each tier models a particular class, such as some type of BSs, mobile device, relays, and sensors. Components across tiers may differ in regard to transmit power, spatial density, and supported data rate. We assume that components of the i^{th} tier are spatially located according to some point process (e.g., a homogeneous PPP or GPP) ϕ_i of density λ_i in the Euclidean plane, with transmit power P_i . The target SINR for successful communication in each tier is θ_i . That is, a mobile user can communicate with a BS reliably in the i^{th} tier as long as its downlink SINR with regard to that BS is greater than θ_i . Therefore, each tier can be uniquely represented as a tuple $\{P_i, \theta_i, \lambda_i\}$.

2.1.2 Related Work and Motivation

During the past years, there has been a proliferation of efforts on cellular network modeling. For example, Wyner model is a common analytical model for performance analysis of mobile multicell networks [92]. It assumes that channel gains from all (mostly less than three) interfering BSs over the entire cell are equal. The limitation of Wyner model is that cell edge and interior users equally are treated equally. Besides, due to fixed deterministic SINR assumption, outage can not be evaluated in most cases. There are also some other approaches such as modeling average metrics [93], or only considering a small number of interfering cells and then modeling the associated and interfering BSs as an interference channel [94, 95]. Before stochastic geometry method has attracted the interest of both academia and industry, perhaps the most prevalent model for system-level simulations is the two-dimensional hexagonal grid model. However, with the increasing network heterogeneity, both the scalability and accuracy of grid model become questionable.

Stochastic geometry-based models for wireless networks have recently received widespread attention. They not only capture the irregularity and variability of the real network configuration, but also provide theoretical insights. Such a model seems more reasonable for femtocells with locations unknown and unplanned. But for the centrally planned higher tiers, as [92] show, the difference between randomly placed and actual planned locations might be smaller than expected, even for a one-tier macro BS networks. [96] also show that for a homogeneous network, the PPP network model is about as accurate as the traditional grid model compared to an actual 4G network. Importantly, such a model allows for an analytical model both tractable and accurate using stochastic geometry.

Most works that use such stochastic geometry models for wireless networks assume a PPP for the node distribution due to its tractability. However, node locations are spatially correlated in many actual wireless networks, i.e., there exists repulsion or attraction between nodes. In that case, GPP can be a model for wireless networks whose nodes exhibit repulsion.

2.2 Point Processes

2.2.1 Point Process

A point process is defined as a type of random process used to describe a collection of points, each of which is localized in time, geographical space, or even more general spaces. It is powerful in statistics for modeling and analysis of spatial data in telecommunications as well as in diverse disciplines such as astronomy, economics and geography. In the analysis of wireless networks it is a tool used to statistically describe the patterns produced by points existing in a two-dimensional space \mathbb{R}^2 .

To characterize a point process, a usual approach is to define a random process Φ as a random measure on a complete separable metric space S taking values in the non-negative integer set Z^+ (or infinity). In this framework the measure $\Phi(A)$ represents the number of points falling in the subset A of S . Attention is typically restricted to the case where Φ may

contain only finite points on any bounded subset of S .

There are several classes of point process, such as PPP, Cox point process and determinantal point processes (DPPs). In this work, we are going to use PPP, DPPs and GPP, which will be introduced in later sections.

2.2.2 Poisson Point Process

In probability theory, PPP is the simplest and most ubiquitous example of a point process, which is a simple point process N such that the number of points $N(A)$ in any bounded region A is a Poisson random variable. The simplicity of PPP lies in the independence between the points of the process. The PPP can be characterized by the following properties:

- The number of isolated points falling within two regions A and B are independent random variables if A and B do not intersect each other.
- The expected number of isolated points falling within a region A is the measure of the region A . This “measure” is often proportional to the area of A (sometimes more elaborate measures are used). The measure must be defined in a way that the measure of the union of regions that do not intersect each other is simply the sum of their measures.

A consequence of these characterizing properties is that the probability distribution of the number $N(A)$ of isolated points falling within any region A is given by

$$\mathbb{P}(N(A) = k) = \frac{(\mu(A))^k e^{-\mu(A)}}{k!} \quad (2.1)$$

where $\mu(A)$ is the measure of the region A .

2.2.3 Determinantal Point Processes

A DPP is a stochastic point process with probability distribution characterized as a determinant of some function.

Definition 1. (Determinantal point processes). Consider a Borel set $S \subset \mathbb{R}^2$ and a measurable function $K : S \times S \rightarrow \mathbb{C}$. We say that Φ is a DPP on S with kernel K if it is a simple locally finite point process on S with a joint density function given by

$$\rho_n(x_1, \dots, x_n) = \det[K(x_i, x_j)_{1 \leq i, j \leq n}] \quad (2.2)$$

for every $n \geq 1$ and $x_1, \dots, x_n \in S$.

2.2.4 Ginibre Point Process

The GPP [82] belongs to the class of DPP and it is more regular than the PPP. Besides, GPP has some critical properties allowing for expressions or bounds of performance metrics in wireless networks.

Definition 2. (Ginibre point process). The kernel of GPP is defined as

$$K(x, y) = \rho e^{\pi \rho x \bar{y}} e^{-\frac{\pi \rho}{2}(|x|^2 + |y|^2)}, \quad x, y \in \mathbb{R}^2 \quad (2.3)$$

$\rho > 0$ is the density of GPP.

It is known that the moduli of the points in the GPP has the same distribution as independent Gamma random variables. For the GPP we have the following proposition:

Proposition 1. Let $\Phi = \{X_i\}, i \in N^+$ be a GPP. Then $\{|X_i|^2\}_{i \in N^+}$ has the same distribution as the set $\{Q_k\}_{k \in N^+}$. Q_k is a random variable with probability density function (PDF) [82]:

$$f_{Q_k}(q) = \frac{q^{k-1} e^{-\pi \rho q}}{(\frac{1}{\pi \rho})^k \Gamma(k)}, \quad 0 < q < \infty. \quad (2.4)$$

i.e., $Q_k \sim \text{gamma}(k, \frac{1}{\pi\rho})$ with Q_k independent of Q_j if $k \neq j$.

2.3 Channel Model

2.3.1 Channels in a wireless network

A wireless network can be considered as a collection of channels (also referred to as links) sharing both space and frequency band. In each channel, there are a set of transmitters aiming at transmitting data to a set of receivers. The simplest case is the point-to-point channel with a single transmitter trying to send data to a single receiver (e.g., D2D communication). The broadcast channel is the one-to-many situation with a single transmitter sending different data to different receivers (e.g., the downlink of cellular networks). The multiple access channel is the many-to-one situation with several transmitters sending different data to a single receiver (e.g., the uplink of cellular networks). There are also some other channel types such as the many-to-many situation.

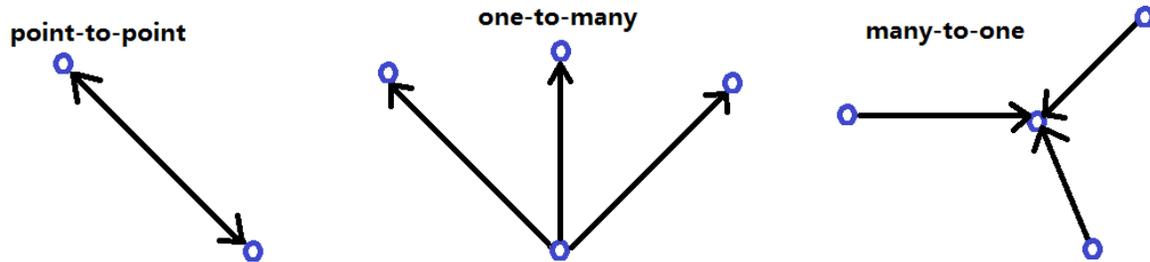


Figure 2.1: *Channel models*

2.3.2 Propagation Models

The propagation model is a building block of the stochastic geometry wireless network model. A common approach is to consider propagation models with two separate parts - the random component and the deterministic component of signal propagation.

The deterministic component is usually represented by some path-loss or attenuation function involving the propagation distance to model the power decay of electromagnetic signals. The common path-loss functions include a simple power-law function, a fast-decaying exponential function, some combination of both, and another decreasing function. Due to its tractability, models usually incorporate the power-law function

$$l(x, y) = \|x - y\|^{-\alpha}, \quad (2.5)$$

where the path-loss exponent $\alpha > 2$, and $\|x - y\|$ denotes the distance between point x and y .

On the other hand, the random component seeks to capture certain types of signal fading associated with absorption and reflections by obstacles. The fading models in use include Rayleigh and Rice distributions, which are introduced in the following sections.

Rayleigh Fading

Rayleigh fading models assume that the magnitude of a signal passing through a communication channel varies randomly according to a Rayleigh distribution. The Rayleigh factor on the received power is an exponential random variable $h \sim \exp(1)$.

Rayleigh fading is viewed as a reasonable model when there is no dominant propagation along a line of sight between the transmitter and receiver. If there is a dominant line of sight, Rician fading may be more applicable.

Rician Fading

Rician fading is a stochastic model for radio propagation when one of the paths, typically a line of sight signal, is much stronger than the others. In Rician fading, the amplitude gain is characterized by a Rician distribution.

A Rician fading channel can be described by two parameters: K_R and Ω . K_R denotes

the ratio between the power in the direct path and that in the other scattered paths. Ω represents the total power from all paths and acts as a scaling factor to the distribution. The PDF of the received signal amplitude is then given by:

$$f(x) = \frac{2(K_R + 1)x}{\Omega} \exp\left(-K_R - \frac{(K_R + 1)x^2}{\Omega}\right) I_0\left(2\sqrt{\frac{K_R(K_R + 1)}{\Omega}}x\right), \quad (2.6)$$

where $I_0(\cdot)$ is the 0th order modified Bessel function of the first kind. The Rice factor g on the received power follows the non-central χ^2 distribution with a PDF of

$$f(g) = e^{-(K_R+g)} I_0(2\sqrt{K_R g}). \quad (2.7)$$

2.4 Key Performance Metrics

In this section, several performance metrics used for analysis in later chapters are introduced.

2.4.1 Signal-to-interference-plus-noise Ratio

In wireless communication, when a collection of channels are active at the same time, the interference from the other channels is considered as noise. This motivates the need for the quantity known as the SINR.

Definition 3. (Signal-to-interference-plus-noise ratio). *If we have a collection of point-to-point channels, the SINR of the channel of a particular transmitter-receiver pair is defined as:*

$$SINR = \frac{S}{I + N} \quad (2.8)$$

where S is the power from the incoming signal to the targeted transmitter, I is the combined power of all the interfering transmitters in the network, and N is the power of thermal noise.

The SINR reduces to SNR when there is no interference (i.e. $I = 0$). In networks where

the noise is negligible, also known as "interference limited" networks, we assume $N = 0$, which gives the SIR.

2.4.2 Coverage Probability

A common goal of wireless network models based on stochastic geometry is to derive expressions for the SINR or functions of the SINR, which determine coverage (or outage) and connectivity. For example, the concept of the outage probability P_{out} (which is informally the probability of not being able to successfully send a signal on a channel) is made more precise in the point-to-point case by defining it as the probability that the SINR of a channel is less than or equal to some network-dependent threshold.

Definition 4. (Coverage probability and Outage probability). *The coverage probability P_C is the probability that the SINR is larger than the SINR threshold. In short, given a SINR threshold t , the outage and coverage probabilities are given by*

$$P_{out} = P(\text{SINR} \leq t), \quad (2.9)$$

and

$$P_C = P(\text{SINR} > t) = 1 - P_{out}. \quad (2.10)$$

2.4.3 Channel Capacity

One aim of the stochastic geometry models is to derive the rate of a typical channel when taking into account the interference created by all other channels.

Definition 5. (Channel capacity). *In the point-to-point channel case, the interference created by other transmitters is considered as noise. When the noise is treated as Gaussian, the law of the typical Shannon channel capacity is then determined by that of the SINR*

through Shannon's formula (in bits per second):

$$C = B \log_2(1 + SINR) \quad (2.11)$$

where B is the bandwidth of the channel in Hertz.

2.4.4 Ergodic Capacity

Definition 6. (Ergodic capacity). *Ergodic capacity is defined as the average of the instantaneous capacity for an additive white Gaussian noise channel, which can be expressed as*

$$\begin{aligned} C &= \mathbb{E}[B \log(1 + SINR)] \\ &= \int_{t=0}^{\infty} \mathbb{P}[B \log(1 + SINR) > t] dt \\ &= \int_{t=0}^{\infty} \mathbb{P}[SINR > e^{t/B} - 1] dt. \end{aligned} \quad (2.12)$$

The integrand in the definition of ergodic capacity is the coverage probability. In other words, there is a direct relationship between the coverage probability and the ergodic capacity.

2.4.5 Average Achievable Throughput

Definition 7. (Average achievable throughput) *We define the average achievable throughput as*

$$\bar{R}(\theta) = \mathbb{E}[B \log(1 + SINR) | SINR > \theta] \quad (2.13)$$

where θ is the SINR threshold. The average achievable throughput is defined under the condition of successful transmission.

2.5 Summary

This chapter was meant to give the readers an introduction of background knowledge regarding stochastic geometry, point processes and channel models used in analysis for different components in 5G networks. Additionally, the definitions of a few performance metrics are briefly covered. They will be used in the network performance analysis presented in later chapters.

Chapter 3

Small Cell Base Station Sleep

Strategies for Energy Efficiency

A small cell [97], with the primary design goal to provide superior cellular coverage in residential, enterprise, or hot spot outdoor environments, is a radio BS with low power and low cost. Examples include microcells, picocells, and femtocells, in order of decreasing cell size. A heterogeneous cellular network [98] is defined as a mixture of macrocells and small cells. Allowing future cellular systems to achieve higher data rates while retaining seamless connectivity as well as mobility, heterogeneous cellular networks (HCNs) can potentially improve spatial reuse and coverage. In addition to providing high data rates, cellular network operators are increasingly concerned about energy efficiency (EE). Technologies to enhance EE have become a critical design and operational consideration due to (1) increasing energy prices; (2) growing attention towards environmental factors, such as climate change and associated carbon emissions.

Primary energy consumers in cellular mobile radio networks include data servers, BSs, and backhaul routers. Approximately 80% of network energy is consumed by BSs [99]. Surging demand for high data rate services, along with increase in number of small cell BSs, is expected to further increase energy consumption of BSs. Hence, most efforts to improve

EE in mobile radio networks focus on BSs [100, 101].

In this work, we aim to quantify the trade-off between energy consumption and throughput in a HCN where small cell BSs have four distinct power saving modes. For such HCNs, based on results from stochastic geometry, we derive expressions for (1) coverage probability, (2) average achievable rate, and (3) EE measured in terms of “bits/joule”. We then attempt to find the optimal trade-off between energy and throughput by maximizing EE under QoS constraints such as coverage probability and average wake-up time. We propose an optimization strategy in which we identify the probability of small cell BSs being in different sleep modes for random sleeping policy. Random sleeping can be viewed as a benchmark against which the performance of strategic sleeping can be compared. Then based on results obtained from random sleeping, an operation scheme with low computational complexity for strategic sleeping is proposed.

This chapter is devoted to quantify the trade-off between energy consumption and throughput in a HCN where small cell BSs have four distinct power saving modes. We start our discussion by first deriving expressions for (1) coverage probability, (2) average achievable rate, and (3) EE measured in terms of “bits/joule.” We then attempt to find the optimal trade-off between energy and throughput by maximizing EE under QoS constraints such as coverage probability and average wake-up time. In later sections, we propose an optimization strategy in which we identify the probability of small cell BSs being in different sleep modes for random sleeping policy. Random sleeping can be viewed as a benchmark against which the performance of strategic sleeping can be compared. Then based on results obtained from random sleeping, an operation scheme with low computational complexity for strategic sleeping is proposed. Section 3.1 describes the system model. Section 3.2 presents random sleeping optimization. Based on that, BS management for strategic sleeping policy is proposed in Section 3.3. In Section 3.4, simulation results obtained are analyzed. Finally, Section 3.5 concludes the discussion and future work.

3.1 System Model

In this section, we first introduce the BS power consumption model. The sleep mode and sleep strategies for small cell BSs are also presented, and, according to a HCN model, we derive the EE metric.

3.1.1 Power Consumption Model

A thorough understanding of small cell hardware is of great importance in designing sleep modes that can utilize the switching-off of certain hardware components in low traffic conditions. It is imperative to scrutinize the limitations of current hardware design with regard to compatibility with sleep mode mechanisms.

The development of the power consumption model is based on the hardware model for a BS presented in [12]. A small cell BS consists of three interacting blocks. The first block is comprised of a microprocessor responsible for the implementation and management of standardized radio protocols as well as associated base band processing. Furthermore, the first block is also in charge of managing backhaul connection to the core network. The second block consists of a power amplifier and radio frequency (RF) transmitter which realizes signal transmissions. The third block is the Field-Programmable Gate Array (FPGA) with other integrated circuitry to support a range of functions including data-encryption, hardware authentication, and network time protocol.

Based on this hardware model, the power consumption P of a small cell (in W) corresponds to,

$$P = P_{\mu p} + P_{trans} + P_{PA} + P_{FPGA}, \quad (3.1)$$

where, $P_{\mu p}$, P_{trans} , P_{PA} , and P_{FPGA} are the power consumption of microprocessor, transmitter, power amplifier and FPGA, respectively.

The largest component with respect to power consumption in a small cell is related to the RF front-end (45%) and the temperature compensated crystal oscillators (TCXO) heater

(7%). Therefore, switching these components off can reduce power consumption by more than 50%. At the same time, waking up the RF front-end takes a few hundred milliseconds. The TCXO requires time to reheat, but tests [28] show that, aside from some induced clock drift, no disruption of small cell operation occurs.

3.1.2 Sleep Modes for Small Cell BSs

Based on the power consumption model, we adopt power-save modes [28] for small cells ordered by ‘depth.’ The deeper a sleep mode, the higher the power saving. However, a deep sleep mode requires a longer time for the small cell BS to wake up. Therefore, an inherent trade-off exists between sleep depth and time to wake up. We consider four modes of BS operation as listed below:

- On: The small cell is in full operation and consuming maximum power.
- Stand-by: The small cell is in “light” sleep and can wake up quickly. The RF and the TCXO heater are shut down.
- Sleep: The small cell is in “deep” sleep and needs longer time to wake up. In this mode, only power supply, backend connection, and generic CPU core are on.
- Off: The small cell is offline but still consumes a certain amount of power to be activated. However, the power consumption is negligible and approximated as zero.

The corresponding wake-up times are summarized in Table 3.1 [28]. We express the overall power consumption as a percentage of active overall power consumption.

We are assuming that a BS always transmits at its maximum power when it is in “on” mode. Note that transmit power of BSs may vary according to different traffic condition. However, in this chapter, we only consider average performance of the network. So, the instantaneous transmit power variation according to instantaneous traffic condition is not taken into consideration in the BS power consumption model we use.

Table 3.1: *Wake-up times and power consumption for different sleep modes*

Sleep Mode	Wake-up Time(s)	Power Consumption
On	0	100%
Stand-by	0.5	50%
Sleep	10	15%
Off	30	0

3.1.3 Sleep Strategies for Small Cell BSs

In this section, we present two sleeping policies for small cell BS sleep mode operation.

Random Sleeping

Each small cell BS fully operates with probability q_{on} , “stands-by” with probability $q_{standby}$, “sleeps” with probability q_{sleep} , and is turned off with probability q_{off} , independent of all other BSs. Note that $q_{on} + q_{standby} + q_{sleep} + q_{off} = 1$.

Strategic Sleeping

Instead of randomly operating small cell BSs, for this strategy, we decide the sleep mode for each small cell BS according to traffic load. In a static traffic model, the number of UE in the coverage of each BS and the distance of the nearest UE to each BS are considered. In a dynamic traffic model, direction and velocity of UE are also taken into account. The objective of strategic sleeping is to maximize usage of BSs in order to serve active users. More details are discussed in Section 3.3. This strategy has relatively larger computational complexity but provides better performance with regard to EE, as shown in Section 3.4.2.

3.1.4 Heterogeneous Cellular Network Model

A HCN consisting of K tiers [92] is considered in which each tier models BSs of a particular class, such as femtocells, picocells, microcells, or macrocells. BSs across tiers may differ in

regard to transmit power, spatial density, and supported data rate. We assume that BSs of the i^{th} tier are spatially located according to a homogeneous PPP ϕ_i of density λ_i in the Euclidean plane, with transmit power P_i . The target SINR for successful communication in each tier is θ_i . That is, a mobile UE can communicate with a BS reliably in the i^{th} tier as long as its downlink SINR with regard to that BS is greater than θ_i . Therefore, each tier can be uniquely represented as a tuple $\{P_i, \theta_i, \lambda_i\}$. In this chapter, small cell BSs with sleep modes constitute the first tier.

The UE is also distributed according to a different independent PPP ϕ_u of density λ_u . Without loss of generality, we conduct analysis on a typical UE at the origin. The downlink transmission fading power between a BS at point x and typical UE is denoted as h_x and assumed to be i.i.d exponential distributed. The standard path loss function is expressed as $l(x) = \|x\|^{-\alpha}$, where $\alpha > 2$ denotes the path loss exponent. Therefore, the received power at a typical UE from a BS located at point x_i (belonging to the i^{th} tier) is $P_i \eta_i h_{x_i} \|x_i\|^{-\alpha}$, where η_i denotes the power consumption fraction corresponding to different sleep modes, and $h_{x_i} \sim \exp(1)$. The SINR expression of a UE connecting to this BS is:

$$SINR_{x_i} = \frac{P_i \eta_i h_{x_i} \|x_i\|^{-\alpha}}{\sum_{j=1}^K \sum_{x \in \phi_j \setminus x_i} P_j \eta_j h_x \|x\|^{-\alpha} + \sigma^2}, \quad (3.2)$$

where, σ^2 denotes the noise power. Applying sleep modes on the first tier only, we have $\eta_i \in \{1, 0\}$ for the first tier, and $\eta_i = 1$ for other tiers.

The instantaneous throughput can be calculated as

$$R(x_i) = \log_2(1 + SINR_{x_i}). \quad (3.3)$$

3.1.5 Energy Efficiency Metric

The focus of our work is on energy efficient communication. Therefore, our metric of interest is bit/joule, a widely used metric for EE [102–104]. The overall EE measured in terms of

bit/joule corresponds to,

$$EE \triangleq \frac{\mathcal{R}}{\mathcal{P}}. \quad (3.4)$$

where \mathcal{R} and \mathcal{P} denotes overall throughput and overall power consumption, respectively.

A user is said to be *in coverage* if the user is able to connect to a BS with SINR above its threshold. The *coverage probability* is then defined as the probability that a user is in coverage, i.e., it is the complementary cumulative distribution function (CCDF) of SINR. Based on that, the expected overall achievable rate per unit area can be obtained as [24]

$$\mathcal{R} = \mathbb{E}_{\Phi, N_c} \left[\sum_{j=1}^{N_c} R_j \right] \quad (3.5)$$

Here $\Phi = \{\phi_1, \phi_2, \dots, \phi_K\}$, N_c and R_j denote PPP of the HCN, number of UEs in coverage per unit area and the rate of the j^{th} UE, respectively.

3.2 Random Sleeping

In this section, the expression of average achievable rate and coverage probability for random sleeping are deduced based on the system model described in Section II. We then formulate an EE maximization problem, taking both delay constraint and coverage probability threshold into consideration.

3.2.1 Average Achievable Rate for Random Sleeping

We assume that a typical UE connects to the strongest candidate BS (i.e., UE connects to the BS that provides the highest SINR greater than the threshold). The primary result for the average achievable rate random sleeping strategy is given in Theorem 1. For simplicity, we restrict our attention to the case of $\sigma^2 = 0$ in an interference limited network, i.e., self-interference dominates thermal noise. Results can be extended in a straightforward manner

to the general case with noise.

Theorem 1. *With the introduction of sleep modes, the average achievable rate by a randomly chosen mobile when it is in coverage can be obtained by*

$$\bar{R} = \frac{\lambda_1 q_{on} P_1^{\frac{2}{\alpha}} \mathcal{A}(\alpha, \theta_1, \theta_{min}) + \sum_{k=2}^K \lambda_k P_k^{\frac{2}{\alpha}} \mathcal{A}(\alpha, \theta_k, \theta_{min})}{\lambda_1 q_{on} P_1^{2/\alpha} \theta_1^{-2/\alpha} + \sum_{k=2}^K \lambda_k P_k^{2/\alpha} \theta_k^{-2/\alpha}} + \log(1 + \theta_{min}), \quad (3.6)$$

$$\mathcal{A}(\alpha, \theta_k, \theta_{min}) = \int_{\theta_{min}}^{\infty} \frac{\max(\theta_k, x)^{-2/\alpha}}{1+x} dx, \quad (3.7)$$

where $\theta_{min} = \min\{\theta_1, \dots, \theta_K\}$ and we assume that $\theta_{min} > 1$.

Proof. For HCNs without sleep modes, the average rate achievable by a randomly chosen mobile in open access when it is in coverage is [92]

$$\bar{R} = \log(1 + \theta_{min}) + \frac{\sum_{i=1}^K \lambda_i P_i^{2/\alpha} \mathcal{A}(\alpha, \theta_i, \theta_{min})}{\sum_{i=1}^K \lambda_i P_i^{2/\alpha} \theta_i^{-2/\alpha}}, \quad (3.8)$$

where, $\mathcal{A}(\alpha, \theta_k, \theta_{min}) = \int_{\theta_{min}}^{\infty} \frac{\max(\theta_k, x)^{-2/\alpha}}{1+x} dx$.

Data transmission/reception only exists between BSs in “on” mode. In other words, if a BS is not in “on” mode, then its RF module is switched off so it cannot contribute to interference at the UE. Therefore, the achievable rate expression with sleep modes is equivalent to that without sleep modes when the density of operated small cell BSs is $\lambda_1 q_{on}$, leading to (3.6). \square

3.2.2 Coverage Probability Consideration

We can define the success probability from a UE at x to a BS at y in the i^{th} tier as $\mathbb{P}(SINR_i(x \rightarrow y) \geq \theta_i) \geq \omega$, where ω is a prescribed QoS threshold. By averaging the success probability over the distance to the connected BS, we obtain the coverage probability of a typical UE. Corollary 1 gives the primary result for coverage probability.

Corollary 1. *In interference limited network, the coverage probability constraint of a typical UE can be expressed as*

$$P_c(\{\lambda_i\}, \{\theta_i\}, \{P_i\}, q_{on}) = \frac{(\lambda_1 q_{on} P_1^{2/\alpha} \theta_1^{-2/\alpha} + \sum_{j=2}^K \lambda_j P_j^{2/\alpha} \theta_j^{-2/\alpha}) \text{sinc}(\frac{2}{\alpha})}{\lambda_1 q_{on} P_1^{2/\alpha} + \sum_{j=2}^K \lambda_j P_j^{2/\alpha}}, \quad (3.9)$$

Here $i = 1, \dots, K$, and $\theta_i > 1, i = 1, \dots, K$.

Proof. In interference limited network, the coverage probability of a typical mobile user without sleep mode can be expressed as [92]

$$P_c(\{\lambda_i\}, \{\theta_i\}, \{P_i\}) = \frac{\sum_{i=1}^K \lambda_i P_i^{2/\alpha} \theta_i^{-2/\alpha}}{\sum_{i=1}^K \lambda_i P_i^{2/\alpha}} \cdot \text{sinc}(\frac{2}{\alpha}), \theta_i > 1. \quad (3.10)$$

Since data transmission/reception only exists for fully operational small cells, when adopting sleep modes, the coverage probability expression with sleep modes is equivalent to that without sleep modes when density of small cell BSs is $\lambda_1 q_{on}$, which leads to (3.9). \square

3.2.3 Delay Constraint Consideration

We consider the average delay for small cell BSs to wake up. That is, while the controller can turn off multiple small cell BSs to save energy, it is important that these BSs wake up in a timely manner to serve new UE demanding service. If we constrain the average waking time for UE, we arrive at the following constraint:

$$t_{on} q_{on} + t_{standby} q_{standby} + t_{sleep} q_{sleep} + t_{off} q_{off} \leq t_{min}. \quad (3.11)$$

3.2.4 Energy Efficiency for Random Sleeping Strategy

As discussed earlier, the total power consumption in a BS includes transmit power and circuit power consumed by modules such as microprocessor, PA and FPGA. In our analysis,

we assume that the transmit power is a constant fraction η of total power consumption for all the tiers. That is, overall power consumption of a BS in tier i is P_i/η .

Lemma 1. *Energy Efficiency for random sleeping strategy can be obtained as*

$$\begin{aligned}
& EE(q_{on}, q_{standby}, q_{sleep}, q_{off}) \\
&= \frac{\lambda_u P_c(\{\lambda_i\}, \{\theta_i\}, \{P_i\}) \cdot \bar{R}}{\lambda_1(q_{on} + 0.5q_{standby} + 0.15q_{sleep})P_1/\eta + \sum_{i=2}^K \frac{\lambda_i P_i}{\eta}} \\
&= \frac{1}{\lambda_1(q_{on} + 0.5q_{standby} + 0.15q_{sleep})P_1 + \sum_{j=2}^K \lambda_j P_j} \frac{\lambda_1 q_{on} P_1^{2/\alpha} \theta_1^{-2/\alpha} + \sum_{j=2}^K \lambda_j P_j^{2/\alpha} \theta_j^{-2/\alpha}}{\lambda_1 q_{on} P_1^{2/\alpha} + \sum_{j=2}^K \lambda_j P_j^{2/\alpha}} \text{sinc}\left(\frac{2}{\alpha}\right) \lambda_u \eta \cdot \\
& \left[\frac{\lambda_1 q_{on} (P_1)^{\frac{2}{\alpha}} \mathcal{A}(\alpha, \theta_1, \theta_{min}) + \sum_{j=2}^K \lambda_j P_j^{\frac{2}{\alpha}} \mathcal{A}(\alpha, \theta_j, \theta_{min})}{\lambda_1 q_{on} (P_1)^{2/\alpha} \theta_1^{-2/\alpha} + \sum_{j=2}^K \lambda_j P_j^{2/\alpha} \theta_j^{-2/\alpha}} + \log(1 + \theta_{min}) \right].
\end{aligned} \tag{3.12}$$

Proof. The expected overall power consumption per unit area is computed as

$$\mathcal{P} = \lambda_1(q_{on} + 0.5q_{standby} + 0.15q_{sleep})P_1/\eta + \sum_{i=2}^K \frac{\lambda_i P_i}{\eta}. \tag{3.13}$$

And the expected overall achievable rate per unit area can be obtained as

$$\begin{aligned}
\mathcal{R} &= \mathbb{E}_{\Phi, N_c} \left[\sum_{j=1}^{N_c} R_j \right] \\
&= \mathbb{E}_{N_c} \left[\mathbb{E}_{\Phi} \left(\sum_{j=1}^{N_c} R_j \mid N_c \right) \right] \\
&= \mathbb{E}_{N_c} \left[\sum_{j=1}^{N_c} \bar{R} \right] \\
&= \mathbb{E}_{N_c} [N_c \bar{R}] \\
&= \mathbb{E}[N_c] \bar{R} \\
&= \lambda_u A P_c(\{\lambda_i\}, \{\theta_i\}, \{P_i\}) / A \cdot \bar{R} \\
&= \lambda_u P_c(\{\lambda_i\}, \{\theta_i\}, \{P_i\}) \cdot \bar{R}.
\end{aligned} \tag{3.14}$$

where A denotes the total considered area.

Thus, the EE can be calculated as

$$\begin{aligned}
& EE(q_{on}, q_{standby}, q_{sleep}, q_{off}) \\
&= \frac{\lambda_u P_c(\{\lambda_i\}, \{\theta_i\}, \{P_i\}) \cdot \bar{R}}{\lambda_1(q_{on} + 0.5q_{standby} + 0.15q_{sleep})P_1/\eta + \sum_{i=2}^K \frac{\lambda_i P_i}{\eta}} \\
&= \frac{1}{\lambda_1(q_{on} + 0.5q_{standby} + 0.15q_{sleep})P_1 + \sum_{j=2}^K \lambda_j P_j} \frac{\lambda_1 q_{on} P_1^{2/\alpha} \theta_1^{-2/\alpha} + \sum_{j=2}^K \lambda_j P_j^{2/\alpha} \theta_j^{-2/\alpha}}{\lambda_1 q_{on} P_1^{2/\alpha} + \sum_{j=2}^K \lambda_j P_j^{2/\alpha}} \text{sinc}\left(\frac{2}{\alpha}\right) \lambda_u \eta \cdot \\
& \left[\frac{\lambda_1 q_{on} (P_1)^{\frac{2}{\alpha}} \mathcal{A}(\alpha, \theta_1, \theta_{min}) + \sum_{j=2}^K \lambda_j P_j^{\frac{2}{\alpha}} \mathcal{A}(\alpha, \theta_j, \theta_{min})}{\lambda_1 q_{on} (P_1)^{2/\alpha} \theta_1^{-2/\alpha} + \sum_{j=2}^K \lambda_j P_j^{2/\alpha} \theta_j^{-2/\alpha}} + \log(1 + \theta_{min}) \right].
\end{aligned} \tag{3.15}$$

□

3.2.5 Problem Formulation

Considering both the delay constraint above and the coverage probability constraint, the goal of maximizing EE of the HCNs for random sleeping can be formulated as an optimization problem. That is,

Maximize

$$\begin{aligned}
& EE(q_{on}, q_{standby}, q_{sleep}, q_{off}) \\
&= \frac{\lambda_u P_c(\{\lambda_i\}, \{\theta_i\}, \{P_i\}) \cdot \bar{R}}{\lambda_1(q_{on} + 0.5q_{standby} + 0.15q_{sleep})P_1/\eta + \sum_{i=2}^K \frac{\lambda_i P_i}{\eta}} \\
&= \frac{1}{\lambda_1(q_{on} + 0.5q_{standby} + 0.15q_{sleep})P_1 + \sum_{j=2}^K \lambda_k P_k} \frac{\lambda_1 q_{on} P_1^{2/\alpha} \theta_1^{-2/\alpha} + \sum_{j=2}^K \lambda_j P_j^{2/\alpha} \theta_j^{-2/\alpha}}{\lambda_1 q_{on} P_1^{2/\alpha} + \sum_{j=2}^K \lambda_j P_j^{2/\alpha}} \text{sinc}\left(\frac{2}{\alpha}\right) \lambda_u \eta \cdot \\
& \left[\frac{\lambda_1 q_{on} (P_1)^{\frac{2}{\alpha}} \mathcal{A}(\alpha, \theta_1, \theta_{min}) + \sum_{j=2}^K \lambda_k P_k^{\frac{2}{\alpha}} \mathcal{A}(\alpha, \theta_k, \theta_{min})}{\lambda_1 q_{on} (P_1)^{2/\alpha} \theta_1^{-2/\alpha} + \sum_{j=2}^K \lambda_k P_k^{2/\alpha} \theta_i^{-2/\alpha}} + \log(1 + \theta_{min}) \right].
\end{aligned} \tag{3.16}$$

subject to,

$$0.5q_{standby} + 10q_{sleep} + 30q_{off} \leq t_{min}, \tag{3.17}$$

$$P_c(q_{on}) = \frac{\lambda_1 q_{on} P_1^{\frac{2}{\alpha}} \theta_1^{-\frac{2}{\alpha}} + \sum_{j=2}^K \lambda_j P_j^{\frac{2}{\alpha}} \theta_j^{-\frac{2}{\alpha}}}{\lambda_1 q_{on} P_1^{2/\alpha} + \sum_{j=2}^K \lambda_j P_j^{2/\alpha}} \cdot \text{sinc}\left(\frac{2}{\alpha}\right) \geq \omega, \tag{3.18}$$

$$q_{on} + q_{standby} + q_{sleep} + q_{off} = 1, \tag{3.19}$$

$$0 \leq q_{on}, q_{standby}, q_{sleep}, q_{off} \leq 1. \quad (3.20)$$

Here, the variables of interest are q_{on} , $q_{standby}$, q_{sleep} , and q_{off} . The optimal solution q_{on}^* , $q_{standby}^*$, q_{sleep}^* , and q_{off}^* are used in BS operation in order to maximize EE while satisfying QoS requirements.

3.2.6 Analysis of Optimization Design

We aim at simultaneously optimizing EE over q_{on} , $q_{standby}$, q_{sleep} , and q_{off} . It can be eliminated to three variables (e.g. q_{on} , $q_{standby}$ and q_{sleep}) because of the equality constraint (3.19). However, a global optimal solution is difficult to obtain due to the non-convex structure of (3.16). Therefore, in this section, we first analyze the behavior of the optimization problem as a function of each variable while considering the other variables to be fixed. Then an approach to obtain a suboptimal solution is presented.

For variables $q_{standby}$ or q_{sleep} , when the other two variables are fixed, it is straightforward to see that the objective function (3.16) is a monotone function and easy to optimize. However, when $q_{standby}$, q_{sleep} are fixed, $EE(q_{on})$ is the ratio of a quadratic function over a cubic function of q_{on} , which is non-convex and not tractable. Therefore, we attempt to obtain a suboptimal solution of q_{on}^* .

From constraint (3.18), we can obtain the lower bound of EE, that is,

$$EE(q_{on}) \geq \frac{\lambda_1 q_{on} (P_1)^{\frac{2}{\alpha}} \mathcal{A}(\alpha, \theta_1, \theta_{min}) + \sum_{j=2}^K \lambda_k P_k^{\frac{2}{\alpha}} \mathcal{A}(\alpha, \theta_k, \theta_{min})}{\lambda_1 q_{on} (P_1)^{2/\alpha} \theta_1^{-2/\alpha} + \sum_{j=2}^K \lambda_k P_k^{2/\alpha} \theta_i^{-2/\alpha}} + \log(1 + \theta_{min}) \lambda_u \eta \omega$$

$$\frac{1}{\lambda_1 (q_{on} + 0.5q_{standby} + 0.15q_{sleep}) P_1 + \sum_{j=2}^K \lambda_k P_k}.$$

(3.21)

If the lower bound of EE can be maximized, the overall EE performance of the HCN can be improved as well. So instead of maximizing (3.16), we attempt to maximize the lower bound of EE, that is

$$EE_{lb}(q_{on}) = \left[\frac{\lambda_1 q_{on} (P_1)^{\frac{2}{\alpha}} \mathcal{A}(\alpha, \theta_1, \theta_{min}) + \sum_{j=2}^K \lambda_j P_j^{\frac{2}{\alpha}} \mathcal{A}(\alpha, \theta_j, \theta_{min})}{\lambda_1 q_{on} (P_1)^{2/\alpha} \theta_1^{-2/\alpha} + \sum_{j=2}^K \lambda_j P_j^{2/\alpha} \theta_j^{-2/\alpha}} + \log(1 + \theta_{min}) \right] \lambda_u \eta \omega$$

$$\frac{1}{\lambda_1 (q_{on} + 0.5q_{standby} + 0.15q_{sleep}) P_1 + \sum_{j=2}^K \lambda_j P_j}.$$
(3.22)

Removing the constants, maximizing (3.22) is equivalent to maximizing

$$EE_{lbe}(q_{on}) = \left[\frac{\lambda_1 q_{on} (P_1)^{\frac{2}{\alpha}} \mathcal{A}(\alpha, \theta_1, \theta_{min}) + \sum_{j=2}^K \lambda_j P_j^{\frac{2}{\alpha}} \mathcal{A}(\alpha, \theta_j, \theta_{min})}{\lambda_1 q_{on} (P_1)^{2/\alpha} \theta_1^{-2/\alpha} + \sum_{j=2}^K \lambda_j P_j^{2/\alpha} \theta_j^{-2/\alpha}} + \log(1 + \theta_{min}) \right].$$

$$\frac{1}{\lambda_1 (q_{on} + 0.5q_{standby} + 0.15q_{sleep}) P_1 + \sum_{j=2}^K \lambda_j P_j}.$$
(3.23)

Since $EE_{lbe}(q_{on}) > 0$, maximizing this objective function (3.23) is equivalent to Minimize

$$f(q_{on}) = 1/EE_{lbe}(q_{on}).$$
(3.24)

Proposition 2. $f(q_{on})$ is a quasiconvex function of q_{on} .

Proof. As a function of q_{on} , expression (3.24) has the form $f(q_{on}) = \frac{aq_{on}^2 + bq_{on} + c}{dq_{on} + e}$, where a , b , c , d , and e denote different positive constants as given below:

$$a = \lambda_1^2 P_1^{2/\alpha+1} \theta_1^{-2/\alpha},$$
(3.25)

$$b = \lambda_1 \left(\sum_{i=2}^K \lambda_i P_i \right) P_1^{2/\alpha} \theta_1^{-2/\alpha} + \lambda_1 P_1 \left(\sum_{i=2}^K \lambda_i P_i^{2/\alpha} \theta_i^{-2/\alpha} \right), \quad (3.26)$$

$$c = \left(\sum_{i=2}^K \lambda_i P_i \right) \left(\sum_{i=2}^K \lambda_i P_i^{2/\alpha} \theta_i^{-2/\alpha} \right), \quad (3.27)$$

$$d = \lambda_1 P_1^{2/\alpha} (\log(1 + \theta_{min}) \theta_1^{-2/\alpha} + \mathcal{A}(\alpha, \theta_1, \theta_{min})), \quad (3.28)$$

and

$$e = \log(1 + \theta_{min}) \left(\sum_{i=2}^K \lambda_i P_i^{\frac{2}{\alpha}} \theta_i^{-\frac{2}{\alpha}} \right) + \left(\sum_{i=2}^K \lambda_i P_i^{\frac{2}{\alpha}} \mathcal{A}(\alpha, \theta_i, \theta_{min}) \right). \quad (3.29)$$

Define $p(q_{on}) = aq_{on}^2 + bq_{on} + c$ and $q(q_{on}) = dq_{on} + e$. It is easy to prove that $p(q_{on})$ is a convex function, $q(q_{on})$ is a concave function, with $p(q_{on}) \geq 0$ and $q(q_{on})$ on a convex set $C = [0, 1]$. Then according to [105], the function f defined by $f(q_{on}) = p(q_{on})/q(q_{on})$ on C , is quasiconvex and can always be solved by quasiconvex optimization. \square

Based on the analysis above, (3.16) can be solved by an alternating iterative algorithm [106, 107] where variables are computed one at a time while fixing all others. Algorithm 1 shows the proposed alternating algorithm for the multivariable EE optimization problem under QoS constraints. The desired optimization $EE(q_{on}(k), q_{standby}(k), q_{sleep}(k))$ in each iterative step is gradually maximized as an iteration index k is increased in Algorithm 1. The difference between $EE(q_{on}(k-1), q_{standby}(k-1), q_{sleep}(k-1))$ and $EE(q_{on}(k), q_{standby}(k), q_{sleep}(k))$ can be used as a criterion to stop the alternating algorithm with a designed positive scalar ϵ in Step 3). This is because the desired optimization $EE(q_{on}(k), q_{standby}(k), q_{sleep}(k))$ has positive and convergent properties at all times [107].

Having obtained q_{on}^* , $q_{standby}^*$, q_{sleep}^* via the alternating algorithm stated above, each small cell BS switches to a mode with corresponding probability, independent of all other BSs.

Algorithm 1 Alternating Algorithm for Multivariable EE Optimization under QoS Constraints

1. Initialization: $k = 0$
 $q_{standby}(k) = 0, q_{sleep}(k) = 0, EE(q_{on}(0), q_{standby}(0), q_{sleep}(0)) = 0$
 2. Iteration: $k + 1 \leftarrow k$
 $q_{on}(k) = \arg \max_{q_{on}} EE(q_{standby}(k-1), q_{sleep}(k-1)), \text{ s.t. (3.18)-(3.20)}$
 $(q_{standby}(k), q_{sleep}(k)) = \arg \max_{q_{standby}, q_{sleep}} EE(q_{on}(k)), \text{ s.t. (3.17), (3.19) and (3.20)}$
 3. If $0 < EE(q_{on}(k), q_{standby}(k), q_{sleep}(k)) - EE(q_{on}(k-1), q_{standby}(k-1), q_{sleep}(k-1)) < \epsilon$, go to 4) and stop, otherwise go back to 2) ($\epsilon = 0.0001$).
 4. $q_{on}^* = q_{on}(k), q_{standby}^* = q_{standby}(k)$ and $q_{sleep}^* = q_{sleep}(k)$.
-

3.3 Strategic Sleeping

Instead of randomly operating small cell BSs, a strategic scheme can be designed to maximize the usage of the small cell BSs to serve active users according to UE information.

In a given setup, ideal strategic sleeping policy could be obtained by actually calculating rate and power empirically based on simulations. However, as the scale of the network increases, the computational complexity of such an approach can become extremely large making it undesirable. Therefore, in this work, we propose a suboptimal strategic sleeping policies for both static and dynamic traffic models, which make use of results calculated from the benchmark random sleeping strategy.

The strategic sleeping approach for a static traffic model is presented in Algorithm 2. Here, the static traffic model refers to the case when all UE are static (with known positions). The sleep scheduling for this case is based on: (1) the number of UE in the coverage of each small cell, and (2) the distance between each BS and its corresponding nearest UE. First, operating probabilities in random sleeping are converted into fractions, i.e., we decide the number of BSs in each state based on corresponding probability and total number of small cell BSs. The small cell BSs with most users in coverage are turned on. Then the sleep modes of other BSs are determined according to the distances between a BS to UE, that

is, a BS sleeps lighter with nearer UE. More detail, the strategic scheme is scheduled as discussed in Algorithm 2.

Algorithm 2 Strategic Sleeping for Static Traffic Model

1. Decide the numbers N_{on} , $N_{standby}$, N_{sleep} , and N_{off} of small cell BSs in on mode, standby mode, sleep mode and off mode by rounding Nq_{on} , $Nq_{standby}$, Nq_{sleep} , and Nq_{off} to an integer, respectively. Here N denotes the total number of operated small cell BSs.
 2. Calculate the number of users in the coverage area of each small cell BS based on radius R calculated according to the SINR target. To simplify the problem, when calculating radii, we assume that all the BSs are switched on.
 3. Sort associated UE numbers in descending order for the entire operated small cell BSs.
 4. Choose the first N_{on} small cell BSs to turn on.
 5. Calculate distances from the nearest UE to each of the undecided small cell BSs, respectively.
 6. Sort distances in ascendant order and make the first $N_{standby}$ small cell BSs standby; then the first N_{sleep} small cell BSs sleep. The rest small cell BSs are turned off.
-

When UE movement is considered, the strategic sleeping approach can be modified as described in Algorithm 3. In this case, the sleep scheduling is based on: (1) the number of UE in coverage of each small cell, and (2) the minimum time of a UE moving to a BS. Identical to static strategy, the operating probabilities in random sleeping are converted into fractions first. The small cell BSs with most users in coverage are chosen to be turned on. Then the sleep modes of other BSs are determined according to the time UE moving to a BS takes, that is, a BS sleep lighter with faster UE moving in. More precisely, the dynamic scheduling for strategic sleeping is described in Algorithm 3.

Algorithm 3 Strategic Sleeping for Dynamic Traffic Model

1. Decide the numbers N_{on} , $N_{standby}$, N_{sleep} , and N_{off} of small cell BSs in on mode, standby mode, sleep mode, and off mode, respectively, by rounding Nq_{on} , $Nq_{standby}$, Nq_{sleep} , and Nq_{off} to an integer. Again, N denotes the total number of operated small cell BSs.
 2. Calculate the number of users in the coverage area of each small cell BS based on radius R calculated according to the SINR target. To simplify the calculation of radii, we assume that all BSs are switched on.
 3. Sort associated UE numbers in descending order for the entire operated small cell BSs.
 4. Choose the first N_{on} small cell BSs to turn on.
 5. For each undecided small cell BS, identify which UE is moving towards it. Then calculate distances between undecided BSs and their corresponding UE.
 6. Calculate the time t_{ij} required for the i^{th} UE moving to the j^{th} small cell BS by $t_{ij} = d_{ij}/v_{ij}$. Here v_{ij} denotes the “relative” velocity, that is, projection of absolute velocity on d_{ij} . The schematic diagram is shown in Fig. 3.1.
 7. For the j^{th} small cell BS, $t_j = \min(t_{ij}), i \in \phi_j$, where ϕ_j denotes the set of all UE moving toward the j^{th} small cell BS or already in coverage.
 8. Sort t_j in ascendant order and make the first $N_{standby}$ small cell BSs standby; then the first N_{sleep} small cell BSs sleep. The rest small cell BSs are turned off.
-

3.4 Simulation Results

In this section, we provide simulation results to demonstrate the proposed approach for deciding sleep modes for small cell BSs. The primary simulation parameters used are summarized in Table 4.3 [108]. Note that the noise power (σ^2) is assigned as zero since self-interference dominates noise in typical HCNs [92].

3.4.1 Random Sleeping Strategy

Based on the coverage probability curve shown in Fig. 3.2, eight different cases are discussed in Table III. In the first case $\omega = 0.5$, the time constraint (3.18) is always satisfied when

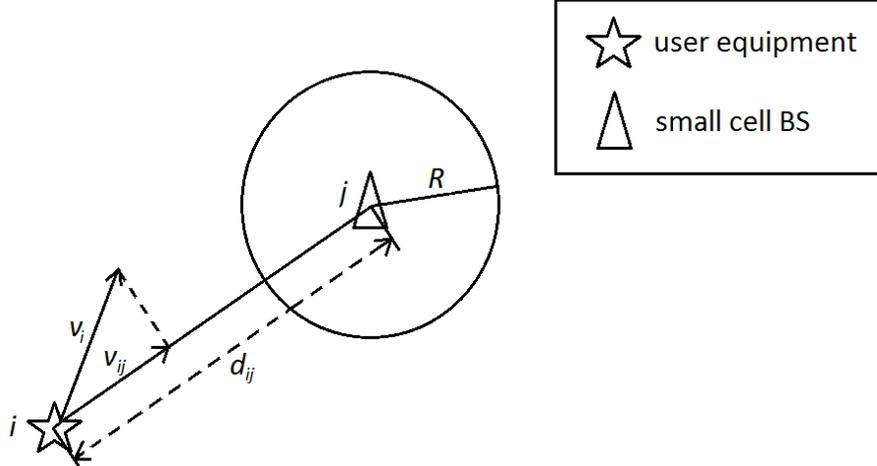


Figure 3.1: Schematic diagram of d_{ij} , v_i and v_{ij}

Table 3.2: Simulation parameters

Pico/Micro/Macro/UE distribution	PPP
Density of picos (λ_1)	$4/500^2 m^2$
Density of micros (λ_2)	$2/500^2 m^2$
Density of macros (λ_3)	$1/500^2 m^2$
Power consumption of picos (P_1)	40W
Power consumption of micros (P_2)	80W
Power consumption of macros (P_3)	400W
SINR threshold of picos (θ_1)	1.1
SINR threshold of micros (θ_2)	1.2
SINR threshold of macros (θ_3)	1.3
Path loss exponent (α)	4

$\omega = 0.58$, the time constraint (3.18) cannot be satisfied unless $q_{on} \geq 0.6038$. We ignore the unsolvable case that the time constraint cannot be satisfied, i.e., $\omega \geq P_c(1) = 0.5843$. On the other hand, for $t_{min} = 0.3s$, the delay constraint can only be satisfied when $q_{on} \geq 0.4$, for $t_{min} = 0.5s$, $t_{min} = 15s$ and $t_{min} = 30s$ it can be satisfied by some $q_{on} \in [0, 1]$, thus parameters $q_{standby}$, q_{sleep} , and q_{off} can be made to meet the delay constraint.

Table 3.3 clearly shows that this method first decide q_{on}^* to be as small as possible while

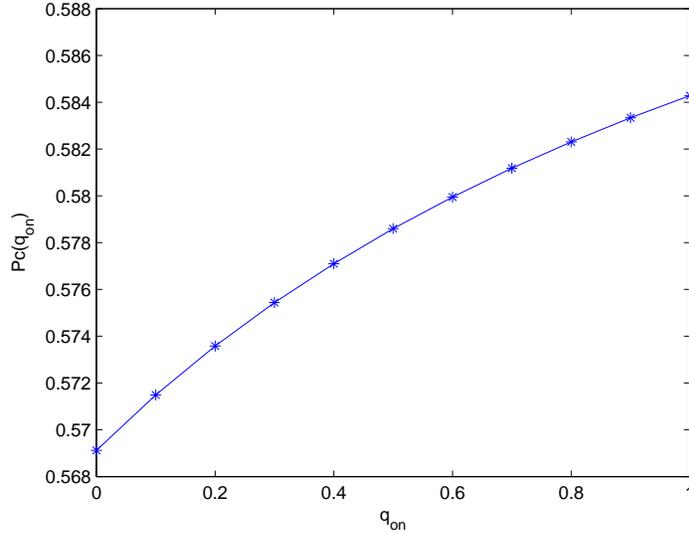


Figure 3.2: q_{on} vs coverage probability

ω	t_{min}	q_{on}	$q_{standby}$	q_{sleep}	q_{off}
0.5	0.3	0.4000	0.6000	0	0
0.5	0.5	0	1	0	0
0.5	10	0	0	1	0
0.5	30	0	0	0	1
0.58	0.3	0.6038	0.3854	0.0107	0
0.58	0.5	0.6038	0.3644	0.0318	0
0.58	10	0.6038	0	0.0942	0.3019
0.58	30	0.6038	0	0	0.3962

Table 3.3: Simulation results of optimal sleep mode probabilities

satisfying coverage probability (3.18), then decide $q_{standby}^*$, q_{sleep}^* , $q_{offline}^*$ to make BSs to sleep “deeper” as much as possible while preserving the required average waiting time constraint (3.17).

EE improvement is defined as the ratio of EE with sleep modes to EE without sleep modes, and is employed as a performance measure. Fig. 3.3 shows EE improvement in four cases with $\omega = 0.5$, $\omega = 0.57$, $\omega = 0.575$, and $\omega = 0.58$, respectively. When $\omega = 0.5$ coverage probability can be satisfied by any q_{on} , and for the other three cases coverage probability

condition can only be fulfilled for a certain range of q_{on} .

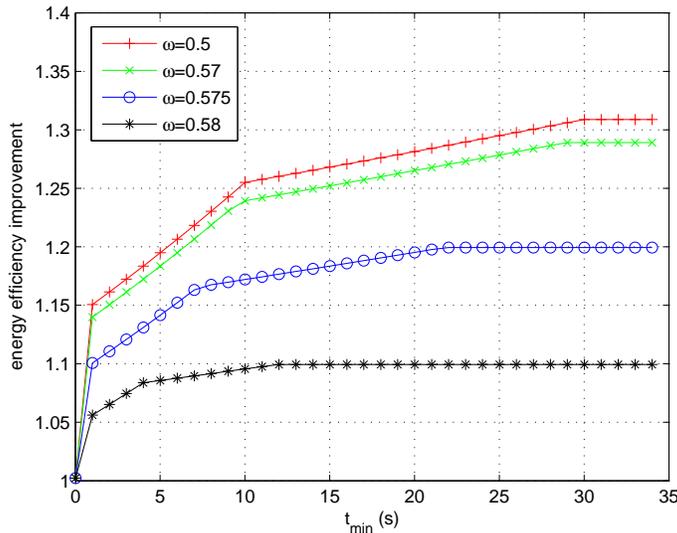


Figure 3.3: Energy efficiency improvement vs t_{min}

Fig. 3.3 shows that EE can be improved approximately 30% in the case $\omega = 0.5$, meaning the coverage probability can be met even if all the operated small cell BSs are switched off. Improvement increases with the allowed average wake-up time. When the wake-up time is long enough, except for the BSs turned on to ensure certain coverage probability, all other operated small cell BSs are switched off and the EE gain is saturated. Furthermore, Fig. 3.3 shows that EE improvement decreases with q_{on}^* , further confirming the rationality of this optimizing method.

3.4.2 Strategic Sleeping Strategy

In order to evaluate the performance of different sleeping schemes, Monte Carlo simulation analysis is employed. The performance of strategic sleeping strategy is studied by a randomly generated HCN. As a performance metric, we calculate the ratio of EE using strategic sleeping strategy to EE using random sleeping strategy. The average EE can be greatly improved with random sleeping policy, and instantaneous EE can be further improved by

strategic sleeping policy.

In this case, since the fraction of BSs of each sleep mode is fixed, the whole network power consumption stays the same. Then maximizing (3.16) is equivalent to a throughput maximization problem, and the EE improvement ratio is the throughput improvement ratio indeed. The throughput between an UE and BS can be obtained by (3.3). Furthermore, it can be concluded that EE improvement will only be affected by q_{on} , or there is no difference when studying EE improvement performance using static model or dynamic model. After calculating rate and power empirically to get EE values, simulation results of the improvements of both the operated tier and whole networks are shown in Fig. 3.4.

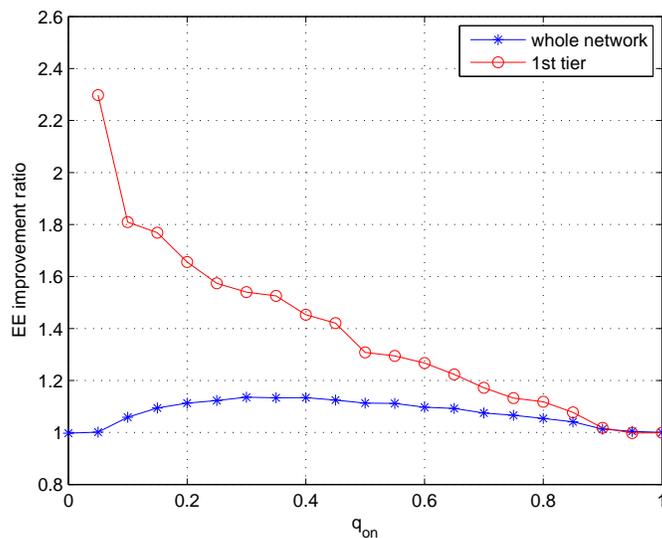


Figure 3.4: Ratio of EE with strategic sleeping policy to EE with random sleeping policy

Fig. 3.4 shows that the proposed strategic sleeping management can effectively improve EE of the network. Considering only the operated tier, EE improvement ratio monotonically decreases with q_{on} , implying that a smaller q_{on} may lead to a smaller probability of finding the “right” small cell BS that should be in “on” mode. The superiority of the strategic sleeping strategy is verified.

In regards to an entire network EE can be improved by approximately 15% by employing

strategic sleeping policy rather than random sleeping policy. This performance would be further improved as density of the operated smaller cell BSs increases, which is a predominant trend in wireless communications. On the other hand, the EE improvement ratio achieves maximum value for some $q_{on} \in [0, 1]$. If q_{on} is too small, EE of the entire network would be dominated by other tiers which remain the same because they are always full operated.

3.5 Summary

We investigate the design of energy efficient HCNs employing small cell BS sleep modes and sleep strategies. Using a stochastic geometry based model, we derive coverage probability, average achievable throughput, and EE in K-tier HCNs. Then we formulate EE maximization problem under either random sleeping policy or strategic sleeping policy and determine optimal operating probability for each sleep mode of small cell BSs. Numerical results confirm the effectiveness of the scheme. Apart from improvements of approximately 30% in EE with random sleeping policy, the simulation indicates that instantaneous EE can be further improved by 15% with a strategic sleeping policy.

Besides sleep mode, renewable energy is also considered as a promising green technique for cellular BSs. In the next chapter, we explore smart energy management of energy harvesting cellular BSs.

Chapter 4

Power Management in Heterogeneous Networks with Energy Harvesting Base Stations

Energy harvesting [109] refers to the aggregation of renewable energy (e.g., solar and wind) from the operational environment. Therefore, utilizing harvested energy to supplement conventional on-grid power in powering BSs can serve as a candidate solution for power savings. However, how to best operate BSs with energy harvesting is not trivial due to limited availability of harvested energy, as well as uncertainty about timing and quantity of energy collected. This chapter provides some key insight on managing BSs powered by both renewable energy and grid power.

In this work, we aim to minimize on-grid power cost of HCNs with energy harvesting BSs by managing both transmit power and stored energy. Taking QoS into account, a lattice model is considered to obtain coverage probability in region of interest (ROI). We formulate an optimization problem to find optimal transmit power and stored energy under a coverage probability constraint. Then, we propose a two-stage BS operation scheme which first optimizes transmit power and then manages stored energy usage.

Unique aspects of our work can be summarized as follows:

- With the introduction of a lattice model, we provide a structured method to obtain coverage probability of HCNs with an irregular deployment of BSs in ROI.
- We conduct a techno-economic analysis of on-grid power consumption in HCNs with hybrid energy supply.
- Including a coverage probability constraint, we provide an approach to minimize energy consumption of a HCN with hybrid energy supply via controlling transmit power of each BS.
- A novel treatment of stored energy management problem is proposed in the context of NMPC theory.

This chapter includes several distinct sections. Section 4.1 presents the system model. Section 4.2 first describes the problem formulation, then decomposes the problem into 2 subproblems. Transmit power optimization is discussed first in Section 4.3. Proposed control solution for stored energy management based on NMPC is discussed in Section 4.4. In Section 4.5, simulation results are obtained and analyzed. Finally, Section 4.6 concludes the discussion and future work.

4.1 System Model

We consider a HCN with BSs powered jointly by an energy harvesting device and the power grid. The operation time line (e.g., a period of 24 hours) is divided into N_T time slots. The harvested renewable energy model, real-time electricity price model, BS management and lattice model are detailed in this section.

4.1.1 Harvested Renewable Energy Model

Stochastic modeling of renewable energy generation has been an active research topic. Renewable energy sources include wind, solar, geothermal, tides, hydropower, and various forms of biomass. We will focus on energy models of wind and solar since they are most commonly used by telecommunication service providers to supplement conventional power in powering BSs, mainly due to the fact that conversion of wind or solar energy into electricity is feasible and easy as compared to other sources. In this chapter, we adopt the model proposed for wind energy in [110]-[111]. It can also be directly used for solar energy as solar energy generation can be estimated by using typical annual meteorological weather data for a given geolocation [112]. This model may also be further extended to other renewable energy sources.

We consider a HCN consisting of K tiers of different classes of BSs such as femtocells, picocells, or macrocells. BSs of all the tiers can be powered by both on-grid energy and renewable energy sources. BSs across tiers may differ in regard to transmit power, spatial density, and supported data rate. We assume that BSs of the k^{th} tier are denoted as ϕ_k with number N_k and transmit power P_k . The target SINR for successful communication in each tier is θ_k . That is, a mobile UE can communicate with a BS reliably in the k^{th} tier as long as its downlink SINR with regard to that BS is greater than θ_k . We assume that the harvesting power of the i^{th} BS of tier k during time slot t , denoted as $P_{Hk}(i, t)$, is modeled as the sum of an accurately predicted mean value $\hat{P}_{Hk}(i, t)$ (based on estimation/prediction methods), and i.i.d. zero-mean random variables $err_H(t)$, i.e.,

$$P_{Hk}(i, t) = \hat{P}_{Hk}(i, t) + err_H(i, t). \quad (4.1)$$

Our formulation allows err_H to obey any distribution. In this work, we assume that it follows a zero mean uniform distribution $err_H \sim Uniform[-a, a]$.

4.1.2 Real-time Electricity Price Model

The real-time price of electricity can be forecasted using price forecast techniques such as the ones proposed in [113]. Similar to the renewable energy model, we assume that the real-time price of energy of time slot t , denoted as $M(t)$, is the sum of an accurately predicted mean value $\hat{M}(t)$ and i.i.d. associated errors $err_M(t)$ with known distribution, i.e.,

$$M(t) = \hat{M}(t) + err_M(t). \quad (4.2)$$

In fact, it has been verified in [114] that the forecasted error can be modeled as either a zero-mean Gaussian distribution $err_M \sim N(0, \sigma_M^2)$ or a zero mean uniform distribution $err_M \sim Uniform[-b, b]$.

4.1.3 Base Station Management

In this chapter, our goal is to minimize cost of power across the entire HCN by: (1) intelligently adjusting transmit powers of each BS, and (2) controlling renewable energy usage at each BS.

Adjusting Transmit Power

One way to minimize on-grid power consumption is to make transmit power of BSs as low as possible while preserving the QoS experienced by users. Since traffic condition is not considered in this work, we assume that the transmit power $P_k(i)$ of a k^{th} tier BS i is time-invariant and can range between P_{kmin} and P_{kmax} , where P_{kmin} and P_{kmax} denote the minimum and maximum transmit power of a k^{th} tier BS, respectively.

Controlling Renewable Energy Usage

In order to minimize the power drawn from the grid over time, harvested energy may be reserved in the battery for future use depending on expected pricing. Renewable energy

usage is denoted by $\eta_k(i, t) \in [0, 1]$, which refers to the fraction of stored renewable energy used to support the k^{th} tier BS i in the beginning of time slot t .

4.1.4 Lattice Model

Coverage probability can be defined as the CCDF of SINR as $P_C = \mathbb{P}[SINR > \theta_k]$. A lattice model is employed in this work, i.e., the ROI is modeled as a grid \mathcal{G} of points. We set a coverage probability requirement \mathbf{P}_C^{req} as a design parameter for each grid point, making sure that each point, $n = (x_n, y_n)$, is served by a k^{th} tier BS with an SINR no less than θ_k at least $P_C^{req}(n)$ of the time. Without loss of generality, we assume that \mathcal{G} is rectangular with dimensions N_x and N_y . BSs are operated in the ROI with the goal of satisfying coverage requirements $\mathbf{P}_C^{req} = \{P_C^{req}(n)\}, \forall n \in \mathcal{G}$.

For simplicity, the noise power is ignored [40] and each point is associated with the strongest BS (that is, the BS from which the received power is highest serves a user at the point). The fading power between a UE n and the associated strongest BS located at point s is denoted by h_{sn} and is assumed to be i.i.d. exponential (Rayleigh fading), i.e., $h_{sn} \sim \exp(1)$. The received SINR of the user n in the coverage of a k^{th} tier BS s is

$$SIR_k(n) = \frac{P_k(s)h_{sn}||s - n||^{-\alpha}}{\sum_{j=1}^K \sum_{b \in \phi_j \setminus s} P_j(b)h_{bn}||b - n||^{-\alpha}}, \quad (4.3)$$

where α denotes the path loss coefficient.

Table 4.1 summarizes the symbols introduced in this chapter.

4.2 On-grid Power Cost Minimization Formulation

In this section, we describe how one can optimize on-grid power cost via controlling both transmit power and renewable energy usage successively. The basic formulation is introduced in this section.

symbol	meaning	symbol	meaning
k	index of tier	i, j	index of BS
t	index of time	N_T	number of time windows
ϕ_k	k^{th} tier BSs	K	number of tiers
N	number of BSs	P	transmit power
θ	target SINR	P_H	harvesting power
\hat{P}_H	predicted mean value of P_H	err	random error
M	real-time electricity price	\hat{M}	predicted mean value of M
P_{min}	minimum transmit power	P_{max}	maximum transmit power
η	energy usage fraction	\mathcal{G}	grid model of ROI
n, s	points in \mathcal{G}	P_C	coverage probability
\mathbf{P}_C^{req}	required coverage probability	N_x, N_y	dimensions of \mathcal{G}
h	fading power	α	path loss coefficient
P_{grid}	on-grid power consumption	τ	time slot length
l	path loss	T	prediction horizon
E	energy storage	e	initial value of energy storage
η_{in}	charging efficiency	η_{out}	discharging efficiency
ϵ	designed positive scalar	\mathbf{u}	control variable
\mathbf{x}	state variable	J	objective function

Table 4.1: *Notation Table*

The problem can be stated as follows: given coverage probability requirement \mathbf{P}_C^{req} , how should BSs be operated (i.e., what are $\{P_k(i)\}$ and $\{\eta_k(i, t)\}$) such that on-grid power cost is minimized while attempting to satisfy coverage requirements. Here static power consumption which is power consumed even when a BS serves no user is not taken into account since it is a non-controllable constant.

We assume that: (1) charging rate is always less than the maximum charging rate; (2) energy storage is not enough to meet the power demand, and (3) capacity of battery is infinite. These assumptions are reasonable as the harvested energy is generally not sufficient for independent reliable network operation. Hence, in a real system, even though the battery capacity and charging rate are finite, there is a very low probability of battery overflow.

We can then mathematically state the decision problem as below:

$$\min_{\eta_k(i,t), P_k(i)} \mathbb{E}\left[\sum_{k=1}^K \sum_{i=1}^{N_k} \sum_{t=0}^{N_T-1} P_{gridk}(i, t)M(t)\tau\right], \quad (4.4)$$

subject to,

$$C1 : E_k(i, t) = (1 - \eta_k(i, t - 1))E_k(i, t - 1) + \eta_{in}\tau \cdot P_{Hk}(i, t - 1) \text{ [battery storage dynamics]},$$

$$C2 : P_{gridk}(i, t) = P_k(i) - \eta_k(i, t)\eta_{out}E_k(i, t)/\tau \text{ [grid power consumption]},$$

$$C3 : 0 \leq \eta_k(i, t) \leq 1 \text{ [constraint on energy usage]},$$

$$C4 : \mathbb{P}[SIR_k(n) > \theta_k] \geq P_C^{req}(n), \forall n \in \mathcal{G} \text{ [outage probability]},$$

$$C5 : P_k(i) \in [P_{kmin}, P_{kmax}], k = 1, \dots, K \text{ [transmit power]}$$

$$C6 : E_k(i, 0) = e_{ki} \text{ [battery storage initial condition]},$$

$$C7 : M(t) = \hat{M}(t) + err_M \text{ [electricity price]},$$

$$C8 : P_{Hk}(i, t) = \hat{P}_{Hk}(i, t) + err_H \text{ [harvested energy]}.$$

Here, N_T denotes the total time window of interest; $P_{gridk}(i, t)$ denotes on-grid power con-

sumption of the k^{th} tier BS i in time slot t ; τ denotes the length of the time slot; $E_k(i, t)$ denotes the amount of stored energy in the battery of the k^{th} tier BS i in the beginning of time slot t , and e_{ki} is a constant denoting the initial value of $E_k(i)$. η_{in} and η_{out} denote charging efficiency and discharging efficiency, respectively. The efficiency terms capture the loss of energy when harvested energy is stored and used from the battery.

Frequently adjusting transmit power of a BS can incur cost and instability. Also, since the BSs are centrally controlled, it is not economical to frequently send harvested energy data to the controller. We can avoid frequent switching by adjusting the value of τ .

Constraint C2 also imposes that $P_k(i) - \eta_k(i, t)\eta_{out}E_k(i, t)/\tau \geq 0$, i.e., $\eta_k(i, t) < \frac{\tau P_k(i)}{\eta_{out}E_k(i, t)}$. Therefore, C3 can be rewritten as

$$0 \leq \eta_k(i, t) \leq \min\left\{1, \frac{\tau P_k(i)}{\eta_{out}E_k(i, t)}\right\}. \quad (4.5)$$

Formulation (4.4) involves two key control parameters - (1) BS transmit power, and (2) fraction of stored energy used by each BS. In order to develop a simple, practical approach to address (4.4), we suggest a two-stage operation. First, we minimize energy consumption of the entire HCN via controlling transmit power of each BS. Second, we manage stored energy usage for each BS. Although these two stages can be completely decorrelated, and both power saving and cost reduction have been studied in literature separately, our work is distinguished by several points listed as below.

In general, our work considers HCNs modeled by PPP, which has been proved as a tractable yet accurate model. Power saving and cost reduction have not been well studied within this framework. In stage 1, unlike [44],[25]-[15], we treat BS power as a continuous parameter that is optimized to satisfy coverage probability constraints. Prior efforts related to power optimization [112]-[48] do not incorporate coverage probability constraints and do not deal with heterogeneous networks. Furthermore, for stage 2, most existing literature focus on reducing costs from deployment or maintenance of BSs (e.g., reduction in capital and operational expenditure [50] and deployment cost [51]). To the best of our knowledge,

[115] is the only work related to on-grid energy cost minimization. However, [115] is focused on active resource block allocation for a homogeneous macrocell network.

4.3 Stage 1 - Optimization of Transmit Power

At a single time slot, let P_{ki} be the transmit power for the i^{th} BS in the k^{th} tier to satisfy coverage probability. Our goal is to

$$\min_{\{P_{ki}\}} F(\{P_{ki}\}) = \sum_{k=1}^K \sum_{i=1}^{N_k} P_{ki}, \quad (4.6)$$

subject to,

$$\mathbb{P}[SIR_k(n) > \theta_k] \geq P_C^{req}(n), \forall n \in \mathcal{G}, \quad (4.7)$$

$$P_{ki} \in [P_{kmin}, P_{kmax}], i = 1, \dots, N_k, k = 1, \dots, K. \quad (4.8)$$

We aim to simultaneously minimize transmit power, i.e., $\{P_{ki}\}$, $i = 1, \dots, N_k$ and $k = 1, \dots, K$. A globally optimal solution is difficult to obtain when N_k is large. The problem in (4.6), however, is convex with respect to P_{ki} for a fixed i and k . Therefore, a locally optimal solution of this problem can be obtained by using the alternating algorithm where variables are computed one at a time while fixing all others [107]. Algorithm 4 shows the proposed alternating algorithm for the multi-variable transmit power minimization problem under QoS constraints.

4.3.1 Simplification of Signal-to-interference Ratio Constraint

The main difficulty of this approach is that the probabilities in (4.7) do not have simple closed-form expressions. Therefore, in this section we will describe methods to simplify the coverage constraint.

In Algorithm 4, we assume that variables in $\{P_{ki}\}$ are computed one at a time while fixing all others. Since the objective function (4.6) is the summation of all variables in $\{P_{ki}\}$,

Algorithm 4 Alternating Algorithm for Transmit Power Minimization under QoS Constraints

1. Initialization: $r = 0$
 $P_{ki}(0) = P_{kmax}, F(\{P_{ki}(0)\}) = 0, k = 1, \dots, K$ and $i = 1, \dots, N_k$.
 2. Iteration: $r + 1 \leftarrow r$
 Suppose $\mathbf{P}_{-ki}(r) = \{P_{11}(r), \dots, P_{k,i-1}(r)\}$ and $\mathbf{P}_{+ki}(r) = \{P_{k,i+1}(r), \dots, P_{KN_K}(r)\}$,
 $P_{11}(r) = \arg \min_{P_{11}} F(\mathbf{P}_{+11}(r-1))$ s.t. (4.7) and (4.8) are satisfied.

 $P_{ki}(r) = \arg \min_{P_{ki}} F(\mathbf{P}_{-ki}(r), \mathbf{P}_{+ki}(r-1))$ s.t. (4.7) and (4.8) are satisfied.

 $P_{KN_K}(r) = \arg \min_{P_{KN_K}} F(\mathbf{P}_{-KN_K}(r))$ s.t. (4.7) and (4.8) are satisfied.
 3. If $0 < F(\{P_{ki}(r-1)\}) - F(\{P_{ki}(r)\}) < \epsilon$, go to step 4 and stop, otherwise go back to step 2 (e.g., $\epsilon = 0.0001$).
 4. $\{P_{ki}^{opt}\} = \{P_{ki}(r)\}$.
-

the problem at each iteration in Algorithm 4 is equivalent to

$$\min_{P_{ki}} P_{ki} \tag{4.9}$$

subject to

$$P_{ki} \in [P_{kmin}, P_{kmax}], \tag{4.10}$$

$$\mathbb{P}[SIR_n = \frac{P_s h_s l_{sn}}{\sum_{j \neq s} P_j h_j l_{jn}} \geq \theta_s] \geq P_C^{req}(n), \forall n \in \mathcal{G}. \tag{4.11}$$

where s denotes the strongest BS serving user n , θ_s denotes the corresponding SIR threshold of the BS s and $l_{sn} = \|s - n\|^{-\alpha}$ denotes the path loss between user n and BS s . Similarly, j denotes an interfering BS and $l_{jn} = \|j - n\|^{-\alpha}$ denotes the path loss between user n and BS j .

According to [116],

$$\mathbb{P}[SIR_n \geq \theta_s] \geq P_C^{req}(n) \tag{4.12}$$

can be written as

$$\prod_{j \neq s} \left(1 + \frac{\theta_s P_j l_{jn}}{P_s l_{sn}}\right) \leq \frac{1}{P_C^{req}(n)} \quad (4.13)$$

That is, problem (4.9) can be expressed in closed form as

$$\min_{P_{ki}} P_{ki} \quad (4.14)$$

subject to

$$P_{ki} \in [P_{kmin}, P_{kmax}], \quad (4.15)$$

$$T(P_{ki}) = \prod_{j \neq s} \left(1 + \frac{\theta_s P_j l_{jn}}{P_s l_{sn}}\right) \leq \frac{1}{P_C^{req}(n)}, \forall n \in \mathcal{G}. \quad (4.16)$$

Based on $\{P_{ki}\}$ values in the previous iteration in Algorithm 4, we suppose that \mathcal{G} can be separated into \mathcal{G}_1 and \mathcal{G}_2 , which are the sets of users connecting and not connecting to the i^{th} BS in the k^{th} tier, respectively.

It can be shown that when $n \in \mathcal{G}_1$, i.e., $s = ki$, $T(P_{ki}) = \prod_{j \neq ki} \left(1 + \frac{\theta_s P_j l_{jn}}{P_{ki} l_{kin}}\right)$ is a monotone decreasing function of P_{ki} . We can obtain an optimal $P_{ki;n}^*$ by solving

$$T(P_{ki}) = \prod_{j \neq ki} \left(1 + \frac{\theta_s P_j l_{jn}}{P_{ki} l_{kin}}\right) = \frac{1}{P_C^{req}(n)}, n \in \mathcal{G}_1. \quad (4.17)$$

We then define P_{ki}^* as

$$P_{ki}^* = \begin{cases} P_{min}, & \max_{n \in \mathcal{G}_1} \{P_{ki;n}^*\} < P_{min} \\ \max_{n \in \mathcal{G}_1} \{P_{ki;n}^*\}, & P_{min} \leq \max_{n \in \mathcal{G}_1} \{P_{ki;n}^*\} \leq P_{max} \\ \phi, & \max_{n \in \mathcal{G}_1} \{P_{ki;n}^*\} > P_{max} \end{cases} \quad (4.18)$$

When $n \in \mathcal{G}_2$, i.e., $s \neq ki$, $T(P_{ki})$ can be written as

$$T(P_{ki}) = \left[\prod_{j \neq s, j \neq ki} \left(1 + \frac{\theta_s P_j l_{jn}}{P_s l_{sn}}\right) \right] \left(1 + \frac{\theta_s P_{ki} l_{kin}}{P_s l_{sn}}\right) \leq \frac{1}{P_C^{req}(n)}, n \in \mathcal{G}_2. \quad (4.19)$$

Note that (4.19) is a monotone increasing function of P_{ki} . We should point out that (4.19) actually provides an upper bound on P_{ki} . Therefore, substituting P_{ki}^* into P_{ki} in (4.19), if all the inequalities hold, P_{ki}^* is the optimal solution for (4.14). Otherwise there is no feasible solution, suggesting that the coverage condition cannot be satisfied and may have to be modified (i.e., θ_s or $P_C^{req}(n)$ may have to be reduced).

4.3.2 Convergence of Algorithm 4

The convergence of Algorithm 4 can be proved by induction presented as below:

1. For loop 1, $P_{ki}(1) \leq P_{ki}(0) = P_{kmax}$, $k = 1, \dots, K$ and $i = 1, \dots, N_k$.
2. For loop $r = 2, 3, 4, \dots$, $k = 1, \dots, K$ and $i = 1, \dots, N_k$.

If

- a) $P_{k'i'}(r) \leq P_{k'i'}(r-1)$, $k' < k$ or $k' = k, i' < i$.
- b) $P_{k'i'}(r-1) \leq P_{k'i'}(r-2)$, $k' > k$ or $k' = k, i' > i$.

then according to (4.17) and (4.18),

$$\begin{aligned} P_{ki}(r) &= \arg \min_{P_{ki}} F(\mathbf{P}_{-ki}(r), \mathbf{P}_{+ki}(r-1)) \\ &\leq \arg \min_{P_{ki}} F(\mathbf{P}_{-ki}(r-1), \mathbf{P}_{+ki}(r-2)) = P_{ki}(r-1) \end{aligned} \quad (4.20)$$

3. Therefore, $P_{ki}(r)$, $k = 1, \dots, K$ and $i = 1, \dots, N_k$ gradually decreases as r increases. Since the objective function $F(P_{ki})$ is the sum of all variables P_{ki} , the desired optimization $F(P_{ki})$ is also gradually minimized as iteration index r is increased.

As [107] suggests, since the update of P_{ki} , $i = 1, \dots, N_k$ and $k = 1, \dots, K$ may either decrease or remain the same but cannot increase the objective in (4.6), monotonic convergence of P_{ki} follows. The difference between $F(\{P_{ki}(r-1)\}) - F(\{P_{ki}(r)\})$ can be used as a criterion to stop the alternating algorithm with a designed positive scalar ϵ in Step 3.

θ_1	θ_2	$P_C^{req}(n_L)$	$P_C^{req}(n_R)$	$\frac{F(\{P_{ki}^*\})}{F(\{P_{kmax}\})}$
-15dB	-15dB	70%	90%	18.38%
-15dB	-15dB	80%	90%	18.38%
-15dB	-15dB	90%	90%	20.15%
-10dB	-10dB	70%	70%	18.41%
-10dB	-10dB	70%	80%	30.47%
-10dB	-10dB	80%	80%	34.02%

Table 4.2: *Power Optimization Simulation Results*

4.3.3 Simulation Result

Consider a 10×10 grid as our ROI. Without loss of generality, we consider a HCN consisting of macrocells and picocells and assume that the coverage probability requirements of the left half of the ROI $P_C^{req}(n_L)$ and that of the right half $P_C^{req}(n_R)$ can be different. The locations of BSs are randomly generated according to uniform distributions. Setting $\alpha = 3$, $N_1 = 2$, $N_2 = 16$, $P_{1max} = 1000W$, $P_{1min} = 150W$, $P_{2max} = 60W$ and $P_{2min} = 15W$ [117]-[119], simulation results are listed in Table 4.3.3.

From simulation results in Table 4.3.3, the effectiveness of the proposed transmit power optimization algorithm is evident. For example, power savings of nearly 20% is feasible for practical SINR threshold and coverage requirements. The exact transmit power reduction depends on several factors such as (1) the allowed power range for BSs, (2) the SIR threshold, and (3) the coverage requirements.

4.3.4 Problem Transformation

After calculating optimal transmit power of each BS and expressing the index of BSs as j for simplicity, the problem introduced in (4.4) can be rewritten as:

$$\min_{\eta(j,t)} \mathbb{E} \left[\sum_{t=0}^{N_T-1} M(t) \sum_{j=1}^{N_1+\dots+N_K} P_{grid}(j,t)\tau \right], \quad (4.21)$$

subject to,

$$\begin{aligned}
P_{grid}(j, t) &= P_j^* - \eta(j, t)E(j, t)\eta_{out}/\tau, \\
E(j, t) &= (1 - \eta(j, t - 1))E(j, t - 1) + \eta_{in}P_H(j, t)\tau, \\
0 \leq \eta(j, t) &\leq \min\left\{1, \frac{\tau P_j^*}{\eta_{out}E_k(j, t)}\right\}, \\
E(j, 0) &= e_j, \\
M(t) &= \hat{M}(t) + err_M, \\
P_H(j, t) &= \hat{P}_H(j, t) + err_H.
\end{aligned}$$

where e_j denotes the initial value of energy storage of the j^{th} BS and $P_H(j, t)$ denotes the harvested energy of the j^{th} BS at the t^{th} time slot.

Since P_j^* is a constant, (4.21) can be further simplified as

$$\max_{\eta(j, t)} \mathbb{E}\left[\sum_{t=0}^{N_T-1} M(t) \sum_{j=1}^{N_1+\dots+N_K} \eta(j, t)E(j, t)\right], \quad (4.22)$$

subject to,

$$\begin{aligned}
E(j, t) &= (1 - \eta(j, t - 1))E(j, t - 1) + \eta_{in}P_H(j, t)\tau, \\
0 \leq \eta(j, t) &\leq \min\left\{1, \frac{\tau P_j^*}{\eta_{out}E_k(j, t)}\right\}, \\
E(j, 0) &= e_j, \\
M(t) &= \hat{M}(t) + err_M, \\
P_{Hk}(i, t) &= \hat{P}_{Hk}(i, t) + err_H.
\end{aligned}$$

Note that there is no interaction between different BSs when optimizing $\eta_k(i, t)$. Therefore, (4.22) can be directly decomposed into subproblems for individual BSs. For a single

BS, the formulation simplifies as,

$$\max_{\eta(t)} \mathbb{E} \left[\sum_{t=0}^{N_T-1} M(t)\eta(t)E(t) \right], \quad (4.23)$$

subject to,

$$E(t) = (1 - \eta(t - 1))E(t - 1) + \eta_{in}P_H(t)\tau,$$

$$0 \leq \eta(t) \leq \min \left\{ 1, \frac{\tau P^*}{\eta_{out}E_k(t)} \right\},$$

$$E(0) = e_0,$$

$$M(t) = \hat{M}(t) + err_M,$$

$$P_H(t) = \hat{P}_H(t) + err_H.$$

where e_0 denotes the initial value of battery stored energy and P^* denotes the optimal transmit power of the considered BS.

In the next section, we solve for the renewable energy usage for each BS considering the simplified problem (4.23).

4.4 Stage 2 - Controlling Stored Energy Usage - NMPC Framework

NMPC is a variant of model predictive control (MPC) that is characterized by the use of nonlinear system models. As in linear MPC, NMPC requires the iterative solution of optimal control problems on a finite prediction horizon. In this section, a convenient form of representing the process to be controlled by NMPC is presented first followed by a discussion of the solution.

4.4.1 NMPC Formulation

In the problem stated in (4.23), it is straightforward to define the state variable $x(t)$ as stored energy $E(t)$ and control variable $u(t)$ as renewable energy usage $\eta(t)$. The cost function can be expressed as $J = -f[u(t)] = -\sum_{t=0}^{N_T-1} M(t)u(t)x(t)$. Equation (4.23) can be rewritten as

$$\max_{\mathbf{u}} \mathbb{E}\left[\sum_{t=0}^{N_T-1} M(t)u(t)x(t)\right], \quad (4.24)$$

subject to,

$$x(t+1) = (1 - u(t))x(t) + \eta_{in}P_H(t)\tau,$$

$$0 \leq u(t) \leq \min\left\{1, \frac{\tau P^*}{\eta_{out}x(t)}\right\},$$

$$x(0) = e_0,$$

$$M(t) = \hat{M}(t) + err_M,$$

$$P_H(t) = \hat{P}_H(t) + err_H.$$

It is clear that equation (4.24) corresponds to a nonlinear, nonconvex constrained optimization problem. Finding a globally optimal solution to such problems are prohibitively complex. With the goal of a simple practical solution for BS energy management, we propose an approach that will provide an effective suboptimal solution. Our method is based on linearizing the constraints and then solving (4.24) via an iterative optimal control approach for a finite prediction horizon. In the constraint

$$x(t+1) = x(t) - u(t)x(t) + \eta_{in}P_H(t)\tau, \quad (4.25)$$

we can linearize $u(t)x(t)$ around $(\hat{u}(t), \hat{x}(t))$:

$$u(t)x(t) \approx \hat{u}(t)\hat{x}(t) + \hat{x}(t)(u(t) - \hat{u}(t)) + \hat{u}(t)(x(t) - \hat{x}(t)), \quad (4.26)$$

which results in the linearized constraint

$$x(t+1) = (1 - \hat{u}(t))x(t) + \hat{x}(t)\hat{u}(t) - \hat{x}(t)u(t) + \eta_{in}P_H(t)\tau. \quad (4.27)$$

Here,

$$\hat{x}(t) = \begin{cases} e_0, & t = 0, \\ x^*(t), & t = 1, 2, \dots \end{cases} \quad (4.28)$$

where $x^*(t) = x^*(t-1) - u^*(t-1)x^*(t-1) + \eta_{in}P_H(t)\tau$, and

$$\hat{u}(t) = \begin{cases} \eta_0, & t = 0 \\ u^*(t-1), & t = 1, 2, \dots \end{cases} \quad (4.29)$$

$\mathbf{x}^*(t)$ denotes states calculated in previous iterates; $\mathbf{u}^*(t)$ denotes obtained optimal actions, and η_0 refers to the initial choice of renewable energy usage fraction.

4.4.2 Solution of the NMPC Problem

The basic idea behind NMPC is to solve a finite horizon optimal control problem at every time step. Since err_H and err_C are zero-mean variables, for a prediction horizon T , we ignore parameter variation and solve the ‘‘certainty-equivalent’’ planning problem [105] at each time t :

$$\max \sum_{\tau=t}^{t+T-1} \hat{M}(\tau)u(\tau)x(\tau), \quad (4.30)$$

subject to,

$$u(\tau) \in \mathcal{U} = [0, \min\{1, \frac{\tau P^*}{\eta_{out} x(t)}\}], \quad \tau = t, \dots, t + T, \quad (4.31)$$

$$x(\tau + 1) = (1 - \hat{u}(t))x(\tau) + \hat{x}(t)\hat{u}(t) + \eta_{in}\hat{P}_H(\tau)\tau_0 - \hat{x}(t)u(\tau), \quad \tau = t, \dots, t + T - 1, \quad (4.32)$$

$$x(0) = e_0, \quad (4.33)$$

$$\hat{x}(t) = \begin{cases} e_0, & t = 0 \\ x^*(t), & t = 1, 2, \dots \end{cases} \quad (4.34)$$

$$\hat{u}(t) = \begin{cases} u_0, & t = 0 \\ u^*(t - 1), & t = 1, 2, \dots \end{cases} \quad (4.35)$$

The law of evolution for the state variable $x(\tau)$ is

$$x(\tau + 1) - x(\tau) = -\hat{u}(t)x(\tau) + \hat{x}(t)\hat{u}(t) - \hat{x}(t)u(\tau) + \eta_{in}\hat{P}_H(\tau)\tau_0, \quad \tau = t, \dots, t + T - 1. \quad (4.36)$$

Then, we form the Hamiltonian and differentiate:

$$\mathcal{H} = \hat{M}(\tau)u(\tau)x(\tau) + \lambda(\tau)(-\hat{u}(t)x(\tau) + \hat{x}(t)\hat{u}(t) - \hat{x}(t)u(\tau) + \eta_{in}\hat{P}_H(\tau)\tau_0), \quad (4.37)$$

$$\frac{\partial \mathcal{H}}{\partial u(\tau)} = \hat{M}(\tau)x(\tau) - \lambda(\tau)\hat{x}(t) = 0, \quad (4.38)$$

$$\lambda(\tau + 1) - \lambda(\tau) = -\frac{\partial \mathcal{H}}{\partial x(\tau)} = -\hat{M}(\tau)u(\tau) + \lambda(\tau)\hat{u}(t). \quad (4.39)$$

Using the above equations, it is easy to solve for the state and co-state evolution ($x(\tau)$ and $\lambda(\tau)$). First, we need to express $x(\tau)$ and $u(\tau)$ as functions of $\lambda(\tau)$, respectively.

From (4.38),

$$x(\tau) = \frac{\lambda(\tau)\hat{x}(t)}{\hat{M}(\tau)}. \quad (4.40)$$

And from (4.39),

$$u(\tau) = \frac{(1 + \hat{u}(t))\lambda(\tau) - \lambda(\tau + 1)}{\hat{M}(\tau)}. \quad (4.41)$$

Then substituting (4.40) and (4.41) into (4.36), we have

$$\begin{aligned} & \frac{\lambda(\tau + 1)\hat{x}(t)}{\hat{M}(\tau + 1)} - \frac{\lambda(\tau)\hat{x}(t)}{\hat{M}(\tau)} \\ &= -\frac{\hat{u}(t)\lambda(\tau)\hat{x}(t)}{\hat{M}(\tau)} - \hat{x}(t)\frac{(1 + \hat{u}(t))\lambda(\tau) - \lambda(\tau + 1)}{\hat{M}(\tau)} + \hat{x}(t)\hat{u}(t) + \eta_{in}\hat{P}_H(\tau)\tau_0. \end{aligned} \quad (4.42)$$

Here (4.42) can be simplified as

$$\begin{aligned} \lambda(\tau) &= \{[\hat{x}(t)\hat{M}(\tau + 1) - \hat{x}(t)\hat{M}(\tau)]\lambda(\tau + 1) + \hat{M}(\tau) \cdot \hat{x}(t)\hat{u}(t)\hat{M}(\tau + 1) + \\ & \hat{M}(\tau)\hat{M}(\tau + 1)\eta_{in}\hat{P}_H(\tau)\tau_0\}[2\hat{u}(t) \cdot \hat{M}(\tau + 1)\hat{x}(t)]^{-1}. \end{aligned} \quad (4.43)$$

We can easily see that

$$\begin{aligned} x(\tau + 1) &= (1 - \hat{u}(t))x(\tau) + \eta_{in}\hat{P}_H(\tau)\tau_0 - \frac{\lambda(\tau + 1) - (1 - \hat{u}(t))\lambda(\tau)}{\hat{M}(\tau)}\hat{x}(t) \\ &+ \hat{x}(t)\hat{u}(t), \tau = t, \dots, t + T - 1. \end{aligned} \quad (4.44)$$

The control action $u^*(t)$ which corresponds to the fraction of stored energy used at each time instant t can be determined via the following steps:

1. From $\lambda(t + T) = 0$, we can solve $\lambda(t + T - 1)$, $\lambda(t + T - 2)$, ..., $\lambda(t)$ sequentially by (4.43).
2. $x(t + 1)$, $x(t + 2)$, ..., $x(t + T - 1)$ can be solved sequentially by (4.44).
3. Having obtained $\lambda(t + T - 1)$, $\lambda(t + T - 2)$, ..., $\lambda(t)$, and $x(t + 1)$, $x(t + 2)$, ..., $x(t + T - 1)$, $u^{opt}(t)$ can be solved by (4.41).
4. By projecting $u^{opt}(t)$ onto the set in (4.31), we can get the optimal control action $u^*(t)$.

Initial value of stored energy (e_0)	1 kWh
Initial guess of action (η_0)	0.9
Number of time slots (N_T)	24
Charging efficiency (η_{in})	0.9
Time steps in the time window (T)	4

Table 4.3: *Simulation parameters*

Thus, by obtaining $u^*(0), \dots, u^*(N_T - 1)$, NMPC problem (4.24) is solved. Simulation results to verify the performance are presented in the next section.

4.5 Simulation Results

We present numerical results to demonstrate the performance of the proposed approach. The basic simulation parameters are listed in Table 4.3.

Three different scenarios are studied in this section. A detailed discussion of the simulation results related to these cases is provided first. Then objective function values $f[\mathbf{u}(t)]$ in three scenarios are presented in Table 4.4 to verify/compare the performance of the proposed approach.

4.5.1 Case-1

The real-time electric price is assumed to vary from 0.15 to 0.25 dollars/kWh with an estimation error distribution $err_M \sim Uniform[-0.02, 0.02]$. Using wind energy model, we assume that harvested energy of each time slot varies from 2 to 1.12 kWh [120] with an estimation error distribution $err_H \sim Uniform[-0.2, 0.2]$. Due to the complexity of the profile, it is difficult to predict apriori any clear trend of energy usage. However, the effectiveness of this algorithm can be verified by objective function values in Table 4.4 and simulation results are presented in Fig. 4.1. Note that we plot a scaled version of electricity price to better show its variation and relationship with other variables on the same plot.

Curve $\mathbf{x}^*\mathbf{u}^*$ represents energy usage. From Fig. 4.1 it is evident that $\mathbf{x}^*\mathbf{u}^*$ primarily follows harvested energy P_H , especially in the period from $t = 5$ to $t = 23$, where electricity price doesn't change much.

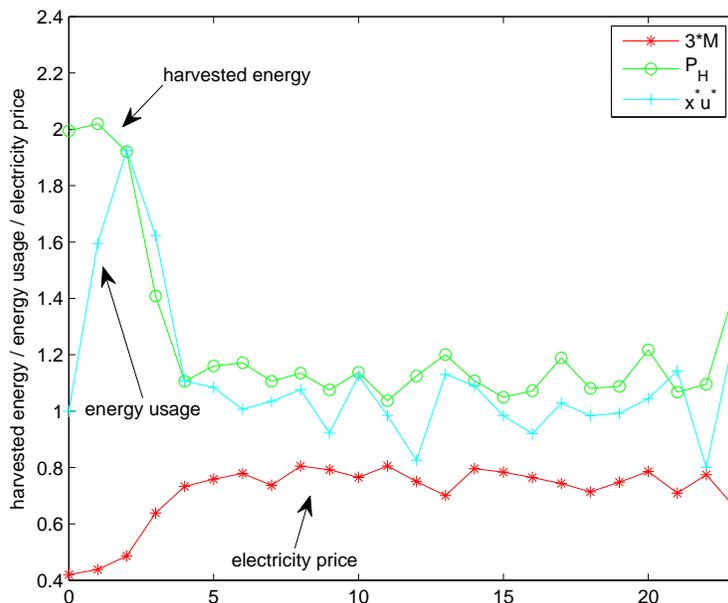


Figure 4.1: Simulation Results of Case-1

4.5.2 Case-2

To clearly illustrate the effect of electricity price on $\mathbf{u}^*(t)$, the real-time electric price is assumed to vary from approximately 0.2 to 1.2 dollars/kWh with an estimation error distribution $err_H \sim Uniform[-0.01, 0.01]$. And we assume that harvested energy of each time slot stays at 1 kWh with an estimation error distribution $err_H \sim Uniform[-0.01, 0.01]$. When P_H stays about the same and err_H is relatively negligible, it is expected that when on-grid electricity cost is high, we would want to use more renewable energy from storage. When cost is low, we tend to reserve more harvested energy for future use. From Fig. 4.2 it is clear that $\mathbf{x}^*\mathbf{u}^*$ curve which represents energy usage follows M with a small delay, which

is a result of the prediction time window in NMPC. Another factor to remember is that energy usage based on storage has an efficiency factor that limits the amount of useful reserve energy. This is reflected by the lower delayed peak of $\mathbf{x}^*\mathbf{u}^*$ relative to the cost profile. Initially, the electricity cost is low, so there is no incentive to use all harvested energy. As the cost goes up, the energy usage also goes up to save on-grid power consumption. Also, the effectiveness of this algorithm can be verified by objective function values in Table 4.4.

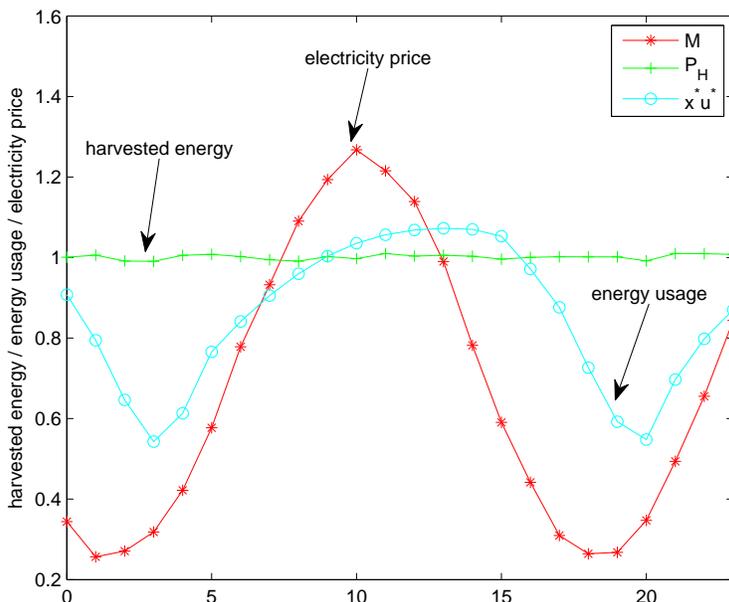


Figure 4.2: Simulation Results of Case-2

4.5.3 Case-3

To clearly show the impact of harvested energy on $\mathbf{u}^*(t)$, the real-time electric price is assumed to stay at 0.5 dollars/kWh with an estimation error distribution $err_M \sim Uniform[-0.02, 0.02]$. We assume that harvested energy at each time slot varies from 0.2 to 1.2 kWh with an estimation error distribution $err_H \sim Uniform[-0.01, 0.01]$. When predicted electricity price \hat{M} stays about the same and the error of it err_M is relatively negligible, it is expected that

it doesn't matter how we use stored energy. Detailed simulation results are presented in Fig. 4.3. From Fig. 4.3 it is shown that the curve of $\mathbf{x}^*\mathbf{u}^*$ follows with P_H with a small delay (that is, energy storage is fully used at each time slot), which is exactly what we expect. Also, it can be verified by objective function values in Table 4.4.

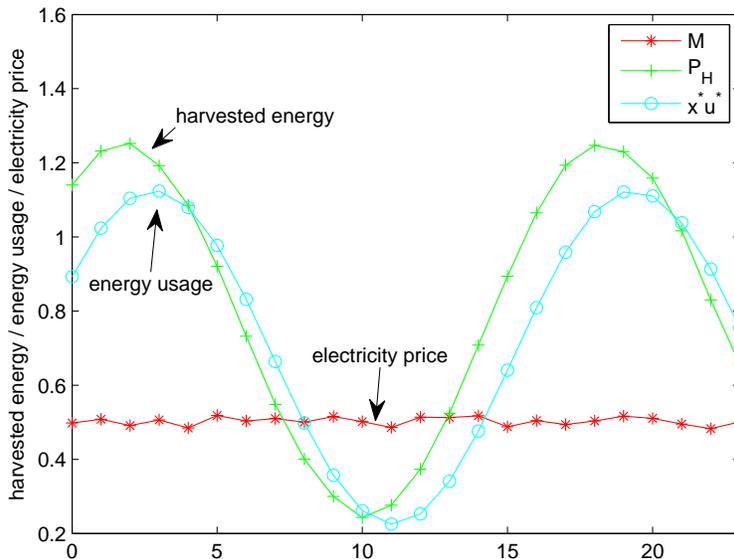


Figure 4.3: *Simulation Results of Case-3*

We compare obtained objective function values $f[\mathbf{u}^*(t)]$ with $f[\mathbf{u}_1(t)]$ and $f[\mathbf{u}_0(t)]$. Here $\mathbf{u}_1(t) = \mathbf{1}_{1 \times N_T}$, which represents the case that harvested energy is always used immediately in the next time slot. And $\mathbf{u}_0(t) = [\mathbf{0}_{1 \times N_T-1} \ 1]$, which represents the case that harvested renewable energy is never used until the last time slot. Table 4.4 summarizes the simulation results of $f[\mathbf{u}^*(t)]$ in comparison of $f[\mathbf{u}_1(t)]$ and $f[\mathbf{u}_0(t)]$. From Table 4.4 we can see that with $\mathbf{u}^*(t)$ obtained by this approach, $f[\mathbf{u}^*(t)]$ has better performance than $f[\mathbf{u}_1(t)]$ and $f[\mathbf{u}_0(t)]$ in general. In Case-3 $f[\mathbf{u}^*(t)]$ is equivalent to $f[\mathbf{u}_1(t)]$ since in that case $\mathbf{u}^* = \mathbf{u}_1$.

	$f[\mathbf{u}^*(t)]$	$f[\mathbf{u}_1(t)]$	$f[\mathbf{u}_0(t)]$
Case-1	6.0428	6.0147	5.2548
Case-2	20.9079	20.1316	20.3989
Case-3	10.2012	10.2012	10.1510

Table 4.4: *Simulation Results*

4.5.4 Case-4: Effect of Time Window

In Fig. 4.4, simulation results of Case-2 with $T = 2$, $T = 3$ and $T = 4$ are presented to analyze the effect of time window T . From Fig. 4.4 we can see that higher time steps in the time window T can enhance the performance of the approach. In this case, $T = 2$ is not sufficient as decisions for managing storage is myopic. That is, there is not enough knowledge of future trends of cost/harvested energy to guide our current behavior. Time window lengths of $T = 3$ and $T = 4$ both seem to provide reasonably intuitive results. As T increases beyond 4, we observe that there is no significant improvement in performance. It is important to realize that computational cost does increase with increasing T . Thus, identifying the most suitable time window for a specific system is a design parameter that needs to be carefully selected.

In summary, the effectiveness of the proposed two-stage BS operation is illustrated by numerical results of Case-1, Case-2 and Case-3. Simulation results also indicate that while a larger time window T can enhance the performance of the approach, the improvement would have a diminishing return as well.

4.6 Summary

This chapter investigates, for the first time, management of energy harvesting BSs in HCNs from a techno-economic point of view. We consider real-time electricity price, QoS of users and harvested energy profile in identifying methods to minimize energy cost. To solve the formulated problem in a simple and practical manner, we divide it into two subproblems and

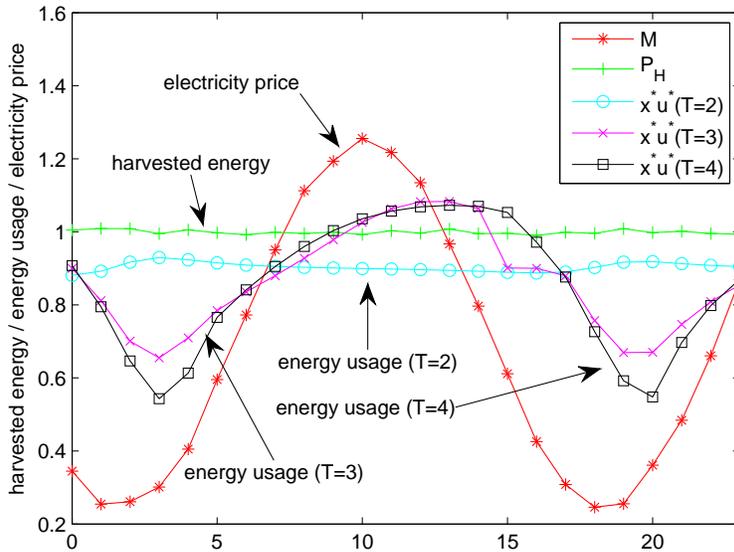


Figure 4.4: *Effect of Time Window*

sequentially solve them via an optimization approach and a control algorithm, respectively. The proposed two-stage optimization/control approach provides a method to manage both transmit power and stored energy usage of HCN BSs to reduce on-grid electricity expenditure. Numerical results confirm the effectiveness of the scheme.

With D2D technology, mobile users in proximity can directly communicate to each other without BSs. This can significantly bring down the energy cost. However, interference brought by D2D users has become a challenging problem. In next chapter, we address energy-aware capacity evaluation and optimization for D2D underlay cellular networks.

Chapter 5

Power-aware Performance Analysis and Optimization in D2D Underlay Networks

Device-to-device (D2D) technology exploits direct communication between two users within a short range [121] and is regarded as one of the key technologies for 5G wireless communication system. It has great potential to improve both spectral and energy efficiency due to spectral reuse gain and the proximity of communication parties, respectively. However, introducing an underlay of D2D users presents many challenges to the long-standing cellular architecture. One of the main consequences involves new sources of interference: intra-cell and inter-cell interference between cellular users (CUs) and D2D users (DUs), and interference among D2D users. Therefore, it is critical to carefully design system parameters (density of D2D users, cellular BSs, transmit power etc.) to make D2D underlay operation beneficial while guaranteeing the performance of cellular networks.

This chapter provides some unique and novel perspectives in terms of modeling, design and analysis of D2D underlay networks. Our goal is to analyze a realistic D2D network. Unlike the work in [72], we consider a network model, where distances between D2D trans-

mitting user (DTU) and D2D receiving user (DRU) are treated as a random variable. First of all, we seek to address a fundamental open question that is important to address as we transition to 5G wireless systems. Is there a critical set of system parameters (density of D2D users, cellular BSs, transmit power etc.) that can ensure that the benefits of D2D underlay operation can outweigh its drawbacks? In the quest to answer this question for a realistic D2D network model, we also uncover new analytical results. With D2D links characterized as Rician fading channels, we derive, for the first time, upper and lower bounds for ergodic capacity of a cellular network, and recursive closed-form expression and closed-form approximation for ergodic capacity of D2D communication for the case of path loss coefficient 4. These results are shown to better approximate ergodic capacity related to the previous result presented in [72]. Additionally, we identify the D2D user density and transmit power that maximizes global ergodic capacity of the network. Specifically, a two-phase scheme is proposed to optimize ergodic capacity while minimizing overall power consumption. Results from this work provide a framework to uncover desirable system design parameters that offer the best gains in terms of ergodic capacity. Secondly, we analyze the average achievable throughput of D2D underlay cellular networks. Closed-form analytical results are provided for average achievable throughput in a D2D underlay cellular network. Finally, we provide closed-form ergodic capacity results (when path loss exponent is 4) for two other cases of the distribution of DRU: (1) distance between a DU pair follows a uniform distribution and (2) a DRU is distributed uniformly in the circular area around its serving DTU.

This chapter includes several distinct sections. Section 5.1 describes the system model. Closed-form results of ergodic capacity are derived in Section 5.2. Section 5.3 discusses optimization of ergodic capacity and total power consumption of a D2D underlay cellular network. Then Section 5.4 analyzes the average achievable throughput of D2D underlay cellular networks. Section 5.5 provides closed-form ergodic capacity results for two other cases of the distribution of DRU. Finally, conclusions and future work are presented in Section 5.6.

5.1 System Model

In this section, the D2D underlay cellular network model is introduced. Then, based on the SIR metric, we provide success probability expressions for both CUs and D2D users.

5.1.1 Network Model

We consider a downlink D2D underlay cellular network. CUs are assumed to communicate with BSs while DUs communicate with each other directly. We assume that the whole bandwidth is divided into a number of subchannels. The DTUs share all of the downlink subchannels (referred to as “shared channels”) with BSs, which causes co-channel interference between cellular networks and D2D systems. The whole bandwidth is assumed to be sufficiently large so multiple DTUs won’t share spectrum. Then we model the spatial locations of cellular BSs (or DTUs) as a PPP Φ_C (Φ_T) with density λ_C (λ_T) and transmission power P_C (P_T). Locations of CUs (or DRUs) are also assumed to follow a PPP and each CU (DRU) is associated with its nearest BS (DTU). The probability density function of the distance X_D from a DRU to its serving DTU is expressed as

$$f_{X_D}(r) = 2\pi\lambda_T r \exp(-\lambda_T \pi r^2), \quad (5.1)$$

which is derived based on the null probability of a two dimensional PPP [96].

We assume the downlink channel to be a combination of large-scale attenuation and fading for both cellular networks and D2D communication. Since DU pairs are usually close to each other with a LOS component, it is practical to characterize the D2D links via a Rician fading model along with a scattering component. The LOS fading power follows a non-central χ^2 distribution with a Rician K_R -factor (defined as the ratio of the LOS component power to the scattered component power). On the other hand, we assume Rayleigh fading for the BS-CU and interfering links such as BS-DRU and DTU-CU based on the rich scattering environment with NLOS channels. Therefore, the interfering link

fading power follows an exponential distribution. We model large-scale attenuation as the standard path loss propagation $l_\alpha(tr, re) = (R_{tr,re})^{-\alpha}$, where $R_{tr,re}$ denotes the distance between transmitter tr and receiver re , and α represents the path loss coefficient for all links.

Note that the analysis in this chapter can also be applied to an uplink D2D underlay cellular network when the DTUs share uplink spectrum with CUs. The only difference is that in this case λ_C and P_C represent the density and transmit power of CUs, respectively.

5.1.2 Success Probabilities

Focusing on spectrum sharing, we assume that (1) the hybrid network is interference-limited, (2) noise is negligible [96] and (3) scheduling amongst D2D users is not taken into consideration. The success probability is defined as the probability of the SIR ξ exceeding a threshold θ , i.e., $P_s(\theta) = \mathbb{P}[\xi > \theta]$ (CCDF of SIR). The success probability for the cellular users is presented in the following lemma.

Lemma 2. *For a D2D underlay cellular network, the cellular success probability can be obtained as:*

$$P_{s,C}(\theta_C) = \frac{1}{\rho(\theta_C, \alpha) + C(\alpha) \frac{\lambda_T}{\lambda_C} \left(\frac{P_T}{P_C} \theta_C\right)^{\frac{2}{\alpha}} + 1}, \quad (5.2)$$

where $\rho(\theta_C, \alpha) = \theta_C^{\frac{2}{\alpha}} \int_{\theta_C^{-\frac{2}{\alpha}}}^{\infty} \frac{1}{1+u^{\frac{\alpha}{2}}} du$, and $C(\alpha) = \frac{2\pi}{\alpha \sin(2\pi/\alpha)}$.

Proof. See [122]. □

From (5.2) we can observe that the success probability of CUs in a shared channel decreases with the increase of λ_T/λ_C and P_T/P_C . The success probability for DUs is presented in Lemma 3 as below.

Lemma 3. *Suppose the D2D signal experiences Rician fading with Rician factor K_R , the*

success probability of a D2D pair at a given distance X_D is expressed as:

$$P_{s,D}(\theta_D|X_D) = \sum_{k=1}^{\infty} \sum_{m=0}^{k-1} (-1)^{k-m} J(m, k) \exp(-C_d \theta_D^{2/\alpha}) \cdot \sum_{j=1}^{k-m} \frac{\beta_j^{k-m}}{j!} (C_d \theta_D^{2/\alpha})^j + \sum_{k=0}^{\infty} J(k, k) \exp(-C_d \theta_D^{2/\alpha}), \quad (5.3)$$

$$\text{where } J(m, k) = \frac{\frac{K_C^k m!}{e^{K_R}} \binom{k}{m}}{(k!)^2}, C_d = \pi C(\alpha) X_D^2 (\lambda_C (\frac{P_C}{P_T})^{\frac{2}{\alpha}} + \lambda_T), \beta_j^n = \sum_{i=1}^j (-1)^j \binom{j}{i} (\frac{2i}{\alpha})_n,$$

$j = 1, \dots, n$ and $(x)_n \triangleq (x)(x-1)\dots(x-n+1)$.

Proof. See [72]. □

5.2 Ergodic Capacity Analysis

Ergodic capacity in a subchannel is defined as $C = \mathbb{E}[\log(1+\xi)] = \int_{t=0}^{\infty} Pr[\log(1+\xi) > t] dt = \int_{t=0}^{\infty} Pr[\xi > e^t - 1] dt$. The integrand in the definition of ergodic capacity is success probability. However, maximizing C is not trivial, since there does not exist any closed-form results for realistic D2D underlay cellular networks. The success probability results in both Lemma 2 and Lemma 3 (from [72]) are based on Rician fading model for D2D links. To evaluate ergodic capacity, the key limitations of the results in [72] are: (1) they are not in closed form, and (2) the distances between all D2D pairs (X_D) are assumed to be fixed, which reduces the applicability of the derived results.

In this section, we aim to overcome these two drawbacks and evaluate the ergodic capacity in closed form for a practical case of path loss coefficient $\alpha = 4$. Although it is a special case, the closed-form results can provide us theoretical insights, such as the behavior of performance metrics and the effect of relevant parameters. Therefore, $\alpha = 4$ case study has been a useful tool and widely utilized. For example, [96] evaluates the coverage probability and average achievable rate of a cellular network when $\alpha = 4$. [117] studies the feasibility of simultaneous information and energy transfer in LTE-A small cell networks in

a interference-limit case with $\alpha = 4$. [123] derives the asymptotic transmission capacities for $\alpha = 4$ of two coexisting wireless networks that operate in the same geographic region and share the same spectrum. [124] derives accurate expressions for the power, channel and total outage probability of ad hoc networks and study the effect of the system parameters on the outage performance for the special case of $\alpha = 4$.

5.2.1 Ergodic Capacity of CUs

In this section, we will first introduce Lemma 4 to numerically calculate the ergodic capacity of CUs for a general case. Then we present closed-form results of cellular capacity for a practical case of path loss exponent $\alpha = 4$.

Lemma 4. *The ergodic capacity of cellular networks in a shared channel is given by:*

$$C_C = \int_0^\infty \frac{dt}{\rho(e^t - 1, \alpha) + \frac{\lambda_T}{\lambda_C} C(\alpha) \left(\frac{P_T}{P_C}\right)^{\frac{2}{\alpha}} (e^t - 1)^{\frac{2}{\alpha}} + 1}. \quad (5.4)$$

Proof. See [72]. □

From (5.4) we can observe that the ergodic capacity of CUs in a shared channel decreases with the increase of λ_T/λ_C and P_T/P_C . This is due to stronger interference from DTUs.

When $\alpha = 4$, we derive both upper and lower bounds for ergodic capacity of CUs in closed form as given in Theorem 2:

Theorem 2. *For the special case of $\alpha = 4$, the cellular ergodic capacity in a shared channel is bounded by:*

$$C_C^L < C_C < C_C^U, \quad (5.5)$$

where

$$C_C^U = \frac{\sqrt[3]{3\zeta}}{3\zeta} \left\{ \log \frac{(\frac{\pi}{2} + \sqrt[3]{3\zeta})^2}{\frac{\pi^2}{4} - \frac{\pi}{2}\sqrt[3]{3\zeta} + (3\zeta)^{2/3}} + 2\sqrt{3} \arctan\left(\frac{\frac{\pi}{2}\sqrt{3}}{2\sqrt[3]{3\zeta} - \frac{\pi}{2}}\right) \right\}, \quad (5.6)$$

$$C_C^L = \frac{2}{\zeta} \arctan \zeta + \left(\frac{\arctan \zeta}{\zeta}\right)^2 - \frac{2}{\zeta^2} \log \sqrt{1 + \zeta^2} + \log \frac{\sqrt{1 + \zeta^2}}{\zeta}, \quad (5.7)$$

and $\zeta = \frac{\pi^2\gamma}{4-2\pi\gamma} + \frac{\pi}{2}$, $\gamma = \frac{2}{\pi(1+\frac{\lambda_C}{\lambda_T}(\frac{P_C}{P_T})^{1/2})}$.

Proof. The ergodic capacity $C_C = \int_{t=0}^{\infty} \mathbb{P}[\xi > (e^t - 1)]dt$. Therefore, the ergodic capacity can be evaluated using the result in Lemma 2 with θ_C replaced by $e^t - 1$.

$$C_C = \int_{t=0}^{\infty} P_{s,C}(e^t - 1)dt = \int_0^{\infty} \frac{dt}{\rho(e^t - 1, \alpha) + C(\alpha)(e^t - 1)^{\frac{2}{\alpha}} \frac{\lambda_T}{\lambda_C} (\frac{P_T}{P_C})^{\frac{2}{\alpha}} + 1}. \quad (5.8)$$

Let $\gamma = \frac{1}{C(\alpha)(\frac{\lambda_C}{\lambda_T}(\frac{P_C}{P_T})^{\frac{2}{\alpha}} + 1)}$, $\zeta = \frac{C^2(\alpha)\gamma}{1-C(\alpha)\gamma} + \frac{\pi}{2}$ and $x = \arctan((e^t - 1)^{-\frac{1}{2}})$. When $\alpha = 4$,

$$\begin{aligned} C_C &= \int_0^{\infty} \frac{dt}{(e^t - 1)^{\frac{1}{2}} \int_{\frac{1}{(e^t-1)^{\frac{1}{2}}} }^{\infty} \frac{du}{1+u^2} + \frac{C(4)}{C(4)\gamma-1} (e^t - 1)^{\frac{1}{2}} + 1} \\ &= \int_0^{\frac{\pi}{2}} \frac{dt}{\zeta(e^t - 1)^{\frac{1}{2}} - (e^t - 1)^{\frac{1}{2}} \arctan[(e^t - 1)^{-\frac{1}{2}}] + 1} \\ &= \int_0^{\frac{\pi}{2}} \frac{2 \sec^2 x dx}{(\tan^2 x + 1)(\tan x - x + \zeta)} = \int_0^{\frac{\pi}{2}} \frac{2dx}{\tan x - x + \zeta}. \end{aligned} \quad (5.9)$$

Substituting for $\tan x > x + \frac{x^3}{3}$, we can get an upper bound for C_C as

$$C_C < 6 \int_0^{\frac{\pi}{2}} \frac{dx}{x^3 + 3\zeta} = \frac{\sqrt[3]{3\zeta}}{3\zeta} \left\{ \log \frac{(\frac{\pi}{2} + \sqrt[3]{3\zeta})^2}{\frac{\pi^2}{4} - \frac{\pi}{2}\sqrt[3]{3\zeta} + (3\zeta)^{2/3}} + 2\sqrt{3} \arctan\left(\frac{\frac{\pi}{2}\sqrt{3}}{2\sqrt[3]{3\zeta} - \frac{\pi}{2}}\right) \right\}. \quad (5.10)$$

On the other hand, when $0 < \gamma < \frac{2}{\pi}$, $\frac{\pi}{2} < \zeta < \infty$, we can conclude that

$$\begin{aligned}
C_C &= \int_0^{\frac{\pi}{2}} \frac{2dx}{\tan x - x + \zeta} \\
&> \int_0^{\arctan \zeta} \frac{2dx}{\tan x - x + \zeta} + \int_{\arctan \zeta}^{\pi/2} \frac{2dx}{2 \tan x} \\
&= \int_0^{\arctan \zeta} \frac{2/\zeta}{1 + \frac{\tan x - x}{\zeta}} dx + \int_{\arctan \zeta}^{\pi/2} \cot x dx \\
&\stackrel{(a)}{>} \frac{2}{\zeta} \int_0^{\arctan \zeta} \left(1 - \frac{\tan x - x}{\zeta}\right) dx + \int_{\arctan \zeta}^{\pi/2} \cot x dx \\
&= \frac{2}{\zeta} \arctan \zeta - \frac{2}{\zeta^2} \int_0^{\arctan \zeta} (\tan x - x) dx + \int_{\arctan \zeta}^{\pi/2} \cot x dx \\
&= \frac{2}{\zeta} \arctan \zeta - \frac{2}{\zeta^2} \left(-\log \cos x - \frac{x^2}{2}\right) \Big|_0^{\arctan \zeta} + \log \sin x \Big|_{\arctan \zeta}^{\pi/2} \\
&= \frac{2}{\zeta} \arctan \zeta - \frac{2}{\zeta^2} \left[-\log \cos \arctan \zeta - \frac{(\arctan \zeta)^2}{2}\right] - \log \sin \arctan \zeta \\
&= \frac{2}{\zeta} \arctan \zeta + \left(\frac{\arctan \zeta}{\zeta}\right)^2 - \frac{2}{\zeta^2} \log \sqrt{1 + \zeta^2} + \log \frac{\sqrt{1 + \zeta^2}}{\zeta}.
\end{aligned} \tag{5.11}$$

To achieve the inequality (a), note that when $0 < x < \arctan \zeta$, $0 < \frac{\tan x - x}{\zeta} < \frac{\zeta - \arctan \zeta}{\zeta} < 1$, then $\frac{1}{1 + \frac{\tan x - x}{\zeta}} > 1 - \frac{\tan x - x}{\zeta}$. \square

It is important to understand the effect of γ ($\gamma \in [0, \frac{2}{\pi}]$) on ergodic capacity. When $\gamma \rightarrow 0$, $\lambda_C \sqrt{P_C} \gg \lambda_T \sqrt{P_T}$, cellular interference dominates the D2D performance. On the other hand, when $\gamma \rightarrow \frac{2}{\pi}$, $\lambda_C \sqrt{P_C} \ll \lambda_T \sqrt{P_T}$, D2D interference dominates cellular user performance. For the case of $\gamma \rightarrow 0$, as D2D interference is negligible, the network can be considered as a single tier cellular network. Therefore, C_C corresponding to (5.8) simplifies to a form independent of the cellular BS density. This observation is consistent with [96] which has a similar result for a single tier network with nearest neighbor association.

5.2.2 Ergodic Capacity of DUs

The main drawback of [72] is that the distance between a D2D pair is assumed to be a constant. In this section, D2D ergodic capacity is evaluated based on $P_{s,D}(\theta_D|X_D)$ assuming D2D pair distance to be a random quantity. We will first present Lemma 5 to numerically calculate the ergodic capacity of DUs for a general case. Then we provide closed-form results of D2D capacity for a practical case $\alpha = 4$.

Lemma 5. *The ergodic capacity of D2D networks in a shared channel is given by:*

$$C_D = \sum_{k=1}^{\infty} \sum_{m=0}^{k-1} \sum_{j=1}^{k-m} (-1)^{k-m} J(m, k) \beta_j^{k-m} \frac{\pi \lambda_T \alpha}{2} I_j + \sum_{k=0}^{\infty} J(k, k) \frac{\alpha \pi \lambda_T}{2} I_0, \quad (5.12)$$

where $I_j = \int_0^{\infty} \frac{t^{j+\frac{\alpha}{2}-1} dt}{(\lambda_T \pi + t)^{j+1} (A^{\frac{\alpha}{2}} + t^{\frac{\alpha}{2}})}, j = 0, 1, 2, \dots$ and $A = \pi C(\alpha) (\lambda_C (\frac{P_C}{P_T})^{2/\alpha} + \lambda_T)$.

Proof. From [72], the ergodic capacity for given X_D can be presented as:

$$R_d(X_D) = \int_{z=0}^{\infty} \frac{P_{s,D}(z|X_D)}{1+z} dz. \quad (5.13)$$

The ergodic capacity averaged over X_D can be evaluated as

$$C_D = \int_0^{\infty} R_d(r) f_{X_D}(r) dr = \int_0^{\infty} \frac{[\int_0^{\infty} f_{X_D}(r) P_{s,D}(z|r) dr]}{1+z} dz. \quad (5.14)$$

Let $A = \pi C(\alpha) (\lambda_C (\frac{P_C}{P_T})^{2/\alpha} + \lambda_T)$ and $t = Az^{2/\alpha}$ for simplicity, then $C_d z^{2/\alpha} = tr^2$,

$$C_D = \sum_{k=1}^{\infty} \sum_{m=0}^{k-1} \sum_{j=1}^{k-m} (-1)^{k-m} J(m, k) \frac{\beta_j^{k-m}}{j!} \int_0^{\infty} [\int_0^{\infty} \frac{r^{2j} f_{X_D}(r)}{e^{tr^2}} dr] \cdot \frac{t^{j\frac{\alpha}{2}} t^{\frac{\alpha}{2}-1} dt}{A^{\frac{\alpha}{2}} + t^{\frac{\alpha}{2}}} \\ + \sum_{k=0}^{\infty} J(k, k) \int_0^{\infty} [\int_0^{\infty} \frac{f_{X_D}(r)}{e^{tr^2}} dr] \frac{\frac{\alpha}{2} t^{\frac{\alpha}{2}-1} dt}{A^{\frac{\alpha}{2}} + t^{\frac{\alpha}{2}}}. \quad (5.15)$$

The integration over r can be calculated as $\int_0^\infty e^{-tr^2} r^{2j} f_{X_D}(r) dr = \frac{\pi \lambda_T j!}{(\lambda_T \pi + t)^{j+1}}$. Thus,

$$C_D = \sum_{k=1}^{\infty} \sum_{m=0}^{k-1} \sum_{j=1}^{k-m} (-1)^{k-m} J(m, k) \beta_j^{k-m} \frac{\pi \lambda_T \alpha}{2} I_j + \sum_{k=0}^{\infty} J(k, k) \frac{\alpha \pi \lambda_T}{2} I_0, \quad (5.16)$$

where $I_j = \int_0^\infty \frac{t^{j+\frac{\alpha}{2}-1} dt}{(\lambda_T \pi + t)^{j+1} (A \frac{\alpha}{2} + t \frac{\alpha}{2})}$, $j = 0, 1, 2, \dots$ \square

From (5.12) we can observe that the ergodic capacity of CUs in a shared channel decreases with the increase of λ_C and P_C/P_T , which is due to stronger BS interference. However, the relationship between λ_T and C_D is nontrivial.

When $\alpha = 4$, the D2D ergodic capacity expression is derived in Theorem 3 in terms of B_j , which can be calculated by a recursive expression provided in Lemma 6.

Lemma 6. Define $B_j(\gamma) = \int_0^{\frac{\pi}{2}} \frac{d\theta}{(1+\gamma \cot \theta)^{j+1}}$. Then $B_j(\gamma)$ can be represented in a recursive form as

$$B_j(\gamma) = \frac{2B_{j-1}(\gamma) - B_{j-2}(\gamma) + \gamma/j}{1 + \gamma^2}, j \geq 2. \quad (5.17)$$

where $B_0(\gamma) = \frac{\pi}{2(1+\gamma^2)} + \frac{\gamma}{1+\gamma^2} \log \gamma$ and $B_1(\gamma) = \frac{\pi(1-\gamma^2)}{2(1+\gamma^2)^2} + \frac{2\gamma}{(1+\gamma^2)^2} \log \gamma + \frac{\gamma}{1+\gamma^2}$.

Proof. See Appendix A.1. \square

Theorem 3. For the case of $\alpha = 4$, the D2D ergodic capacity in a shared channel corresponds to:

$$C_D = \sum_{k=1}^{\infty} \sum_{m=0}^{k-1} \sum_{j=1}^{k-m} 2(-1)^{k-m} J(m, k) \beta_j^{k-m} \gamma B_j(\gamma) + \sum_{k=0}^{\infty} 2J(k, k) \gamma B_0(\gamma), \quad (5.18)$$

where, γ is given in Theorem 2 and a recursive expression of $B_j(\gamma)$ is provided in Lemma 6.

Proof. When $\alpha = 4$, from (5.12), let $t = A \tan \theta$,

$$I_j = \frac{1}{A} \int_0^{\frac{\pi}{2}} \frac{1}{[1 + \frac{\pi \lambda_T}{A} \cot \theta]^{j+1}} d\theta = \frac{1}{A} B_j\left(\frac{\pi \lambda_T}{A}\right). \quad (5.19)$$

Note that $\frac{\pi\lambda_T}{A} = \frac{2}{\pi(1+\frac{\lambda_C}{\lambda_T}(\frac{P_C}{P_T})^{1/2})} = \gamma$. When $\alpha = 4$,

$$C_D = \sum_{k=1}^{\infty} \sum_{m=0}^{k-1} \sum_{j=1}^{k-m} 2(-1)^{k-m} J(m, k) \beta_j^{k-m} \gamma B_j(\gamma) + \sum_{k=0}^{\infty} 2J(k, k) \gamma B_0(\gamma). \quad (5.20)$$

□

Corollary 2. *For the case of $\alpha = 4$, the D2D ergodic capacity in a shared channel can be approximated as:*

$$\hat{C}_D = 2\gamma B_0(\gamma) + \gamma B_1(\gamma)(1 - e^{-K_R}), \quad (5.21)$$

Proof. Since

$$J(k, k) = e^{-K_R} K_R^k k! \binom{k}{k} / (k!)^2 = e^{-K_R} \frac{K_R^k}{k!}. \quad (5.22)$$

the second term in (5.20) is

$$\sum_{k=0}^{\infty} 2\gamma B_0(\gamma) J(k, k) = 2\gamma B_0(\gamma) \sum_{k=0}^{\infty} e^{-K_R} \frac{K_R^k}{k!} = 2\gamma B_0(\gamma). \quad (5.23)$$

The first term in (5.20) can be rewritten as

$$\begin{aligned} & \sum_{k=1}^{\infty} \sum_{n=1}^k \sum_{j=1}^n 2\gamma (-1)^n J(k-n, k) \beta_j^n B_j(\gamma) \\ &= \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} \sum_{j=1}^n 2\gamma (-1)^n J(k-n, k) \beta_j^n B_j(\gamma) \\ &= 2\gamma (-1)^1 \beta_1^1 B_1(\gamma) \left[\sum_{k=1}^{\infty} J(k-1, k) \right] + \sum_{n=2}^{\infty} 2\gamma (-1)^n \left[\sum_{j=1}^n \beta_j^n B_j(\gamma) \right] \left[\sum_{k=n}^{\infty} J(k-n, k) \right] \\ &\approx 2\gamma (-1) \beta_1 B_1(\gamma) \left[\sum_{k=1}^{\infty} J(k-1, k) \right]. \end{aligned} \quad (5.24)$$

Since

$$\beta_1 = (-1) \left(\frac{1}{2} \right) = -\frac{1}{2}, \quad (5.25)$$

$$\sum_{k=1}^{\infty} J(k-1, k) = \sum_{k=1}^{\infty} e^{-K_R} K_R^k (k-1)! \binom{k}{k-1} / (k!)^2 = \sum_{k=1}^{\infty} e^{-K_R} \frac{K_R^k}{k!} = 1 - e^{-K_R}, \quad (5.26)$$

the ergodic capacity can be approximated as

$$C_D \approx \hat{C}_D = 2\gamma B_0(\gamma) + \gamma B_1(\gamma)(1 - e^{-K_R}). \quad (5.27)$$

□

Once again, it is interesting to interpret the result in (5.21) for the cases when $\gamma \rightarrow \frac{2}{\pi}$ and $\gamma \rightarrow 0$. When $\gamma \rightarrow \frac{2}{\pi}$, the network can be considered as a single tier D2D network as cellular interference becomes negligible. Therefore, C_D corresponding to (5.21) simplifies to a form independent of the DU density, consistent with [96] which has a similar result for a single tier network with nearest neighbor association.

Corollary 3. *For the case of $\gamma \rightarrow 0$, i.e., the cellular interference dominates the D2D performance, $C_D \rightarrow 0$, specifically, $C_D = o(\gamma)$.*

Proof. When $\gamma \rightarrow 0$,

$$B_0(\gamma) \approx \frac{\pi}{2}(1 - \gamma^2) + \gamma \log \gamma(1 - \gamma^2), \quad (5.28)$$

$$B_1(\gamma) \approx \frac{\pi}{2}(1 - 3\gamma^2) + 2\gamma \log \gamma(1 - 2\gamma^2) + \gamma(1 - \gamma^2). \quad (5.29)$$

Then according to Corollary 2,

$$\begin{aligned} C_D &\approx 2\gamma(1 - \gamma^2)\frac{\pi}{2} + 2\gamma^2 \log \gamma(1 - \gamma^2) + (1 - e^{-K_R})\left[\frac{\pi}{2}\gamma(1 - 3\gamma^2) + 2\gamma^2 \log \gamma(1 - 2\gamma^2) + \gamma^2(1 - \gamma^2)\right] \\ &\approx (3 - e^{-K_R})\frac{\pi}{2}\gamma + (1 - e^{-K_R})\gamma^2 + 2(2 - e^{-K_R})\gamma^2 \log \gamma \\ &= o(\gamma). \end{aligned} \quad (5.30)$$

□

5.2.3 Simulation Results

In this subsection, we validate the derived analytical results via simulations. The primary simulation parameters used in this and later sections are listed in Table 5.1 [72, 125].

Table 5.1: *Simulation parameters*

Simulation Parameter	Value
Transmit power of cellular BSs (P_C)	20 W
Maximum transmit power of cellular BSs (P_{CMax})	40 W
Minimum transmit power of cellular BSs (P_{CMin})	10 W
Transmit power of DTUs (P_T)	500 mW
Maximum transmit power of DTUs (P_{TMax})	1 W
Minimum transmit power of DTUs (P_{TMin})	1 mW
Density of cellular BSs (λ_C)	$5 \times 10^{-6} m^{-2}$
Density of DTUs (λ_T)	$1 \times 10^{-4} m^{-2}$
Maximum density of DTUs (λ_{TMax})	$1 \times 10^{-2} m^{-2}$
Minimum density of DTUs (λ_{TMin})	$5 \times 10^{-6} m^{-2}$
Rician factor (K_R)	10 dB
D2D link distance using method in [72] (X_D)	40 m

In Fig. 5.1 and Fig. 5.2, it is shown that the ergodic capacity of a cellular network in a shared channel C_C increases with P_C or λ_C (in other words, decreases with γ since γ corresponds to $\frac{2}{\pi(\frac{\lambda_C}{\lambda_T}(\frac{P_C}{P_T})^{\frac{1}{2}}+1)}$). The fact that C_C increases with P_C or λ_C is a direct result of stronger BS signal compared with interference from DTUs. Furthermore, C_C is shown to be well bounded by its upper bound C_C^U and lower bound C_C^L . These bounds get tighter as P_C or λ_C decreases.

In Fig. 5.3 and Fig. 5.4, it is shown that the ergodic capacity of a D2D network in a shared channel C_D increases with P_T or λ_T . In other words, C_D increases with γ since γ corresponds to $\frac{2}{\pi(\frac{\lambda_C}{\lambda_T}(\frac{P_C}{P_T})^{\frac{1}{2}}+1)}$. The fact that C_D increases with P_T or λ_T is a direct result of stronger DTU signals compared with interference from BSs. Furthermore, C_D is shown to be well approximated by \hat{C}_D . For comparison, we show the ergodic capacity of D2D communication C_D obtained by the results in [72] fixing D2D pair distance X_D as the

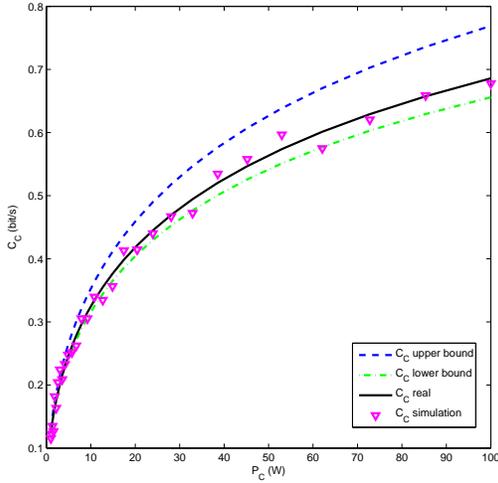


Figure 5.1: C_C vs. P_C

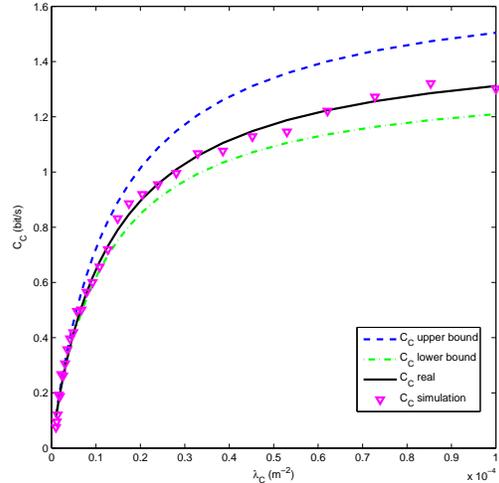


Figure 5.2: C_C vs. λ_C

expected value according to its distribution. The result derived from our model is more useful than [72], because in reality the distance between D2D pairs is expected to vary. When X_D and λ_T are fixed, and P_T increases, DTU signals are stronger compared with BS signals, which leads to an increase of C_D as shown in Fig. 5.3. However, when X_D and P_T are fixed, C_D decreases with λ_T as shown in Fig. 5.4. This behavior can be attributed to the fact that the signal strength of the direct D2D link does not change while the interference from other DTUs grows as λ_T increases.

5.3 Optimization of Ergodic Capacity and Power Consumption

In this section, we present a two-stage approach to maximize the ergodic capacity and minimize the total power consumption. Ergodic capacity in a subchannel of a D2D underlay cellular network is defined as $C \triangleq C_C + C_D$. As is shown in Fig. 5.1-Fig. 5.4, C_C can be well estimated by its lower bound $\hat{C}_C = C_C^L$ and C_D can be well approximated by \hat{C}_D , so we maximize the approximated ergodic capacity $\hat{C} = \hat{C}_C + \hat{C}_D$ instead. In the first stage, we

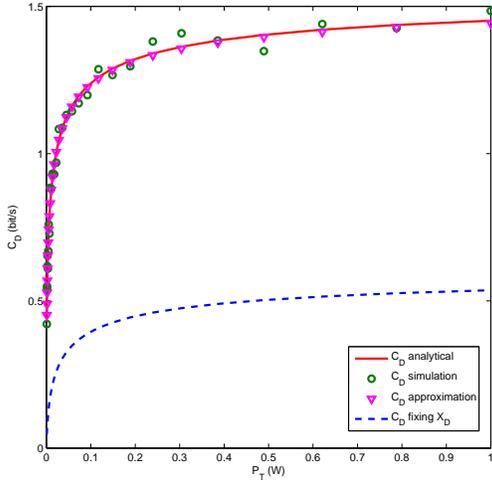


Figure 5.3: C_D vs. P_T

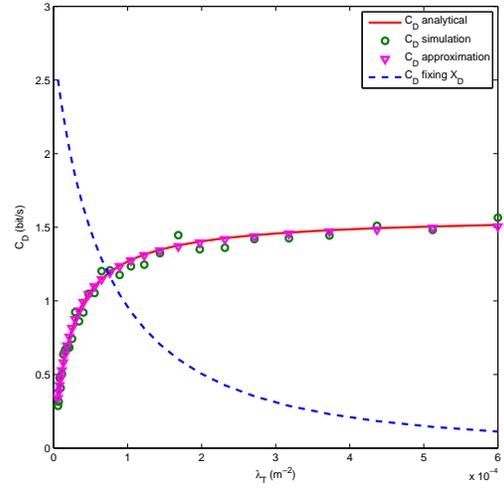


Figure 5.4: C_D vs. λ_T

aim to maximize the ergodic capacity by adjusting DTU transmit power P_T or DTU density λ_T . In other words, we try to investigate research questions such as how should DTUs adjust their transmit power to improve \hat{C} while achieving certain power consumption constraint, or what would be the optimal density of DTU λ_T when \hat{C} is maximized. Furthermore, in the second stage, we aim to minimize the total power consumption of the network P_{total} while this approach provides the best balance guaranteeing the optimal ergodic capacity \hat{C}^* .

5.3.1 Stage-1: Ergodic Capacity Optimization

In this subsection, considering a total power consumption constraint, we aim to maximize the ergodic capacity \hat{C} as:

$$\max_{\gamma} \hat{C} = \hat{C}_C + \hat{C}_D \quad (5.31)$$

subject to,

$$P_{TMin} \leq P_T \leq P_{TMax} \quad (5.32)$$

$$\lambda_C P_C + \lambda_T P_T \leq P_{total} \quad (5.33)$$

Theorem 4. When $\alpha = 4$, the approximated cellular and D2D ergodic capacity in a shared channel \hat{C}_C and \hat{C}_D are concave functions of γ , and therefore the total ergodic capacity in a shared channel $\hat{C} = \hat{C}_C + \hat{C}_D$ is a concave function of γ .

Proof. To prove concavity of \hat{C}_C we prove the second order derivative $\frac{d^2\hat{C}_C}{d\gamma^2} < 0$. We first derive $\frac{d^2\hat{C}_C}{d\gamma^2}$ and then find an upper bound of it by replacing the arctan and log function by segments. Since this upper bound takes a polynomial form, it is easy to find roots and to show this upper bound is negative over the parameter region we consider. Similarly we can prove $\frac{d^2\hat{C}_D}{d\gamma^2} < 0$ by replacing the log function by segments and obtain an upper bound of $\frac{d^2\hat{C}_D}{d\gamma^2}$ that is negative over the parameter region. Detailed proof is included in appendix A.2 and not presented here due to tedious calculations involved. \square

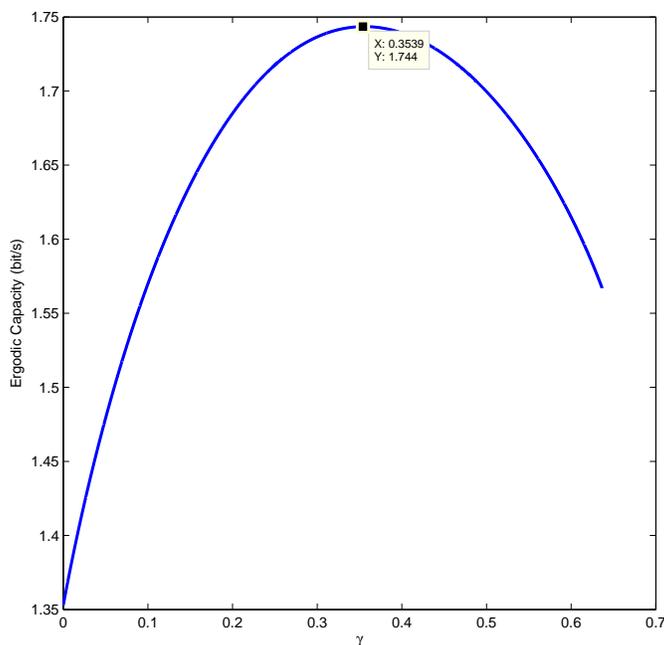


Figure 5.5: γ vs. \hat{C}

If we do not consider the constraints (5.32) and (5.33), then according to Theorem 4, we can find the optimal solution $\gamma^* = 0.3539$ and $\hat{C}^* = 1.744$ by efficient convex optimization

algorithms [105], as is shown in Fig. 5.5. Based on the optimal γ^* , we can solve the corresponding transmit power P_T^* when λ_C , λ_T and P_C are fixed. However the parameter range of P_T is confined by (5.32) and (5.33) such that P_T^* may not always be achievable. If P_T^* lies outside the available P_T range, \hat{C} will monotonically increase or decrease with P_T in this range resulting in the optimal P_T at the boundary. This optimization procedure enables the DTUs to adjust their transmit power by taking into account the expected interference to and from other users. Thus this approach (1) avoids excessive power consumption and interference, and (2) maximizes the ergodic capacity of a D2D underlay cellular network.

We can find the optimal DTU density λ_T for maximum total ergodic capacity \hat{C} when λ_C , P_C , P_T are fixed and consider constraint (5.33) imposed by P_{total} . Fig. 5.6 shows the impact of DTU density λ_T on the ergodic capacity \hat{C} with parameters $\lambda_C = 5 \times 10^{-6}$, $P_C = 20W$, $P_T = 0.5W$ and $P_{total} = 1W$. We can observe that as λ_T increases, \hat{C} improves at first, then has a diminishing return and stays as a constant. This behavior is expected as when DTU density λ_T increases from $1 \times 10^{-5}m^{-2}$, ergodic capacity of D2D network is much enhanced because of stronger signal of DTUs, while cellular network is not much affected due to sparsity of DTUs. After λ_T becomes larger than $1 \times 10^{-4}m^{-2}$, the interference from D2D network to cellular network becomes more significant so that the total ergodic capacity decreases. Then when λ_T is larger than $1 \times 10^{-1}m^{-2}$, this network can be considered as a single tier D2D network, and its ergodic capacity becomes independent of λ_T .

5.3.2 Stage-2: Optimization of Ergodic Capacity and Power Consumption

Since the density of cellular BSs can be considered to be fixed in the short term while that of DTUs varies with time, we consider λ_C as a constant and λ_T as a variable. P_C and P_T are also assumed to be variables since the transmit power of both cellular BSs and DTUs can be adjusted. We aim to find the best situation that the total power consumption of the network P_{total} is minimized while the optimal ergodic capacity \hat{C}^* is achieved. This

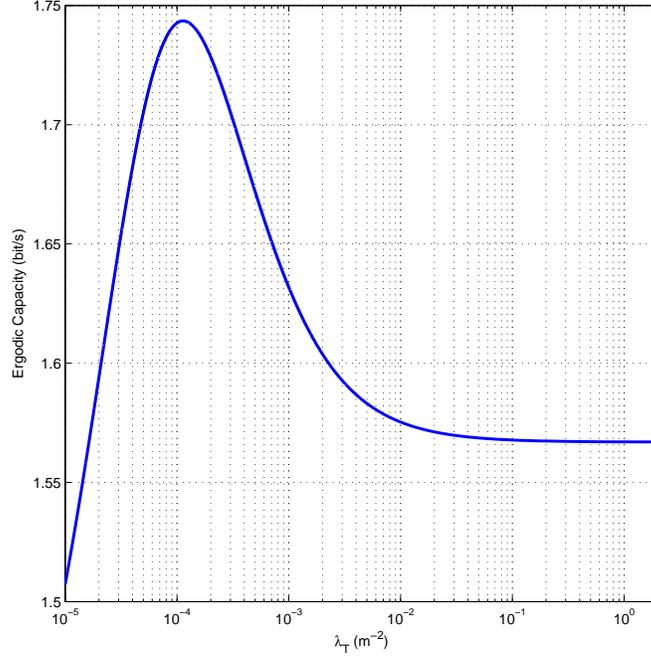


Figure 5.6: *DTU density λ_T vs. approximated ergodic capacity \hat{C}*

optimization problem can be presented in (5.34) as follows.

$$\min_{\lambda_T, P_C, P_T} P_{total} = \lambda_C P_C + \lambda_T P_T, \quad (5.34)$$

subject to,

$$\hat{C} = \hat{C}^*, \quad (5.35)$$

$$P_{TMin} \leq P_T \leq P_{TMax}, \quad (5.36)$$

$$P_{CMin} \leq P_C \leq P_{CMax}, \quad (5.37)$$

$$0 \leq \lambda_T \leq \lambda_{TMax}. \quad (5.38)$$

Here, P_{TMin} and P_{TMax} denote the minimum and maximum transmit power of DTUs, respectively. P_{CMin} and P_{CMax} denote the minimum and maximum transmit power of

cellular BSs, respectively. And λ_{TMax} denotes the maximum density of DTUs. Typically $P_{TMin} < P_{CMin}$ and $P_{TMax} < P_{CMax}$.

Constraint (5.35) is equivalent to $\frac{\lambda_C}{\lambda_T} \left(\frac{P_C}{P_T}\right)^{1/2} = x_{opt} = \frac{2}{\pi\gamma^*} - 1 = 0.7999$. So from (5.35) we can obtain $\frac{P_C}{P_T} = \frac{x_{opt}^2}{\lambda_C^2} \lambda_T^2$. Then from (5.38),

$$0 < \frac{P_C}{P_T} < \frac{x_{opt}^2}{\lambda_C^2} \lambda_{TMax}^2 \quad (5.39)$$

Constraint (5.36), (5.37) and (5.39) determine the area of (P_C, P_T) that can achieve \hat{C}^* . According to the value of P_{CMin} , P_{CMax} , P_{TMin} , P_{TMax} and λ_{TMax} , there are three possibilities listed as below.

- (a) Case-1: $\frac{P_{CMin}}{P_{TMax}} < \frac{x_{opt}^2}{\lambda_C^2} \lambda_{TMax}^2 < \frac{P_{CMin}}{P_{TMin}}$.

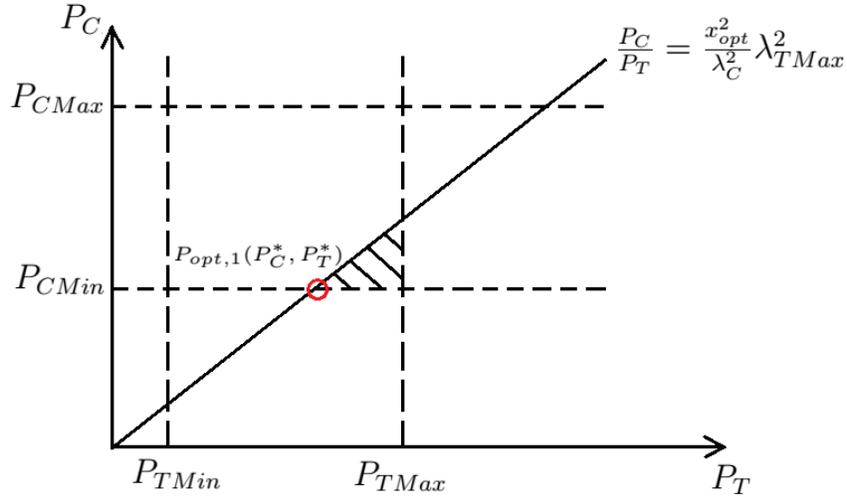


Figure 5.7: Achievable Area in Case 1

For point (P_C, P_T) in the shadowed area, according to (5.35),

$$\lambda_T = \frac{\lambda_C}{x_{opt}} \left(\frac{P_C}{P_T}\right)^{1/2}, \quad (5.40)$$

$$\begin{aligned}
P_{total} &= \lambda_C P_C + \lambda_T P_T \\
&= \lambda_C P_C + \frac{\lambda_C}{x_{opt}} \left(\frac{P_C}{P_T}\right)^{1/2} P_T \\
&= \lambda_C \left(P_C + \frac{1}{x_{opt}} \sqrt{P_C P_T}\right).
\end{aligned} \tag{5.41}$$

Obviously P_{total} increases with P_C or P_T . As is shown in Fig. 5.7, the optimal choice of (P_C, P_T) is point $P_{opt,1}$ in the lower left corner of this shadowed area. That is, $\lambda_T = \lambda_{TMax}$, $P_C = P_{CMin}$ and $P_T = \frac{P_{CMin}\lambda_C^2}{x_{opt}^2\lambda_{TMax}^2}$.

(b) Case-2: $\frac{x_{opt}^2}{\lambda_C^2}\lambda_{TMax}^2 > \frac{P_{CMin}}{P_{TMin}}$.

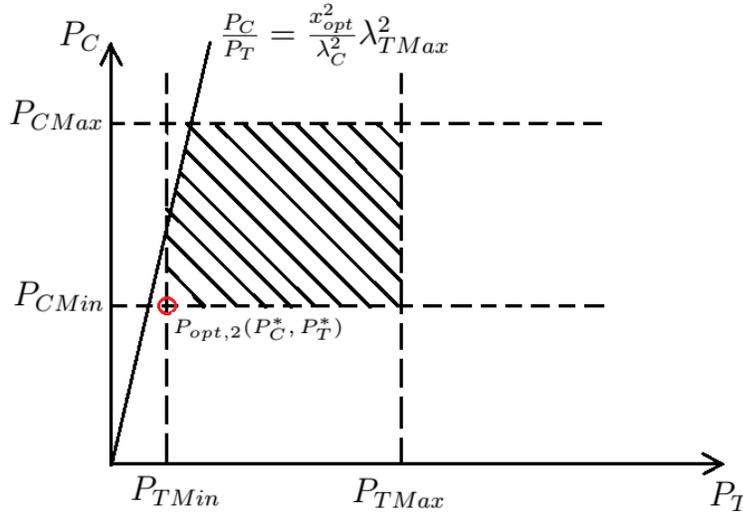


Figure 5.8: Achievable Area in Case 2

In the shadowed area shown in Fig. 5.8, (5.40) and (5.41) also holds. Therefore the optimal point (P_C, P_T) is point $P_{opt,2}$ in the lower left corner of this shadowed area as shown in Fig. 5.8, namely $P_C = P_{CMin}$, $P_T = P_{TMin}$ and $\lambda_T = \frac{\lambda_C}{x_{opt}} \left(\frac{P_{CMin}}{P_{TMin}}\right)^{1/2}$.

(c) Case-3: $\frac{x_{opt}^2}{\lambda_C^2}\lambda_{TMax}^2 < \frac{P_{CMin}}{P_{TMax}}$.

In this case, R^* can not be achieved as is shown in Fig. 5.9. Our goal has to be changed to find the maximum achievable \hat{C} . Note that in the square area, $\frac{P_C}{P_T} > \frac{x_{opt}^2}{\lambda_C^2}\lambda_{TMax}^2$, then $x_{opt} < \frac{\lambda_C}{\lambda_{TMax}} \left(\frac{P_C}{P_T}\right)^{1/2} < \frac{\lambda_C}{\lambda_T} \left(\frac{P_C}{P_T}\right)^{1/2}$, i.e., $\gamma < \gamma^*$. In this case from Fig. 5.5, the

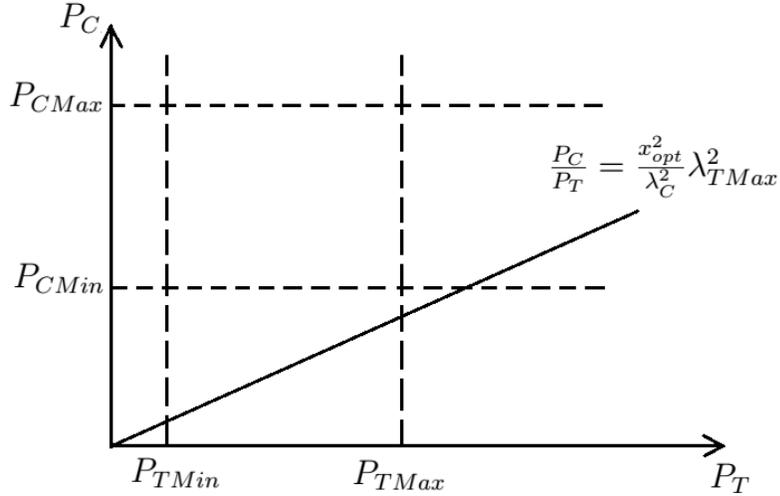


Figure 5.9: *Achievable Area in Case 3*

maximum achievable \hat{C} is determined by the maximum achievable γ , which is equivalent to the minimum achievable $\frac{\lambda_C}{\lambda_T} (\frac{P_C}{P_T})^{1/2}$. Therefore we get $\lambda_T = \lambda_{TMax}$, $P_C = P_{CMin}$ and $P_T = P_{TMax}$.

These results can be verified by comparing both the ergodic capacity and the total power consumption in the optimal case with other cases as shown in Fig. 5.10. We set simulation parameter values as listed in Table 5.1. Since $\frac{P_{CMin}}{P_{TMax}} = 10 < \frac{P_{CMin}}{P_{TMin}} = 10000 < \frac{x_{opt}^2}{\lambda_C^2} \lambda_{TMax}^2 = 2559360$, this scenario meets the condition of case-2 and the optimal case $(P_C, P_T, \lambda_T) = (P_{CMin}, P_{TMin}, \frac{\lambda_C}{x_{opt}} (\frac{P_{CMin}}{P_{TMin}})^{1/2})$. We can see that the optimal case has the highest ergodic capacity with only 0.49% of the maximum power consumption (Here, maximum power consumption corresponds to the case of maximum DTU density and all BSs and DTUs transmitting at their maximum transmit power). By comparison of the optimal case with the lowest power consumption scenario (lowest power consumption as happened to the case of minimum DTU density and all BSs and DTUs transmitting at their minimum transmit power), we can see that the optimal case has 27.11% higher ergodic capacity but only 0.07% higher power consumption. This demonstrates the potency of the proposed approach.

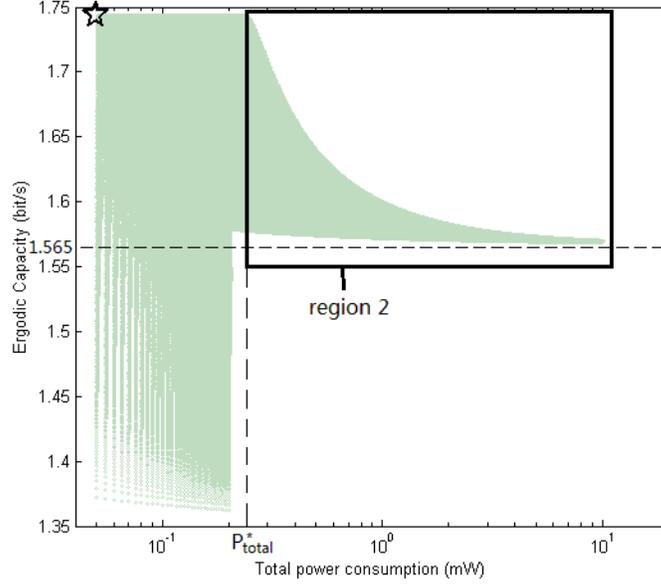


Figure 5.10: Ergodic capacity and total power consumption for different choices of parameters (P_C, P_T, λ_T) . The black pentagram represents the optimal case while green circles represent other cases.

Note that for a given (P_C, P_T) , the optimal ergodic capacity C^* can be obtained when $\lambda_T = \lambda_T^* = \frac{\lambda_C}{x_{opt}} \left(\frac{P_C}{P_T} \right)^{1/2}$ and then $P_{total} = P_{total}^* = \lambda_C (P_C + \frac{1}{x_{opt}} \sqrt{P_C P_T})$. As λ_T increases such that $\lambda_T > \lambda_T^*$, $P_{total} > P_{total}^*$ and $\gamma > \gamma^*$, $1.565 < C < 1.744$. We conclude that each (P_C, P_T) has its $P_{total}^* = \lambda_C (P_C + \frac{1}{x_{opt}} \sqrt{P_C P_T})$, when $P_{total} > P_{total}^*$, $C > 1.565$. We can find an overall $\bar{P}_{total}^* = \max_{P_C, P_T} P_{total}^* = \lambda_C (P_{CMax} + \frac{1}{x_{opt}} \sqrt{R_{CMax} P_{TMax}})$ so that if $P_{total} > P_{total}^*$, it can be guaranteed that the ergodic capacity $C > 1.565$. This is illustrated in region 2 of Fig. 5.10.

5.4 Average Achievable Throughput Analysis

Definition 8. We define the average achievable throughput as $\bar{R}(\theta) = \mathbb{E}[\log(1 + \xi) | \xi > \theta] \cdot \mathbb{P}(\xi > \theta)$ in which ξ is the signal to interference ratio (SIR) and θ is the SIR threshold for successful transmission.

The average achievable throughput can be expressed in terms of the success probability $P_s(\theta)$, which is given in the following corollary:

Corollary 4. *The average achievable throughput*

$$\bar{R}(\theta) = \mathbb{P}(\xi > \theta) \log(1 + \theta) + \int_{\theta}^{\infty} \frac{P_s(\xi)}{1 + \xi} d\xi. \quad (5.42)$$

Proof. Assume the PDF and cumulative density function (CDF) of the SIR ξ are $f(\xi)$ and $F(\xi)$ respectively,

$$\begin{aligned} \bar{R}(\theta) &= \mathbb{E}[\log(1 + \xi) | \xi > \theta] \cdot P(\xi > \theta) = \int_{\theta}^{\infty} \log(1 + \xi) f(\xi) d\xi \\ &= \log(1 + \theta) P_s(\theta) + \int_{\theta}^{\infty} \frac{P_s(\xi)}{1 + \xi} d\xi. \end{aligned} \quad (5.43)$$

If we introduce $t = \log(1 + \xi)$ in the above equation, we will have another expression of $\bar{R}(\theta)$:

$$\bar{R}(\theta) = P_s(\theta) \log(1 + \theta) + \int_{\log(1+\theta)}^{\infty} P_s(e^t - 1) dt. \quad (5.44)$$

□

5.4.1 Average Achievable Throughput of CUs

Based on Lemma 2 we derive the average achievable throughput of CUs when $\alpha = 4$ as the following:

Theorem 5. *When $\alpha = 4$, the CU average achievable throughput for a given SIR threshold θ_C is*

$$\bar{R}_C(\theta_C) = \frac{\log(1 + \theta_C)}{\eta \theta_C^{1/2} - \arctan(\theta_C^{-1/2}) \theta_C^{1/2} + 1} + U(\theta_C, \eta), \quad (5.45)$$

in which the term $U(\theta_C, \eta) = \int_0^{\arctan \theta_C^{-1/2}} \frac{2dx}{\tan x - x + \eta}$. The parameter $\eta = \frac{1}{\frac{2}{\pi} - \gamma}$ and $\gamma = \frac{2}{\pi(1 + \frac{\lambda_C}{\lambda_T} (\frac{P_C}{P_T})^{1/2})}$.

Proof. From Lemma 2 when $\alpha = 4$,

$$\begin{aligned}
P_{s,C}(\theta_C) &= \frac{1}{\theta_C^{1/2} \int_{\theta_C^{-1/2}}^{\infty} \frac{du}{1+u^2} + \frac{\pi}{2} \frac{\lambda_T}{\lambda_C} \left(\frac{P_T}{P_C} \theta_C\right)^{1/2} + 1} \\
&= \frac{1}{\theta_C^{1/2} \left[\frac{\pi}{2} - \arctan(\theta_C^{-1/2})\right] + \frac{\pi}{2} \frac{\lambda_T}{\lambda_C} \left(\frac{P_T}{P_C}\right)^{1/2} \theta_C^{1/2} + 1} \\
&= \frac{1}{\eta \theta_C^{1/2} - \arctan(\theta_C^{-1/2}) \theta_C^{1/2} + 1}.
\end{aligned} \tag{5.46}$$

Here we introduce a parameter $\eta = \frac{\pi}{2} + \frac{\pi}{2} \frac{\lambda_T}{\lambda_C} \left(\frac{P_T}{P_C}\right)^{1/2}$. We can also introduce another parameter $\gamma = \frac{2}{\pi(1 + \frac{\lambda_C}{\lambda_T} (\frac{P_C}{P_T})^{1/2})}$ such that $\eta = \frac{1}{\frac{2}{\pi} - \gamma}$. Then substituting (5.46) into (5.44) we get

$$\begin{aligned}
\bar{R}_C(\theta_C) &= \frac{\log(1 + \theta_C)}{P_{s,C}(\theta_C)} + \int_{\log(1+\theta_C)}^{\infty} P_{s,C}(e^t - 1) dt \\
&= \frac{\log(1 + \theta_C)}{\eta \theta_C^{1/2} - \arctan(\theta_C^{-1/2}) \theta_C^{1/2} + 1} + U(\theta_C, \eta).
\end{aligned} \tag{5.47}$$

Where $U(\theta_C, \eta)$ is given as:

$$U = \int_{\log(1+\theta_C)}^{\infty} \frac{dt}{\eta(e^t - 1)^{1/2} - \arctan((e^t - 1)^{-1/2})(e^t - 1)^{1/2} + 1}. \tag{5.48}$$

Letting $x = \arctan((e^t - 1)^{-1/2})$, $(e^t - 1)^{1/2} = \cot x$, $e^t = \csc^2 x$, $t = 2 \log(\csc x)$, and $dt = -2 \cot x dx$, then

$$\begin{aligned}
U(\theta_C, \eta) &= \int_{\arctan \theta_C^{-\frac{1}{2}}}^0 \frac{-2 \cot x dx}{\eta \cot x - x \cot x + 1} \\
&= \int_0^{\arctan \theta_C^{-\frac{1}{2}}} \frac{2 dx}{\tan x - x + \eta}.
\end{aligned} \tag{5.49}$$

□

By approximating the term $U(\theta_C, \eta)$ in Theorem 5, we can obtain an upper bound and a lower bound of $\bar{R}_C(\theta_C)$.

Corollary 5. *The average achievable throughput of CUs is upper bounded by $\bar{R}_C^U(\theta_C)$, where*

$$\begin{aligned} \bar{R}_C^U(\theta_C) &= \frac{\log(1 + \theta_C)}{\eta\theta_C^{\frac{1}{2}} - \arctan(\theta_C^{-\frac{1}{2}})\theta_C^{\frac{1}{2}} + 1} + \left\{ \frac{\sqrt[3]{3\eta}}{3\eta} \left[\log \frac{(\arctan \theta_C^{-\frac{1}{2}} + \sqrt[3]{3\eta})^2}{(\arctan \theta_C^{-\frac{1}{2}})^2 - (\arctan \theta_C^{-\frac{1}{2}})\sqrt[3]{3\eta} + (3\eta)^{\frac{2}{3}}} \right. \right. \\ &\quad \left. \left. + 2\sqrt{3} \arctan\left(\frac{\sqrt{3} \arctan \theta_C^{-\frac{1}{2}}}{2\sqrt[3]{3\eta} - \arctan \theta_C^{-\frac{1}{2}}}\right) \right] \right\}. \end{aligned} \quad (5.50)$$

Proof. Because $\tan x > x + \frac{x^3}{3}$, then from (5.49),

$$\begin{aligned} U(\theta_C, \eta) &< \int_0^{\arctan \theta_C^{-1/2}} \frac{2dx}{\frac{x^3}{3} + \eta} = 6 \int_0^{\arctan \theta_C^{-1/2}} \frac{dx}{x^3 + 3\eta} \\ &= \frac{\sqrt[3]{3\eta}}{3\eta} \left[\log \frac{(\arctan \theta_C^{-\frac{1}{2}} + \sqrt[3]{3\eta})^2}{(\arctan \theta_C^{-\frac{1}{2}})^2 - (\arctan \theta_C^{-\frac{1}{2}})\sqrt[3]{3\eta} + (3\eta)^{\frac{2}{3}}} + 2\sqrt{3} \arctan\left(\frac{\sqrt{3} \arctan \theta_C^{-\frac{1}{2}}}{2\sqrt[3]{3\eta} - \arctan \theta_C^{-\frac{1}{2}}}\right) \right]. \end{aligned} \quad (5.51)$$

□

Corollary 6. *The average achievable throughput of CUs is lower bounded by $\bar{R}_C^L(\theta_C)$ where*

$$\begin{aligned} \bar{R}_C^L(\theta_C) &= \frac{\log(1 + \theta_C)}{\eta\theta_C^{1/2} - \arctan(\theta_C^{-1/2})\theta_C^{1/2} + 1} + \left\{ \frac{2}{\frac{\theta_C^{-\frac{1}{2}}}{2} - 1} \log \left[\frac{\frac{\theta_C^{-\frac{1}{2}}}{2} - \arctan(\frac{\theta_C^{-\frac{1}{2}}}{2}) + \eta}{\eta} \right] \right. \\ &\quad \left. + \frac{2}{\frac{\theta_C^{-\frac{1}{2}}}{2} - 1} \log \left[\frac{g(\arctan(\theta_C^{-\frac{1}{2}}))}{g(\arctan(\frac{\theta_C^{-\frac{1}{2}}}{2}))} \right] \right\}. \end{aligned} \quad (5.52)$$

where

$$g(x) = \left[\frac{\theta_C^{-1/2}/2}{\arctan(\theta_C^{-1/2}) - \arctan(\theta_C^{-1/2}/2)} - 1 \right] x + \left[1 - \frac{\arctan(\theta_C^{-1/2}/2)}{\arctan(\theta_C^{-1/2}) - \arctan(\theta_C^{-1/2}/2)} \right] \cdot (\theta_C^{-1/2}/2) + \eta. \quad (5.53)$$

Proof. Note that for $0 < x < \arctan(\theta_C^{-1/2}/2)$, $\tan x < \frac{\theta_C^{-1/2}/2}{\arctan(\theta_C^{-1/2}/2)}x$, and for $\arctan(\theta_C^{-1/2}/2) < x < \arctan \theta_C^{-1/2}$, $\tan x < \theta_C^{-1/2} + \frac{\theta_C^{-1/2}/2}{\arctan \theta_C^{-1/2} - \arctan(\theta_C^{-1/2}/2)}(x - \arctan(\theta_C^{-1/2}/2))$. Then from (5.49)

$$\begin{aligned} & \int_0^{\arctan(\theta_C^{-\frac{1}{2}})} \frac{2dx}{\tan x - x + \eta} \\ &= \int_0^{\arctan(\theta_C^{-\frac{1}{2}}/2)} \frac{2dx}{\tan x - x + \eta} + \int_{\arctan(\theta_C^{-\frac{1}{2}}/2)}^{\arctan(\theta_C^{-\frac{1}{2}})} \frac{2dx}{\tan x - x + \eta} \\ &> \int_0^{\arctan(\theta_C^{-\frac{1}{2}}/2)} \frac{2dx}{\frac{\theta_C^{-\frac{1}{2}}/2}{\arctan(\theta_C^{-\frac{1}{2}}/2)}x - x + \eta} + \int_{\arctan(\theta_C^{-\frac{1}{2}}/2)}^{\arctan(\theta_C^{-\frac{1}{2}})} \frac{2dx}{\theta_C^{-\frac{1}{2}}/2 + \frac{(\theta_C^{-\frac{1}{2}}/2)(x - \arctan(\theta_C^{-\frac{1}{2}}/2))}{\arctan(\theta_C^{-\frac{1}{2}}) - \arctan(\theta_C^{-\frac{1}{2}}/2)} - x + \eta} \\ &= \frac{2(\tan^{-1}(\theta_C^{-\frac{1}{2}}) - \tan^{-1}(\frac{\theta_C^{-\frac{1}{2}}}{2}))}{\frac{\theta_C^{-\frac{1}{2}}}{2} - \tan^{-1}(\theta_C^{-\frac{1}{2}}) + \tan^{-1}(\frac{\theta_C^{-\frac{1}{2}}}{2})} \log \left[\frac{g(\tan^{-1}(\theta_C^{-\frac{1}{2}}))}{g(\tan^{-1}(\frac{\theta_C^{-\frac{1}{2}}}{2}))} \right] + \log \left[\frac{\frac{\theta_C^{-\frac{1}{2}}}{2} - \tan^{-1}(\frac{\theta_C^{-\frac{1}{2}}}{2}) + \eta}{\eta} \right] \\ &\cdot \frac{2 \tan^{-1}(\frac{\theta_C^{-\frac{1}{2}}}{2})}{\frac{\theta_C^{-\frac{1}{2}}}{2} - \tan^{-1}(\frac{\theta_C^{-\frac{1}{2}}}{2})}. \end{aligned} \quad (5.54)$$

where

$$g(x) = \left[\frac{\theta_C^{-1/2}/2}{\arctan(\theta_C^{-1/2}) - \arctan(\theta_C^{-1/2}/2)} - 1 \right] x + \left[1 - \frac{\arctan(\theta_C^{-1/2}/2)}{\arctan(\theta_C^{-1/2}) - \arctan(\theta_C^{-1/2}/2)} \right] \cdot (\theta_C^{-1/2}/2) + \eta. \quad (5.55)$$

□

5.4.2 Average Achievable Throughput of DUs

The main drawback of [72] is that the distance between a D2D pair is assumed to be a constant. In this section, D2D average achievable throughput is evaluated based on $P_{s,D}(\theta_D|X_D)$ assuming D2D pair distance to be a random quantity. By using the results in Corollary 4, the D2D average achievable throughput is provided in Theorem 6.

Theorem 6. *When $\alpha = 4$, the D2D average achievable throughput for a given SIR threshold θ_D is*

$$\bar{R}_D(\theta_D) = \log(1 + \theta_D)W(\theta_D) + V(\theta_D). \quad (5.56)$$

Where

$$W(\theta_D) = \sum_{k=1}^{\infty} \sum_{m=0}^{k-1} \sum_{j=1}^{k-m} (-1)^{k-m} J(m, k) \beta_j^{k-m} \frac{(\frac{\theta_D^{1/2}}{\gamma})^j}{(1 + \frac{\theta_D^{1/2}}{\gamma})^{j+1}} + \sum_{k=0}^{\infty} J(k, k) \frac{1}{1 + \frac{\theta_D^{1/2}}{\gamma}}, \quad (5.57)$$

$$V(\theta_D) = \sum_{k=1}^{\infty} \sum_{m=0}^{k-1} \sum_{j=1}^{k-m} (-1)^{k-m} J(m, k) \beta_j^{k-m} A_j(\theta_D) + \sum_{k=0}^{\infty} J(k, k) A_0(\theta_D), \quad (5.58)$$

and

$$A_j(\theta_D) = \int_0^{\frac{\gamma}{\theta_D^{1/2}}} \frac{2\gamma^2}{(1+t)^{j+1} (t^2 + \gamma^2)} dt. \quad (5.59)$$

The parameter γ is the same as in Theorem 5.

Proof.

$$\begin{aligned}
\bar{R}_D(\theta_D) &= \int_0^\infty f_{X_D}(r) \{P_{s,D}(\theta_D|r) \log(1 + \theta_D) + \int_{\theta_D}^\infty \frac{P_{s,D}(\xi|r)}{1 + \xi} d\xi\} dr \\
&= \log(1 + \theta_D) \int_0^\infty f_{X_D}(r) P_{s,D}(\theta_D|r) dr + \int_0^\infty f_{X_D}(r) \int_{\theta_D}^\infty \frac{P_{s,D}(\xi|r)}{1 + \xi} d\xi dr \\
&= \log(1 + \theta_D) W(\theta_D) + V(\theta_D).
\end{aligned} \tag{5.60}$$

where $W(\theta_D) = \int_0^\infty f_{X_D}(r) P_{s,D}(\theta_D|r) dr$ and $V(\theta_D) = \int_{\theta_D}^\infty \frac{W(\xi)}{1 + \xi} d\xi$. From (5.1) and (5.3),

$$\begin{aligned}
W(\xi) &= \int_0^\infty f_{X_D}(r) P_{s,D}(\xi|r) dr \\
&= \sum_{k=1}^\infty \sum_{m=0}^{k-1} \sum_{j=1}^{k-m} (-1)^{k-m} J(m, k) \frac{\beta_j^{k-m}}{j!} I_j(\xi) + \sum_{k=0}^\infty J(k, k) I_0(\xi),
\end{aligned} \tag{5.61}$$

where

$$\begin{aligned}
I_j(\xi) &= \int_0^\infty 2\pi\lambda_T r \left(\frac{\pi\lambda_T}{\gamma} \xi^{1/2} r^2\right)^j e^{-\frac{\pi\lambda_T}{\gamma} \xi^{1/2} r^2 - \lambda_T \pi r^2} dr \\
&\stackrel{s=\pi\lambda_T r^2}{=} \int_0^\infty \left(\frac{\xi^{1/2}}{\gamma}\right)^j s^j e^{-(\frac{\xi^{1/2}}{\gamma} + 1)s} ds \\
&\stackrel{t=(\frac{\xi^{1/2}}{\gamma} + 1)s}{=} \int_0^\infty \frac{(\frac{\xi^{1/2}}{\gamma})^j}{(1 + \frac{\xi^{1/2}}{\gamma})^{j+1}} t^j e^{-t} dt \\
&= (j!) \frac{(\frac{\xi^{1/2}}{\gamma})^j}{(1 + \frac{\xi^{1/2}}{\gamma})^{j+1}}.
\end{aligned} \tag{5.62}$$

So

$$W(\xi) = \sum_{k=1}^\infty \sum_{m=0}^{k-1} \sum_{j=1}^{k-m} (-1)^{k-m} J(m, k) \beta_j^{k-m} \frac{(\frac{\xi^{1/2}}{\gamma})^j}{(1 + \frac{\xi^{1/2}}{\gamma})^{j+1}} + \sum_{k=0}^\infty J(k, k) \frac{1}{1 + \frac{\xi^{1/2}}{\gamma}}, \tag{5.63}$$

and

$$\begin{aligned}
V(\theta_D) &= \int_{\theta_D}^{\infty} \frac{W(\xi)}{1+\xi} d\xi \\
&= \sum_{k=1}^{\infty} \sum_{m=0}^{k-1} \sum_{j=1}^{k-m} (-1)^{k-m} J(m, k) \beta_j^{k-m} A_j(\theta_D) + \sum_{k=0}^{\infty} J(k, k) A_0(\theta_D),
\end{aligned} \tag{5.64}$$

where

$$\begin{aligned}
A_j(\theta_D) &= \int_{\theta_D}^{\infty} \frac{\left(\frac{\xi^{1/2}}{\gamma}\right)^j}{\left(1 + \frac{\xi^{1/2}}{\gamma}\right)^{j+1} (1 + \xi)} d\xi \\
&\stackrel{s=\frac{\xi^{1/2}}{\gamma}}{=} \int_{\frac{\theta_D^{1/2}}{\gamma}}^{\infty} \frac{s^j 2\gamma^2 s ds}{(1+s)^{j+1} (1+\gamma^2 s^2)} \\
&= \int_{\frac{\theta_D^{1/2}}{\gamma}}^{\infty} \frac{2\gamma^2 ds}{\left(1 + \frac{1}{s}\right)^{j+1} (1+\gamma^2 s^2)} \\
&\stackrel{t=\frac{1}{s}}{=} \int_{\frac{\gamma}{\theta_D^{1/2}}}^0 \frac{2\gamma^2}{(1+t)^{j+1}} \frac{-\frac{1}{t^2} dt}{1 + \frac{\gamma^2}{t^2}} \\
&= \int_0^{\frac{\gamma}{\theta_D^{1/2}}} \frac{2\gamma^2}{(1+t)^{j+1}} \frac{dt}{(t^2 + \gamma^2)}.
\end{aligned} \tag{5.65}$$

□

Theorem 7. *When $\alpha = 4$, the D2D average achievable throughput for a given SIR threshold θ_D can be approximated as*

$$\bar{R}_D(\theta_D) \approx \log(1 + \theta_D) \left(\frac{1}{1 + \frac{\theta_D^{1/2}}{\gamma}} + \frac{1 - e^{-K}}{2} \frac{\frac{\theta_D^{1/2}}{\gamma}}{\left(1 + \frac{\theta_D^{1/2}}{\gamma}\right)^2} \right) + A_0(\theta_D) + \frac{1 - e^{-K}}{2} A_1(\theta_D), \tag{5.66}$$

where

$$A_0(\theta_D) = \frac{\gamma^2}{1 + \gamma^2} \left[2 \log\left(1 + \frac{\gamma}{\theta_D^{1/2}}\right) - \log\left(\frac{1 + \theta_D}{\theta_D}\right) + \frac{2 \arctan(\theta_D^{-1/2})}{\gamma} \right], \tag{5.67}$$

and

$$\begin{aligned}
A_1(\theta_D) &= -\frac{2\gamma^4 + 2\gamma^2 + 2\gamma(\gamma^2 - 1)(\frac{\gamma}{\theta_D^{1/2}} + 1) \arctan(\frac{1}{\theta_D^{1/2}})}{(\gamma^2 + 1)^2(\gamma/\theta_D^{1/2} + 1)} + \frac{2\gamma^2(2 \log \gamma + \gamma^2 + 1)}{(\gamma^2 + 1)^2} \\
&+ \frac{-2\gamma^2(\frac{\gamma}{\theta_D^{1/2}} + 1) \log(\gamma^2 + \frac{\gamma^2}{\theta_D}) + 4\gamma^2(\frac{\gamma}{\theta_D^{1/2}} + 1) \log(\frac{\gamma}{\theta_D^{1/2}} + 1)}{(\gamma^2 + 1)^2(\gamma/\theta_D^{1/2} + 1)}.
\end{aligned} \tag{5.68}$$

Proof. Since

$$J(k, k) = e^{-K} K^k k! \binom{k}{k} / (k!)^2 = e^{-K} \frac{K^k}{k!}, \tag{5.69}$$

the second term in (5.63) is

$$\sum_{k=0}^{\infty} \frac{1}{1 + \frac{\xi^{1/2}}{\gamma}} J(k, k) = \frac{1}{1 + \frac{\xi^{1/2}}{\gamma}} \sum_{k=0}^{\infty} e^{-K} \frac{K^k}{k!} = \frac{1}{1 + \frac{\xi^{1/2}}{\gamma}}. \tag{5.70}$$

The first term in (5.63) can be rewritten as

$$\begin{aligned}
&\sum_{k=1}^{\infty} \sum_{n=1}^k \sum_{j=1}^n (-1)^n J(k-n, k) \beta_j^n \frac{(\frac{\xi^{1/2}}{\gamma})^j}{(1 + \frac{\xi^{1/2}}{\gamma})^{j+1}} \\
&= \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} \sum_{j=1}^n (-1)^n J(k-n, k) \beta_j^n \frac{(\frac{\xi^{1/2}}{\gamma})^j}{(1 + \frac{\xi^{1/2}}{\gamma})^{j+1}} \\
&= (-1)^1 \beta_1^1 \frac{\frac{\xi^{1/2}}{\gamma}}{(1 + \frac{\xi^{1/2}}{\gamma})^2} \left[\sum_{k=1}^{\infty} J(k-1, k) \right] + \sum_{n=2}^{\infty} (-1)^n \left[\sum_{j=1}^n \beta_j^n \frac{(\frac{\xi^{1/2}}{\gamma})^j}{(1 + \frac{\xi^{1/2}}{\gamma})^{j+1}} \right] \left[\sum_{k=n}^{\infty} J(k-n, k) \right] \\
&\approx (-1)^1 \beta_1^1 \frac{\frac{\xi^{1/2}}{\gamma}}{(1 + \frac{\xi^{1/2}}{\gamma})^2} \left[\sum_{k=1}^{\infty} J(k-1, k) \right].
\end{aligned} \tag{5.71}$$

Since

$$\beta_1^1 = (-1) \left(\frac{1}{2}\right) = -\frac{1}{2}, \tag{5.72}$$

$$\begin{aligned}
\sum_{k=1}^{\infty} J(k-1, k) &= \sum_{k=1}^{\infty} e^{-K} K^k (k-1)! \binom{k}{k-1} / (k!)^2 \\
&= \sum_{k=1}^{\infty} e^{-K} \frac{K^k}{k!} = 1 - e^{-K}.
\end{aligned} \tag{5.73}$$

Therefore,

$$W(\theta_D) \approx \frac{1}{1 + \frac{\theta_D^{1/2}}{\gamma}} + \frac{1 - e^{-K}}{2} \frac{\frac{\theta_D^{1/2}}{\gamma}}{(1 + \frac{\theta_D^{1/2}}{\gamma})^2}. \tag{5.74}$$

Similarly,

$$V(\theta_D) \approx A_0(\theta_D) + \frac{1 - e^{-K}}{2} A_1(\theta_D), \tag{5.75}$$

where

$$\begin{aligned}
A_0(\theta_D) &= \int_0^{\gamma/\theta_D^{1/2}} \frac{2\gamma^2}{(1+t)(t^2 + \gamma^2)} dt \\
&= \frac{\gamma^2}{1 + \gamma^2} \left[2 \log\left(1 + \frac{\gamma}{\theta_D^{1/2}}\right) - \log\left(\frac{1 + \theta_D}{\theta_D}\right) + \frac{2 \arctan(\theta_D^{-1/2})}{\gamma} \right],
\end{aligned} \tag{5.76}$$

and

$$\begin{aligned}
A_1(\theta_D) &= \int_0^{\gamma/\theta_D^{1/2}} \frac{2\gamma^2}{(1+t)^2(t^2 + \gamma^2)} dt \\
&= \frac{-2\gamma^2(\frac{\gamma}{\theta_D^{1/2}} + 1) \log(\gamma^2 + \frac{\gamma^2}{\theta_D}) + 4\gamma^2(\frac{\gamma}{\theta_D^{1/2}} + 1) \log(\frac{\gamma}{\theta_D^{1/2}} + 1)}{(\gamma^2 + 1)^2(\gamma/\theta_D^{1/2} + 1)} + \frac{2\gamma^2(2 \log \gamma + \gamma^2 + 1)}{(\gamma^2 + 1)^2} \\
&\quad - \frac{2\gamma^4 + 2\gamma^2 + 2\gamma(\gamma^2 - 1)(\frac{\gamma}{\theta_D^{1/2}} + 1) \arctan(\frac{1}{\theta_D^{1/2}})}{(\gamma^2 + 1)^2(\gamma/\theta_D^{1/2} + 1)}.
\end{aligned} \tag{5.77}$$

□

5.4.3 Simulation Results

In this section, we present simulation results to validate the derived analytical results. The primary simulation parameters used are listed in Table 5.2 [72, 125].

Table 5.2: *Simulation parameters*

Simulation Parameter	Value
Transmit power of cellular BSs (P_C)	20 W
Transmit power of DTUs (P_T)	500 mW
Density of cellular BSs (λ_C)	$5 \times 10^{-5} \text{ m}^{-2}$
Density of DTUs (λ_T)	$1 \times 10^{-4} \text{ m}^{-2}$
Rician factor (K_R)	10 dB
SINR threshold of cellular networks (θ_C)	0 dB

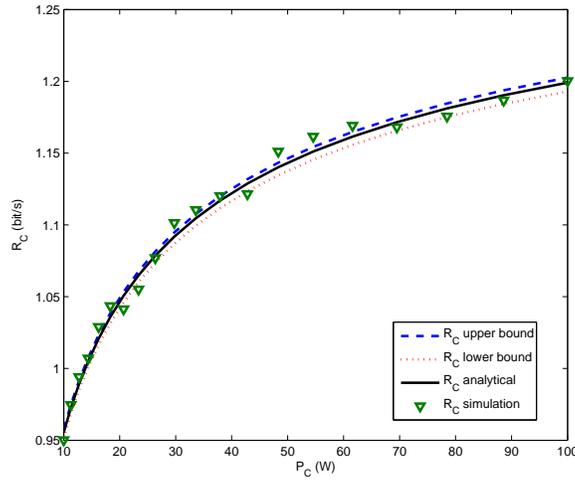


Figure 5.11: \bar{R}_C vs P_C

In Fig.5.11 and Fig.5.12, it is shown that the average achievable throughput of cellular network in a shared channel \bar{R}_C is shown to be tightly bounded by its upper bound \bar{R}_C^U and lower bound \bar{R}_C^L . Furthermore, \bar{R}_C increases with P_C or λ_C , in other words, decreases with γ since γ corresponds to $\frac{2}{\pi(\frac{\lambda_C}{\lambda_T}(\frac{P_C}{P_T})^{\frac{1}{2}}+1)}$. This is a direct result of stronger BS signal compared with interference from DTUs.

In Fig.5.13 and Fig.5.14, it is shown that the average achievable throughput of cellular network in a shared channel \bar{R}_D is shown to be well approximated by \hat{R}_D . Furthermore, \bar{R}_D increases with P_T or λ_T , in other words, increases with γ since γ corresponds to $\frac{2}{\pi(\frac{\lambda_C}{\lambda_T}(\frac{P_C}{P_T})^{\frac{1}{2}}+1)}$. This is a direct result of stronger D2D signal compared with interference from cellular BSs.

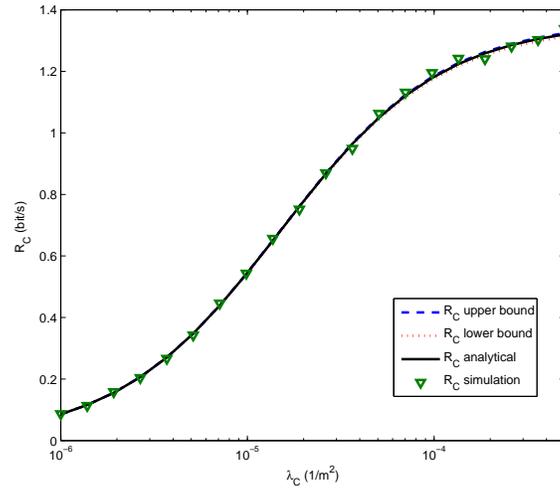


Figure 5.12: \bar{R}_C vs λ_C

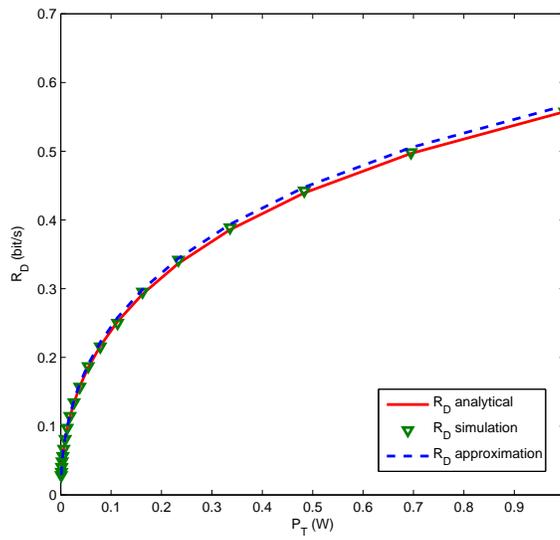


Figure 5.13: \bar{R}_D vs P_T

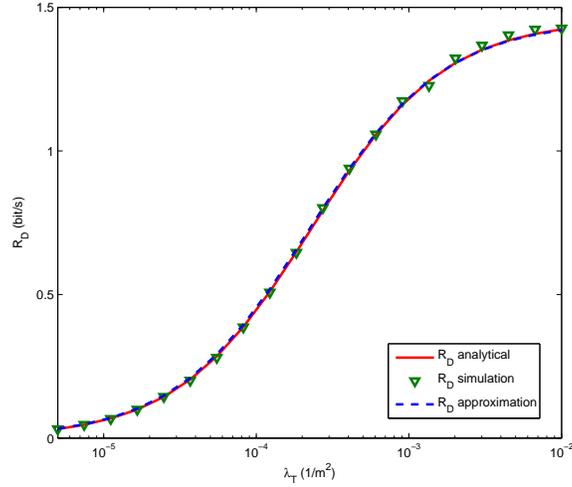


Figure 5.14: \bar{R}_D vs λ_T

5.5 Ergodic Capacity with Varying User Distributions

5.5.1 Network Model

In this section, DRU is assumed to be randomly distributed around its DTU, with a maximum distance R_{max} . Then we consider two cases: uniform distribution in distance (1D) and uniform distribution in area (2D). The probability density function of the distance X_D from a DRU to its serving DTU in the first case is expressed as:

$$f_{X_D}(r) = \frac{1}{R_{max}}, \quad 0 \leq r \leq R_{max}, \quad (5.78)$$

and in the second case:

$$f_{X_D}(r) = \frac{2r}{R_{max}^2}, \quad 0 \leq r \leq R_{max}. \quad (5.79)$$

5.5.2 Ergodic Capacity of DUs

The results in Lemma 3 (from [72]) are based on Rician fading model for D2D links. To evaluate ergodic capacity, the key limitations of the results in [72] are: (1) they are not in

closed form, and (2) the distances between all D2D pairs (X_D) are assumed to be fixed, which reduces the applicability of the derived results. In this section, for the case of $\alpha = 4$ the D2D ergodic capacity is derived based on $P_{s,D}(\theta_D|X_D)$ assuming D2D pair distance to be a random quantity.

Theorem 8. *When $\alpha = 4$, the D2D ergodic capacity in a shared channel corresponds to:*

$$C_D = \sum_{k=1}^{\infty} \sum_{m=0}^{k-1} \sum_{j=1}^{k-m} (-1)^{k-m} J(m, k) \frac{\beta_j^{k-m}}{j!} I_j + \sum_{k=0}^{\infty} J(k, k) I_0, \quad (5.80)$$

where I_j is given by the integral:

$$I_j = \int_0^{\infty} \frac{2t^{j+1} dt}{(\frac{\pi\lambda_T}{\gamma})^2 + t^2} \int_0^{\infty} r^{2j} e^{-tr^2} f_{X_D}(r) dr, \quad j = 0, 1, 2, \dots, \quad (5.81)$$

and the parameter

$$\gamma = \frac{2}{\pi(1 + \frac{\lambda_C}{\lambda_T} (\frac{P_C}{P_T})^{1/2})}. \quad (5.82)$$

Proof. We assume that the PDF and CDF of the DRU SIR are $f(\xi|X_D)$ and $F(\xi|X_D)$, respectively. The ergodic capacity for given X_D can be rewritten as:

$$\begin{aligned} C_D(X_D) &= \mathbb{E}[\log(1 + \xi)|X_D] = \int_0^{\infty} \log(1 + \xi) f(\xi|X_D) d\xi \\ &= \log(1 + \xi)(1 - F(\xi|X_D)) \Big|_0^{\infty} + \int_0^{\infty} \frac{1 - F(\xi|X_D)}{1 + \xi} d\xi \\ &= \int_0^{\infty} \frac{P_{s,D}(\xi|X_D)}{1 + \xi} d\xi. \end{aligned} \quad (5.83)$$

The ergodic capacity averaged over X_D can be evaluated as

$$C_D = \int_0^{\infty} C_D(r) f_{X_D}(r) dr = \int_0^{\infty} \frac{d\xi}{1 + \xi} \int_0^{\infty} P_{s,D}(\xi|r) f_{X_D}(r) dr. \quad (5.84)$$

When $\alpha = 4$, applying Lemma 3 to equation (5.84) and introducing parameter γ ,

$$\begin{aligned}
C_D &= \sum_{k=1}^{\infty} \sum_{m=0}^{k-1} \sum_{j=1}^{k-m} (-1)^{k-m} J(m, k) \frac{\beta_j^{k-m}}{j!} \times \int_0^{\infty} \frac{d\xi}{1+\xi} \int_0^{\infty} \left(\frac{\pi\lambda_T}{\gamma} \xi^{1/2}\right)^j r^{2j} e^{-\frac{\pi\lambda_T}{\gamma} r^2 \xi^{1/2}} f_{X_D}(r) dr \\
&+ \sum_{k=0}^{\infty} J(k, k) \int_0^{\infty} \frac{d\xi}{1+\xi} \int_0^{\infty} e^{-\frac{\pi\lambda_T}{\gamma} r^2 \xi^{1/2}} f_{X_D}(r) dr.
\end{aligned} \tag{5.85}$$

After changing the variable $\delta = \frac{\pi\lambda_T}{\gamma} \xi^{1/2}$,

$$\begin{aligned}
C_D &= \sum_{k=1}^{\infty} \sum_{m=0}^{k-1} \sum_{j=1}^{k-m} (-1)^{k-m} J(m, k) \frac{\beta_j^{k-m}}{j!} \times \int_0^{\infty} \frac{2\delta^{j+1} dt}{\left(\frac{\pi\lambda_T}{\gamma}\right)^2 + \delta^2} \int_0^{\infty} r^{2j} e^{-\delta r^2} f_{X_D}(r) dr \\
&+ \sum_{k=0}^{\infty} J(k, k) \int_0^{\infty} \frac{2\delta d\delta}{\left(\frac{\pi\lambda_T}{\gamma}\right)^2 + \delta^2} \int_0^{\infty} e^{-\delta r^2} f_{X_D}(r) dr \\
&= \sum_{k=1}^{\infty} \sum_{m=0}^{k-1} \sum_{j=1}^{k-m} (-1)^{k-m} J(m, k) \frac{\beta_j^{k-m}}{j!} I_j + \sum_{k=0}^{\infty} J(k, k) I_0.
\end{aligned} \tag{5.86}$$

□

If the distance X_D between a DRU and its serving DTU follows an uniform distribution, then the term I_j can be calculated by the following corollary.

Corollary 7. *When the distance between a DU pair follows a uniform distribution, i.e., the PDF of the distance X_D is given in (5.78),*

$$I_0 = \sqrt{\pi} \int_0^{\infty} \frac{\sqrt{x} \operatorname{erf}(\sqrt{x})}{x^2 + A^2} dx \tag{5.87}$$

$$I_j = -D_j + (j - 1/2)I_{j-1}, \tag{5.88}$$

$$D_j = (j - 2)! - A^2 D_{j-2}, \tag{5.89}$$

$$D_0 = \frac{1}{A} [ci(A) \sin A - si(A) \cos A], \tag{5.90}$$

$$D_1 = -[ci(A) \cos A + si(A) \sin A], \quad (5.91)$$

where the parameter

$$A = \frac{\pi \lambda_T R_{max}^2}{\gamma} = \frac{\pi^2}{2} (\lambda_T + \lambda_C (\frac{P_C}{P_T})^{1/2}) R_{max}^2, \quad (5.92)$$

and the special function $erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-\delta^2} d\delta$, $ci(x) = -\int_x^\infty \frac{\cos \delta}{\delta} d\delta$ and $si(x) = -\int_x^\infty \frac{\sin \delta}{\delta} d\delta$.

Moreover, I_0 can be approximated by

$$\hat{I}_0 = \sqrt{\pi} \left[\frac{\arctan\left(\frac{2\sqrt[4]{-A^2}}{\sqrt{\pi}}\right) + \tanh^{-1}\left(\frac{2\sqrt[4]{-A^2}}{\sqrt{\pi}}\right)}{\sqrt[4]{-A^2}} \right] + \log\left(1 + \frac{\pi^2}{16A^2}\right) \quad (5.93)$$

Proof. Starting from (5.81),

$$I_j = \int_0^\infty \frac{2\delta^{j+1} d\delta}{\delta^2 + (\frac{\pi \lambda_T}{\gamma})^2} \int_0^{R_{max}} r^{2j} e^{-\delta r^2} \frac{1}{R_{max}} dr. \quad (5.94)$$

By changing the variable $s = \delta r^2$,

$$\begin{aligned} I_j &= \frac{1}{R_{max}} \int_0^\infty \frac{2t^{j+1} d\delta}{\delta^2 + (\frac{\pi \lambda_T}{\gamma})^2} \int_0^{\delta R_{max}^2} \left(\frac{s}{\delta}\right)^j e^{-s} \frac{1}{\sqrt{\delta}} \frac{ds}{2\sqrt{s}} \\ &= \frac{1}{R_{max}} \int_0^\infty \frac{\sqrt{\delta} d\delta}{\delta^2 + (\frac{\pi \lambda_T}{\gamma})^2} \int_0^{\delta R_{max}^2} s^{j-\frac{1}{2}} e^{-s} ds. \end{aligned} \quad (5.95)$$

We can introduce the incomplete gamma function $\Gamma(\alpha, x) = \int_0^x \delta^{\alpha-1} e^{-\delta} d\delta$,

$$I_j = \frac{1}{R_{max}} \int_0^\infty \frac{\sqrt{\delta} \Gamma(j + \frac{1}{2}, \delta R_{max}^2) d\delta}{\delta^2 + (\frac{\pi \lambda_T}{\gamma})^2}. \quad (5.96)$$

Employing a change of variable $x = \delta R_{max}^2$ results in

$$I_j = \int_0^\infty \frac{\sqrt{x} \Gamma(j + \frac{1}{2}, x) dx}{x^2 + (\frac{\pi \lambda_T R_{max}^2}{\gamma})^2} = \int_0^\infty \frac{\sqrt{x} \Gamma(j + \frac{1}{2}, x) dx}{x^2 + A^2}. \quad (5.97)$$

According to the property of an incomplete gamma function,

$$\Gamma(\alpha, x) = -x^{\alpha-1}e^{-x} + (\alpha - 1)\Gamma(\alpha - 1, x), \quad \alpha > 1. \quad (5.98)$$

When $j = 1, 2, 3, \dots$,

$$\begin{aligned} I_j &= \int_0^\infty \frac{\sqrt{x}\Gamma(j + \frac{1}{2}, x)}{x^2 + A^2} dx \\ &= \int_0^\infty \frac{\sqrt{x}[-x^{j-1/2}e^{-x} + (j - 1/2)\Gamma(j - 1/2, x)]}{x^2 + A^2} dx \\ &= -\int_0^\infty \frac{x^j e^{-x}}{x^2 + A^2} dx + (j - 1/2) \int_0^\infty \frac{\sqrt{x}\Gamma(j - 1/2, x)}{x^2 + A^2} dx \\ &= -D_j + (j - 1/2)I_{j-1}. \end{aligned} \quad (5.99)$$

The term D_j is defined as

$$D_j = \int_0^\infty \frac{x^j e^{-x}}{x^2 + A^2} dx. \quad (5.100)$$

We can obtain that

$$D_0 = \int_0^\infty \frac{e^{-x}}{x^2 + A^2} dx = \frac{1}{A} [ci(A) \sin A - si(A) \cos A], \quad (5.101)$$

$$D_1 = \int_0^\infty \frac{x e^{-x}}{x^2 + A^2} dx = -[ci(A) \cos A + si(A) \sin A]. \quad (5.102)$$

For $j = 2, 3, 4, \dots$,

$$\begin{aligned} D_j &= \int_0^\infty \frac{x^j e^{-x}}{x^2 + A^2} dx = \int_0^\infty \frac{x^{j-2}(x^2 + A^2 - A^2)e^{-x}}{x^2 + A^2} dx \\ &= \int_0^\infty x^{j-2} e^{-x} dx - A^2 \int_0^\infty \frac{x^{j-2} e^{-x}}{x^2 + A^2} dx \\ &= (j - 2)! - A^2 D_{j-2}. \end{aligned} \quad (5.103)$$

Finally we can derive that

$$I_0 = \int_0^\infty \frac{\sqrt{x}\Gamma(\frac{1}{2}, x)}{x^2 + A^2} dx = \sqrt{\pi} \int_0^\infty \frac{\sqrt{x}erf(\sqrt{x})}{x^2 + A^2} dx. \quad (5.104)$$

Next we use the approximation:

$$erf(x) \approx \begin{cases} \frac{2}{\sqrt{\pi}}x & 0 < x < \frac{\sqrt{\pi}}{2} \\ 1 & x > \frac{\sqrt{\pi}}{2} \end{cases} \quad (5.105)$$

that is,

$$erf(\sqrt{x}) \approx \begin{cases} \frac{2}{\sqrt{\pi}}\sqrt{x} & 0 < x < \frac{\pi}{4} \\ 1 & x > \frac{\pi}{4}. \end{cases} \quad (5.106)$$

Since $A \geq 0$, we can conclude that

$$\begin{aligned} I_0 &\approx \sqrt{\pi} \int_0^{\frac{\pi}{4}} \frac{\sqrt{x} \frac{2}{\sqrt{\pi}} \sqrt{x}}{x^2 + A^2} dx + \sqrt{\pi} \int_{\frac{\pi}{4}}^\infty \frac{\sqrt{x}}{x^2 + A^2} dx \\ &= \int_0^{\frac{\pi}{4}} \frac{2x}{x^2 + A^2} dx + \sqrt{\pi} \int_{\frac{\pi}{4}}^\infty \frac{\sqrt{x}}{x^2 + A^2} dx \\ &= \sqrt{\pi} \left[\frac{\arctan\left(\frac{2\sqrt[4]{-A^2}}{\sqrt{\pi}}\right) + \tanh^{-1}\left(\frac{2\sqrt[4]{-A^2}}{\sqrt{\pi}}\right)}{\sqrt[4]{-A^2}} \right] + \log\left(1 + \frac{\pi^2}{16A^2}\right) = \hat{I}_0. \end{aligned} \quad (5.107)$$

□

We define \hat{C}_D in which I_0 is replaced by \hat{I}_0 , which is a closed-form approximation of C_D .

On the other hand, if a DRU is distributed uniformly in the circular area around its serving DTU, the term I_j can be given by the following corollary.

Corollary 8. *When a DRU is distributed uniformly in the circular area around its serving DTU, i.e., the probability density function of the distance X_D is given in (5.79),*

$$I_0 = \frac{\pi}{A} - 2D_0, \quad (5.108)$$

$$I_j = -2D_j + jI_{j-1}, \quad (5.109)$$

The term D_j and the parameter A are the same as that in corollary 7.

Proof. Starting from (5.81),

$$I_j = \int_0^\infty \frac{2\delta^{j+1}d\delta}{\delta^2 + (\frac{\pi\lambda_T}{\gamma})^2} \int_0^{R_{max}} r^{2j} e^{-\delta r^2} \frac{2r}{R_{max}^2} dr. \quad (5.110)$$

By changing the variable $s = \delta r^2$,

$$\begin{aligned} I_j &= \int_0^\infty \frac{2\delta^{j+1}d\delta}{\delta^2 + (\frac{\pi\lambda_T}{\gamma})^2} \int_0^{\delta R_{max}^2} (\frac{s}{\delta})^j e^{-s} \frac{1}{R_{max}^2} \frac{ds}{\delta} \\ &= \frac{2}{R_{max}^2} \int_0^\infty \frac{d\delta}{\delta^2 + (\frac{\pi\lambda_T}{\gamma})^2} \int_0^{\delta R_{max}^2} s^j e^{-s} ds \\ &= \frac{2}{R_{max}^2} \int_0^\infty \frac{\Gamma(j+1, \delta R_{max}^2)}{\delta^2 + (\frac{\pi\lambda_T}{\gamma})^2} d\delta. \end{aligned} \quad (5.111)$$

Changing variable again $x = \delta R_{max}^2$,

$$I_j = 2 \int_0^\infty \frac{\Gamma(j+1, x)dx}{x^2 + (\frac{\pi\lambda_T R_{max}^2}{\gamma})^2} = 2 \int_0^\infty \frac{\Gamma(j+1, x)dx}{x^2 + A^2}. \quad (5.112)$$

When $j = 1, 2, 3, \dots$,

$$\begin{aligned} I_j &= 2 \int_0^\infty \frac{\Gamma(j+1, x)}{x^2 + A^2} dx = 2 \int_0^\infty \frac{-x^j e^{-x} + j\Gamma(j, x)}{x^2 + A^2} dx \\ &= -2 \int_0^\infty \frac{x^j e^{-x}}{x^2 + A^2} dx + j \int_0^\infty \frac{2\Gamma(j, x)}{x^2 + A^2} dx \\ &= -2D_j + jI_{j-1}. \end{aligned} \quad (5.113)$$

D_j is defined in the same way as in corollary 7, and finally

$$\begin{aligned}
I_0 &= 2 \int_0^\infty \frac{\Gamma(1, x)}{x^2 + A^2} dx = 2 \int_0^\infty \frac{1 - e^{-x}}{x^2 + A^2} dx \\
&= 2 \int_0^\infty \frac{1}{x^2 + A^2} dx - 2 \int_0^\infty \frac{e^{-x}}{x^2 + A^2} dx \\
&= \frac{2}{A} \arctan\left(\frac{x}{A}\right)\Big|_0^\infty - 2D_0 = \frac{\pi}{A} - 2D_0.
\end{aligned} \tag{5.114}$$

□

5.5.3 Simulation Results

Simulation results are presented to validate the derived analytical results. The primary simulation parameters used are listed in Table 5.3 [72, 125].

Table 5.3: *Simulation parameters*

Simulation Parameter	Value
Transmit power of cellular BSs (P_C)	20 W
Transmit power of DTUs (P_T)	500 mW
Maximum distance of DU pairs (R_{max})	100 m
Density of cellular BSs (λ_C)	$5 \times 10^{-6} \text{ m}^{-2}$
Density of DTUs (λ_T)	$1 \times 10^{-4} \text{ m}^{-2}$
Rician factor (K_R)	10 dB

Fig. 5.15 shows that the ergodic capacity of a cellular network in a shared channel C_D increases with P_T , in other words, decreases with γ since γ corresponds to $\frac{2}{\pi(\frac{\lambda_C}{\lambda_T}(\frac{P_C}{P_T})^{\frac{1}{2}}+1)}$. The fact that C_D increases with P_T is a direct result of stronger BS signal compared with interference from DTUs. Furthermore, C_D is shown to be well approximated by \hat{C}_D . For comparison, we show the ergodic capacity of D2D communication C_D obtained by the results in [72] fixing D2D pair distance X_D as the expected value according to its distribution: $R_{max}/2$ when distance follows a uniform distribution and $2R_{max}/3$ when a DRU is distributed uniformly in the circular area around its serving DTU. We can see that the ergodic

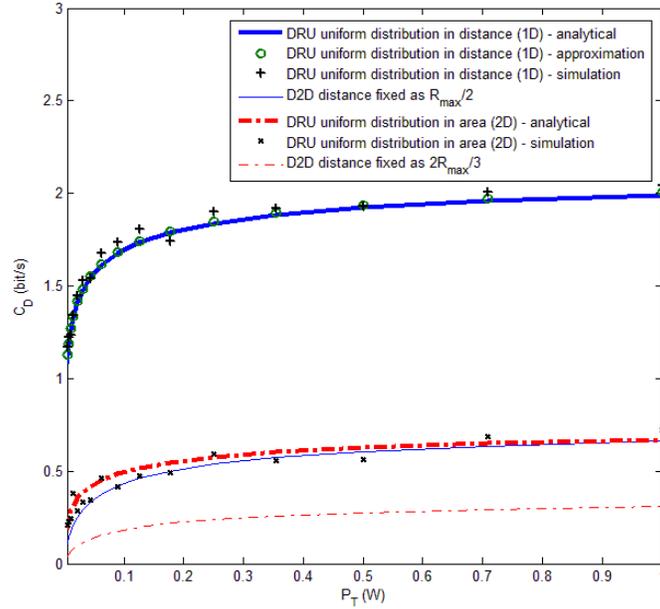


Figure 5.15: P_T vs C_D

capacity C_D obtained by [72] greatly underestimates actual capacity. The result derived from our model is more useful than [72], because in reality the distance between D2D pairs is expected to vary.

Fig. 5.16 shows that the ergodic capacity of a cellular network in a shared channel C_D decreases with R_{max} , which is a direct result of stronger BS signal with less fading distance. We also show the ergodic capacity of D2D communication C_D obtained by the results in [72] fixing D2D pair distance X_D as the expected value according to its distribution for comparison. The ergodic capacity C_D obtained by [72] is shown to be greatly underestimated compared with our model. Again the result derived from our model is more useful than [72] because of expected variation in D2D pair distances.

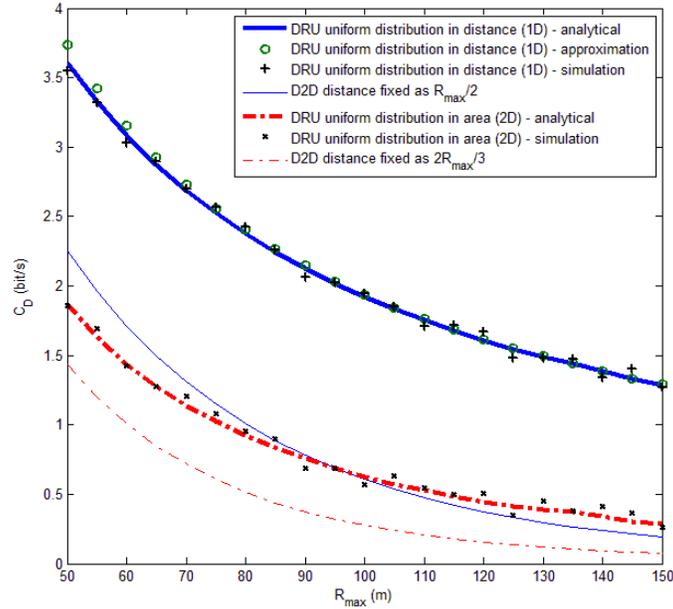


Figure 5.16: R_{max} vs C_D

5.6 Summary

We study the performance evaluation and maximization of D2D underlay cellular networks. First of all, we aim to find a critical set of system parameters to ensure the benefits of D2D underlay operation to outweigh its drawbacks in the context of trade-off between ergodic capacity and power consumption. To achieve that, we propose a network model to understand the behavior of D2D underlay cellular networks using tools from stochastic geometry. Unlike prior efforts, D2D direct links are modeled by Rician fading and the distances between D2D pairs are treated as random quantities. For a special case of path loss exponent of 4, our work derives (1) tight lower and upper bounds for ergodic capacity of a cellular network, and (2) recursive closed-form expressions and closed-form approximation for ergodic capacity of D2D networks. It is shown that our analytical results better approximate the actual capacity compared with prior approaches. Then we investigate ergodic capacity maximization and the impact of both transmit power and density of D2D users. A two-phase

scheme is proposed to maximize the ergodic capacity of D2D underlay cellular networks as well as to minimize the overall power consumption. We evaluate both ergodic capacity and total power consumption of the optimal situation comparing them with some extreme cases as well. Secondly, we analyze the average achievable throughput of D2D underlay cellular networks. Assuming a path loss exponent of 4, we derive closed-form results including: (1) tight upper and tight lower bounds for average achievable throughput of cellular networks, and (2) good approximation for average achievable throughput of D2D networks. Finally, we provide closed-form ergodic capacity results (when path loss exponent is 4) for two other cases of the distribution of DRU: (1) distance between a DU pair follows a uniform distribution and (2) a DRU is distributed uniformly in the circular area around its serving DTU.

D2D technique is a key enabler for emerging IoT. In the meantime, SIET is attracting widespread attention for wireless networks. Next chapter will focus on the performance evaluation of simultaneous information and energy transfer in IoT.

Chapter 6

Modeling and Analysis of Simultaneous Information and Energy Transfer

Nowadays, nearly all devices in our daily lives are capable of some form of computation and communication. It is predicted that the devices will interact with each other sharing common goals through a network, which is called the Internet of Things (IoT). The IoT nodes, which are referred to as “things”, are typically equipped with sensors, controlling processors, wireless transceivers, and an energy source (e.g., a battery) to monitor their environment and transmit data. Applications envisioned for IoT include home automation, surveillance, healthcare, real-time supply chain monitoring, transportation, and smart environments.

However, operation of traditional battery-powered IoT nodes can be periodically interrupted because of finite battery lives, especially when battery charging or replacement is inconvenient or even impossible in certain cases. Therefore, simultaneous information and energy transfer (SIET) has become a promising solution to improve the energy efficiency and longevity of IoT nodes.

In this chapter, we study SIET in the context of wireless energy harvesting IoT networks.

First of all, we analyze the network performances using Ginibre point process to capture the geometric characteristics of IoT gateway locations. Based on that, we present, for the first time, valid closed-form upper bounds for both power outage probability and transmission outage probability in IoT networks assuming a practical case when the path-loss coefficient is 4. We also provide a tight upper bound for the transmission outage probability in a computable integral representation. Then we try to balance both outage probabilities by minimizing the maximum outage probability via looking for the optimal power split ratio. The effects of several parameters (e.g., density and transmit power of gateways, and SINR threshold) on the optimal power split ratio are analyzed as well.

This chapter includes several distinct sections. Section 6.1 describes the system model. Then Network performance analysis in terms of transmission outage probability and power outage probability are derived in Section 6.2. Based on the obtained closed-form results, Section 6.3 presents the convexity proof of the maximum outage probability. Numerical results and discussion are given in Section 6.4 and the chapter is finally concluded in Section 6.5.

6.1 System Model

We consider an IoT network consisting of IoT gateways (e.g., wireless routers and cellular mobiles) and IoT nodes (sensors [85], actuators [63], etc.), as is shown in Fig. 6.1. We assume that the IoT gateways can also be deemed as RF energy sources for the IoT nodes. In this chapter, we focus on SIET on the downlink. For simplicity and tractability, we assume that: (1) all gateways transmit at the same power P_G ; and (2) each node is associated with its nearest gateway. To capture repulsion between IoT gateways, we propose the Ginibre point process (GPP) as a model for their locations. The kernel of GPP is defined as

$$K(x, y) = \rho e^{\pi\rho(x,y)} e^{-\frac{\pi\rho}{2}(|x|^2+|y|^2)}, \quad x, y \in \mathbb{R}^2 \quad (6.1)$$

where $\rho > 0$ denotes the density of GPP.

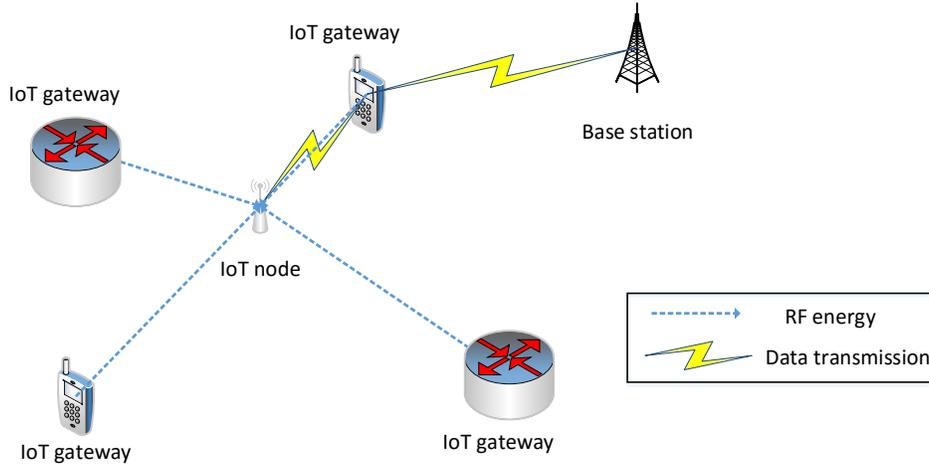


Figure 6.1: *IoT Network Model*

For the GPP, from Theorem 4.7.1 in [126], we have the following result:

Proposition 3. *Let $\Phi = \{X_i\}, i \in N^+$ be a GPP. Then $\{|X_i|^2\}_{i \in N^+}$ has the same distribution as the set $\{Q_k\}_{k \in N^+}$. Q_k is a random variable with PDF [82]:*

$$f_{Q_k}(q) = \frac{q^{k-1} e^{-\pi\rho q}}{(\frac{1}{\pi\rho})^k \Gamma(k)}, \quad 0 < q < \infty. \quad (6.2)$$

i.e., $Q_k \sim \text{gamma}(k, \frac{1}{\pi\rho})$ with Q_k independent of Q_j if $k \neq j$.

The IoT nodes are considered to adopt the power-splitting architecture. Each node is equipped by an information decoder and a RF energy harvester, with independent antennas functioning separately and observing different channel gains. This enables the nodes to perform data transmission and RF energy harvesting simultaneously, as is shown in Fig. 6.2. After the power splitting, the portion of RF signals split to the information receiver is denoted by ϵ ($0 \leq \epsilon \leq 1$), and that to the energy harvester is $1 - \epsilon$. The white Gaussian noise introduced by the receiving antenna $n \sim \mathcal{CN}(0, \sigma^2)$. We also assume that the energy harvester has a conversion efficiency η .

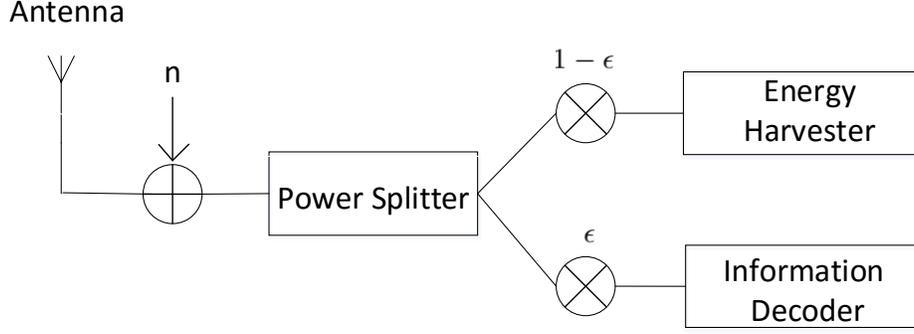


Figure 6.2: *IoT Network Model*

Therefore, the harvested energy of an IoT node is

$$P_H = \eta(1 - \epsilon) \sum_k P_G Q_k^{-\alpha/2}. \quad (6.3)$$

The received power from a tagged gateway to an IoT node is $P_G \{\min_k Q_k\}^{-\alpha/2}$. Therefore, the SINR expression for this node connecting to its IoT gateway is:

$$SINR = \frac{\epsilon P_G \{\min_k Q_k\}^{-\alpha/2}}{\epsilon (\sum_k P_G Q_k^{-\alpha/2} - P_G \{\min_k Q_k\}^{-\alpha/2}) + \sigma^2}. \quad (6.4)$$

We assume that the IoT nodes are active and sending signals to its associated gateway. So they not only need power to operate the sensor circuit, but need extra energy provided for information transmission to its associated gateway as well. Thus, the power consumption at each node is given by

$$P_D = P_O + C_{TX} \{\min_k Q_k\}^{\alpha/2}, \quad (6.5)$$

where P_O denotes the static operational power expenditure and C_{TX} denotes the proportionality constant coefficient of the RF transmit power.

We consider battery-free IoT nodes powered solely by the instant harvested RF energy.

Then we evaluate downlink transmission outage probability and power outage probability of nodes. We will analyze these performance metrics in the following section 6.2.

6.2 Performance Analysis

In this work, we consider a practical case of path loss coefficient $\alpha = 4$ and aim to derive closed-form analytical results for both transmission outage probability and power outage probability.

6.2.1 Transmission Outage Probability

Given a SINR threshold θ , the transmission outage probability is defined as $\mathbb{P}_{outage}^T = \mathbb{P}(SINR < \theta)$. First, we present a tight upper bound for the transmission outage probability in a computable integral representation in Theorem 9.

Theorem 9. *When the SINR threshold $\theta \geq 1$, the transmission outage probability of an IoT node can be upper bounded as*

$$\mathbb{P}_{outage}^T < \mathbb{P}_{outage}^{Tu1} = F_W\left(\frac{\theta\sigma^2}{\epsilon P_G}\right) + \int_{\frac{\theta\sigma^2}{\epsilon P_G}}^{\infty} f_W(u) \left(1 - F_Z\left(\frac{u - \frac{\theta\sigma^2}{\epsilon P_G}}{\theta}\right)\right) du, \quad (6.6)$$

where

$$F_W(u) = (\pi\rho)^2 \int_0^{\infty} x e^{-\pi\rho x \left(1 + \frac{1}{\sqrt{\theta+ux^2}}\right)} dx + \pi\rho \int_0^{\infty} e^{-\pi\rho z \left(1 + \frac{1}{\sqrt{\theta+uz^2}}\right)} \left(1 + \pi\rho z \frac{1}{\sqrt{\theta+uz^2}}\right) dz - 1, \quad (6.7)$$

$$f_W(u) = \frac{dF_W(u)}{du}, \quad (6.8)$$

$$F_Z(u) = \int_0^u f_{W_3}(u_3) \int_0^{u-u_3} f_{W_4}(u_4) \cdots \int_0^{u-u_3-u_4-\cdots-u_{N-1}} f_{W_N}(u_N) du_N du_{N-1} \cdots du_4 du_3, \quad (6.9)$$

and

$$f_{W_k}(u) = \frac{(\pi\rho)^k}{2\Gamma(k)} \sqrt{\frac{1}{u^{k+2}}} e^{-\pi\rho\sqrt{\frac{1}{u}}}. \quad (6.10)$$

Proof. Based on the definition of transmission outage probability and (6.4),

$$\begin{aligned} \mathbb{P}_{outage}^T &= \mathbb{P}(SINR < \theta) \\ &= \mathbb{P}\left[\frac{\{\min_k Q_k\}^{-2}}{\sum_k Q_k^{-2} - \{\min_k Q_k\}^{-2} + \frac{\sigma^2}{\epsilon P_G}} < \theta\right] \\ &= \mathbb{P}\left[(1 + \theta)\{\min_k Q_k\}^{-2} - \theta \sum_k Q_k^{-2} < \frac{\theta\sigma^2}{\epsilon P_G}\right]. \end{aligned} \quad (6.11)$$

Since $\min_k Q_k < \min(Q_1, Q_2)$, we can find an upper bound of \mathbb{P}_{outage}^T :

$$\begin{aligned} \mathbb{P}_{outage}^T &< \mathbb{P}\left[(1 + \theta)\{\min(Q_1, Q_2)\}^{-2} - \theta(Q_1^{-2} + Q_2^{-2}) - \theta \sum_{k>2} Q_k^{-2} < \frac{\theta\sigma^2}{\epsilon P_G}\right] \\ &= \mathbb{P}\left(W - \theta Z < \frac{\theta\sigma^2}{\epsilon P_G}\right) \\ &= \mathbb{P}\left(W < \frac{\theta\sigma^2}{\epsilon P_G}\right) + \mathbb{P}\left(W \geq \frac{\theta\sigma^2}{\epsilon P_G}, Z > \frac{W - \frac{\theta\sigma^2}{\epsilon P_G}}{\theta}\right) \\ &= F_W\left(\frac{\theta\sigma^2}{\epsilon P_G}\right) + \int_{\frac{\theta\sigma^2}{\epsilon P_G}}^{\infty} f_W(u) \left(1 - F_Z\left(\frac{u - \frac{\theta\sigma^2}{\epsilon P_G}}{\theta}\right)\right) du. \end{aligned} \quad (6.12)$$

Here we define the random variable

$$W = (1 + \theta)\{\min(Q_1, Q_2)\}^{-2} - \theta(Q_1^{-2} + Q_2^{-2}), \quad (6.13)$$

and

$$Z = \sum_{k>2} Q_k^{-2}. \quad (6.14)$$

When $\theta \geq 1$, we can calculate the CDF of W as

$$\begin{aligned}
F_W(u) &= \mathbb{P}((1 + \theta)\{\min(Q_1, Q_2)\}^{-2} - \theta Q_1^{-2} - \theta Q_2^{-2} < u) \\
&= \mathbb{P}(Q_1^{-2} - \theta Q_2^{-2} < u, Q_1 < Q_2) + \mathbb{P}(Q_2^{-2} - \theta Q_1^{-2} < u, Q_2 < Q_1) \\
&= \mathbb{P}(W_1 - \theta W_2 < u, W_1 > W_2) + \mathbb{P}(W_2 - \theta W_1 < u, W_2 > W_1) \\
&= \mathbb{P}(0 < W_2 < \infty, W_2 < W_1 < u + \theta W_2) + \mathbb{P}(0 < W_1 < \infty, W_1 < W_2 < u + \theta W_1) \\
&= \int_0^\infty f_{W_2}(v)[F_{W_1}(u + \theta v) - F_{W_1}(v)]dv + \int_0^\infty f_{W_1}(v)[F_{W_2}(u + \theta v) - F_{W_2}(v)]dv.
\end{aligned} \tag{6.15}$$

Here we introduce the random variables $W_1 = Q_1^{-2}$, $W_2 = Q_2^{-2}$. Since $Q_k \sim \text{gamma}(k, \frac{1}{\pi\rho})$, we have

$$F_{Q_1}(x) = 1 - e^{-\pi\rho x}, \tag{6.16}$$

$$F_{Q_2}(x) = 1 - (1 + \pi\rho x)e^{-\pi\rho x}. \tag{6.17}$$

Then we can derive the CDF of W_1 and W_2 as

$$F_{W_1}(x) = 1 - F_{Q_1}\left(\sqrt{\frac{1}{x}}\right) = e^{-\pi\rho\sqrt{\frac{1}{x}}}, \tag{6.18}$$

and

$$F_{W_2}(x) = 1 - F_{Q_2}\left(\sqrt{\frac{1}{x}}\right) = (1 + \pi\rho\sqrt{\frac{1}{x}})e^{-\pi\rho\sqrt{\frac{1}{x}}}, \tag{6.19}$$

and the PDF of W_1 and W_2 as

$$f_{W_1}(x) = \frac{\pi\rho}{2}x^{-3/2}e^{-\pi\rho\sqrt{\frac{1}{x}}}, \tag{6.20}$$

and

$$f_{W_2}(x) = \frac{(\pi\rho)^2}{2}x^{-2}e^{-\pi\rho\sqrt{\frac{1}{x}}}. \tag{6.21}$$

By inserting the above results into (6.15) and calculate the whole integral separately we can get

$$F_W(u) = A_0(u) + B_0(u) + C_0(u) + D_0(u), \quad (6.22)$$

where

$$\begin{aligned} A_0(u) &= \int_0^\infty f_{W_2}(v)F_{W_1}(u + \theta v)dv = \frac{(\pi\rho)^2}{2} \int_0^\infty v^{-2}e^{-\pi\rho(\sqrt{\frac{1}{v}}+\sqrt{\frac{1}{u+\theta v}})}dv \\ &\stackrel{x=\sqrt{\frac{1}{v}}}{=} (\pi\rho)^2 \int_0^\infty xe^{-\pi\rho x(1+\frac{1}{\sqrt{\theta+ux^2}})}dx, \end{aligned} \quad (6.23)$$

$$\begin{aligned} B_0(u) &= - \int_0^\infty f_{W_2}(v)F_{W_1}(v)dv = -\frac{(\pi\rho)^2}{2} \int_0^\infty v^{-2}e^{-2\pi\rho\sqrt{\frac{1}{v}}}dv \\ &\stackrel{y=(2\pi\rho)\sqrt{\frac{1}{v}}}{=} -\frac{1}{4} \int_0^\infty e^{-y}ydy = -\frac{1}{4}, \end{aligned} \quad (6.24)$$

$$\begin{aligned} C_0(u) &= \int_0^\infty f_{W_1}(v)F_{W_2}(u + \theta v)dv = \frac{\pi\rho}{2} \int_0^\infty v^{-3/2}e^{-\pi\rho(\sqrt{\frac{1}{v}}+\sqrt{\frac{1}{u+\theta v}})}(1 + \pi\rho\sqrt{\frac{1}{u + \theta v}})dv \\ &\stackrel{z=\sqrt{\frac{1}{v}}}{=} \pi\rho \int_0^\infty e^{-\pi\rho z(1+\frac{1}{\sqrt{\theta+uz^2}})}(1 + \pi\rho z\frac{1}{\sqrt{\theta + uz^2}})dz, \end{aligned} \quad (6.25)$$

and

$$\begin{aligned} D_0(u) &= - \int_0^\infty f_{W_1}(v)F_{W_2}(v)dv = -\frac{\pi\rho}{2} \int_0^\infty v^{-3/2}e^{-2\pi\rho\sqrt{\frac{1}{v}}}(1 + \pi\rho\sqrt{\frac{1}{v}})dv \\ &\stackrel{g=(2\pi\rho)\sqrt{\frac{1}{v}}}{=} -\frac{1}{2} \int_0^\infty (1 + \frac{g}{2})e^{-g}dg = -\frac{3}{4}. \end{aligned} \quad (6.26)$$

From results above, we can obtain the first term in (6.12) and $f_W(u)$ in the second term.

The second term in (6.12) also requires the CDF of Z , which can be calculated by

$$\begin{aligned}
F_Z(u) &= \mathbb{P}\left(\sum_{k>2} Q_k^{-2} < u\right) \\
&= \mathbb{P}\left(\sum_{k=3}^N W_k < u\right) \\
&= \int_0^u f_{W_3}(u_3) \int_0^{u-u_3} f_{W_4}(u_4) \cdots \int_0^{u-u_3-u_4-\cdots-u_{N-1}} f_{W_N}(u_N) du_N du_{N-1} \cdots du_4 du_3.
\end{aligned} \tag{6.27}$$

Here we introduce the random variables $W_k = Q_k^{-2}$, $k = 3, \dots, N$. Since

$$F_{Q_k}(u) = 1 - \frac{\Gamma(k, \pi\rho u)}{\Gamma(k)}, \tag{6.28}$$

where $\Gamma(k, x) = \int_x^\infty t^{k-1} e^{-t} dt$ is the incomplete Gamma function, we have

$$F_{W_k}(u) = 1 - F_{Q_k}\left(\sqrt{\frac{1}{u}}\right) = \frac{\Gamma(k, \pi\rho\sqrt{\frac{1}{u}})}{\Gamma(k)}, \tag{6.29}$$

and then the PDF of W_k reads

$$\begin{aligned}
f_{W_k}(u) &= \frac{d}{du} F_{W_k}(u) = \frac{1}{\Gamma(k)} \frac{d}{du} \left(\int_{\pi\rho\sqrt{\frac{1}{u}}}^\infty t^{k-1} e^{-t} dt \right) \\
&= \frac{(\pi\rho)^k}{2\Gamma(k)} \sqrt{\frac{1}{u^{k+2}}} e^{-\pi\rho\sqrt{\frac{1}{u}}}.
\end{aligned} \tag{6.30}$$

Using the results (6.22) and (6.27), we can calculate the upper bound of the transmission outage probability by (6.12), in which $f_W(u)$ is the derivative of $F_W(u)$. \square

Next we derive a closed-form upper bound for the transmission outage probability as given in Theorem 10:

Theorem 10. *The transmission outage probability \mathbb{P}_{outage}^T for $\theta \geq 1$ can be upper bounded*

as

$$\mathbb{P}_{outage}^T \approx F_W\left(\frac{\theta\sigma^2}{\epsilon P_G}\right) < \mathbb{P}_{outage}^{Tu2} = \tilde{A}_0\left(\frac{\theta\sigma^2}{\epsilon P_G}\right) + \tilde{C}_0\left(\frac{\theta\sigma^2}{\epsilon P_G}\right) - 1, \quad (6.31)$$

where

$$\tilde{A}_0(u) = \frac{1}{\left(1 + \frac{1}{\sqrt{2\theta}}\right)^2} \left[1 - \left(1 + \pi\rho\sqrt{\frac{\theta}{u}}\left(1 + \frac{1}{\sqrt{2\theta}}\right)\right) e^{-\pi\rho\sqrt{\frac{\theta}{u}}\left(1 + \frac{1}{\sqrt{2\theta}}\right)} \right] + e^{-\frac{\pi\rho}{\sqrt{2u}}} \left(1 + \pi\rho\sqrt{\frac{\theta}{u}}\right) e^{-\pi\rho\sqrt{\frac{\theta}{u}}}, \quad (6.32)$$

and

$$\begin{aligned} \tilde{C}_0(u) &= \frac{1}{1 + \sqrt{\frac{1}{2\theta}}} \left[1 - e^{-\pi\rho\sqrt{\frac{\theta}{u}}\left(1 + \frac{1}{\sqrt{2\theta}}\right)} \right] + \frac{1}{\sqrt{\theta}} \frac{1}{\left(1 + \frac{1}{\sqrt{2\theta}}\right)^2} \left\{ 1 - \left(1 + \pi\rho\sqrt{\frac{\theta}{u}}\left(1 + \frac{1}{\sqrt{2\theta}}\right)\right) e^{-\pi\rho\sqrt{\frac{\theta}{u}}\left(1 + \frac{1}{\sqrt{2\theta}}\right)} \right\} \\ &+ \left(1 + \frac{\pi\rho}{\sqrt{u}}\right) e^{-\frac{\pi\rho}{\sqrt{2u}}} e^{-\pi\rho\sqrt{\frac{\theta}{u}}}. \end{aligned} \quad (6.33)$$

Proof. Suppose Q_k^{-2} ($k > 2$) are much smaller than Q_1^{-2} and Q_2^{-2} , then in Eq.(6.12) we neglect the variable Z :

$$\mathbb{P}_{outage}^T \approx \mathbb{P}(W < \frac{\theta\sigma^2}{\epsilon P_G}) = F_W\left(\frac{\theta\sigma^2}{\epsilon P_G}\right). \quad (6.34)$$

Furthermore, from (6.22), (6.23) and (6.25) we can find upper bounds for the integral A_0 and C_0 :

$$\begin{aligned} A_0(u) &< \tilde{A}_0(u) = (\pi\rho)^2 \left\{ \int_0^{\sqrt{\frac{\theta}{u}}} s e^{-\pi\rho s\left(1 + \frac{1}{\sqrt{2\theta}}\right)} ds + \int_{\sqrt{\frac{\theta}{u}}}^{\infty} s e^{-\pi\rho s\left(1 + \frac{1}{\sqrt{2us^2}}\right)} ds \right\} \\ &\stackrel{k=\pi\rho\left(1 + \frac{1}{\sqrt{2\theta}}\right)s, l=\pi\rho s}{=} \frac{1}{\left(1 + \frac{1}{\sqrt{2\theta}}\right)^2} \int_0^{\pi\rho\sqrt{\frac{\theta}{u}}\left(1 + \frac{1}{\sqrt{2\theta}}\right)} k e^{-k} dk + e^{-\frac{\pi\rho}{\sqrt{2u}}} \int_{\pi\rho\sqrt{\frac{\theta}{u}}}^{\infty} l e^{-l} dl \\ &= \frac{1}{\left(1 + \frac{1}{\sqrt{2\theta}}\right)^2} \left[1 - \left(1 + \pi\rho\sqrt{\frac{\theta}{u}}\left(1 + \frac{1}{\sqrt{2\theta}}\right)\right) e^{-\pi\rho\sqrt{\frac{\theta}{u}}\left(1 + \frac{1}{\sqrt{2\theta}}\right)} \right] + e^{-\frac{\pi\rho}{\sqrt{2u}}} \left(1 + \pi\rho\sqrt{\frac{\theta}{u}}\right) e^{-\pi\rho\sqrt{\frac{\theta}{u}}}, \end{aligned} \quad (6.35)$$

and

$$\begin{aligned}
C_0(u) &< \tilde{C}_0(u) = \pi\rho \int_0^{\sqrt{\frac{\theta}{u}}} e^{-\pi\rho s(1+\frac{1}{\sqrt{2\theta}})} \left(1 + \pi\rho s \frac{1}{\sqrt{\theta}}\right) ds + \pi\rho \int_{\sqrt{\frac{\theta}{u}}}^{\infty} e^{-\pi\rho s(1+\frac{1}{\sqrt{2us^2}})} \left(1 + \pi\rho s \frac{1}{\sqrt{us^2}}\right) ds \\
&= \pi\rho \int_0^{\sqrt{\frac{\theta}{u}}} e^{-\pi\rho(1+\frac{1}{\sqrt{2\theta}})s} ds + \frac{(\pi\rho)^2}{\sqrt{\theta}} \int_0^{\sqrt{\frac{\theta}{u}}} s e^{-\pi\rho s(1+\frac{1}{\sqrt{2\theta}})} ds + \pi\rho \left(1 + \frac{\pi\rho}{\sqrt{u}}\right) e^{-\frac{\pi\rho}{\sqrt{2u}}} \int_{\sqrt{\frac{\theta}{u}}}^{\infty} e^{-\pi\rho s} ds \\
&= \frac{1}{1 + \sqrt{\frac{1}{2\theta}}} \left[1 - e^{-\pi\rho\sqrt{\frac{\theta}{u}}(1+\frac{1}{\sqrt{2\theta}})}\right] + \frac{1}{\sqrt{\theta}} \frac{1}{(1 + \frac{1}{\sqrt{2\theta}})^2} \left\{1 - \left(1 + \pi\rho\sqrt{\frac{\theta}{u}}\left(1 + \frac{1}{\sqrt{2\theta}}\right)\right) e^{-\pi\rho\sqrt{\frac{\theta}{u}}(1+\frac{1}{\sqrt{2\theta}})}\right\} \\
&+ \left(1 + \frac{\pi\rho}{\sqrt{u}}\right) e^{-\frac{\pi\rho}{\sqrt{2u}}} e^{-\pi\rho\sqrt{\frac{\theta}{u}}}.
\end{aligned} \tag{6.36}$$

Therefore we reach Theorem 10

$$\mathbb{P}_{outage}^T \approx F_W\left(\frac{\theta\sigma^2}{\epsilon P_G}\right) < \tilde{A}_0\left(\frac{\theta\sigma^2}{\epsilon P_G}\right) + \tilde{C}_0\left(\frac{\theta\sigma^2}{\epsilon P_G}\right) - 1, \tag{6.37}$$

□

6.2.2 Power Outage Probability

The power outage probability is defined as $\mathbb{P}_{outage}^P = \mathbb{P}(P_H < P_D)$. Then we derive an upper bound for the power outage probability in closed form as given in Theorem 11:

Theorem 11. *The power outage probability of IoT nodes can be upper bounded as*

$$\mathbb{P}_{outage}^P < \mathbb{P}_{outage}^{Pu} = e^{-\pi\rho\sqrt{\frac{\sqrt{P_O^2 + 4C_{TX}\eta(1-\epsilon)P_G - P_O}}{2C_{TX}}}}. \tag{6.38}$$

Proof. The power outage probability is

$$\begin{aligned}
\mathbb{P}_{outage}^P &= \mathbb{P}(P_H < P_D) \\
&= \mathbb{P}[\eta(1 - \epsilon) \sum_k P_G Q_k^{-2} < P_O + C_{TX} \{\min_k Q_k\}^2] \\
&= \mathbb{P}(\eta(1 - \epsilon) P_G \sum_k Q_k^{-2} - C_{TX} \{\min_k Q_k\}^2 < P_O).
\end{aligned} \tag{6.39}$$

Since $\min_k Q_k < Q_1$, we can find an upper bound of the power outage probability:

$$\begin{aligned}
\mathbb{P}_{outage}^P &< \mathbb{P}(\eta(1 - \epsilon) P_G Q_1^{-2} - C_{TX} Q_1^2 < P_O) \\
&= \mathbb{P}(C_{TX} Q_1^4 + P_O Q_1^2 - \eta(1 - \epsilon) P_G > 0) \\
&= \mathbb{P}(C_{TX} V_1^2 + P_O V_1 - \eta(1 - \epsilon) P_G > 0) \\
&= \mathbb{P}(V_1 > \frac{\sqrt{P_O^2 + 4C_{TX}\eta(1 - \epsilon)P_G} - P_O}{2C_{TX}}) \\
&= 1 - F_{V_1}(\frac{\sqrt{P_O^2 + 4C_{TX}\eta(1 - \epsilon)P_G} - P_O}{2C_{TX}}) \\
&= e^{-\pi\rho\sqrt{\frac{\sqrt{P_O^2 + 4C_{TX}\eta(1 - \epsilon)P_G} - P_O}{2C_{TX}}}}.
\end{aligned} \tag{6.40}$$

In the above derivation we introduce $V_1 = Q_1^2$. Its CDF is

$$F_{V_1}(x) = F_{Q_1}(\sqrt{x}) = 1 - e^{-\pi\rho\sqrt{x}} \tag{6.41}$$

□

6.3 Convexity of the Maximum Outage Probability

Upper Bound

Based on the closed-form results we present in Section 6.2, one may choose to balance the trade-off between transmission outage probability and power outage probability by careful design of the power split ratio ϵ . Therefore, it would be interesting to identify the optimal ϵ^* to minimize the maximum outage probability upper bound, i.e., $\max\{P_{outage}^{Tu2}(\epsilon), P_{outage}^{Pu}(\epsilon)\}$.

Theorem 12. *The maximum outage probability upper bound $F(\epsilon) = \max\{P_{outage}^{Tu2}(\epsilon), P_{outage}^{Pu}(\epsilon)\}$ is a convex function of ϵ .*

Proof. According to [105] (Section 3.2.3 in page 80), if $P_{outage}^{Tu2}(\epsilon)$ and $P_{outage}^{Pu}(\epsilon)$ are both convex functions of ϵ , then $F(\epsilon) = \max\{P_{outage}^{Tu2}(\epsilon), P_{outage}^{Pu}(\epsilon)\}$ is also convex.

We know that $P_O > 0$, $C_{TX} > 0$, $\eta > 0$, $P_G > 0$ and $0 < \epsilon < 1$. First, we prove that $\mathbb{P}_{outage}^{Tu2}$ is convex. We can derive that

$$\begin{aligned} \mathbb{P}_{outage}^{Tu2} &= \tilde{A}_0\left(\frac{\theta\sigma^2}{\epsilon P_G}\right) + \tilde{C}_0\left(\frac{\theta\sigma^2}{\epsilon P_G}\right) - 1 \\ &= f_T(\epsilon) + \frac{1}{1 + \frac{1}{\sqrt{2\theta}}} + \frac{1 + \frac{1}{\sqrt{\theta}}}{\left(1 + \frac{1}{\sqrt{2\theta}}\right)^2} - 1 \end{aligned} \quad (6.42)$$

where

$$f_T(\epsilon) = (A_1 + B_1\sqrt{\epsilon})e^{-C_1\sqrt{\epsilon}}, \quad (6.43)$$

$$A_1 = \frac{\frac{1}{\theta} + \frac{1}{\sqrt{\theta}}(2\sqrt{2} - 1 - \frac{1}{\sqrt{2}})}{\left(1 + \frac{1}{\sqrt{2\theta}}\right)^2} > 0, \quad (6.44)$$

$$B_1 = \pi\rho\sqrt{\frac{P_G}{\sigma^2}}\left(1 + \frac{1}{\sqrt{\theta}}\right)\left(\frac{\frac{1}{\sqrt{2\theta}}}{1 + \frac{1}{\sqrt{2\theta}}}\right) > 0, \quad (6.45)$$

and

$$C_1 = \pi\rho\sqrt{\frac{P_G}{\sigma^2}}\left(1 + \frac{1}{\sqrt{2\theta}}\right) > 0. \quad (6.46)$$

To prove $\mathbb{P}_{outage}^{Tu2}$ to be convex is equivalent to prove that the second order derivative

$$f_T''(\epsilon) > 0,$$

$$f_T''(\epsilon) = e^{-C_1\sqrt{\epsilon}} \left(A_1 + \frac{C_1}{2\sqrt{\epsilon}} \left(\frac{B_1 C_1}{2} - A_1 \right) + \left(\frac{C_1}{4\epsilon} + \frac{1}{4}\epsilon^{-3/2} \right) (A_1 C_1 - B_1) \right). \quad (6.47)$$

On one hand, when $\epsilon \rightarrow 0$, we can get

$$f_T''(\epsilon) \approx e^{-C_1\sqrt{\epsilon}} \frac{A_1 C_1 - B_1}{4\epsilon^{3/2}}. \quad (6.48)$$

We can prove that

$$\begin{aligned} A_1 C_1 - B_1 &= \pi\rho\sqrt{\frac{P_G}{\sigma^2}} \left(\frac{\frac{1}{\theta} + \frac{1}{\sqrt{\theta}}(2\sqrt{2} - 1 - \frac{1}{\sqrt{2}})}{1 + \frac{1}{\sqrt{2\theta}}} - \frac{(1 + \frac{1}{\sqrt{\theta}})\frac{1}{\sqrt{2\theta}}}{1 + \frac{1}{\sqrt{2\theta}}} \right) \\ &= \frac{\pi\rho\sqrt{\frac{P_G}{\sigma^2}}}{1 + \frac{1}{\sqrt{2\theta}}} \left(\frac{1}{\theta} \left(1 - \frac{1}{\sqrt{2}} \right) + \frac{1}{\sqrt{\theta}} (\sqrt{2} - 1) \right) > 0. \end{aligned} \quad (6.49)$$

Therefore, when $\epsilon \rightarrow 0$,

$$f_T''(\epsilon) > 0. \quad (6.50)$$

On the other hand, when $\epsilon \rightarrow \infty$,

$$f_T''(\epsilon) \approx e^{-C_1\sqrt{\epsilon}} A_1 > 0. \quad (6.51)$$

Solving $f_T''(\epsilon) = 0$, we can get two imaginary roots and a real root. Therefore, combined with (6.50) and (6.51), we can conclude that $f_T''(\epsilon) > 0$ when $\epsilon \in (0, 1)$. That is, the transmission outage probability $\mathbb{P}_{outage}^{Tu2}$ is proved to be convex.

Secondly, according to [105] (Example 3.13 on page 86), if g is convex then $\exp g(x)$ is convex. Therefore, to prove the power outage probability upper bound

$$\mathbb{P}_{outage}^{Pu} = e^{-\pi\rho\sqrt{\frac{\sqrt{P_O^2 + 4C_{TX}\eta(1-\epsilon)P_G - P_O}}{2C_{TX}}}} \quad (6.52)$$

to be convex is equivalent to prove

$$f_P(\epsilon) = \sqrt{\sqrt{P_O^2 + 4C_{TX}\eta(1-\epsilon)P_G} - P_O} \quad (6.53)$$

to be concave.

We can calculate the second order derivative of $f_P(\epsilon)$ as

$$\begin{aligned} f_P''(\epsilon) = & - \frac{C_{TX}^2 \eta^2 P_G^2}{(\sqrt{P_O^2 + 4C_{TX}\eta P_G(1-\epsilon)} - P_O)^{3/2} (P_O^2 + 4C_{TX}\eta P_G(1-\epsilon))} \\ & - \frac{2C_{TX}^2 \eta^2 P_G^2}{\sqrt{\sqrt{P_O^2 + 4C_{TX}\eta P_G(1-\epsilon)} - P_O} (P_O^2 + 4C_{TX}\eta P_G(1-\epsilon))^{3/2}} < 0. \end{aligned} \quad (6.54)$$

Then $\mathbb{P}_{outage}^{Pu}(\epsilon)$ is proved to be convex.

Therefore, we can prove that $\max\{P_{outage}^{Tu2}(\epsilon), P_{outage}^{Pu}(\epsilon)\}$ is a convex function of ϵ . \square

Based on Theorem 12, one can easily find the optimal power split ratio ϵ^* to minimize the maximum outage probability upper bound.

6.4 Simulation Results

In this section, we present simulation results to validate the derived analytical results. The primary simulation parameters used are listed in Table 6.1 [72].

Fig. 6.3 and Fig. 6.4 show that the transmission outage probability can be reduced by increasing ϵ (i.e., a larger fraction of power is sent into the information decoder) or increasing the transmit power of gateways P_G , which is reasonable intuitive. The results verify that the transmission outage probability is well bounded by the closed-form bounds we provide in (6.31). Although they do not appear to be tight, an explicit and closed-form expression for the transmission outage probability is always desirable. On the other hand, the numerical calculable upper bound we provide for the transmission outage probability in (6.6) is rather tight.

Table 6.1: *Simulation parameters*

Simulation Parameter	Value
Harvesting efficiency (η)	0.9
Operation power of nodes (P_O)	0.03 W
Transmit power of gateways (P_G)	0.25 W
Power Split Ratio (ϵ)	0.2
Proportionality constant for transmit power (C_{TX})	0.01
Noise power (σ^2)	0.01 W
GPP Density (ρ)	0.3
SINR threshold (θ)	0 dB

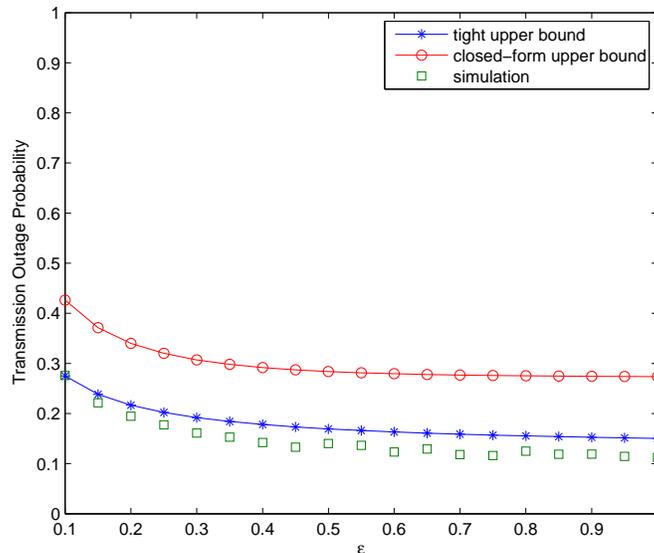


Figure 6.3: ϵ vs. P_{outage}^T

Fig. 6.5 and Fig. 6.6 show that the power outage probability can be reduced by decreasing ϵ (i.e., a larger fraction of power is sent into the energy harvester) or increasing the transmit power of gateways P_G , which is expected intuitively. The results verify that the power outage probability is well bounded by the closed-form bounds we provide in (6.38). This bound appears tighter when ϵ gets larger (i.e., a smaller fraction of power is sent into the energy harvester) or P_G gets smaller.

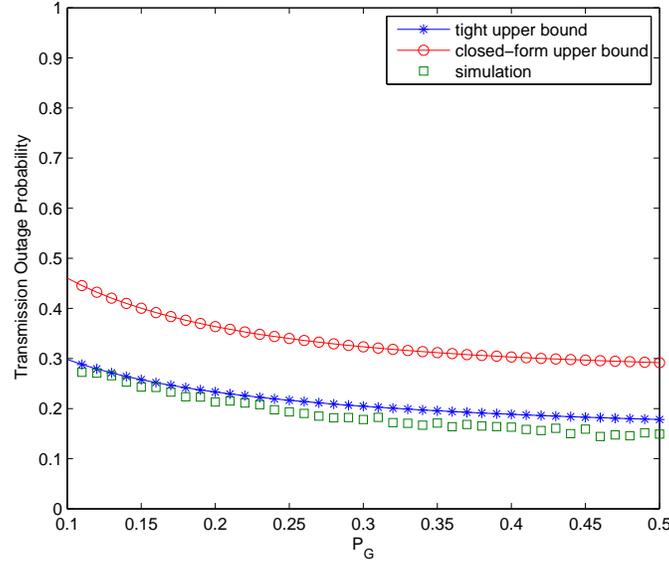


Figure 6.4: P_G vs. P_{outage}^T

In Fig. 6.7 - Fig. 6.9 we investigate the impacts of θ , P_G and ρ on the maximum outage probability upper bound, respectively.

From Fig. 6.7, we can see that the optimal ϵ^* increases with θ . In other words, a larger fraction of power should be sent into the information decoder to meet the SINR threshold, which is expected.

From Fig. 6.8, we can see that as P_G increases, ϵ^* first decreases and then increases. When P_G is small, the SINR of IoT node is significantly reduced by relatively strong noise. Therefore, a larger fraction of power should be fed into the information decoder as P_G gets smaller. In other words, ϵ^* decreases as P_G increases. On the other hand, when P_G is above a certain value, the noise power becomes negligible. As P_G increases, both the desired signal and interference increase according to (6.4), i.e., the power outage probability is brought down by increasing P_G much faster than the transmission outage probability. Therefore, to minimize the maximum outage probability, a larger fraction of power should be sent into the information decoder as P_G increases.

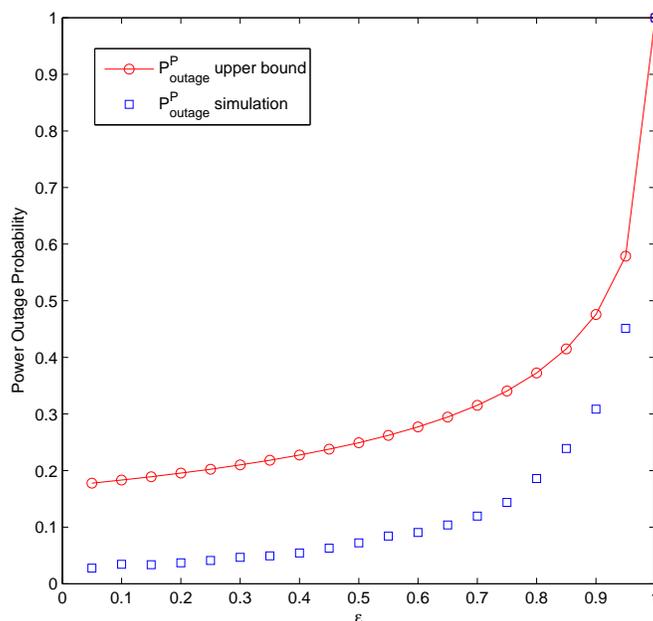


Figure 6.5: ϵ vs. P_{outage}^P

Similarly, from Fig. 6.9, we can see that as ρ increases, ϵ^* first decreases and then increases. When ρ is small, the SINR of IoT node is significantly reduced by relatively strong noise. Therefore, a larger fraction of power should be fed into the information decoder as ρ gets smaller. In other words, ϵ^* decreases as ρ increases. On the other hand, when ρ is above a certain value, the noise power becomes negligible. As ρ increases, the expected value of Q_k decreases, then both the desired signal and interference increase according to (6.4), i.e., the power outage probability is brought down by increasing ρ much faster than the transmit outage probability. Therefore, to minimize the maximum outage probability, a larger fraction of power should be sent into the information decoder as ρ increases.

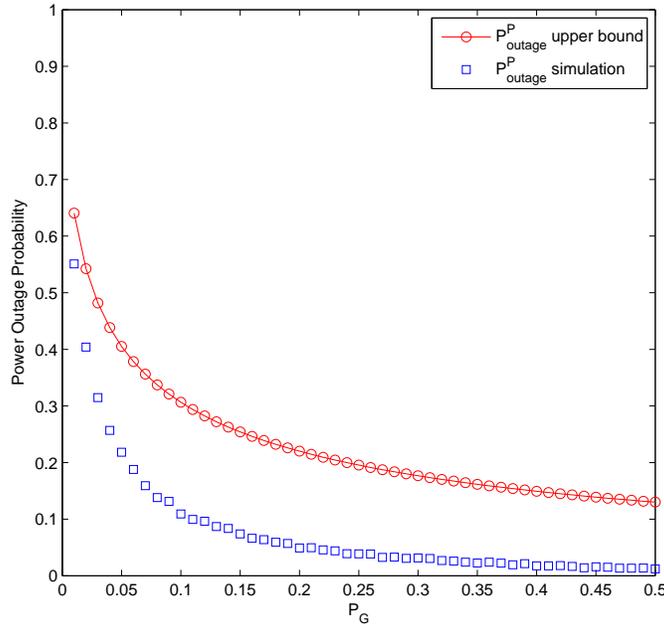


Figure 6.6: P_G vs. P_{outage}^P

6.5 Summary

This chapter focuses on the modeling and analysis of simultaneous information and energy transfer in Internet of Things. To capture the repulsion between nodes, the IoT networks are modeled according to a Ginibre point process. Based on that, we derive: (1) a closed-form upper bounds for the power outage probability, (2) a closed-form upper bound and a tight numerically calculable upper bound for the transmission outage probability. Then to identify the optimal power split ratio to balance the trade-off between outage probabilities, we prove the convexity of the maximum outage probability upper bounds. We also analyze the effects of network parameters (e.g., density and transmit power of gateways, power split ratio and SINR threshold) on the outage performances.

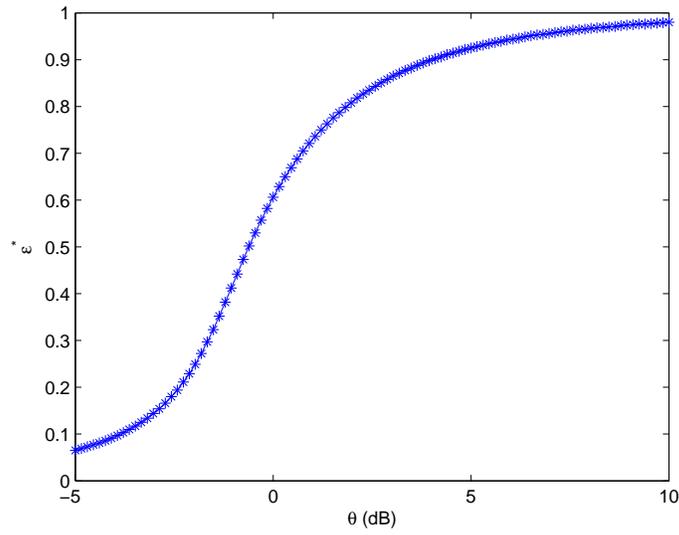


Figure 6.7: ϵ^* for Maximum outage probability vs. different θ

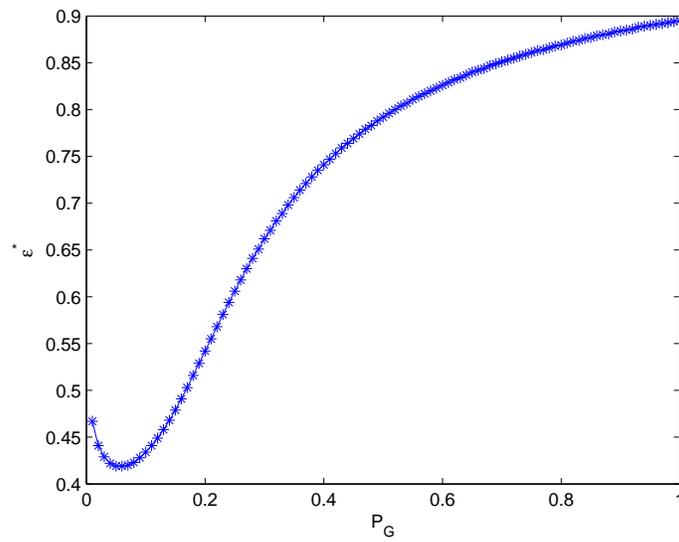


Figure 6.8: ϵ^* for Maximum outage probability vs. different P_G

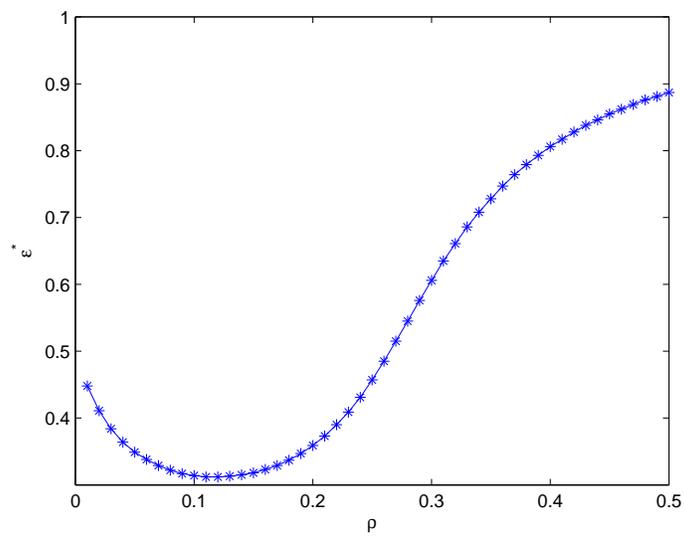


Figure 6.9: ϵ for *Maximum outage probability vs. different ρ*

Chapter 7

Conclusion

In this chapter, we provide concluding remarks on the results of our inquiries, explain how the loose ends will wrap up, and propose possible future research directions based on this work.

7.1 Conclusion

This dissertation has addressed the energy efficiency improvement of 5G networks via several key techniques. Specifically, four fundamental research questions are identified based on the techniques and architectural enhancement of 5G networks: (1) small cells, (2) energy harvesting base stations, (3) device-to-device communication, and (4) Internet of Things. We summarize the key research contributions of this dissertation below:

- We design small cell BS sleep modes and sleep strategies for HCNs. Using a stochastic geometry based model, we derive coverage probability, average achievable throughput, and EE in K-tier HCNs. Then we formulate EE maximization problem under either random sleeping policy or strategic sleeping policy and determine optimal operating probability for each sleep mode of small cell BSs. Numerical results confirm the effectiveness of the scheme. Apart from improvements of approximately 30% in EE

with random sleeping policy, the simulation indicates that instantaneous EE can be further improved by 15% with a strategic sleeping policy.

- We investigate, for the first time, management of energy harvesting BSs in HCNs from a techno-economic point of view. We consider real-time electricity price, QoS of users and harvested energy profile in identifying methods to minimize energy cost. To solve the formulated problem in a simple and practical manner, we divide it into two subproblems and sequentially solve them via an optimization approach and a control algorithm, respectively. The proposed two-stage optimization/control approach provides a method to manage both transmit power and stored energy usage of HCN BSs to reduce on-grid electricity expenditure. Numerical results confirm the effectiveness of the scheme.
- We study the performance evaluation and maximization of D2D underlay cellular networks. First of all, we aim to find a critical set of system parameters to ensure the benefits of D2D underlay operation to outweigh its drawbacks in the context of trade-off between ergodic capacity and power consumption. To achieve that, we propose a network model to understand the behavior of D2D underlay cellular networks using tools from stochastic geometry. Unlike prior efforts, D2D direct links are modeled by Rician fading and the distances between D2D pairs are treated as random quantities. For a special case of path loss exponent of 4, our work derives (1) tight lower and upper bounds for ergodic capacity of a cellular network, and (2) recursive closed-form expressions and closed-form approximation for ergodic capacity of D2D networks. It is shown that our analytical results better approximate the actual capacity compared with prior approaches. Then we investigate ergodic capacity maximization and the impact of both transmit power and density of D2D users. A two-phase scheme is proposed to maximize the ergodic capacity of D2D underlay cellular networks as well as to minimize the overall power consumption. We evaluate both ergodic capacity and total power consumption of the optimal situation comparing them with some extreme

cases as well. Secondly, we analyze the average achievable throughput of D2D underlay cellular networks. Assuming a path loss exponent of 4, we derive closed-form results including: (1) tight upper and tight lower bounds for average achievable throughput of cellular networks, and (2) good approximation for average achievable throughput of D2D networks. Finally, we provide closed-form ergodic capacity results (when path loss exponent is 4) for two other cases of the distribution of DRU: (1) distance between a DU pair follows a uniform distribution and (2) a DRU is distributed uniformly in the circular area around its serving DTU.

- We investigate the modeling and analysis of simultaneous information and energy transfer in Internet of Things. To capture the repulsion between nodes, the IoT networks are modeled according to a Ginibre point process. Based on that, we derive: (1) a closed-form upper bounds for the power outage probability, (2) a closed-form upper bound and a tight numerically calculable upper bound for the transmission outage probability. Then to identify the optimal power split ratio to balance the trade-off between outage probabilities, we prove the convexity of the maximum outage probability upper bounds. We also analyze the effects of some factors (e.g., density and transmit power of gateways, power splitting ratio, and SINR threshold) on the outage performances.

Based on the research accomplished in this dissertation, some future research directions are highlighted in the next section.

7.2 Future Work

In this section, we present possible future research direction in energy efficient design of 5G networks. We first propose extensions to, and more complete explorations, of our existing research:

- In Chapter 3, the sleep mode operation applies to a single small cell tier. BSs of other tiers are assumed to be fully operational all the time. In the future, this work may be extended to simultaneous sleep mode optimization for multiple tiers.
- Recently several projects show that it is technically feasible to have self-powered BSs. For example, China Mobile has established over 800 BSs in Tibet powered solely by solar energy [127]. Nokia-Siemens has also built renewable energy powered BS sites in Germany [112]. Self-powered BS management has attracted a lot of attention in academia as well [128–130]. Therefore, the work in Chapter 4 might be extended to self-powered HCNs.
- The results from Chapter 5 may also be used in a wide range of endeavors related to power control, mode selection and resource allocation in D2D underlay cellular networks. Besides, one interesting extension would be how to deal with the node mobility problem. Since different mobility will actually influence the capacity and system throughput, the mobility model into the analysis could be included.
- In Chapter 6 only single-tier gateways is considered. The modeling and analysis of SIET in multi-tier gateway networks can be an interesting extension of this work. Also, when using mobile equipment as gateways, how to capture the dynamical properties about the system would be interesting.

Next, we present some new areas of investigation in 5G networks. A subset of possible future research works is discussed below.

- Relay enhanced cellular networks is a key enabling infrastructure redesign to enhance the performance of 5G networks. In 5G, advanced cellular relay can: (1) improve the cellular topology, (2) improve the network robustness, and (3) decrease power consumption. A multi-hop structure can efficiently support tremendous access of D2D terminals. Therefore, relay deployment from the perspective of energy efficiency is an interesting problem which has not been deeply addressed.

- Massive MIMO has been recognized as a promising technology for 5G networks. Since massive MIMO employs a large number of antennas to serve a much smaller number of users, it can enhance energy efficiency by exploiting a large array gain. However, the circuit power consumption of massive MIMO scales with the tremendous number of antennas. Recently, the performance of massive MIMO on energy saving has drawn significant attention and is still in its infancy.
- Cognitive cellular networks employ cognitive radio to lease additional spectrum outside the licensed cellular bands, which is a key enabling technology to tackle challenges in 5G cellular networks such as unevenly distributed and increasingly diverse mobile traffic. To jointly utilize licensed and cognitive radio resources to optimize the overall system performance, understanding the trade-off between energy efficiency and spectral efficiency in cognitive cellular networks is of fundamental importance for system design and optimization.

Bibliography

- [1] Cisco, *Visual Networking Index*, white paper at Cisco.com, Feb. 2014.
- [2] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, “Cellular architecture and key technologies for 5g wireless communication networks,” *Communications Magazine, IEEE*, vol. 52, no. 2, pp. 122–130, 2014.
- [3] M. Etoh, T. Ohya, and Y. Nakayama, “Energy consumption issues on mobile network systems,” in *Applications and the Internet, 2008. SAINT 2008. International Symposium on*, 2008, pp. 365–368.
- [4] J. Louhi, “Energy efficiency of modern cellular base stations,” in *Telecommunications Energy Conference, 2007. INTELEC 2007. 29th International*, 2007, pp. 475–476.
- [5] Z. Niu, Y. Wu, J. Gong, and Z. Yang, “Cell zooming for cost-efficient green cellular networks,” *Communications Magazine, IEEE*, vol. 48, no. 11, pp. 74–79, 2010.
- [6] J. Dong, Z. Ou, A. Yla-Jaaski, and Y. Cui, “Mobile hotspots cooperation towards better energy efficiency,” in *Globecom Workshops (GC Wkshps), 2013 IEEE*, 2013, pp. 760–765.
- [7] “IEEE standard for local and metropolitan area networks part 16: Air interface for fixed and mobile broadband wireless access systems amendment 2: Physical and medium access control layers for combined fixed and mobile operation in licensed bands and corrigendum 1,” *IEEE Std 802.16e-2005 and IEEE Std 802.16-2004/Cor 1-2005 (Amendment and Corrigendum to IEEE Std 802.16-2004)*, pp. 801–822, 2006.

- [8] *Public Safety Tech Topic #23 - Femtocells*, <http://www.fcc.gov/help/public-safety-tech-topic-23-femtocells>.
- [9] 3GPP, “Potential solutions for energy saving for E-UTRAN (Release 10),” 3rd Generation Partnership Project (3GPP), TS 36.927 V2.0.0, 2011.
- [10] Y. S. Chen and C. Y. Wu, “A green handover protocol in two-tier ofdma macrocell-femtocell networks,” *Mathematical and Computer Modelling*, pp. 1–18, 2012.
- [11] I. Ashraf, F. Boccardi, and L. Ho, “Sleep mode techniques for small cell deployments,” *Communications Magazine, IEEE*, vol. 49, no. 8, pp. 72–79, 2011.
- [12] —, “Power savings in small cell deployments via sleep mode techniques,” in *Personal, Indoor and Mobile Radio Communications Workshops (PIMRC Workshops), 2010 IEEE 21st International Symposium on*, 2010, pp. 307–311.
- [13] M. Arshad, A. Vastberg, and T. Edler, “Energy efficiency gains through traffic offloading and traffic expansion in joint macro pico deployment,” in *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*, 2012, pp. 2203–2208.
- [14] Y. Li, H. Celebi, M. Daneshmand, C. Wang, and W. Zhao, “Energy-efficient femtocell networks: challenges and opportunities,” *Wireless Communications, IEEE*, vol. 20, no. 6, pp. 99–105, December 2013.
- [15] M. Wildemeersch, T. Quek, C. Slump, and A. Rabbachin, “Cognitive small cell networks: Energy efficiency and trade-offs,” pp. 1–14, 2013.
- [16] W. Guo and T. O’Farrell, “Green cellular network: Deployment solutions, sensitivity and tradeoffs,” in *Wireless Advanced (WiAd), 2011*, 2011, pp. 42–47.
- [17] W. Li, W. Zheng, Y. Xie, and X. Wen, “Clustering based power saving algorithm for self-organized sleep mode in femtocell networks,” in *Wireless Personal Multimedia*

- Communications (WPMC), 2012 15th International Symposium on*, 2012, pp. 379–383.
- [18] L. Falconetti, P. Frenger, H. Kallin, and T. Rimhagen, “Energy efficiency in heterogeneous networks,” in *Online Conference on Green Communications (GreenCom), 2012 IEEE*, 2012, pp. 98–103.
- [19] L. Saker, S.-E. Elayoubi, R. Combes, and T. Chahed, “Optimal control of wake up mechanisms of femtocells in heterogeneous networks,” *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 3, pp. 664–672, 2012.
- [20] R. Torrea-Duran, C. Desset, S. Pollin, and A. Dejonghe, “Adaptive energy efficient scheduling algorithm for lte pico base stations,” in *Future Network Mobile Summit (FutureNetw), 2012*, 2012, pp. 1–8.
- [21] L. Falconetti, L. Hevizi, and I. Godor, “Sleep mode control for low power nodes in heterogeneous networks,” in *Wireless Communication Systems (ISWCS 2013), Proceedings of the Tenth International Symposium on*, Aug 2013, pp. 1–5.
- [22] S. Morosi, P. Piunti, and E. Del Re, “Improving cellular network energy efficiency by joint management of sleep mode and transmission power,” in *Digital Communications - Green ICT (TIWDC), 2013 24th Tyrrhenian International Workshop on*, Sept 2013, pp. 1–6.
- [23] S. Wang and W. Guo, “Energy and cost implications of a traffic aware and quality-of-service constrained sleep mode mechanism,” *Communications, IET*, vol. 7, no. 18, pp. 2092–2101, December 2013.
- [24] S. Mukherjee and H. Ishii, “Energy efficiency in the phantom cell enhanced local area architecture,” in *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*, 2013, pp. 1267–1272.

- [25] Y. S. Soh, T. Q. Quek, M. Kountouris, and H. Shin, "Energy efficient heterogeneous cellular networks," *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 5, pp. 840–850, 2013.
- [26] "IEEE standard for local and metropolitan area networks part 16: Air interface for broadband wireless access systems amendment 3: Advanced air interface," *IEEE Std 802.16m-2011(Amendment to IEEE Std 802.16-2009)*, pp. 1–1112, 2011.
- [27] G. Wu, G. Feng, and S. Qin, "Cooperative sleep-mode and performance modeling for heterogeneous mobile network," in *Wireless Communications and Networking Conference Workshops (WCNCW), 2013 IEEE*, 2013, pp. 6–11.
- [28] W. Vereecken, L. Haratcherev, M. Deruyck, W. Joseph, M. Pickavet, L. Martens, and P. Demeester, "The effect of variable wake up time on the utilization of sleep modes in femtocell mobile access networks," in *Wireless On-demand Network Systems and Services (WONS), 2012 9th Annual Conference on*, 2012, pp. 63–66.
- [29] "Sustainable energy use in mobile communications," [EricssonInc., WhitePaper](#), 2007.
- [30] "E-plus, nokia siemens networks build germany's first off-grid base station," <http://www.nokiasiemensnetworks.com>, 2011.
- [31] W. Yu and X. Qian, "Design of 3kw wind and solar hybrid independent power supply system for 3g base station," in *Knowledge Acquisition and Modeling, 2009. KAM '09. Second International Symposium on*, vol. 3, 2009, pp. 289–292.
- [32] P. Nema, S. Rangnekar, and R. Nema, "Pre-feasibility study of pv-solar / wind hybrid energy system for gsm type mobile telephony base station in central india," in *Computer and Automation Engineering (ICCAE), 2010 The 2nd International Conference on*, vol. 5, 2010, pp. 152–156.

- [33] S. Hashimoto, M. Nitta, T. Tani, and T. Yachi, “A stand-alone wind turbine generator system for a small-scale radio base station,” in *Telecommunications Energy Conference, 2003. INTELEC '03. The 25th International*, 2003, pp. 404–409.
- [34] S. Hashimoto, T. Yachi, and T. Tani, “A new stand-alone hybrid power system with wind generator and photovoltaic modules for a radio base station,” in *Telecommunications Energy Conference, 2004. INTELEC 2004. 26th Annual International*, 2004, pp. 254–259.
- [35] H. Huang and V. Lau, “Decentralized delay optimal control for interference networks with limited renewable energy storage,” *Signal Processing, IEEE Transactions on*, vol. 60, no. 5, pp. 2552–2561, May 2012.
- [36] C. K. Ho and R. Zhang, “Optimal energy allocation for wireless communications with energy harvesting constraints,” *Signal Processing, IEEE Transactions on*, vol. 60, no. 9, pp. 4808–4818, Sept 2012.
- [37] K. Tutuncuoglu and A. Yener, “Sum-rate optimal power policies for energy harvesting transmitters in an interference channel,” *Communications and Networks, Journal of*, vol. 14, no. 2, pp. 151–161, April 2012.
- [38] —, “Optimum transmission policies for battery limited energy harvesting nodes,” *Wireless Communications, IEEE Transactions on*, vol. 11, no. 3, pp. 1180–1189, March 2012.
- [39] L. Huang and M. Neely, “Utility optimal scheduling in energy-harvesting networks,” *Networking, IEEE/ACM Transactions on*, vol. 21, no. 4, pp. 1117–1130, Aug 2013.
- [40] H. Dhillon, R. Ganti, and J. Andrews, “Load-aware modeling and analysis of heterogeneous cellular networks,” *Wireless Communications, IEEE Transactions on*, vol. 12, no. 4, pp. 1666–1677, 2013.

- [41] J. Zhou, M. Li, L. Liu, X. She, and L. Chen, “Energy source aware target cell selection and coverage optimization for power saving in cellular networks,” in *Green Computing and Communications (GreenCom), 2010 IEEE/ACM Int’l Conference on Int’l Conference on Cyber, Physical and Social Computing (CPSCoM)*, 2010, pp. 1–8.
- [42] B. Lindemark and G. Oberg, “Solar power for radio base station (rbs) sites applications including system dimensioning, cell planning and operation,” in *Telecommunications Energy Conference, 2001. INTELEC 2001. Twenty-Third International*, 2001, pp. 587–590.
- [43] J. Yang and S. Ulukus, “Optimal packet scheduling in an energy harvesting communication system,” *Communications, IEEE Transactions on*, vol. 60, no. 1, pp. 220–230, 2012.
- [44] J. Gong, J. Thompson, and Z. S. Niu, “Base station sleeping and resource allocation in renewable energy powered cellular networks,” *Wireless Communications, IEEE Transactions on*, 2013.
- [45] T. Han and N. Ansari, “Heuristic relay assignments for green relay assisted device to device communications,” in *Globecom 2013 - Ad Hoc and Sensor Networking Symposium*, Dec 2013, pp. 1–5.
- [46] —, “Ice: Intelligent cell breathing to optimize the utilization of green energy,” *Communications Letters, IEEE*, vol. 16, no. 6, pp. 866–869, 2012.
- [47] —, “Optimizing cell size for energy saving in cellular networks with hybrid energy supplies,” in *Global Communications Conference (GLOBECOM), 2012 IEEE*, 2012, pp. 5189–5193.
- [48] D. Ng, E. Lo, and R. Schober, “Energy-efficient resource allocation in ofdma systems with hybrid energy harvesting base station,” *Wireless Communications, IEEE Transactions on*, vol. 12, no. 7, pp. 3412–3427, 2013.

- [49] A. Lalitha, S. Mondal, S. K. V, and V. Sharma, “Power-optimal scheduling for a green base station with delay constraints,” in *Communications (NCC), 2013 National Conference on*, 2013, pp. 1–5.
- [50] M. Zheng, P. Pawelczak, S. Stanczak, and H. Yu, “Planning of cellular networks enhanced by energy harvesting,” *Communications Letters, IEEE*, vol. 17, no. 6, pp. 1092–1095, 2013.
- [51] P. Pawelczak, M. Zheng, S. Stanczak, and H. Yu, “Enriching cellular networks with dynamic spectrum access and energy harvesting: A network planning case,” in *Dynamic Spectrum Access Networks (DYSPAN), 2012 IEEE International Symposium on*, 2012, pp. 285–285.
- [52] X. Lin, J. Andrews, A. Ghosh, and R. Ratasuk, “An overview of 3GPP device-to-device proximity services,” *Communications Magazine, IEEE*, vol. 52, no. 4, pp. 40–48, April 2014.
- [53] D. Feng, L. Lu, Y. Yuan-Wu, G. Li, S. Li, and G. Feng, “Device-to-device communications in cellular networks,” *Communications Magazine, IEEE*, vol. 52, no. 4, pp. 49–55, April 2014.
- [54] A. Asadi, Q. Wang, and V. Mancuso, “A survey on device-to-device communication in cellular networks,” *Communications Surveys Tutorials, IEEE*, vol. 16, no. 4, pp. 1801–1819, Fourthquarter 2014.
- [55] Y. Zhao, B. Pelletier, P. Marinier, and D. Pani, “D2D neighbor discovery interference management for LTE systems,” in *Globecom Workshops (GC Wkshps), 2013 IEEE*, Dec 2013, pp. 550–554.
- [56] B. Guo, S. Sun, and Q. Gao, “Downlink interference management for D2D communication underlying cellular networks,” in *Communications in China - Workshops (CIC/ICCC), 2013 IEEE/CIC International Conference on*, Aug 2013, pp. 193–196.

- [57] H. Elkotby, K. Elsayed, and M. Ismail, “Shrinking the reuse distance: Spectrally-efficient radio resource management in D2D-enabled cellular networks with interference alignment,” in *Wireless Days (WD), 2012 IFIP*, Nov 2012, pp. 1–6.
- [58] H. Min, J. Lee, S. Park, and D. Hong, “Capacity enhancement using an interference limited area for device-to-device uplink underlaying cellular networks,” *Wireless Communications, IEEE Transactions on*, vol. 10, no. 12, pp. 3995–4000, December 2011.
- [59] H. Kwon, J. Lee, and I. Kang, “Interference-aware interference mitigation for device-to-device communications,” in *Vehicular Technology Conference (VTC Spring), 2014 IEEE 79th*, May 2014, pp. 1–5.
- [60] Y. Chai, Q. Du, and P. Ren, “Partial time-frequency resource allocation for device-to-device communications underlaying cellular networks,” in *Communications (ICC), 2013 IEEE International Conference on*, June 2013, pp. 6055–6059.
- [61] Y. V. de Melo, R. L. Batista, T. F. Maciel, C. F. Silva, J. M. B. da Silva, and F. R. Cavalcanti, “Power control with variable target SINR for D2D communications underlying cellular networks,” in *European Wireless 2014; 20th European Wireless Conference; Proceedings of*, May 2014, pp. 1–6.
- [62] Q. Duong, Y. Shin, and O.-S. Shin, “Resource allocation scheme for device-to-device communications underlaying cellular networks,” in *Computing, Management and Telecommunications (ComManTel), 2013 International Conference on*, Jan 2013, pp. 66–69.
- [63] C. Lee, S.-M. Oh, and A.-S. Park, “Interference avoidance resource allocation for d2d communication based on graph-coloring,” in *Information and Communication Technology Convergence (ICTC), 2014 International Conference on*, 2014, pp. 895–896.

- [64] A. Hamdi and E. Hossain, “Cognitive and energy harvesting-based D2D communication in cellular networks: Stochastic geometry modeling and analysis,” *Communications, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2015.
- [65] X. Lin, J. Andrews, and A. Ghosh, “Spectrum sharing for device-to-device communication in cellular networks,” *Wireless Communications, IEEE Transactions on*, vol. 13, no. 12, pp. 6727–6740, Dec 2014.
- [66] H. Feng, H. Wang, X. Xu, and C. Xing, “A tractable model for device-to-device communication underlying multi-cell cellular networks,” in *Communications Workshops (ICC), 2014 IEEE International Conference on*, June 2014, pp. 587–591.
- [67] Z.-S. Syu and C.-H. Lee, “Spatial constraints of device-to-device communications,” in *Communications and Networking (BlackSeaCom), 2013 First International Black Sea Conference on*, July 2013, pp. 94–98.
- [68] X. Chen, R. Hu, and Y. Qian, “Coverage study of dense device-to-device communications underlying cellular networks,” in *Global Communications Conference (GLOBECOM), 2014 IEEE*, Dec 2014, pp. 4353–4358.
- [69] W. Zhibo, T. Hui, C. Nannan, and H. Yao, “Device-to-device resource allocation for qos support using a graphic theory,” in *Consumer Communications and Networking Conference (CCNC), 2014 IEEE 11th*, Jan 2014, pp. 525–530.
- [70] D. Malak and M. Al-Shalash, “Optimal caching for device-to-device content distribution in 5G networks,” in *Globecom Workshops (GC Wkshps), 2014*, Dec 2014, pp. 863–868.
- [71] E. Zihan, K. Choi, and D. Kim, “Distributed random access scheme for collision avoidance in cellular device-to-device communication,” *Wireless Communications, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2015.

- [72] Y. Li, J. Li, J. Jiang, and M. Peng, "Performance analysis of device-to-device underlay communication in rician fading channels," in *Globecom Workshops (GC Wkshps), 2013 IEEE*, Dec 2013, pp. 4465–4470.
- [73] L.-C. Wang and Y.-H. Cheng, "A statistical mobile-to-mobile rician fading channel model," in *Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st*, vol. 1, May 2005, pp. 63–67 Vol. 1.
- [74] N. Kaur and S. Sood, "An energy-efficient architecture for the internet of things (iot)," *Systems Journal, IEEE*, vol. PP, no. 99, pp. 1–10, 2015.
- [75] J.-M. Liang, J.-J. Chen, H.-H. Cheng, and Y.-C. Tseng, "An energy-efficient sleep scheduling with qos consideration in 3gpp lte-advanced networks for internet of things," *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on*, vol. 3, no. 1, pp. 13–22, 2013.
- [76] K. Cho, B.-G. Lee, K. Lee, and D. H. Lee, "Energy-efficient replica detection for resource-limited mobile devices in the internet of things," *Communications, IET*, vol. 7, no. 18, pp. 2141–2150, 2013.
- [77] Y. Liu and X. Wang, "Information and energy cooperation in OFDM relaying: Protocols and optimization," *Vehicular Technology, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2015.
- [78] R. Zhang and C. K. Ho, "Mimo broadcasting for simultaneous wireless information and power transfer," *Wireless Communications, IEEE Transactions on*, vol. 12, no. 5, pp. 1989–2001, May 2013.
- [79] D. Ng and R. Schober, "Secure and green swipt in distributed antenna networks with limited backhaul capacity," *Wireless Communications, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2015.

- [80] H. Ju and R. Zhang, "Optimal resource allocation in full-duplex wireless-powered communication network," *Communications, IEEE Transactions on*, vol. 62, no. 10, pp. 3528–3540, Oct 2014.
- [81] I. Flint, L. Xiao, N. Privault, D. Niyato, and P. Wang, "Performance analysis of ambient rf energy harvesting with repulsive point process modeling," *Wireless Communications, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2015.
- [82] N. Deng, W. Zhou, and M. Haenggi, "The ginibre point process as a model for wireless networks with repulsion," *Wireless Communications, IEEE Transactions on*, vol. 14, no. 1, pp. 107–121, Jan 2015.
- [83] J. Gomez, A. Vasseur, A. Vergne, P. Martins, L. Decreusefond, and W. Chen, "A case study on regularity in cellular network deployment," *Wireless Communications Letters, IEEE*, vol. 4, no. 4, pp. 421–424, 2015.
- [84] T. Kobayashi and N. Miyoshi, "Uplink cellular network models with ginibre deployed base stations," in *Teletraffic Congress (ITC), 2014 26th International*, 2014, pp. 1–7.
- [85] X. Lu, I. Flint, D. Niyato, N. Privault, and P. Wang, "Performance analysis of simultaneous wireless information and power transfer with ambient rf energy harvesting," in *Wireless Communications and Networking Conference (WCNC), 2015 IEEE*, March 2015, pp. 1303–1308.
- [86] C. Liu, B. Natarajan, and H. Xia, "Small cell base station sleep strategies for energy efficiency," *Vehicular Technology, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2015.
- [87] C. Liu and B. Natarajan, "Power management in heterogeneous networks with energy harvesting base stations," *Physical Communication*, vol. 16, pp. 14–24, 2015.

- [88] —, “Maximizing ergodic capacity in d2d underlay networks,” *Vehicular Technology, IEEE Transactions on (under review)*, 2015.
- [89] —, “Ergodic capacity in d2d underlay networks with varying user distribution,” 2016.
- [90] —, “Average achievable throughput in d2d underlay networks,” 2016.
- [91] —, “Modeling and analysis of simultaneous information and energy transfer in internet of things.”
- [92] H. Dhillon, R. Ganti, F. Baccelli, and J. Andrews, “Modeling and analysis of k-tier downlink heterogeneous cellular networks,” *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 3, pp. 550–560, 2012.
- [93] J. Xu, J. Zhang, and J. Andrews, “On the accuracy of the wyner model in cellular networks,” *Wireless Communications, IEEE Transactions on*, vol. 10, no. 9, pp. 3098–3109, 2011.
- [94] I. H. R. W. H. J. Chae, Chan-Byoung. and V. Tarokh., “Interference aware-coordinated beamforming system in a two-cell environment,” *IEEE Journal of Selected Areas in Communications*, to appear.
- [95] S. Jing, D. Tse, J. Soriaga, J. Hou, J. Smee, and R. Padovani, “Multicell downlink capacity with coordinated processing,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2008, no. 1, p. 586878, 2008. [Online]. Available: <http://jwcn.urasipjournals.com/content/2008/1/586878>
- [96] J. Andrews, F. Baccelli, and R. Ganti, “A tractable approach to coverage and rate in cellular networks,” *Communications, IEEE Transactions on*, vol. 59, no. 11, pp. 3122–3134, November 2011.

- [97] B. Mawlawi, E. Bastug, C. Nerguizian, S. Azarian, and M. Debbah, “Non-invasive green small cell network,” in *Signals, Systems and Computers (ASILOMAR), 2012 Conference Record of the Forty Sixth Asilomar Conference on*, 2012, pp. 68–73.
- [98] S. Mukherjee, “Distribution of downlink SINR in heterogeneous cellular networks,” *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 3, pp. 575–585, 2012.
- [99] H. Holtkamp, G. Auer, S. Bazzi, and H. Haas, “Minimizing base station power consumption,” *Selected Areas in Communications, IEEE Journal on*, vol. PP, no. 99, pp. 1–10, 2013.
- [100] C. Han, T. Harrold, S. Armour, I. Krikidis, S. Videv, P. M. Grant, H. Haas, J. Thompson, I. Ku, C.-X. Wang, T. A. Le, M. Nakhai, J. Zhang, and L. Hanzo, “Green radio: radio techniques to enable energy-efficient wireless networks,” *Communications Magazine, IEEE*, vol. 49, no. 6, pp. 46–54, June 2011.
- [101] L. Correia, D. Zeller, O. Blume, D. Ferling, Y. Jading, I. Godor, G. Auer, and L. Van der Perre, “Challenges and enabling technologies for energy aware mobile radio networks,” *Communications Magazine, IEEE*, vol. 48, no. 11, pp. 66–72, November 2010.
- [102] X. Xiao, X. Tao, and J. Lu, “Energy efficiency optimization in multi-user cellular systems with radio resource constraints,” in *Global Communications Conference (GLOBECOM), 2012 IEEE*, 2012, pp. 3490–3495.
- [103] G. Miao, N. Himayat, Y. Li, and D. Bormann, “Energy efficient design in wireless ofdma,” in *Communications, 2008. ICC '08. IEEE International Conference on*, 2008, pp. 3307–3312.
- [104] S. Wen and F. Yu, “Predictive control for energy efficiency in wireless cellular net-

- works,” in *Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th*, 2012, pp. 1–5.
- [105] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [106] T. Wang, R. De Lamare, and A. Schmeink, “Joint receiver design and power allocation strategies for multihop wireless sensor networks,” in *Wireless Communication Systems (ISWCS), 2011 8th International Symposium on*, 2011, pp. 86–90.
- [107] Y. Rong and Y. Hua, “Optimality of diagonalization of multi-hop mimo relays,” *Wireless Communications, IEEE Transactions on*, vol. 8, no. 12, pp. 6068–6077, 2009.
- [108] S. Tombaz, M. Usman, and J. Zander, “Energy efficiency improvements through heterogeneous networks in diverse traffic distribution scenarios,” in *Communications and Networking in China (CHINACOM), 2011 6th International ICST Conference on*, 2011, pp. 708–713.
- [109] Y. Cui, V. K. N. Lau, and Y. Wu, “Delay-aware bs discontinuous transmission control and user scheduling for energy harvesting downlink coordinated mimo systems,” *Signal Processing, IEEE Transactions on*, vol. 60, no. 7, pp. 3786–3795, 2012.
- [110] C. Wu, H. Mohsenian-Rad, and J. Huang, “Wind power integration via aggregator-consumer coordination: A game theoretic approach,” in *Innovative Smart Grid Technologies (ISGT), 2012 IEEE PES*, 2012, pp. 1–6.
- [111] M. He, S. Murugesan, and J. Zhang, “Multiple timescale dispatch and scheduling for stochastic reliability in smart grids with wind generation integration,” in *INFOCOM, 2011 Proceedings IEEE*, 2011, pp. 461–465.
- [112] T. Han and N. Ansari, “On optimizing green energy utilization for cellular net-

- works with hybrid energy supplies,” *Wireless Communications, IEEE Transactions on*, vol. 12, no. 8, pp. 3872–3882, 2013.
- [113] T. Niimura, “Forecasting techniques for deregulated electricity market prices - extended survey,” in *Power Systems Conference and Exposition, 2006. PSCE '06. 2006 IEEE PES*, 2006, pp. 51–56.
- [114] M. Zhou, Z. Yan, Y. X. Ni, G. Li, and Y. Nie, “Electricity price forecasting with confidence-interval estimation through an extended arima approach,” *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 153, no. 2, pp. 187–195, 2006.
- [115] J. Gong, J. Thompson, S. Zhou, and Z. Niu, “Base station sleeping and resource allocation in renewable energy powered cellular networks,” *Communications, IEEE Transactions on*, vol. 62, no. 11, pp. 3801–3813, Nov 2014.
- [116] S. Kandukuri and S. Boyd, “Optimal power control in interference-limited fading wireless channels with outage-probability specifications,” *Wireless Communications, IEEE Transactions on*, vol. 1, no. 1, pp. 46–55, Jan 2002.
- [117] H. Xia, B. Natarajan, and C. Liu, “Feasibility of simultaneous information and energy transfer in lte-a small cell networks,” in *Consumer Communications and Networking Conference (CCNC), 2014 IEEE*, Jan 2014, pp. 216–221.
- [118] H.-S. Jo, Y. J. Sang, P. Xia, and J. Andrews, “Heterogeneous cellular networks with flexible cell association: A comprehensive downlink sinr analysis,” *Wireless Communications, IEEE Transactions on*, vol. 11, no. 10, pp. 3484–3495, October 2012.
- [119] “Belair networks data sheet,” <http://www.virtualpressoffice.com>, 2014.
- [120] M. Goonewardena and L. B. Le, “Charging of electric vehicles utilizing random wind: A stochastic optimization approach,” in *Globecom Workshops (GC Wkshps), 2012 IEEE*, 2012, pp. 1520–1525.

- [121] X. Ma, R. Yin, G. Yu, and Z. Zhang, “A distributed relay selection method for relay assisted device-to-device communication system,” in *Personal Indoor and Mobile Radio Communications (PIMRC), 2012 IEEE 23rd International Symposium on*, Sept 2012, pp. 1020–1024.
- [122] Y. Wang, Y. Chang, and D. Yang, “An efficient inter-cell interference coordination scheme in heterogeneous cellular networks,” in *Vehicular Technology Conference (VTC Fall), 2012 IEEE*, 2012, pp. 1–5.
- [123] C. Yin, C. Chen, and S. Cui, “Stable distribution based analysis of transmission capacities for overlaid wireless networks,” in *Wireless Communications Signal Processing, 2009. WCSP 2009. International Conference on*, Nov 2009, pp. 1–5.
- [124] J. Guo, S. Durrani, X. Zhou, and H. Yanikomeroglu, “Outage probability of ad hoc networks with wireless information and power transfer,” *Wireless Communications Letters, IEEE*, vol. 4, no. 4, pp. 409–412, Aug 2015.
- [125] A. Doukas and G. Kalivas, “Rician K factor estimation for wireless communication systems,” in *Wireless and Mobile Communications, 2006. ICWMC '06. International Conference on*, July 2006, pp. 69–69.
- [126] J. B. Hough, *Zeros of Gaussian analytic functions and determinantal point processes*. AMS Bookstore, 2009.
- [127] Y. Bao, X. Wang, X. Liu, S. Zhou, and Z. Niu, “Solar radiation prediction and energy allocation for energy harvesting base stations,” in *Communications (ICC), 2014 IEEE International Conference on*, 2014, pp. 3487–3492.
- [128] P. Diamantoulakis, K. Pappi, G. Karagiannidis, and H. Poor, “Autonomous energy harvesting base stations with minimum storage requirements,” *Wireless Communications Letters, IEEE*, vol. 4, no. 3, pp. 265–268, 2015.

- [129] C. Wang, S. Durrani, J. Guo, and X. Zhou, “Call completion probability in heterogeneous networks with energy harvesting base stations,” in *Telecommunications (ICT), 2015 22nd International Conference on*, 2015, pp. 191–197.
- [130] H. Dhillon, Y. Li, P. Nuggehalli, Z. Pi, and J. Andrews, “Fundamentals of heterogeneous cellular networks with energy harvesting,” *Wireless Communications, IEEE Transactions on*, vol. 13, no. 5, pp. 2782–2797, May 2014.
- [131] S. of Research and E. Association, in *Handbook of mathematical, scientific, and engineering formulas, tables, functions, graphs, transforms*, 1988, p. 452.

Appendix A

Proofs of Chapter 5

A.1 Proof of Lemma 6

From [131], we know that

$$B_0(\gamma) = \int_0^{\frac{\pi}{2}} \frac{1}{1 + \gamma \cot \theta} d\theta = \frac{\pi}{2(1 + \gamma^2)} + \frac{\gamma}{1 + \gamma^2} \log \gamma. \quad (\text{A.1})$$

and

$$B_1(\gamma) = \frac{\pi(1 - \gamma^2)}{2(1 + \gamma^2)^2} + \frac{2\gamma}{(1 + \gamma^2)^2} \log \gamma + \frac{\gamma}{1 + \gamma^2}. \quad (\text{A.2})$$

Given that $B_j(\gamma) = \int_0^{\frac{\pi}{2}} \frac{(\sin \theta)^{j+1}}{(\sin \theta + \gamma \cos \theta)^{j+1}} d\theta$, $j > 0$, consider

$$B_{j-1}(\gamma) = \int_0^{\frac{\pi}{2}} \frac{(\sin \theta)^j (\sin \theta + \gamma \cos \theta)}{(\sin \theta + \gamma \cos \theta)^{j+1}} d\theta = B_j(\gamma) + \gamma A_j(\gamma). \quad (\text{A.3})$$

Here, we introduce $A_j(\gamma) = \int_0^{\frac{\pi}{2}} \frac{(\sin \theta)^j \cos \theta}{(\sin \theta + \gamma \cos \theta)^{j+1}} d\theta$.

On the other hand, $A_j(\gamma) = \frac{1}{j+1} + A_{j+1}(\gamma) - \gamma B_{j+1}(\gamma)$. From (A.3), we obtain a recursive relation

$$B_j(\gamma) = \frac{2B_{j-1}(\gamma) - B_{j-2}(\gamma) + \gamma/j}{1 + \gamma^2}, j \geq 2. \quad (\text{A.4})$$

A.2 Proof of Theorem 4

Let

$$\xi = \frac{1}{\eta} = \frac{2}{\pi} - \gamma, \quad 0 < \xi < \frac{2}{\pi}, \quad (\text{A.5})$$

From (5.7),

$$\begin{aligned} \hat{C}_C = \bar{R}_{Cl} &= \frac{2}{\eta} \arctan \eta + \left(\frac{\arctan \eta}{\eta} \right)^2 - \frac{2}{\eta^2} \log \sqrt{1 + \eta^2} + \log \frac{\sqrt{1 + \eta^2}}{\eta} \\ &= 2\xi \arctan \frac{1}{\xi} + \left(\xi \arctan \frac{1}{\xi} \right)^2 + \left(\frac{1}{2} - \xi^2 \right) \log \left(1 + \frac{1}{\xi^2} \right) + \log \xi, \end{aligned} \quad (\text{A.6})$$

then

$$\frac{d^2 \hat{C}_C}{d\gamma^2} = \frac{d^2 R_C}{d\xi^2} = \frac{3}{1 + \xi^2} + 2 \left(\arctan \frac{1}{\xi} \right)^2 - \frac{4\xi(\xi^2 + 2)}{(1 + \xi^2)^2} \arctan \frac{1}{\xi} - 2 \log \left(1 + \frac{1}{\xi^2} \right). \quad (\text{A.7})$$

For a convex function $u(x)$, $a < x < b$, the following inequalities hold:

$$u(x) < u(a) + \frac{u(b) - u(a)}{b - a} (x - a), \quad (\text{A.8})$$

$$u(x) > u(c) + u'(c)(x - c), \quad a < c < b. \quad (\text{A.9})$$

The right hand side of (A.8) is the segment connecting $(a, u(a))$ and $(b, u(b))$, and the right hand side of (A.9) is the tangent of $u(x)$ at $(c, u(c))$.

Similarly, for a concave function $v(x)$, $a < x < b$,

$$v(x) > v(a) + \frac{v(b) - v(a)}{b - a} (x - a), \quad (\text{A.10})$$

$$v(x) < v(c) + v'(c)(x - c), \quad a < c < b. \quad (\text{A.11})$$

Note that $\arctan \frac{1}{\xi}$ is convex and $\log \xi$ is concave in the region $0 < \xi < \frac{2}{\pi}$ because

$$\left(\arctan \frac{1}{\xi}\right)'' = \frac{2\xi}{(\xi^2 + 1)^2} > 0, \quad 0 < \xi < \frac{2}{\pi}, \quad (\text{A.12})$$

and

$$(\log \xi)'' = -\frac{1}{\xi^2} < 0, \quad 0 < \xi < \frac{2}{\pi}. \quad (\text{A.13})$$

Therefore from (A.8),

$$1.003884 = \arctan \frac{\pi}{2} < \arctan \frac{1}{\xi} < \frac{\pi}{2} + \frac{\arctan \frac{\pi}{2} - \frac{\pi}{2}}{\frac{2}{\pi} - 0}(\xi - 0) = \frac{\pi}{2} - 0.8905\xi, \quad (\text{A.14})$$

and from (A.11)

$$\log \xi < \log\left(\frac{1}{3}\right) + \frac{1}{\frac{1}{3}}\left(\xi - \frac{1}{3}\right) = 3\xi - 2.0986. \quad (\text{A.15})$$

Moreover according to the Taylor expansion of $\log(1 + \xi^2)$:

$$\log(1 + \xi^2) > \xi^2 - \frac{\xi^4}{2}, \quad 0 < \xi < \frac{2}{\pi}. \quad (\text{A.16})$$

Let $h(\xi) = \frac{d^2 \hat{C}_C}{d\xi^2}(1 + \xi^2)^2$. From (A.7) we can derive that

$$\begin{aligned} h(\xi) &= \frac{d^2 \hat{C}_C}{d\xi^2}(1 + \xi^2)^2 \\ &= 3(1 + \xi^2) + 2(1 + \xi^2)^2 \left(\arctan \frac{1}{\xi}\right)^2 - 4\xi(\xi^2 + 2) \arctan \frac{1}{\xi} - 2(1 + \xi^2)^2 \log(1 + \xi^2) \\ &\quad + 4(1 + \xi^2)^2 \log \xi \\ &< 3(1 + \xi^2) + 2(1 + \xi^2)^2 \left(\frac{\pi}{2} - 0.8905\xi\right)^2 - 4\xi(\xi^2 + 2) \times 1.003884 - 2(1 + \xi^2)^2 \left(\xi^2 - \frac{\xi^4}{2}\right) \\ &\quad + 4(1 + \xi^2)^2 (3\xi - 2.0986). \end{aligned} \quad (\text{A.17})$$

Solving $h(\xi) = 0$, we can find two roots -1.973 and 0.7145. Since the range of $\xi \in [0, 2/\pi]$

is between two roots and $h(0) = -0.4596 < 0$, we can conclude that $h(\xi) < 0$ when $0 < \xi < \frac{2}{\pi}$, and $\frac{d^2 \hat{C}_C}{d\xi^2} = \frac{d^2 \hat{C}_C}{d\gamma^2} < 0$, thus \hat{C}_C is a concave function with respect to γ .

From Corollary 2,

$$\hat{C}_D = 2\gamma B_0(\gamma) + \gamma B_1(\gamma)(1 - e^{-K}) = f(\gamma) + g(\gamma)(1 - e^{-K}), \quad 0 < \gamma < \frac{2}{\pi}, \quad (\text{A.18})$$

where

$$f(\gamma) = 2\gamma B_0(\gamma) = \frac{\pi\gamma}{1 + \gamma^2} + \frac{2\gamma^2}{1 + \gamma^2} \log \gamma, \quad (\text{A.19})$$

and

$$g(\gamma) = \gamma B_1(\gamma) = \frac{\pi\gamma(1 - \gamma^2)}{2(1 + \gamma^2)^2} + \frac{2\gamma^2}{(1 + \gamma^2)^2} \log \gamma + \frac{\gamma^2}{1 + \gamma^2}. \quad (\text{A.20})$$

We can obtain that

$$f''(\gamma) = -\frac{4\pi\gamma}{(1 + \gamma^2)^2} + \frac{4(1 - \gamma^2)}{(1 + \gamma^2)^2} \log \gamma + \frac{4}{1 + \gamma^2} + \frac{2\pi\gamma(3\gamma^2 - 1)}{(1 + \gamma^2)^3} - \frac{4\gamma^2(3 - \gamma^2)}{(1 + \gamma^2)^3} \log \gamma + \frac{2(1 - 3\gamma^2)}{(1 + \gamma^2)^2}, \quad (\text{A.21})$$

and then

$$\frac{1}{2}f''(\gamma)(1 + \gamma^2)^3 = 2(1 - 3\gamma^2) \log \gamma - (\gamma^2 - 3)(\gamma^2 - \pi\gamma + 1). \quad (\text{A.22})$$

(1) When $\frac{\pi - \sqrt{\pi^2 - 4}}{2} < \gamma < \frac{1}{\sqrt{3}}$ ($\frac{\pi - \sqrt{\pi^2 - 4}}{2} \approx 0.3594$, $\frac{1}{\sqrt{3}} \approx 0.5773$), $1 - 3\gamma^2 > 0$, $\gamma^2 - \pi\gamma + 1 < 0$, $\gamma^2 - 3 < 0$, therefore

$$\frac{1}{2}f''(\gamma)(1 + \gamma^2)^3 = 2(1 - 3\gamma^2) \log \gamma - (\gamma^2 - 3)(\gamma^2 - \pi\gamma + 1) < 0. \quad (\text{A.23})$$

(2) When $\frac{1}{\sqrt{3}} < \gamma < \frac{2}{\pi}$, $\gamma^2 - 3 < 0$, $\gamma^2 - \pi\gamma + 1 < 0$, $1 - 3\gamma^2 < 0$, according to (A.10):

$$\log \gamma > \log\left(\frac{1}{\sqrt{3}}\right) + \frac{\log\left(\frac{2}{\pi}\right) - \log\left(\frac{1}{\sqrt{3}}\right)}{\frac{2}{\pi} - \frac{1}{\sqrt{3}}} \left(\gamma - \frac{1}{\sqrt{3}}\right) = 1.6488\gamma - 1.5012. \quad (\text{A.24})$$

Then

$$\begin{aligned}
-\frac{1}{2}f''(\gamma)(1+\gamma^2)^3 &= (\gamma^2-3)(\gamma^2-\pi\gamma+1) - 2(1-3\gamma^2)\log\gamma \\
&> (\gamma^2-3)(\gamma^2-\pi\gamma+1) - 2(1-3\gamma^2)(1.6488\gamma-1.5012) = f_1(\gamma).
\end{aligned} \tag{A.25}$$

Solving $f_1(\gamma) = 0$, we can find two roots -8.187 and -0.0003. Since there is no root in the range of $\gamma \in [\frac{1}{\sqrt{3}}, \frac{2}{\pi}]$ and $f_1(\frac{1}{\sqrt{3}}) = 1.2812 > 0$, we can conclude that $f_1(\gamma) > 0$ when $\frac{1}{\sqrt{3}} < \gamma < \frac{2}{\pi}$, then $f''(\gamma) < 0$ in this region.

(3) When $0 < \gamma < \frac{\pi-\sqrt{\pi^2-4}}{2}$, $\gamma^2-3 < 0$, $\gamma^2-\pi\gamma+1 > 0$, $1-3\gamma^2 > 0$, according to (A.11):

$$\log\gamma < \log(0.25) + \frac{1}{0.25}(\gamma-0.25) = 4\gamma - 2.3863. \tag{A.26}$$

Then

$$\begin{aligned}
-\frac{1}{2}f''(\gamma)(1+\gamma^2)^3 &= (\gamma^2-3)(\gamma^2-\pi\gamma+1) - 2(1-3\gamma^2)\log\gamma \\
&> (\gamma^2-3)(\gamma^2-\pi\gamma+1) - 2(1-3\gamma^2)(4\gamma-2.3863) = f_2(\gamma).
\end{aligned} \tag{A.27}$$

Solving $f_2(\gamma) = 0$, we can find a root -0.2555. Since there is no root in the range of $\gamma \in [0, \frac{\pi-\sqrt{\pi^2-4}}{2}]$ and $f_2(0) = 1.7726 > 0$, we can conclude that $f_2(\gamma) > 0$ when $0 < \gamma < \frac{\pi-\sqrt{\pi^2-4}}{2}$, and $f''(\gamma) < 0$ in this region. Therefore $f(\gamma)$ is a concave function with respect to γ for $0 < \gamma < \frac{2}{\pi}$.

Similarly, we can obtain that

$$g''(\gamma) = -\frac{2\pi\gamma(3-\gamma^2)}{(1+\gamma^2)^3}\log\gamma + \frac{2(3-\gamma^2)}{(1+\gamma^2)^2} - \frac{3\pi\gamma(\gamma^4-6\gamma^2+1)}{(1+\gamma^2)^4} - \frac{24\gamma^2(1-\gamma^2)}{(1+\gamma^2)^4}\log\gamma + \frac{2(\gamma^4-10\gamma^2+1)}{(1+\gamma^2)^3}, \tag{A.28}$$

and

$$g''(\gamma)(1+\gamma^2)^4 = (4-32\gamma^2+24\gamma^3-12\gamma^4)\log\gamma - (\pi\gamma^5+16\gamma^4-14\pi\gamma^3+8\gamma^2+9\pi\gamma-8). \tag{A.29}$$

(1) When $0.2933 < \gamma < 0.4058$, $4 - 32\gamma^2 + 24\gamma^3 - 12\gamma^4 > 0$, $\pi\gamma^5 + 16\gamma^4 - 14\pi\gamma^3 + 8\gamma^2 + 9\pi\gamma - 8 > 0$, therefore

$$g''(\gamma)(1 + \gamma^2)^4 = (4 - 32\gamma^2 + 24\gamma^3 - 12\gamma^4) \log \gamma - (\pi\gamma^5 + 16\gamma^4 - 14\pi\gamma^3 + 8\gamma^2 + 9\pi\gamma - 8) < 0. \quad (\text{A.30})$$

(2) When $0.4058 < \gamma < \frac{2}{\pi}$, $4 - 32\gamma^2 + 24\gamma^3 - 12\gamma^4 < 0$, $\pi\gamma^5 + 16\gamma^4 - 14\pi\gamma^3 + 8\gamma^2 + 9\pi\gamma - 8 > 0$, according to (A.10):

$$\log \gamma > \log(0.4058) + \frac{\log(\frac{2}{\pi}) - \log(0.4058)}{\frac{2}{\pi} - 0.4058}(\gamma - 0.4058) = 1.9509\gamma - 1.6936. \quad (\text{A.31})$$

Then

$$\begin{aligned} -g''(\gamma)(1 + \gamma^2)^4 &= (\pi\gamma^5 + 16\gamma^4 - 14\pi\gamma^3 + 8\gamma^2 + 9\pi\gamma - 8) - (4 - 32\gamma^2 + 24\gamma^3 - 12\gamma^4) \log \gamma \\ &> (\pi\gamma^5 + 16\gamma^4 - 14\pi\gamma^3 + 8\gamma^2 + 9\pi\gamma - 8) - (4 - 32\gamma^2 + 24\gamma^3 - 12\gamma^4)(1.9509\gamma - 1.6936) = g_1(\gamma). \end{aligned} \quad (\text{A.32})$$

Solving $g_1(\gamma) = 0$, we can find a root 0.072. Since there is no root in the range of $\gamma \in [0.4058, \frac{2}{\pi}]$ and $g_1(0.4058) = 3.2576 > 0$, we can conclude that $g_1(\gamma) > 0$ when $0.4058 < \gamma < \frac{2}{\pi}$, and $g''(\gamma) < 0$ in this region.

(3) When $0 < \gamma < 0.2933$, $4 - 32\gamma^2 + 24\gamma^3 - 12\gamma^4 > 0$, $\pi\gamma^5 + 16\gamma^4 - 14\pi\gamma^3 + 8\gamma^2 + 9\pi\gamma - 8 < 0$, according to (A.11):

$$\log \gamma < \log(0.142857) + \frac{1}{0.142857}(\gamma - 0.142857) = 7\gamma - 2.9459. \quad (\text{A.33})$$

Then

$$\begin{aligned}
-g''(\gamma)(1+\gamma^2)^4 &= (\pi\gamma^5 + 16\gamma^4 - 14\pi\gamma^3 + 8\gamma^2 + 9\pi\gamma - 8) - (4 - 32\gamma^2 + 24\gamma^3 - 12\gamma^4) \log \gamma \\
&> (\pi\gamma^5 + 16\gamma^4 - 14\pi\gamma^3 + 8\gamma^2 + 9\pi\gamma - 8) - (4 - 32\gamma^2 + 24\gamma^3 - 12\gamma^4)(7\gamma - 2.9459) = g_2(\gamma).
\end{aligned}
\tag{A.34}$$

Solving $g_2(\gamma) = 0$, we can find a root -0.1676 . Since there is no root in the range of $\gamma \in [0, 0.2933]$ and $g_2(0) = 3.7836 > 0$, we can conclude that $g_2(\gamma) > 0$ when $0 < \gamma < 0.2933$, and $g''(\gamma) < 0$ in this region. Therefore $g(\gamma)$ is concave with respect to γ for $0 < \gamma < \frac{2}{\pi}$. From (A.18), we can deduce that \hat{C}_D is also a concave function of γ .