GREEN HETEROGENEOUS CELLULAR NETWORKS

A thesis submitted to the University of Manchester
for the degree of Doctor of Philosophy
in the Faculty of Engineering and Physical Sciences

2016

Edwin Mugume

School of Electrical and Electronic Engineering

Contents

List of Tables

List of Figures

Abstract

Data traffic demand has been increasing exponentially and this trend will continue over the foreseeable future. This has forced operators to upgrade and densify their mobile networks to enhance their capacity. Future networks will be characterized by a dense deployment of different kinds of base stations (BSs) in a hierarchical cellular structure. However network densification requires extensive capital and operational investment which limits operator revenues and raises ecological concerns over greenhouse gas emissions. Although networks are planned to support peak traffic, traffic demand is actually highly variable in both space and time which makes it necessary to adapt network energy consumption to inevitable variations in traffic demand.

In this thesis, stochastic geometry tools are used to perform simple and tractable analysis of the coverage, rate and energy performance of homogeneous networks and heterogeneous networks (HetNets). BSs in each tier are located according to independent Poisson Point Processes (PPPs) to generate irregular topologies that fairly resemble practical deployment topologies. The homogeneous network is optimized to determine the optimal BS density and transmit power configuration that minimizes its area power consumption (APC) subject to both coverage and average rate constraints. Results show that optimal transmit power only depends on the BS power consumption parameters and can be predetermined. Furthermore, various sleep mode mechanisms are applied to the homogeneous network to adapt its APC to changes in user density. A centralized strategic scheme which prioritize BSs with the least number of users enhances energy efficiency (EE) of the network. Due to the complexity of such a centralized scheme, a distributed scheme which implements the strategic algorithm within clusters of BSs is proposed and its performance closely matches that of its centralized counterpart.

It is more challenging to model the optimal deployment configuration per tier in a multi-tier HetNet. Appropriate assumptions are used to determine tight approximations of these deployment configurations that minimize the APC of biased and unbiased HetNets subject to coverage and rate constraints. The optimization is performed for three different user association schemes. Similar to the homogeneous network, optimal transmit power per tier also depends on BS power consumption parameters only and can also be predetermined. Analysis of the effect of biasing on HetNet performance shows appropriate biasing can further reduce the deployment configuration (and consequently the APC) compared to an unbiased HetNet. In addition, biasing can be used to offload traffic from congesting and high-power macro BSs to low-power small BSs. If idle BSs are put into sleep mode, more energy is saved and HetNet EE improves. Moreover, appropriate biasing also enhances the EE of the HetNet.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright

Dedication

This thesis is dedicated to my family
Caroline, Lisa and Liam.

Acknowledgements

Over the course of my PhD research, several people have provided me with invaluable support that has enabled me to reach the conclusion of my research. I would like to express my sincerest gratitude to my PhD supervisor Dr. Daniel K. C. So for his tireless support and encouragement in directing and guiding my research over the years. He believed in my academic ability and mentored me to become a better researcher and person and I will forever be grateful.

I would also like to thank my first year and second year internal examiner, Dr. Khairi Hamdi, for his interest in the direction of my research and for his honest opinions and constructive criticisms of my work during the vivas. I learnt a lot from these vivas and incorporated his ideas into my research whenever possible.

I would also like to thank students within the MACS group who helped me from time to time. Special mention goes to Dr. Warit Prawatmuang who helped me a lot at the start of my PhD journey, especially with writing efficient MATLAB code. In equal measure, I would like to thank my neighbor over the years, Mansour Aldosari, for his friendship and brotherhood. We shared and discussed our academic challenges and helped each other whenever possible. In equal measure, I would also like to thank Hanifa Nabuuma, my fellow PhD student and contemporary, for her support, encouragement and help over the years.

I would like to acknowledge the generous financial support from the University of Manchester's Presidents Doctoral Scholar (PDS) award which enabled me to pursue this PhD research without any financial difficulties. I would also like to acknowledge and thank Dr. Daniel So for supporting my application for the PhD scholarship which allowed me to pursue this research.

Last but not least, I would like to thank in a special way my wife Caroline who put her promising engineering career in Uganda in jeopardy by deciding to join me in the United Kingdom for the duration of my PhD research. Bringing our daughter with her gave our family the opportunity to remain and grow stronger together. We also appreciate the love and support from our respective families during this period.

The Author

Edwin Mugume received the Bachelor of Science degree in Electrical Engineering (First Class Honors) from Makerere University, Uganda in 2007, and the Master of Science degree in Communication Engineering (with Distinction) from The University of Manchester, United Kingdom in 2011. Since 2012, he has been pursuing a PhD in Electrical and Electronic Engineering in the Microwave and Communication Systems (MACS) group at the School of Electrical and Electronic Engineering, The University of Manchester, United Kingdom. He previously worked as a Teaching Assistant and Assistant Lecturer at the Department of Electrical Engineering, Makerere University. He also has industry experience in cellular network planning and optimization from his previous roles at Zain, Nokia Siemens Networks and Bharti Airtel. In 2003, he was awarded the National Merit Scholarship from the Government of Uganda to pursue undergraduate study at Makerere University. In 2010, he was awarded the UK Commonwealth Scholarship to study MSc Communication Engineering at The University of Manchester. In 2011, he won the Agilent Top Student Award as the best student in his graduating class. In 2012, he won the prestigious President's Doctoral Scholar (PDS) award from The University of Manchester to pursue a PhD in Electrical and Electronic Engineering. His PhD research has investigated the performance of small cell-based mobile cellular networks with particular emphasis on energy performance aspects. His main research interests are in the areas of green cellular communications, dense heterogeneous cellular networks and 5G technologies.

List of Abbreviations

3G	Third Generation mobile network
4G	Fourth Generation mobile network
5G	Fifth Generation mobile network
ABRP	Average Biased Received Power
ABR	Average Blocking Ratio
AC	Alternating Current
ACI	Adjacent Channel Interference
ACR	Average Connectivity Ratio
AMC	Adaptive Modulation and Coding
AP	Access Point
APC	Area Power Consumption
AWGN	Additive White Gaussian Noise
BB	Base band
BPF	Band Pass Filter
BS	Base Station
BTD	Biased Transmission Distance
CAPEX	Capital Expenditure
CQI	Channel Quality Indicator
CR	Cognitive Radio
C-RAN	Cloud Radio Access Network
D2D	Device-to-device
DC	Direct Current
DL	Downlink
EB	Exabytes
DSL	Digital Subscriber Line
ECG	Energy Consumption Gain
ECR	Energy Consumption Ratio
EE	Energy Efficiency
ERG	Energy Reduction Gain
ETSI	European Telecommunications Standards Institute
HetNet	Heterogeneous Network
i-SINR	Instantaneous SINR
ICIC	Inter-Cell Interference Coordination

ICT	Information and Communication Technology
ISI	Intersymbol Interference
LOS	Line of sight
LTE	Long Term Evolution
MIMO	Multiple-input multiple-output
NLOS	Non-line-of-sight
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OPEX	Operational Expenditure
PA	Power Amplifier
PI	Performance Indicator
PPP	Poisson Point Process
PV	Poisson-Voronoi
QoS	Quality of Service
RF	Radio Frequency
RRH	Remote Radio Head
SE	Spectral Efficiency
SINR	Signal-to-Interference-plus-Noise Ratio
SNR	Signal-to-Noise Ratio
SON	Self-organizing network
TRX	Transceiver
UE	User Equipment
UL	Uplink

Symbols and Notations

α	Pathloss exponent
β	Bias value with ABRP connectivity
ν	Bias value with BTD connectivity
Φ	PPP
λ	PPP density or intensity
υ	BS-user density ratio in a homogeneous network
$\mathcal{P}_c$	Coverage probability
$\overline{\mathcal{P}}_c$	Coverage probability in the interference-limited network
$\mathcal{R}_u$	Average user rate
$\overline{\mathcal{R}}_u$	Average user rate in the interference-limited network
$\mathcal{R}_{ch}$	Average subchannel rate
$\overline{\mathcal{R}}_{ch}$	Average subchannel rate in the interference-limited network
$\mathcal{T}$	Average sum rate
$\overline{\mathcal{T}}$	Average sum rate in the interference-limited network
ϵ	Coverage probability constraint
κ	Average user rate constraint
A	Network area
$\mathcal{A}$	Tier association probability
$\mathcal{B}$	System Bandwidth
$\mathcal{B}_\delta$	Subchannel bandwidth
δ	Number of subchannels
F	Noise figure
$\mathcal{H}$	Deployment factor of a network
$\mathcal{K}$	A constant used in cell size distributions
K	Number of tiers of a HetNet
L	Pathloss coefficient
$\mathcal{L}(\cdot)$	Laplace transform
$\mathbb{E}[\cdot]$	Expectation operator
$\mathbb{P}[\cdot]$	Probability operator
σ^2	Additive noise power
N	Number of users in a cell
$\mathcal{N}$	Number of transceiver chains at a BS
T	SINR or SIR threshold
T_a	Ambient temperature

Chapter 1

Introduction

1.1 Background

Mobile network operators are faced with exponentially increasing data traffic demand which has placed extreme demands on existing networks. A forecast of the global mobile data traffic for the period 2015-2020 confirms that this trend will continue over the foreseeable future [1], [2]. In 2015, global mobile data traffic grew by 74%, from 2.1 exabytes (EB) per month at the end of 2014 to 3.7 EB per month at the end of 2015. The report also shows that mobile data traffic will grow to 30.6 EB per month by the end of 2020, a 53% compound annual growth rate. New access technologies such as third-generation (3G) and fourth-generation (4G) systems coupled with advanced end user devices such as smart phones and tablets are responsible for this rapid traffic increase. For example in 2015, a 4G connection generated six times more traffic on average than a non-4G connection. Moreover, 4G traffic exceeded 3G traffic in 2015 for the first time. This is telling especially since 4G connections represented only 14% of total mobile connections in 2015 [1].

The popularity and rapid uptake of advanced terminals such as smart phones and tablets and their associated data-hungry applications has also fueled this data explosion. According to [1], smart phones and tablets will increasingly be the source of most of the year-on-year growth of data traffic up to 2020. For example in 2015, smartphones generated 97% of all global handset traffic although they represented only 43% of global handsets. A smart device generated 14 times more traffic than a non-smart device in 2015. Mobile video, a service that requires high bandwidth, has the highest growth rate of any other mobile traffic category. It accounted for 55% of all mobile traffic in 2015 and will grow 11-fold to account for 75% of all data traffic by the end of 2020 [1]. However, research shows that operator revenues are growing by a mere 23% per annum [3]. In addition, Cisco reports in [4] that operator revenues will begin to shrink from 2018 onwards to the high cost of investment and operation. Compared to the exponential growth of mobile traffic and the inevitable investment in modern networks infrastructure, it will become increasingly difficult for operators to finance these network upgrades and still be able to make profit.

Over time, network capacity has been increased using techniques such as increasing link capacity

or bandwidth. However, radio links are fast approaching their theoretical capacity limits and the usable mobile spectrum is very congested and expensive. The most effective technique of increasing network capacity has been to reduce the size of cells and increase the spatial reuse of frequency bands. By reducing the cell size, the number of subscribers sharing the bandwidth of each base station (BS) reduces which avails more bandwidth to each user [5], [6], [7]. In areas with a sparse deployment of BSs such as rural areas, it is possible to add more BSs and enhance network capacity by effectively managing inter-cell interference. However, this cell splitting strategy may cause significant inter-cell interference in dense urban, urban and sub-urban areas which already have a significantly dense deployment of macro BSs. Besides, acquiring site leases is a very expensive venture especially in urban areas [5], [6]. Other techniques being considered to enhance capacity of future networks include multiple-input and multiple-output (MIMO) and massive MIMO systems [8], [9], cognitive radio (CR) [10], [11], [12], sophisticated user association algorithms [13], [14], etc. Furthermore, future 5G systems are expected to provide up to 1000 times more area spectral efficiency (SE) than current 4G technologies [14], [15], [16].

Mobile networks consume a lot of power which has made the cost of energy one of the major operational expenditures (OPEX) incurred by operators. This problem becomes worse if BSs have to run on diesel generators due to a lack of the electricity grid especially in remote/rural areas. The associated greenhouse gas (CO2) emissions into the atmosphere have caused ecological concerns in a world grappling with the effects of global warming [17]. A report published in 2008 by the Climate Change Group estimated that the ICT industry is currently responsible for 3% of global energy consumption, generating 2% of the total CO2 emissions [18]. In the telecommunications sector alone, mobile networks are predicted to contribute 51% of the total CO2 emissions, up from 43% in 2002. Given the increasing traffic demand, energy consumption and the associated CO2 emissions will continue to increase unless measures are taken to design more energy efficient networks.

Operators are therefore looking for economical, sustainable and environmentally friendly solutions to not only reduce their OPEX such as energy consumption but to also enhance the capacity of networks to handle even larger volumes of data traffic. Various techniques may be used to reduce energy consumption: improved network deployment techniques [3], designing energy efficient network equipment and cooling systems [19], avoiding cooling altogether by using remote radio heads (RRHs) [20], and implementing sleep mode schemes [21]. In RRHs, radio equipment is installed in the tower next to the antennas to avoid feeder losses and take advantage of natural air saturation for its cooling.

Heterogeneous networks (HetNets) are a promising solution towards energy efficient and capacity-enhanced mobile networks compared to traditional homogeneous macrocell networks. A HetNet is a mobile network that combines various types of BSs to provide mobile services to end users [5]. Therefore, HetNets combine macro BSs with small BSs such as microcells, picocells, femtocells and relay nodes. Small BSs transmit low power and therefore cover a relatively small area compared to the high-power macro BSs. They typically transmit 250 mW to 2 W in outdoor environments compared to macro BSs which typically transmit between 5-40W. Femtocells are indoor BSs that can transmit up to 100 mW of power [5], [7], [22], [23].

Small BSs may be deployed over the existing macrocell network to provide targeted coverage and capacity in different locations: in coverage gaps such as indoor environments and underground parking lots, traffic hotspots such as shopping malls and busy streets and in the cell edge region where they boost signal-to-interference-plus-noise ratio (SINR). In the presence of cell-edge small BSs, cell-edge users can connect to nearby BSs which reduces their uplink (UL) transmit power and improves battery life [24]. Small BSs are cheap, easy to install and have much lower energy consumption. They also do not require cooling which gives further energy savings. Therefore, HetNets provide a cost-effective way for operators to improve network capacity and coverage without incurring significant CAPEX and OPEX. However, HetNets generate new research challenges which must be tackled to improve their performance [5].

Some works have also studied the potential for using renewable energy sources such as solar and wind to save grid power and reduce energy bills [25], [26]. This concept called energy harvesting can be used on its own or as part of a hybrid system where it is combined with grid power to guarantee the availability of mobile services. On its own, it enables the real possibility of deploying *drop and play* small BSs as opposed to *plug and play* small BSs which rely on a grid power source. However, the main challenge facing *stand-alone* energy harvesting systems is the random spatial and temporal availability of renewable energy. This makes a hybrid system perhaps more attractive to provide uninterrupted mobile services [14], [27].

Other interesting technologies that can provide energy savings and generally reduce capital expenditures (CAPEX) and OPEX include [14]: (i) self-organizing networks (SONs) which reduce operational costs since they are self-optimizing and self-healing [28]-[29]; (ii) device-to-device (D2D) communications which allow any two devices in close proximity to communicate directly without BS or core network assistance – this improves SE and energy efficiency (EE) of the network [30]; and (iii) Cloud radio access network (C-RAN) which is a new architectural paradigm where all base band (BB) processing is centralized in the cloud and simple, low-cost and low-energy consuming RRHs provide radio access [20].

1.2 Aims and Objectives

The main aim of the research project was to investigate and design energy efficient HetNets consisting of a joint deployment of macro BSs and different kinds of small BSs. The research covered different performance aspects of HetNets such as coverage and rate, energy consumption, network deployment strategies, mechanisms of interference management, etc.

The objectives of the research are:

- To perform a comprehensive literature review and understand several complementary technologies to be applied in the research such as small cell technologies, HetNets, deployment strategies, load balancing strategies, etc.
- To analyze the performance aspects of HetNets and propose mechanisms to improve performance in terms of its capacity, energy consumption, deployment strategies, interference management, etc.

- To study the performance aspects of dense small cell networks and propose novel techniques of improving EE and network capacity.

1.3 List of Publications

The following papers have already been published or submitted.

P.1 Edwin Mugume and Daniel K. C. So, "Optimal Deployment Configuration of Energy-Aware Dense HetNets," *IEEE Transactions on Wireless Communications* (submitted in March 2016).

P.2 Edwin Mugume and Daniel K. C. So, "Energy-Aware Optimization of Small Cell Networks with Sleep Mode," *IEEE Journal on Selected Areas in Communications* (under second review).

P.3 Edwin Mugume, Daniel K. C. So and E. Alsusa, "Energy Efficient Deployment of Dense Heterogeneous Cellular Networks," *2015 IEEE Global Communications Conference (GLOBECOM)*, pp. 1-6, San Diego, CA, 6-10 December 2015.

P.4 Edwin Mugume and Daniel K. C. So, "Capacity and Energy Efficiency Analysis of Dense HetNets with Biasing," *IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC),* pp. 1553-1557, Hong Kong, 30 August - 2 September 2015.

P.5 Edwin Mugume and Daniel K. C. So, "Sleep Mode Mechanisms in Dense Small Cell Networks," *IEEE International Conference on Communications (ICC)*, pp. 192-197, London UK, 8-12 June 2015.

P.6 Edwin Mugume and Daniel K. C. So, "Spectral and Energy Efficiency Analysis of Dense Small Cell Networks," *IEEE 81st Vehicular Technology Conference (VTC Spring)*, pp. 1-5, Glasgow, Scotland, 11-14 May 2015.

P.7 Edwin Mugume, Warit Prawatmuang, and Daniel K. C. So, "Cooperative Spectrum Sensing for Green Cognitive Femtocell Network," *IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC),* pp. 2368-2372, London UK, 8-11 September 2013.

1.4 Contributions of the Thesis

The following major contributions have been accomplished during the course of this research:

1. Developed user connectivity models for both Poisson Point Process (PPP)-based homogeneous networks and HetNets. These connectivity models determine the ability of the network to connect the prevailing user density and avoid congestion or blocking. This

work was published in P.5 and also forms the basis for P.2. In this thesis, user connectivity analysis in a homogeneous network is discussed in Chapter 4.

2. Performed an area power consumption (APC)-minimization framework on a homogeneous network to determine its optimal BS density and associated transmit power configuration subject to both coverage and average rate constraints. This framework captures the effect of the prevailing user density on optimal deployment configuration. In some special cases, the deployment configuration is expressed in closed form. This optimization analysis is published in P.2 and is discussed in Chapter 4 of this thesis.

3. Proposed a novel sleep mode scheme for homogeneous networks called centralized strategic sleep mode which prioritizes BSs with the least number of users. The strategic algorithm searches over the entire network which makes implementation and management potentially challenging especially in large dense networks. To ease complexity, a distributed strategic sleep mode scheme is proposed in which the strategic algorithm is implemented within clusters of BSs all over the network. Centralized strategic sleep mode was published in P.5. More detailed analysis of both centralized and distributed strategic sleep mode is presented in P.2. This work is discussed in Chapter 5 of this thesis.

4. Used appropriate approximations to simplify and understand the analytical relationships between the prevailing user density and various performance measures of the homogeneous network such as coverage probability and average rate. For instance when user and BS densities are comparable, average user rate approximately varies linearly with user density. This work was published in P.6 and is presented in Chapter 5 of this thesis.

5. Performed coverage probability and average rate analysis of a PPP-based HetNet using minimum biased transmission distance (BTD) association scheme (where a user strictly connects to the nearest BS from any tier). Coverage and rate performance of this scheme is compared with two other existing schemes (maximum average biased received power (ABRP) and maximum instantaneous SINR (i-SINR) schemes). The analysis of minimum BTD scheme is published in P.1 and can be found in Chapter 6 of this thesis.

6. The APC-minimization framework is extended to a general K-tier biased HetNet to determine its optimal configuration per tier subject to coverage and average rate constraints. In some cases, the deployment configuration is expressed in closed form. For biased HetNets, a two-tier scenario is assumed to further analysis and draw insights into the effect of biasing on the APC performance of the HetNet. Detailed analysis of this HetNet deployment optimization is published in P.1 and can be found in Chapter 6.

7. Extended the existing cell size distribution analysis of a two-tier PPP-based unbiased HetNet to a general two-tier biased HetNet. Using these cell distributions, idle BS probability per tier is determined to facilitate an investigation of the effect of the prevailing user density on HetNet performance. In addition, user connectivity models are developed for this two-tier biased HetNet to determine its ability to avoid blocking at peak times. This user connectivity analysis in a two-tier biased HetNet is published in P.4 and can be found in Chapter 7 of this thesis.

1.5 Organization of the Thesis

This thesis has eight chapters. Chapter 1 introduces the opportunities and challenges facing the mobile communications industry such as increasing traffic demand and the rising cost of energy. It also discusses the aims and objectives and the main contributions of the thesis.

Chapter 2 discusses the energy consumption challenges facing mobile networks and the green potential of HetNets. The energy consumption model of different types of BSs is discussed and the energy performance metrics of cellular networks are introduced. Existing works on green HetNets such as energy saving approaches are reviewed with particular emphasis on sleep mode mechanisms. HetNet deployment approaches are discussed in contrast to traditional homogeneous networks. Other fundamental theories on wireless communication and Long Term Evolution (LTE) technology are provided in Appendices A and B respectively.

Chapter 3 introduces the stochastic geometry approach to network analysis and makes a complete review of all mathematical preliminaries. It then discusses all existing analytical results of PPP-based homogeneous networks and HetNets that are relevant to the analysis in this thesis.

Presentation of the main contributions of this thesis begins in Chapter 4 which discusses user connectivity in a PPP-based homogeneous network by utilizing existing cell size distributions. Where idle BSs exist, aggregate interference reduces which consequently affects the coverage, rate and energy performance. The network is then optimized to determine the optimal deployment configuration that minimizes its APC subject to coverage and rate constraints.

Chapter 5 discusses sleep mode approaches in a homogeneous network to adapt its energy consumption to changes in user density. Two novel schemes called centralized and distributed strategic schemes are then proposed and compared with existing conventional and random schemes. In addition, appropriate approximations are utilized to simplify the analytical relationship between the prevailing BS-user density ratio and various major performance measures.

Chapter 6 analyzes the coverage probability and average rate performance of a general multi-tier PPP-based HetNet implementing minimum BTD association scheme. It then presents an APC minimization framework to determine the optimal HetNet deployment configuration subject to appropriate coverage probability and average rate constraints. Optimization is performed on a HetNet using minimum BTD and two other existing user association schemes.

Chapter 7 extends existing cell size distributions of an unbiased two-tier HetNet to a biased two-tier HetNet. The distributions are then used to investigate user occupancy of cells as the user density varies spatiotemporally. Presence of idle BSs has an effect on aggregate interference which impacts coverage, average rate and energy performance of the HetNet.

Finally, Chapter 8 provides a general conclusion of the results and discussions in this thesis. It also discusses some ideas for future work. References and appendices follow this chapter.

Chapter 2

Fundamentals of Green HetNets

2.1 Introduction

The power consumption of cellular networks has generated economic and ecological concerns because of the rising cost of energy and associated greenhouse gas emissions which cause global warming. This has necessitated the research community to seek cellular network solutions that can reduce energy bills and CO_2 emissions but this task is very challenging. A comprehensive energy consumption analysis of the network should consider all stages of the process such as manufacture and production, distribution, operation and possible waste treatment. Each stage is considered in isolation to identify the worst offenders [31], [32], [33].

2.2 HetNet Deployment

Traditionally, cellular networks have always consisted of a homogeneous deployment of macro BSs in a planned fashion to provide the required network coverage and capacity per unit area. However due to the explosion of data traffic demand and the high energy consumption associated with installing more macro BSs, such networks are no longer feasible economically and ecologically [18]. Future networks require a strategic combination of different types of BSs to enhance coverage and provide targeted capacity enhancements. A HetNet architecture provides operators with opportunities to manage their network CAPEX and OPEX. The BS types in a typical HetNet deployment include macro BSs, micro BSs, pico BSs and femtocell access points (APs) [5], [22], [23], [34].

Macro BSs transmit the highest power which gives them the widest coverage, typically on the order of 1 km or more. They are normally installed in outdoor locations to provide wide coverage. However they also have the highest power consumption of all BSs, transmitting power in the order of 5-40 W. Micro BSs and pico BSs are similar to regular macro BSs only that they cover relatively smaller coverage areas. They are installed by the operator in planned locations, mostly in traffic hotspots and coverage holes. By design, micro BSs transmit relatively higher power than pico BSs. Pico BSs use omnidirectional antennas and typically transmit

approximately 0.25-2 W in outdoor deployments and 100 mW or less in indoor deployments. Their backhaul connection may be via microwave links or fiber optic [5].

Femtocell APs are low-power, low-cost and small coverage data APs that are installed in indoor locations to enhance indoor coverage and capacity. They are typically plug-and-play devices that are installed by the subscriber and the operator has no control over their location in the network. Femtocells use existing digital subscriber line (DSL) or cable modem for their backhaul connection to the parent network. They use omnidirectional antennas and transmit a power of 100 mW or less. Compared to other BSs, femtocell APs require very low initial investment in hardware and have very low energy consumption [5], [35], [36], [37], [38], [39].

Femtocells are classified according to their user association mechanisms into closed, open and hybrid access [35]. A closed access femtocell restricts access to only registered terminals while an open access FAP connects any terminal of the same operator if it is within range. A femtocell can also be hybrid where any terminal can access it but registered terminals have priority. In the DL, a closed access femtocell appears as a coverage hole and can be a source of significant interference to restricted terminals located nearby [5]. It is estimated that at least 50% of all voice connections and 70% of all data traffic will originate from indoor locations [35], [36]. Therefore femtocells are potentially an effective solution to enhance indoor user experience and boost overall network capacity. It was estimated that nearly 50 million femtocell APs would have been deployed in networks all over the world by the end of 2014 [37], [40].

Relay nodes are installed mainly to extend coverage to an uncovered area or to boost coverage around the cell edge region [41]. A relay receives a signal from the BS and retransmits it over the surrounding area. Thus it appears as a BS to the mobile terminals that it serves while it appears as a mobile terminal to its parent BS. Each relay is equipped with an omnidirectional antenna on the access part and a directional antenna pointing towards the parent BS for the backhaul connection. Relays also transmit approximately 0.25-2 W in outdoor deployments and 100 mW or less in indoor deployments [5], [24]. Relay nodes use the same air interface resources to connect back to the parent BS. If the backhaul frequency band is the same as that used by the relay node to communicate to/from the user on the DL/UL, the relay node is referred to as 'in-band'. Otherwise, the relay node is referred to as 'out-of-band'. Out-of-band relays require dedicated spectrum which reduces overall network SE. In-band relays are more attractive to operators although they present more challenges in the physical layer [5].

2.2.1 BS Power Consumption

Fig. 2.1 (a) shows a breakdown of the percentage contribution of various network elements to the total energy consumption [31], [42]. The result shows that BSs are responsible for about 57% of total consumption which is by far the highest. Therefore efforts to save energy should concentrate on the access part of cellular networks as this is clearly where the biggest energy-saving opportunities lie.

Fig. 2.1 (b) shows the embodied and operational energy consumption of both the BS and mobile handset [31]. It is clear that the cost of operating BSs is much higher than that of

handsets. However, the mobile handset has a much higher manufacturing/embodied energy because they have a very short life time of about 2 years compared to BSs whose operational lifetime reaches 10-15 years. In addition, much fewer pieces of BSs are manufactured compared to the number of handsets. Significant progress has been made in the manufacture of more energy efficient mobile handsets. For instance, carbon footprint per subscriber reduced from a high of 100 kg of CO2 emission per year in the early 1990s to about 25kg in mid-2000s. However, the overall carbon footprint of mobile handsets continues to increase as their volume continues to increase dramatically [43]. To further reduce the embodied energy of mobile handsets, the manufacturing process needs to be more energy efficient and their lifetime needs to be improved, for example by recycling them [31], [42].

To identify where the opportunities for energy saving lie in a BS, it is necessary to breakdown its total consumption into contributions of its constituent elements namely the power amplifier (PA), a radio-frequency (RF) TRX module, a BB unit, a DC-DC power supply, a cooling system, antenna interface and an AC-DC mains supply [19]. Fig. 2.2 shows a simple block diagram of a BS architecture which can be generalized for all types of BSs (only one transmit chain is shown). Each BS consists of at least one transceiver (TRX) where one TRX serves one transmit antenna in the DL. The components that consume the most power are the BB unit, the RF TRX and PA unit, antenna system and the cooling system [19].

The antenna system loss can be modeled using the losses caused by the feeder, antenna bandpass filters, duplexers and matching components [19]. In cases where the BS is physically separated from the antenna, a feeder loss of about $l_{feed} = 3\,\text{dB}$ should be added. Using a RRH in a macrocell removes the need for a feeder because the PA is located at the same location as the antennas (in the tower). Smaller BS types also have negligible feeder losses [19].

The best operating point of the PA is near the saturation point. In LTE however, the PA is forced to operate way below this point in a more linear region (6-12 dB below saturation) due to non-linear effects [19]. Non-linear effects cause signal distortions which result into adjacent channel interference (ACI) and performance degradation at the receiver. Unfortunately, this high operating back-off translates into a poor PA efficiency η_{pa}, increasing its power consumption according to $P_{pa} = P_{out}/\eta_{pa}(1 - l_{feed})$ [19].

The RF TRX consists of components for transmission on the DL and reception on the UL. The BB unit is responsible for carrying out digital operations such as digital up/down conversion, filtering, modulation and demodulation, signal detection, channel coding and decoding etc [19]. The DC-DC power supply, mains supply and cooling introduce further power losses in the BS. However, RRHs and small BSs use natural air circulation for cooling and therefore do not incur cooling losses [19].

2.2.2 BS Linear Power Model

The total amount of power consumed by a BS depends on its type and operating mode [19]. Macro BSs generally consume more power than smaller coverage BSs such as micro and pico

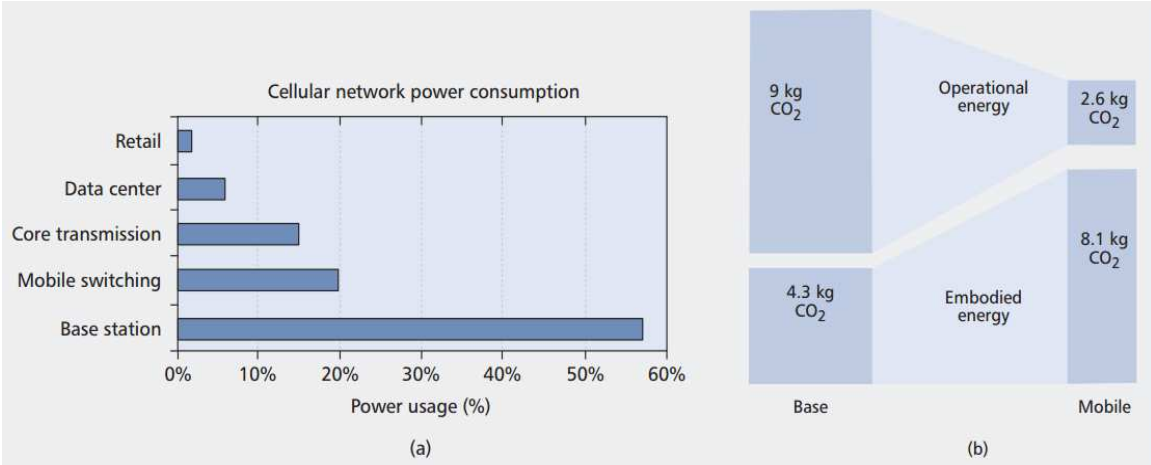


Figure 2.1: (a) Typical energy consumption of a cellular network; (b) CO_2 emissions per subscriber per year for the BS and user.

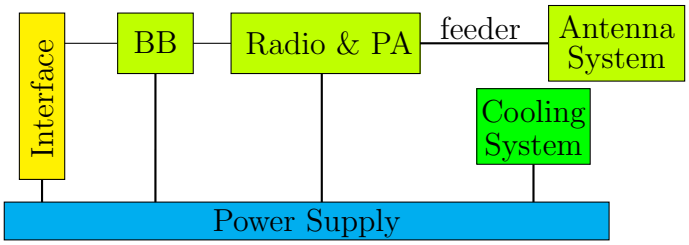


Figure 2.2: Architecture of a general BS showing one transceiver chain.

Table 2.1: Power Parameters for Different BS Types

BS Type	$\mathcal{N}$	P_{max} [W]	P_0 [W]	Δ_p	P_{sleep}
Macro	6	20.0	130.0	4.7	75.0
Micro	2	6.3	56.0	2.6	39.0
Pico	2	0.13	6.8	4.0	4.3
Femto	2	0.05	4.8	8.0	2.9

BSs. Power consumption is highest when the BS is in *active* mode (denote as P_{act}) but considerably reduces in *sleep* mode (denote as P_{sleep}). When a BS is in idle mode, it still consumes a significant but fixed amount of power. Simulations in [19] show that the input power of a BS varies linearly with its output power. For a K-tier HetNet, the power consumption of a k-th tier BS in active and sleep modes is expressed as

$$P_{cons,k} = \begin{cases} P_{act,k} = \mathcal{N}_k P_{0,k} + \Delta_k P_k, & 0 < P_k < \overline{P}_k \\ P_{sleep,k} = \mathcal{N}_k P_{sl_k}, & \overline{P}_k = 0 \end{cases} \tag{2.1}$$

where $\mathcal{N}_k$ is the number of transceiver chains, $P_{0,k}$ is the fixed power consumption at zero load, Δ_k is the slope of the load-dependent power consumption and $P_k \in [0, \overline{P}_k]$ where $\overline{P}_k$ is the maximum transmit power. These parameters are defined in [19] for different BS types as shown in Table 2.1 [19]. This power model also verifies Fig. 2.2 which shows that BS power consumption increases proportionally with the number of transceiver chains.

2.3 Energy-Aware Cellular Deployment

The biggest challenges currently facing mobile cellular network operators are two fold; on one hand, data traffic demand is increasing exponentially which calls for network upgrades and expansion together with BS densification to improve capacity significantly. On the other hand, the associated CAPEX and OPEX have affected operator revenues which are expected to begin shrinking from 2018 onwards [4]. The cost of energy constitutes a large and rising proportion of total OPEX which is a direct consequence of the required BS densification. This has motivated researchers in both industry and academia to seek network solutions that can enhance network capacity and minimize energy consumption of current and future networks.

Furthermore, ecological concerns about greenhouse gas emissions associated with network operation has also led to calls for more energy efficient cellular networks. It was estimated in 2013 that the ICT industry was responsible for about 10% of global electricity consumption and about 4% of CO2 emissions (which is expected to double by 2020) [21], [44]. The increasing prevalence of smartphones and tablets has increased data demand; for example, the year 2012 saw data consumption of 10 GB per subscriber and it is estimated to increase to 82 GB by 2020 [45]. These smart devices and their associated data-hungry applications, which often require latest cellular technologies, have directly increased energy consumption; for instance, an LTE network consumes 60 times more energy than a 2G network to provide the same coverage [46].

2.3.1 Traffic Profile vs Power Consumption

Since mobile cellular networks are traditionally planned to meet peak traffic during the busy hour, their utilization drops during off-peak periods. However, a typical cellular traffic profile shows very deep variations between peak and off-peak traffic levels [21], [47], [48]. Generally on a typical day, these two main factors influence the traffic profile [49]-[50]:

1. The daily traffic profile varies in time, from very low levels (early morning hours) to high peaks (lunch time or evening hours).

2. The daily spatial movement of a large number of mobile subscribers from residential areas to the office and commercial districts in the morning and their return in the evening hours.

These two factors require that network capacity is provided to meet the peak traffic demand at all times and in all geographical locations. However in more than half of observed areas, the maximum-to-minimum traffic ratio is larger than five and it is more than ten in 30% of these areas [51]. Therefore there is a need to closely adapt network energy consumption to the spatiotemporal variations in the traffic profile.

Furthermore, traffic load at a BS is a major determinant of its power consumption [21], [52], [53], [54]. Whereas an idle BS consumes a fixed amount of power [19], its the load-proportional power consumption constitutes more than half of overall consumption at peak load; for instance in LTE, about 60% of power consumption is load-proportional [55]. The rationale of a load-proportional energy profile is that at high load, more transmission resources are required.

Ideally, the fixed power consumption of a BS should be close to *zero* when it is idle; in other words, all energy consumption should be perfectly load proportional [56].

2.3.2 Energy Saving Approaches

Researchers are considering the following categorical approaches to minimize the energy consumption of future high capacity networks without affecting their QoS [21]:

- Using energy efficient hardware
- Using renewable energy sources
- Strategically deploying HetNets
- Optimizing radio transmission techniques
- Sleep mode techniques

Energy efficient hardware: Since BSs are responsible for about 57% of the total network power consumption, energy efficient BS hardware can potentially result in huge savings [57]. For instance the PA, which consumes the most power in a BS, is hugely inefficient, dissipating over 80% of its input power as heat. Whereas the PA power efficiency can potentially be improved to 70%, giving huge energy savings, the implementation cost is very high [57].

Renewable energy sources: renewable sources of energy such as wind and solar do not emit any greenhouse gases and are abundantly available in some areas [49], [50]. In under-developed areas with no grid power, off-grid BSs which use hydrocarbon fuels are ten times more costly to operate [33], [58]. In such scenarios, renewable energy powered BSs would be ideal. However, renewable energy can be intermittent which might affect continuous reliable service provision. Therefore hybrid systems, which combine both renewable sources and fuel-powered generators (or grid power), are perhaps a more desirable solution [59].

HetNet Deployment: Small BSs can be located strategically to reduce the propagation distance of users and potentially reduce energy consumption. The benefits are maximized when small BSs are deployed to boost coverage in the edge region or to boost capacity in hotspots. Small BSs can also benefit the network by reducing the required density of macro BSs which consume significantly more energy. However, deploying too many small BSs may reverse this benefit. Significant energy savings can be achieved by applying sleep mode techniques in such dense HetNets [54], [60], [61], [62]. For instance, closed access femtocell APs should switch off automatically until a registered user needs to use them. Therefore their sleep mode algorithm may be controlled via user activity detection [63]. This algorithm may also be extended to open access femto APs if they are not a major contributor to the overall outdoor coverage.

Radio transmission techniques: This involves techniques such as MIMO, cognitive radio, cooperative transmission, resource allocation and channel coding which can all improve EE of cellular networks [31], [64], [65], [66], [67], [68]. Normally there is a tradeoff between EE and

other performance measures since they are often contradictory. The following four tradeoffs related to EE have been identified in [52], [53] to guide the design of green networks:

1. Deployment efficiency-EE: Deployment efficiency (DE) is essentially the throughput per unit deployment cost where the total cost includes both OPEX and CAPEX. DE and EE are sometimes contradictory; for instance if few BSs are deployed (low deployment cost), the cells are larger which requires high transmit power and can worsen EE [52].

2. SE-EE: SE defines the throughput per unit bandwidth and is a common criterion used to study the performance of cellular networks. In some cases, EE and SE are contradictory; for example if transmit power is increased to enhance SE, EE may worsen as a result. Consider an AWGN channel in which SE is expressed as

$$\eta_{SE} = \log_2(1 + \text{SNR}) \quad [\text{b/s/Hz}] \tag{2.2}$$

where SNR $= \frac{P_t}{No\mathcal{B}}$, P_t is the transmit power, No is the AWGN power spectral density and $\mathcal{B}$ is the system bandwidth. Hence, EE is expressed as

$$\eta_{EE} = \frac{\mathcal{B}\eta_{SE}}{P_{cons}} = \frac{\mathcal{B}\eta_{SE}}{(2^{\eta_{SE}} - 1)NoB + P_c} \quad [\text{b/J}] \tag{2.3}$$

where $P_{cons} = P_c + P_t$ and P_c is the fixed zero-load (or circuit) power consumption. In an ideal system where $P_c = 0$, EE is bounded as follows: (i) as $\eta_{SE} \to \infty$, $\eta_{EE} \to 0$; (ii) as $\eta_{SE} \to 0$, then η_{EE} converges to the constant $1/(No\ln 2)$ [52], [53]. However, when $P_c > 0$, the EE-SE relationship is defined by a bell-shaped curve as shown in Fig. 2.3. Moreover, EE reduces when P_c increases which is intuitive.

3. Bandwidth-power: Using the Shannon capacity equation, the relationship between transmit power and bandwidth expressed as

$$P_t = NoB(2^{R/W} - 1) \tag{2.4}$$

where R is the capacity in b/s. This shows that if bandwidth is increased, transmit power can be reduced to maintain the same target rate. In practice however, besides the fact that bandwidth is limited, other energy costs related to the circuit power consumption actually increase with bandwidth. Analysis in [52] shows that for a target EE in a practical system, the power-bandwidth relationship is non-monotonic.

4. Delay-power: Delay or latency is a QoS measure that measures the composite signal processing and propagation time. In an AWGN channel, a bit of information takes $t_b = 1/R$ seconds to be transmitted. Therefore, the relationship between delay and power is

$$P_t = No\mathcal{B}(2^{1/t_b W} - 1) \tag{2.5}$$

which shows that P_t decreases monotonically with t_b. However in a practical network, the also delay includes queuing time. According to [52], the delay-power relationship is no longer simply monotonic. In addition, various services especially in HetNets have different delay tolerances and this can impact the optimal power usage.

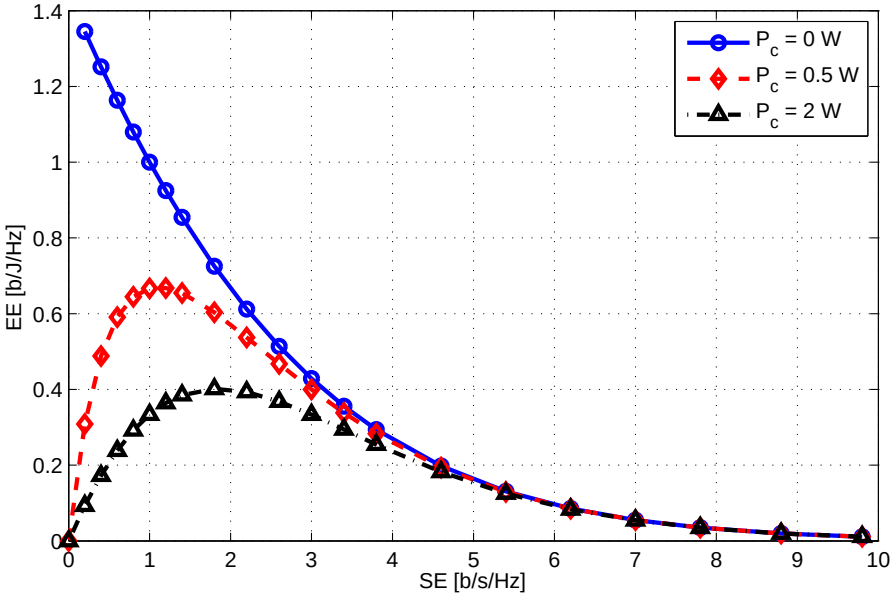


Figure 2.3: Tradeoff between EE and SE in an AWGN channel ($\sigma^2 = 1$).

Sleep mode techniques: This involves switching off or putting some components or the entire BS or network into sleep mode during off-peak hours [21]. Sleep mode algorithms may be centralized at the switching centre, distributed in the BSs or instigated by user activity [60]. Existing works on sleep mode techniques are discussed in the next section.

2.4 Sleep Mode Techniques

Generally as the traffic demand varies in space and time, the network can selectively switch off entire BSs or put them into sleep mode to save energy. Local sleep mode approaches may include switching off individual components such as the PA, base band unit and cooling system to save the bulk of energy consumed by the BS [19]. For instance in [69], the authors investigate an energy saving technique where a 3-sectored macrocell autonomously varies its number of sectors based on the prevailing traffic conditions. At low traffic, omnidirectional operation of the macrocell site can give energy savings of 30%. This dynamic desectorization of the macrocell can be achieved by connecting the three sector antennas through one transmission chain via a 3-way power splitter; the other two transmissions chains are shut down.

However, global (or network-wide) approaches to the analysis of network energy consumption are potentially more beneficial and can give better energy savings. Global techniques involve analyzing the whole traffic profile over a number of sectors or the entire network and switching off an appropriate number of BSs such that the remaining BSs provide the service. Authors in [70] argue that sleep mode decisions should consider the load profile in the wider geographical coverage area of the network as opposed to localized decisions although this makes its management more difficult. Authors in [71] show that coverage and energy performance of the network are improved when cells are switched off based on their levels of activity.

Furthermore, authors in [72] study an energy minimization problem subject to DL coverage and UL power constraints. Some macrocells are switched off and replaced by microcells coupled

with power adjustment techniques to maintain acceptable QoS. In [73], a joint optimization of the BS sleeping control and power matching schemes is performed to achieve flexible tradeoffs between power consumption and QoS. In [74], authors study the effect of sleep mode on average user capacity and optimize the transmit power of remaining BSs to maintain the outage target. In [75], authors proposed a technique of managing power consumption by switching off some BSs and balancing the prevailing traffic load between the remaining active BSs. In [76], author proposed a technique that prioritizes cells with the fewest users for sleep mode and this enhances coverage probability, average rate and EE of the network.

2.4.1 Sleep Mode Enablers

The implementation of sleep mode is not trivial because there are many dynamic parameters that have to be considered. Therefore, sleep mode algorithms have got to rely on several techniques to minimize their disruption of the network. Most of these techniques depend and require cooperation between BSs to maximize their energy saving potential.

Perhaps the biggest challenge to the implementation of sleep mode is the risk of emergence of coverage holes and overall drop in QoS once some BSs are put into sleep mode or switched off. An obvious solution is for the remaining BSs to *zoom out* and cover any coverage gaps. Zooming out requires BSs to increase their transmit power but this increase in transmit power may not significantly increase the total BS input power [19]. In [77], two algorithms are proposed to control the cell zooming process. In the centralized algorithm, a separate server controls the process; the server transmits appropriate messages to BSs that should go into sleep mode and to those that should zoom out. In the distributed algorithm, BSs in a cluster collaborate to switch off/zoom out appropriately so as to save energy while maintaining the QoS. Cell zooming can also be used to offload traffic from congesting cells (which zoom in or shrink their coverage) to cells with excess capacity (which zoom out to enhance their coverage area) [21].

User association is an important function in cellular network because it determines which BS each user connects to. It can also be used to enhance the performance of sleep modes. For example if a BS is put into sleep mode, its affected users connect to any of the other remaining BSs. However, user and BS locations and the prevailing load of all neighboring BSs has to be considered before users are assigned. In other words, an affected user should not necessarily connect to its nearest or 'best-SINR' BS if it receives better capacity (or more bandwidth) from a more distance or 'lower-SINR' BS [13], [21], [78], [79]. Therefore, an optimal user association scheme that utilizes BS cooperation is required to maximize energy savings and maintain acceptable QoS.

2.4.2 Implementation Approaches

The approaches to implementing sleep algorithms differ in two main ways [21]: (i) whether algorithms are dynamic (*dynamic switch schemes*) or run periodically at a fixed frequency (*fixed switch schemes*), and (ii) whether the algorithms are *centralized* or *distributed*. These

four approaches have implications on the complexity of implementation of sleep mode and the realizable energy savings.

In fixed switch schemes, the sleep mode algorithms runs a fixed number of times in a given period while in dynamic switch schemes, the algorithms runs dynamically in response to any changes in the network. Generally dynamic schemes have a better performance than fixed schemes [77]. Moreover, high-frequency fixed switch schemes also give more energy saving than their low-frequency counterparts [80]. Dynamic schemes are even more superior in medium-to-high capacity environments where they respond and adapt to the dynamic traffic profile [81]. It is argued in [81] that in coverage-limited environments (low traffic), fixed switch schemes are more beneficial because the coverage constraint prevails over sleep mode objectives.

However, dynamic schemes generally require more switching operations especially if the traffic is highly fluctuating which can impact energy savings. The costs of switching involve overheads, transient time, power for monitoring and switching, switching delays, and possible impact on the lifetime of BSs [21], [80]. This presents an interesting tradeoff between the cost of frequent switching and the absolute energy savings. Authors in [49] studied the cost of transient periods on overall energy savings and found that transient periods are very short (order of 1 minute) and therefore have no significant effect on the energy savings of sleep mode. However, since most legacy BSs are not designed for frequency switching (on and off), future designs should consider this requirement to further minimize transient period and its energy cost.

On the other hand, both centralized and distributed sleep mode schemes are dynamic in nature [21]. In centralized schemes, the algorithm is controlled by a central server all over the network or within a cluster of BSs [62]. BSs which have more information about the network environment cooperate to maximize energy saving while preventing coverage gaps. Where the algorithm is centralized in clusters, cluster-based controllers should cooperate to further enhance performance. In distributed algorithms, the user initiates sleep mode by deciding which BS it connects to. Then each BS can make an independent decision on whether to remain active or go to sleep depending on its level of association. This scheme has the advantage of having no coordination overheads between BSs. However, centralized schemes generally give better energy performance because they have more information to work with. To get a fair comparison of both approaches, the associated coordination overheads of centralized schemes should be considered [21], [82].

2.5 Energy Performance Metrics

The energy performance of a system can be quantified using a number of metrics. A comparison of the energy performance of different network designs is necessary to allow operators to incorporate such information into their long-term plans as they seek to reduce energy bills. Energy metrics can either be measured at the system level or node level [21]. Node level metrics provide useful insights into the energy savings possible at a single node such as a BS. However it is almost always the case that energy saving in one node is made possible by increased consumption in another node; for instance, if one node goes to sleep mode, another node zooms

out and takes on extra load. Therefore, system level metrics are simulations generally provide a more balanced and accurate performance indicator [83]. There are many metrics in existing works; the main ones are discussed in this thesis.

Among the classical metrics is a common metric called energy efficiency (EE), which quantifies the number of bits transmitted per unit amount of power consumption [84], [85] i.e.

$$\text{EE} = \frac{M}{E} \quad \text{[bits/Joule]} \tag{2.6}$$

where M is the network sum rate in bps and E is the energy consumption in Joules. Its reciprocal, called energy consumption ratio (ECR), measures the power consumption per delivered bit of information [86], [87]. Two other metrics, energy consumption gain (ECG) and energy reduction gain, use the ECR metric to compare two systems as follows [88], [89]

$$\text{ECG} = \frac{\text{ECR}_1}{\text{ECR}_2} \quad \text{and} \quad \text{ERG} = \frac{\text{ECR}_1 - \text{ECR}_2}{\text{ECR}_2} \times 100\% \tag{2.7}$$

where ECR_1 and ECR_2 are the ECRs of any two systems.

In networks with sleep mode schemes, energy savings at node level can be quantified by the fraction of time that a given node spends in sleep mode [21] i.e.

$$\text{Energy saving} = \frac{T_{sleep}}{T_{total}} \tag{2.8}$$

where T_{sleep} is the time spent in sleep mode and T_{total} is the total measurement time. However, this metric depends assumptions such as (i) fixed power consumption in active and sleep modes, and (ii) instantaneous switching with no power consumption.

At system level, a common metric is area power consumption (APC) which quantifies the absolute amount of power consumed by a network per unit area. For instance in a general K-tier HetNet, the APC in (Watts/m^2) is calculated as

$$\text{APC} = \sum_{k=1}^{K} \lambda_k P_{cons,k} \tag{2.9}$$

where λ_k is the BS density and $P_{cons,k}$ is shown in (2.1) as either $P_{act,k}$ or $P_{sleep,k}$.

Another common metric called performance indicator (PI) was proposed by ETSI as [21], [90]

$$\text{PI}_{rural} = \frac{\text{Total coverage area}}{\text{Power consumption}} \quad \text{and} \quad \text{PI}_{urban} = \frac{\text{Number of users in peak hour}}{\text{Power consumption}}. \tag{2.10}$$

In rural areas where user density is low and sparsely distributed, PI_{rural} uses total coverage area. In an urban area, energy consumption per user is a more appropriate metric.

This thesis variously applies EE and APC to quantify the energy performance of homogeneous networks and HetNets.

2.6 HetNets as a Paradigm Shift

Multi-tier HetNets are different from traditional homogeneous cellular networks and their architecture raises important challenges in deployment, mobility support, interference management, load balancing and bandwidth allocation. Finding suitable solutions to these challenges requires a paradigm shift from the homogeneous network in the following ways.

2.6.1 Network Topology

In traditional macrocell networks, cells are commonly represented as hexagons on a grid with a macro BS located at the centre of each cell. The boundaries between neighboring hexagons represent the cell edge regions where handovers are initiated. However, small BSs which typically use omnidirectional antennas cannot be represented by a similar structure. They transmit much lower power and their DL coverage areas are relatively small. Femtocells for example may be scattered over the network area depending on where subscribers install them. Even operator-planned pico BSs that are strategically located in the cell edge or hotspot areas cannot conform to the regular lattice structure [6], [7], [24]. Therefore system models for analyzing HetNets must simulate this non-regular topology to enhance accuracy of predicted results.

2.6.2 User Association

Traditionally users connect to a BS which provides them with the best SINR. Under the assumption that all BSs are fully loaded, such a strategy maximizes network throughput [6]. However, the transmit power of a pico BS may be up to 20 dB lower than that of the macro BS. This causes a large disparity between the DL coverage areas of the macro BS and overlaid pico BSs which pushes the handover boundary on the DL very close to the pico BSs. However, the power disparity on the UL is different since users in all cells have the same power configuration. This creates a mismatch between the UL and DL handover boundaries, making user association more challenging in HetNets. If user association is based on the DL received signal, macro BSs are likely to remain heavily loaded due to their high transmit power and antenna gain while small BSs are likely to remain lightly loaded due to their small coverage areas. This is undesirable because users in the macrocell tier are likely to have much lower bandwidth compared to users in the small cell tier which is unfair and can potentially affect the sum rate. Besides, many small BSs may remain idle which reduces the SE of the network [91]. In HetNets, user association is a major design feature that affects many other performance indicators such as mobility support, interference management, load balancing, etc [14].

2.6.3 Cell Range Extension

Network capacity can be enhanced by harnessing the potential of small BSs to offload traffic from the congesting macrocell tier to the small cell tier. This offloading potentially enhances

the average bandwidth per user in the HetNet and it can be used to implement fairness in the sharing of network resources between users [91]. To ensure traffic offloading to the small cell tier, the coverage of small BSs needs to extend outwards so as to encompass more users from the underlaid macrocell tier. This concept is known as cell range extension [5], [91], [92], [93]. Cell range extension can easily be achieved by a simple technique called *association biasing* where a user is biased by some degree to favor small BSs over macro BSs [6], [92].

For instance, assume a general K-tier HetNet and consider an association scheme where a user connects to the BS from any tier that provides the best instantaneous SINR. Different bias values are applied to different tiers; for example, all BSs in the k-th tier have a bias value $\beta_k \geq 1$. Before cell association, this bias value is multiplied with the SINR received from all k-th tier BSs i.e. $\beta_k \text{SINR}_{k,i}$ where $\text{SINR}_{k,i}$ is the SINR received from the i-th BS in the k-th tier. Therefore if $\beta_k \geq \beta_j$, then cell association is biased to the k-th tier. The biased SINR considered by the user for cell association purposes is expressed as [93]

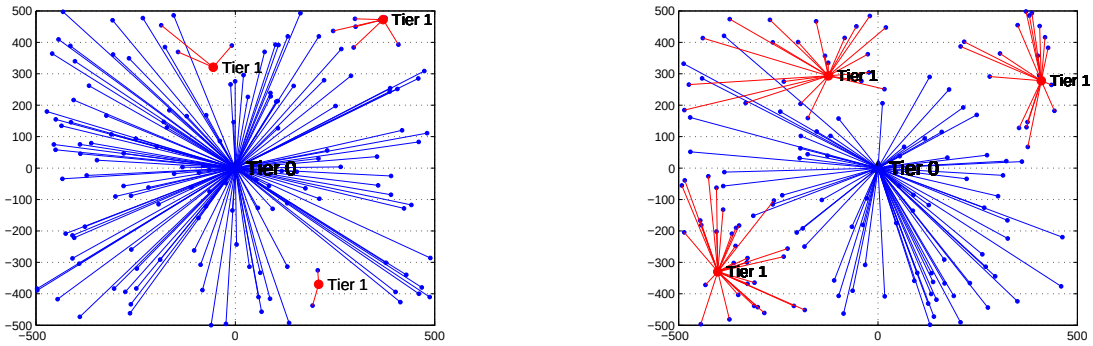
$$\text{SINR}^b_{k,i} = \beta_k \text{SINR}_{k,i} = \frac{\beta_k P_{k,i} G_{k,i}}{\sum_{j \in B_j / \{b_o\}} P_j G_j + \sigma^2} \tag{2.11}$$

where B_j is the set of BSs in the j-th tier, b_o represents the parent BS, $P_{k,i}$ is the transmit power of the parent BS, $G_{k,i}$ is the pathloss of the channel between the parent BS and the user, P_j and G_j are the transmit powers and pathloss of all other interfering BSs respectively and σ^2 is the additive noise power. The user associates with the i-th BS in the k-th tier if $\beta_k \text{SINR}_{k,i} \geq \max_{j, j \neq k} \beta_j \text{SINR}_{j,i}$. Therefore biasing can easily be used to achieve load balancing.

The premise of biasing is that although a user might receive better SINR from a macro BS, a small BS might avail the user with more bandwidth to maximize the rate [92]. To enhance load balancing in such a case, user association is biased towards the small cell tier by assigning small BSs with a larger bias value relative to macro BSs. For instance, simulations in [94] showed that average user rate increases with the bias value up to a certain point and starts to reduce. This presents an optimization problem of choosing a biasing factor that maximizes user throughput. In general, the biasing factor should depend on the prevailing load conditions to keep both macro BSs and small BSs sufficiently loaded so as to maximize the average user rate [6]. In other words, user association should not only consider the received signal or SINR but it should also consider the amount of accessible bandwidth at the BS it connects to. Although biasing is difficult to optimize, it can be used to achieve this objective to some extent [6].

Fig. 2.4a shows user association between a macro BS and three pico BSs based on the "best received SINR" criterion. Without biasing, the macro BS dominates and most users are connected to it. Fig. 2.4b shows more balanced user association when pico BSs are associated with a bias value which is 15 dB greater than that of the macro BS.

Biasing also has significant implications on the energy performance of the HetNet. While biasing is useful, inappropriate biasing is counterproductive and can worsen both the SE and EE performance of the HetNet [95], [96]. These works show that as biasing is increased, the EE also increases up to a certain point beyond which it begins to reduce. Therefore, there is an optimal bias value at which EE is maximized. In addition, analysis on a biased two-tier



(a) Without biasing

(b) With a biasing of 15 dB

Figure 2.4: Impact of cell biasing on cell association.

HetNet in [97] showed that there is an optimal biasing value at which the APC of the HetNet is minimized. Besides the load profile in the network, appropriate biasing should also consider the number of tiers in the HetNet, its deployment configuration in terms of BS densities and transmit powers, and the user association scheme.

2.6.4 Mobility Support

Cellular networks must support mobility by handing over terminals in on-going sessions from one cell to another along their path. HetNets provide a new challenge to mobility support mainly because some BSs have very small coverage areas. For a slow-moving user such as a pedestrian, it may be beneficial to hand it over to a small BS along its path to assist in traffic offloading and also to give a better quality link. For users moving at vehicular speeds, any handovers to small BSs may last a few seconds at most. Thus it may be beneficial to avoid such handovers to minimize costly signaling overheads and potential handover delays and call drops. However a user that does not handover may generate strong interference to the small BS on the UL or suffer strong interference from the small BS on the DL. In OFDMA, intelligent and mobility-aware resource allocation can mitigate this interference. Studies by 3GPP have shown that the average handover failure rate of a macro-pico HetNet may reach as high as 60% which is twice as high as in macrocell only networks [6], [7], [98]. Since biasing is used to artificially increase the small cell coverage area, [99] shows that an optimal *speed-dependent* bias value exists at which the speed-dependent coverage probability of the HetNet is maximized.

2.6.5 Interference Management

Interference is a big challenge that can greatly limit user throughput in HetNets. Due to the high transmit power of macro BSs, the handover boundary between macro BSs and overlaid small BSs is pushed close to the small BSs. As a consequence, small BS users are likely to suffer very high interference from the macro BS on the DL. Similarly, small BSs can cause high interference to nearby macro BS users [87]. On the UL, nearby macro BS users can cause significant interference at the small BS. Similarly, small BS users can cause UL interference at the macro BS [5], [6]. To illustrate, consider the following interference scenarios in a two-tier

HetNet of macro BSs and small BSs:

- Cross-tier interference involves any two BSs in different tiers (either a macro BS causing interference to a small BS or vice versa). Fig. 2.5 shows an example of cross-tier interference suffered by a femtocell user from a nearby macro BS.
- Co-tier interference exists between any two BSs in the same tier such as small BS to a small BS. Fig. 2.5 shows an example of co-tier interference at the user in Picocell 2 caused by the DL transmission of the BS of Picocell 1.

Biasing association to favor small BSs worsens the interference problem since the high-power macro BSs become interferers instead [6], [98]. As Fig. 2.4b showed, a pico user may receive up to 15 dB more interference from a nearby macro BS. This interference, coupled with the low transmit power of small BSs, can easily wipe out the intended gains of cell range extension. It is therefore essential to design robust techniques that can manage or control this interference [91]. Two of such techniques use coordination of resource usage by BSs to manage interference and maximize average capacity of the HetNet.

Inter-Cell Interference Coordination (ICIC): In ICIC, a small BS interacts with a high-power macro BS to coordinate their transmission activities and avoid interference in both control and data channels. Using resource partitioning, the interfering macro BS can use part of the resources and leave the other part to be used by the small BS. For example, the macro BS may reserve subframes within a radio frame for pico BSs, the number of which depends on the loading and bandwidth requests on each BS [7], [91], [100]. In general, this coordination could involve transmission power control to facilitate spatial reuse of resources, beam forming to reduce interference in specific directions, etc.

Slowly-Adaptive Interference Management: In this approach, the coordination of resources takes a long-term view compared to ICIC. The transmit powers of BSs and UEs are negotiated and allocated longer time scales than the radio frame. The negotiation is premised on the objective of maximizing a given network metric such as network capacity, average user rate, fairness, etc. The algorithm to control this negotiation may be centralized or distributed. In the centralized case, the central server must have access to all inputs needed for the optimization algorithm. In case centralized control is not possible or feasible, distributed algorithms can be used where BSs negotiate via the X2 interface and/or over-the-air (OTA) messages. For example, small BSs can send their load and bandwidth requests via the X2 interface and the macro BSs can reply via the same interface. The OTA messages can be used where the X2 interface is not available or does not meet the delay and/or bandwidth requirements [91].

2.6.6 Backhaul Challenge

Dense HetNets are expected to be a major composition of the evolution towards high capacity 5G networks. However high capacity in the access layer must be matched by sufficient capacity in the backhaul to support end-to-end QoS. Most works assume the existence of a perfect

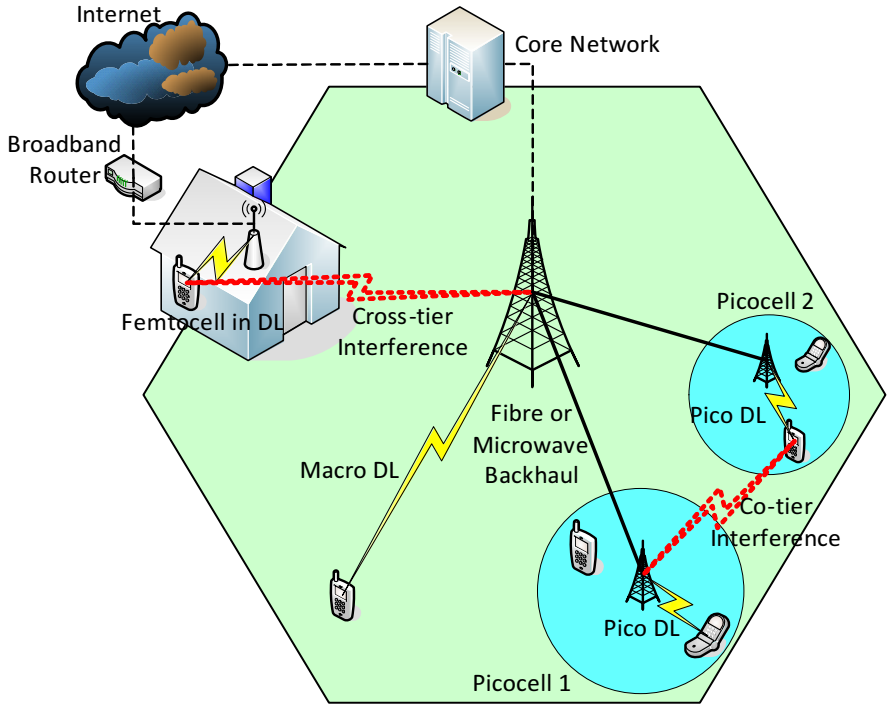


Figure 2.5: HetNet layout showing co-tier and cross-tier interference

backhaul but potential bottlenecks need to be considered in dense and high-capacity HetNets [6]. In addition, current backhaul technologies based on DSL and NLOS microwave links have limited capacities and may be unsuitable [101]. Some works have proposed backhaul-aware user association algorithms to efficiently utilize the available backhaul resources [102], [103].

2.7 Deployment Scenarios of HetNets

In a multi-tier HetNet, bandwidth allocation is a challenging task that may result into increased cross-tier interference and lower network SE. Several techniques can be used to allocate the available frequency band among the tiers with varying consequences on the network coverage and rate performance as discussed next.

2.7.1 Co-channel Deployment

In co-channel deployment, small BSs and macro BSs are deployed to use the same frequency carrier for their transmissions. This technique avoids carrier segmentation and is scalable to all LTE system bandwidths and does not require carrier-aggregation-capable mobile terminals. However, the high potential for cross-tier interference can greatly limit the throughput of end users. Resource partitioning is an effective technique to manage interference in a co-channel deployment by ensuring that different channels are allocated to the macro BSs and overlaid small BSs at any given time [5]. The principles of resource partitioning can be applied to both the DL and UL. In LTE Rel-10, adaptive resource partitioning is enabled by allowing BSs to exchange resource scheduling information via the X2 backhaul [5]. High-power macro BSs avail small BSs with information about their scheduled resources for DL transmissions. This

information also identifies subframes which are to be left unutilized (almost blank subframes (ABS)) so that small BSs can use them during this period.

2.7.2 Multi-carrier Deployment

In multi-carrier deployment, macro BSs and small BSs are deployed on different frequency carriers F_1 and F_2. For example, in a two-tier macro-femto HetNet, macro BSs use carrier F_1 while closed-access femtocells use carrier F_2. Carrier F_2 is necessary to ensure that closed-access femto APs do not cause a lot of interference to nearby co-channel macrocell users on carrier F_1. This technique can also be used with other small BSs (open-access femto APs, pico BSs, micro BSs, etc). Multi-carrier deployment has poor temporal and spatial SE because it restricts each carrier to a single tier of the network [5].

2.7.3 Carrier Aggregation

An improved technique is carrier aggregation where the macrocell tier can access both carriers while the closed-access femto BSs use only one carrier F_2. In this arrangement, macro BSs have complete coverage on carrier F_1 and partial coverage on carrier F_2 (due to closed-access femto BSs appearing as coverage holes to restricted users). However, only carrier-aggregation-capable mobile terminals can exploit both carriers. This technique also requires a lot of spectrum and does not scale down to smaller bandwidths like a single carrier LTE system is capable of doing. This is due to the inefficiencies associated with smaller bandwidths and the limitation of peak throughput for users that are not carrier-aggregation-capable [5].

Carrier aggregation is also applicable to open small BSs. In this case, it is more beneficial that both macro BSs and small BSs have access to both carriers. Macro BSs overwhelm small BSs in terms of coverage area due to their high transmit power and antenna gain. To counter this and improve performance, the macro BS transmits nominal power on carrier F_1 to maintain normal macrocell coverage while open small BSs use carrier F_2. Macro BSs also access carrier F_2 but transmit less power than the nominal value they transmit on carrier F_1. Small BSs are also allowed to access carrier F_1 to enhance their capacity around the small cell BS [5].

2.8 Summary

Energy consumption is a major contributor towards overall OPEX incurred by operators. In remote areas with no grid power, BS equipment is powered by diesel generators which worsens both OPEX and greenhouse gas emissions into the atmosphere. This chapter identified the BS as the highest contributor to overall network energy consumption. Within the BS, the PA contributes about 60% of the total consumption. A linear power model is used to calculate the input power consumption of various BS types. Metrics such as EE and APC facilitate comparison of the energy performance of different mobile network systems.

Due to exponentially increasing traffic demand, operators must continue to expand the capacity and improve the quality of their networks. With falling revenues due to increased competition, a cheaper and energy efficient solution is required to improve network capacities and overall QoS. The HetNet is a new kind of deployment that combines the traditional macrocell layer with low-power and small-coverage BSs to provide targeted capacity and coverage enhancements. However, the HetNet paradigm introduces different engineering challenges that must be overcome to achieve these objectives. Such dense HetNets can become underutilized as traffic demand inevitably varies spatiotemporally. During low traffic conditions, sleep mode schemes can be used to switch off some BSs and adapt energy consumption to changes in traffic demand.

Chapter 3

Stochastic Geometry Approach to Network Analysis

3.1 Introduction

Traditionally, mobile cellular networks have been analyzed by assuming regular topologies where cells are represented as hexagonal, square or circular shapes with a BS node located at the center of each cell. Mobile users are either placed in the cell according to some defined distribution model, or they are also located deterministically [104]. Therefore, grid topologies assume that all cells have the same coverage area, a scenario that is very unlikely especially in urban and suburban areas due to the random clutter. With such topologies, it is very difficult to obtain tractable analytical models to quantify the SINR performance of randomly located users. Monte Carlo simulations are usually performed using software to investigate coverage, capacity and other performance measures. However, Monte Carlo simulations are very intensive and time consuming and the results are always difficult to verify and may not always be reliable due to human error in the coding [105].

Furthermore, due to the densification of current and future networks to support the exponentially increasing data traffic demand, grid topologies will become even more unreliable to predict the performance of real mobile networks. Moreover, HetNets combine several tiers of different BS types which have different transmit powers and deployment strategies. In cases where small BSs (such as femtocells) are deployed by subscribers, grid-based topologies cannot capture the reasonably high probability that interfering BSs may be located relatively close to each other. Hence, grid-based analysis is rather optimistic and usually gives results that are an upper bound on actual performance [106].

Recently, a new tractable approach to network analysis called stochastic geometry has become popular. In the stochastic geometry approach, mobile cellular networks are assumed to be completely random such that both BSs and users are independently located according to appropriate spatial point processes of defined intensity in the Euclidean plane [105], [106], [107], [108]. This analytical approach ensures the lack of edge effects since the network extends indefinitely and therefore considers all possible interference. In addition, it simulates cells with

varying shapes and sizes. Although this approach facilitates very useful network analysis, it also has a few obvious weaknesses: (i) due to the independent distribution of BSs, interfering BSs may be deployed too close to each other which may generate higher aggregate interference than likely in a planned real network; (ii) in reality, BS deployment is not independent due to site planning, places where BSs cannot be deployed such as water bodies, difficulties with site acquisition, etc; (iii) even in customer-driven deployment (such as femtocells), deployment is unlikely to be uniform over the entire plane since deployment decisions are influenced by many other factors such as real-world spatial traffic intensity which is not uniform, distance from the nearest macro BS, etc. As a consequence of these weaknesses, this approach to network modeling and analysis defines a lower bound on the performance of a practical network [106].

This chapter presents a discussion of the mathematical preliminaries required to analyze cellular networks using the stochastic geometry approach. This is followed by a discussion of important and existing analytical results, particularly those describing coverage probability and average rate performance of both homogeneous networks and HetNets.

3.2 Mathematical Preliminaries

Stochastic geometry is simply the study of random spatial processes and is applied in a wide range of fields including communications, astronomy, forestry, medicine, etc. In communication engineering, stochastic geometry has found particular application in the modeling and analysis of spatially-located mobile cellular networks to derive simple and tractable expressions for coverage, capacity and other performance measures [106], [107], [109].

3.2.1 Spatial Point Processes

Consider $\mathbb{N}$ to be the set of all sequences $\phi \subset \mathbb{R}^2$ which satisfy the two conditions [107], [110]:

- Finite: Any bounded set $A \subset \mathbb{R}^2$ contains a finite number of points.
- Simple: No two or more points are in the same location i.e. $x_i \neq x_j$ for $i \neq j$.

If these two characteristics are satisfied, then the point process in $\mathbb{R}^2$ is defined by a random variable taking on the values in the space $\mathbb{N}$. The point process is denoted by Φ, and its instance is denoted by ϕ. Given a point process Φ, the number of points of the point process within a bounded set $A \subset \mathbb{R}^2$ is denoted $\Phi(A)$. By definition, a stationary point process is one whose distribution is invariant to any translation. In other words, if $\Phi = x_n$ is stationary, then $\Phi_x = x_n + x$ has the same distribution for all $x \in \mathbb{R}^2$. Therefore, statistically the point process is similar regardless of where it is viewed from within the space. The density of a stationary point process Φ is obtained as [107], [109], [110]:

$$\lambda = \frac{\mathbb{E}\left[\Phi(A)\right]}{|A|}, \quad A \subset \mathbb{R}^2. \tag{3.1}$$

3.2.2 Stationary PPPs

The stationary Poisson Point Process (PPP) is widely used in literature and is very popular due to its independence property which eases and simplifies network analysis. In addition to being stationary and simple, a PPP is also isotropic. Isotropy defines its invariance to rotation i.e. if $\Phi = x_n$, then $\mathbf{r}\Phi_x = \mathbf{r}x_n$ where $\mathbf{r}$ is the rotation around the origin [107], [109], [110]. For a stationary PPP Φ of density λ, the number of points in a bounded set $A \subset \mathbb{R}^2$ has a Poisson distribution with mean $\lambda|A|$ i.e.

$$\mathbb{P}(\Phi(A) = n) = \frac{(\lambda|A|)^n}{n!} e^{-\lambda|A|}. \tag{3.2}$$

The single value λ completely characterizes a stationary PPP. In addition, the number of points in disjoint sets $A \subset \mathbb{R}^2$, $B \subset \mathbb{R}^2$, ... are independent. Furthermore, the points within a set $A \subset \mathbb{R}^2$ are independent and uniformly distributed [107], [109]. Therefore, the following steps can be followed to simulate a PPP of density λ over the region $A = [-d, d]^2$:

- Generate a Poisson distributed number N which represents the number of points. In MATLAB, this is written as `N=poissrnd`($\lambda|A|$).
- Generate N independent points that are uniformly distributed over the whole area. In the $\mathbb{R}^2$ plane for example, these points may be generated as x_i and y_i which are both uniformly distributed. In MATLAB, the code may be written as `(x,y) = unifrnd(-d,d,N,2)` or `(x,y) = 2d[rand(N,2)-0.5]`.
- For any two points x, $y \subset \mathbb{R}^2$ where $x = (x_1, x_2)$ and $y = (y_1, y_2)$, the Euclidean distance is measured as $\| x - y \| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$.

3.2.3 The Thinning Property of PPPs

Consider a PPP $\Phi(\lambda)$ and assume that a node stays with a probability of p or is removed from the process with a probability of $1 - p$, independently of other nodes. If Φ_p denotes the remaining nodes and Φ_q denotes those that are removed, then $\Phi = \Phi_p \cup \Phi_q$. This gives rise to the following three thinning properties [107], [109], [110]: (i) Φ_p has the density $p\lambda$, (ii) Φ_q has the density $(1 - p)\lambda$, and (iii) Φ_p is independent of Φ_q.

3.2.4 Poisson-Voronoi Tessellation

Stochastic geometric analysis of cellular networks is very suitable particularly where the network density is very high and the locations of BS nodes are highly randomized. Therefore, it is a very useful tool for the analysis of future networks which are likely to comprise a dense deployment of different types of BSs. A homogeneous PPP is one in which the density of the point process is constant over the whole plane. In homogeneous networks, all BSs are assumed to transmit the same amount of power. Therefore, ignoring shadowing and other channel effects, a user always

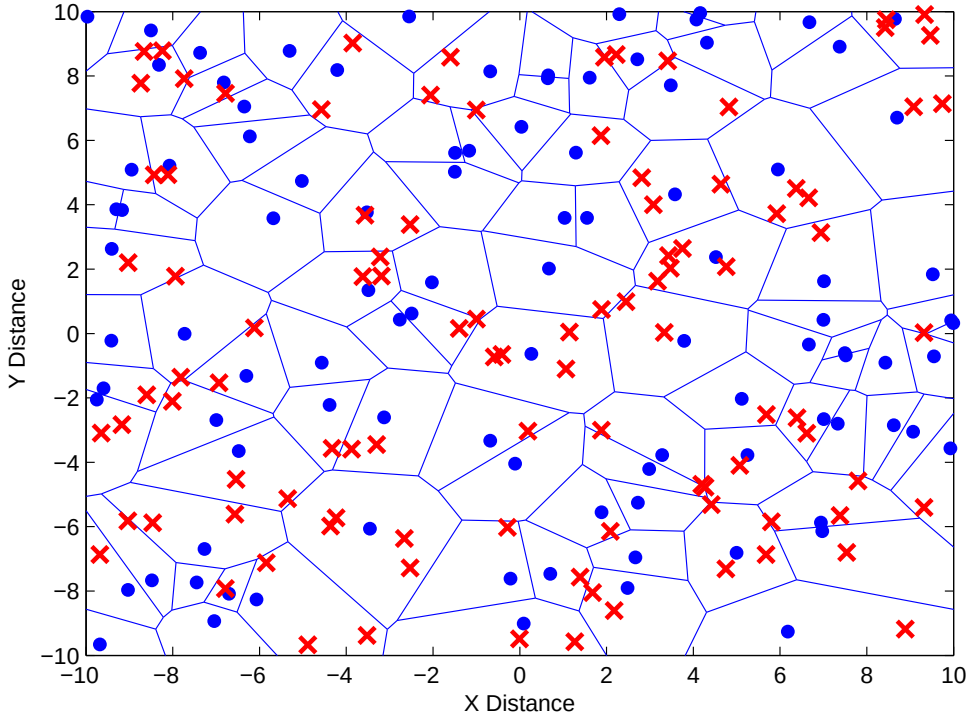


Figure 3.1: Layout of a PPP-based homogeneous network where macro BSs (represented as •) and users (represented as ×) have the same density i.e. $\lambda_b = \lambda_u$.

connects to its closest BS node [106], [109], [111], [112]. With added complexity, some works consider shadowing by introducing random displacement in BS locations [105], [113], [114].

When shadowing is ignored in a homogeneous network, all users connected to a given BS are located in a polygonal cell such that the distance to their parent BS is always less than the distance to all other BSs. This polygonal cell, called a Voronoi cell $V_b \subset \mathbb{R}^2$, is such that

$$V_b = \{x \in X : d(x, S_b) \leq d(x, S_k),\ \forall\, k \neq b\} \tag{3.3}$$

where $d(a, b)$ computes the distance between point a and site b. The set V_b consists of all points which are associated to site S_b and whose distances to site S_b are always less than or equal to their distances to all other sites in the set S_k. When the users and BSs are distributed according to a PPP, the resulting Voronoi tessellation is called a Poisson Voronoi (PV) tessellation and is illustrated in Fig. 3.1. A PV tessellation results when points generated according to a PPP grow at the same isotropic rate until their boundaries get into contact [107]. In the context of homogeneous cellular networks, a PV tessellation is a special case of a *weighted* PV tessellation in which all BSs transmit the same power [106], [115]. Therefore, a weighted PV tessellation defines a typical HetNet in which BSs in different tiers transmit at different power levels and is illustrated in Fig. 3.2 [116].

3.3 Assumptions of the PPP-based Model

The analysis in this thesis is based on several network assumptions that are aimed at either easing network analysis or providing simple and tractable results. The main mathematical assumptions are presented in this section. These assumptions will be justified and any existing

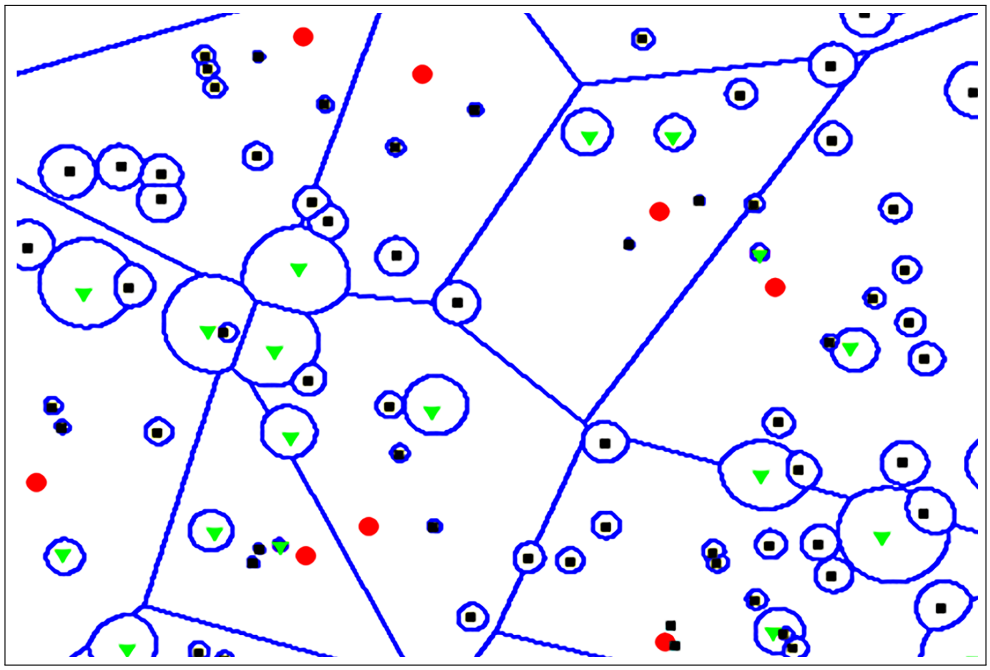


Figure 3.2: Layout of a PPP-based 3-tier HetNet of macrocells (large circles), picocells (triangle shpwing s) and femtocells (squares) where $P_b = 100P_s = 1000P_f$ and $\lambda_f = 4\lambda_b = 8\lambda_b$.

works that make similar assumptions will be referenced. After discussing all major assumptions, a HetNet system model will be presented. Since a homogeneous network is a special case of a HetNet with only one tier, the system model can be specialized accordingly.

3.3.1 Rayleigh Fading

This thesis considers Rayleigh fading to model the small scale fading effects of the wireless channel on signal propagation. Rayleigh fading offers significant simplicity and tractability in the analysis of PPP-based networks. However, Rayleigh fading may not always work especially in small cells where the chance of a direct LOS link is high. Some works have applied more generic fading models at the cost of decreased tractability, for instance in [105] and [106]. In [105], Nakagami-m fading is applied in a general K-tier load-aware HetNet while [106] considers a generalized fading phenomena in a homogenous network.

3.3.2 No Shadowing

Another major assumption made in this thesis is that long term shadowing is ignored. This allows the topology of homogeneous networks and HetNets to resemble a PV and weighted PV tessellation respectively which improves the tractability and simplicity of the analysis. Long term shadowing is often ignored in many existing works; for example [71], [94], [106], [108], [117]. Since shadowing affects cell selection, it potentially has an effect on the coverage, rate and energy performance of a cellular network. However in order to verify the accuracy of this assumption, Monte Carlo simulations in [106] showed that long term shadowing does not significantly alter the analytical results.

Some works consider long-term shadowing in their analysis of PPP-based networks such as

[105], [113], [114], [118]. In general, consider a K-tier HetNet whose k-th tier BSs are located according to a PPP $\Phi_k \subset \mathbb{R}^2$ of intensity λ_k. If the serving k-th tier BS is located at x_k, then the power received is expressed as $P_r(x_k) = P_k h_k \mathcal{X}_k \|x_k\|^{-\alpha}$ where $\mathcal{X}_k$ models long-term shadowing. Long-term shadowing is simulated by introducing random displacement in the original BS locations such that the received power becomes $P_r(x_k) = P_k h_k \|\mathcal{X}_k^{-1/\alpha} x_k\|^{-\alpha}$. Therefore, points $x_k \in \Phi_k$ are transformed to $y_k \in \mathbb{R}^2$ where $y_k = \mathcal{X}_k^{-1/\alpha} x_k$. The new point process defined by y_k, denoted as $\widehat{\Phi}_k \subset \mathbb{R}^2$, is also homogeneous with intensity $\widehat{\lambda}_k = \lambda_k \mathbb{E}[\mathcal{X}_k^{2/\alpha}]$ [105], [113], [114].

3.3.3 Full Buffers

Our analysis of sleep mode assumes that each active BS has a full buffer i.e. all its subchannels are fully loaded. This assumption is commonly used in many works because it offers analytical simplicity and tractability and is generally realistic in highly dense environments. This means that a typical user on a given subchannel receives interference on that subchannel from all active BSs. However it is rather pessimistic and generally defines a lower bound on network performance. Some works attempt to study more realistic traffic models; for example, authors in [119] consider spatial-time PPP-based user arrivals (bursty space-time traffic) and different user-BS assignment schemes in a random PPP-based homogeneous network. In [105], the aggregate interference is computed by considering some notion of network load as opposed to a fully-loaded network. It would interesting to understand how results improve when a practical traffic profile is considered.

In addition, some works assume that interference is independently thinned by switching off BSs randomly (for example random sleep mode) but this ignores possible correlation between BSs to determine which BSs go to sleep mode. The analysis in this thesis uses the interaction between cell sizes and user distribution to determine the number of users in each cell and this is sometimes used to decide which BSs go to sleep mode. For example, idle BSs are put to sleep in the conventional scheme. Even centralized and distributed strategic sleep mode schemes consider the distribution of users in cells. However a more realistic analysis of the effect of possible BS correlation and real traffic demand is necessary to enhance the results.

3.3.4 Universal Frequency Reuse

Universal frequency reuse, where each BS can access the entire available bandwidth, is considered to investigate the worst case scenario in terms of aggregate interference. For instance, if all active BSs are always transmitting over all channels in the DL, then a typical user receives interference from all BSs except its parent BS (it is assumed that there is no intra-cell interference). However, universal frequency reuse is desirable because it maximizes spectrum utilization in both space and time domains. Many existing works perform analysis using universal frequency reuse such as [71], [94], [105], [106], [108].

In the PPP-based approach to network analysis, it is not straightforward to implement any frequency reuse scheme because of the random nature and irregularity of the cellular layout.

In [106], an idea of frequency reuse is implemented by assuming that each user accesses $1/f$-th of the available system bandwidth where f is the number of available bands. However, since allocation of the bands to cells is random, some neighboring cells end up using the same band which is not the design objective of frequency reuse. Therefore, this defines a lower bound on actual frequency reuse where neighboring cells are always allocated different frequency bands.

3.3.5 No MIMO

The system model used in this thesis only considers a single antenna at both transmit and receive ends of the link. A lot of existing works do not consider MIMO to simplify analysis and enhance tractability of results [71], [94], [105], [106], [108]. Although MIMO techniques enhance network capacity, they also consume more energy [16]. However some of the works that analyze MIMO networks using stochastic geometry techniques include [120], [121], [122].

3.4 System Model

The following system model incorporates all the stated assumptions. Consider a general K-tier HetNet consisting of a tier of macro BSs and $(K-1)$ tiers of small BSs, all independently located on the 2-D Euclidean plane. BS locations in the k-th tier are modeled according to a homogeneous PPP Φ_k of density λ_k. In addition, each k-th tier BS transmits the same power P_k and is assigned a bias value of β_k. The pathloss exponent of the k-th tier is α_k where all pathloss exponents $\{\alpha_j\} > 2, \forall j \in K$. Hence, each tier is uniquely described by the tuple $(\lambda_k, P_k, \beta_k, \alpha_k)$. Users are also distributed independently according to a homogeneous PPP Φ_u of intensity λ_u in the same Euclidean plane. Universal frequency reuse is considered such that a typical user receives interference from every active BS other than its parent BS. Long term shadowing is ignored so that the cellular layout resembles a weighted PV tessellation [115].

A homogeneous network is a special case of the HetNet when $K = 1$. Therefore, there is no biasing and the network is described by the tuple (λ_b, P_t, α) where λ_b and P_t are the macro BS density and transmit power respectively. Without shadowing, each user therefore connects to its nearest BS which makes the layout resemble the PV tessellation shown in Fig. 3.1.

Without loss of generality, assume that a typical user is located at the origin and at a distance r_k from its serving k-th tier BS. The pathloss model considered is $l(r_k) = L\|r_k\|^{-\alpha_k}$ where L is a pathloss constant. The fading loss is assumed to be i.i.d exponential i.e. $h_{r_k} \sim \exp(1)$. Therefore the power received by the typical user from its serving BS, denoted as b_o, is expressed as $P_{r,k} = P_k L h_{r_k} \|r_k\|^{-\alpha_k}$. The resulting SINR is expressed as

$$\mathrm{SINR}(r_k) = \frac{P_k h_{r_k} \|r_k\|^{-\alpha_k}}{\frac{\sigma^2}{L} + I_j} \tag{3.4}$$

where $I_j = \sum_{j=1}^{K} \sum_{r_j \in \Phi_j \setminus r_k} P_j h_{r_j} \|r_j\|^{-\alpha_j}$ is the aggregate interference received by the typical

user and σ^2 is the additive noise power. Noise power is expressed as [123]

$$N_0 = FkT_aB_{ch} \tag{3.5}$$

where F is the receiver noise figure, k is the Boltzmann constant, T_a is the ambient temperature and B_{ch} is the channel bandwidth. In the homogeneous network, (3.4) is simplified as

$$\text{SINR} = \frac{hr^{-\alpha}}{\frac{\sigma^2}{P_tL} + \sum_{i\in\Phi_b\setminus\{b_o\}} h_iR_i^{-\alpha}}. \tag{3.6}$$

Therefore when $K > 1$, the SINR experienced by a typical user depends on the deployment configuration $(\lambda_k, P_k, \beta_k)$ of the tier it is associated to relative to the other $K-1$ tiers. For a user connected to the k-th tier, denote the k-th tier coverage probability and average user rate as $\mathcal{P}_{c,k}$ and $\mathcal{R}_{u,k}$ respectively. These values are used to obtain the overall average HetNet performance using the law of total probability as [94]

$$\mathcal{P}_c = \sum_{k=1}^{K} \mathcal{P}_{c,k}\mathcal{A}_k \quad \text{and} \quad \mathcal{R}_u = \sum_{k=1}^{K} \mathcal{R}_{u,k}\mathcal{A}_k \tag{3.7}$$

where $\mathcal{A}_k$ is the k-th tier association probability. The association probability of a given tier can also be described as the fraction of the network area covered BSs belonging to this tier or the fraction of all users covered by BSs belonging to the tier. Therefore, it depends on HetNet deployment parameters such as the number of tiers and their respective BS densities and transmit powers. Intuitively, the higher the BS density and transmit power of a tier, the better its association probability. However, the tier association probability also depends on the user association scheme implemented in the HetNet. Note that the concept of tier association does not arise in a homogeneous network and the steps in (3.7) are not applicable.

3.5 Homogeneous Networks

The homogeneous network is comparatively easier to analyse than a HetNet because it has only one tier of BSs. This thesis concentrates on analyzing DL performance because the network consumes the bulk of its power during DL transmission. Consider a homogeneous network in which BSs are deployed according to an independent homogeneous PPP Φ_b of intensity λ_b in the Euclidean plane $\mathbb{R}^2$. Ignoring shadowing effects, the SINR received by a typical user located at the origin is shown in (3.6). Coverage probability and average user rate analysis follows [106].

3.5.1 Probability of Coverage

Coverage probability, denoted as $\mathcal{P}_c$, is defined as the probability that a typical user receives SINR greater than a predefined threshold value T. In other words, $\mathcal{P}_c$ is the CCDF of SINR

while the probability of outage, denoted as $\mathcal{P}_o$, is its CDF. Hence,

$$\mathcal{P}_c = \mathbb{P}[\text{SINR} > T] \quad \text{and} \quad \mathcal{P}_o = \mathbb{P}[\text{SINR} \leq T]. \tag{3.8}$$

Coverage probability (or outage) may be interpreted as the fraction of network area that is in coverage (or outage) or the fraction of all users that are in coverage (or outage) [106].

The statistical distance between the typical user and its serving BS is derived from the null probability of a 2D PPP in an area A which is expressed as $e^{-\lambda_b A}$ [106]. In other words,

$$\mathbb{P}[r > R] = \mathbb{P}[\text{No other BS closer than } R] = e^{-\lambda_b \pi R^2}. \tag{3.9}$$

However, the CDF is $F_r(r) = \mathbb{P}[r \leq R] = 1 - e^{-\lambda_b \pi R^2}$. Hence, the PDF is expressed as

$$f_r(r) = \frac{\mathrm{d}\, F_r(r)}{\mathrm{d}\, r} = 2\pi\lambda_b\, e^{-\lambda_b \pi r^2}. \tag{3.10}$$

Theorem 3.1. *The coverage probability of a typical user in a homogeneous PPP-based mobile cellular network is expressed as [106]*

$$\mathcal{P}_c(\lambda_b, T, \alpha) = \pi\lambda_b \int_{x>0} e^{-\frac{T\sigma^2}{P_t L} x^{\alpha/2}}\, e^{-\pi\lambda_b(1+\rho(T,\alpha))x}\, dx \tag{3.11}$$

where $\rho(T,\alpha) = T^{2/\alpha} \int_{T^{-2/\alpha}}^{\infty} \frac{1}{1+u^{\alpha/2}} \mathrm{d}u$. *When* $\alpha = 4$, $\rho(T,4) = \sqrt{T}\,\mathrm{atan}\sqrt{T}$.

Proof. See the proof in [106, Theorem 2]. ∎

Corollary 3.1. *When the network is interference-limited, coverage probability becomes [106]*

$$\overline{\mathcal{P}}_c(T,\alpha) = \frac{1}{1+\rho(T,\alpha)}. \tag{3.12}$$

Proof. In the interference-limited network, the effect of noise is negligible. Therefore, substituting $\sigma^2 = 0$ into (3.11) and solving the resulting integral gives the result. ∎

The coverage probability in (3.12) is essentially in closed form as it requires the evaluation of one very simple integral in $\rho(T,\alpha)$. When $\alpha = 4$, the result is fully in closed form. Coverage probability has an inverse relationship with the target threshold T due to the fact that a higher coverage threshold is more difficult to achieve. Remarkably, coverage probability in (3.12) is independent of the BS density which means that both sparse and dense networks provide the same coverage performance. The intuitive explanation is that when the BS density is increased, the gain in received power is counterbalanced by the additional interference power which maintains the average SINR level. In practical networks, interference management schemes such as frequency reuse are used to enhance coverage and capacity performance. Denser networks are generally more desirable to reduce the risk of congestion and enhance average sum rate.

3.5.2 Average User Rate

For the typical user located at the origin and receiving the SINR shown in (3.6), the average user rate in the DL is expressed as [106]

$$\mathcal{R}_u = \mathbb{E}[\log_2(1 + \text{SINR})] \quad [\text{in bps/Hz}]. \tag{3.13}$$

Since rate $R = \log_2(1 + \text{SINR})$ is a non-negative random variable, its expected value can be computed as $\mathbb{E}[R] = \int_{t>0} \mathbb{P}[R > t]\,\mathrm{d}t$. Hence the average user rate becomes

$$\begin{aligned}\mathcal{R}_u &= \int_{r>0} \mathbb{E}[\log_2(1 + \text{SINR})]\, f_r(r)\,\mathrm{d}r \\ &= \int_{t>0}\int_{r>0} \mathbb{P}[\log_2(1 + \text{SINR}) > t]\, f_r(r)\,\mathrm{d}r\,\mathrm{d}t.\end{aligned} \tag{3.14}$$

Theorem 3.2. *The average rate of a typical user in a PPP-based homogeneous network is expressed as [106]*

$$\mathcal{R}_u(\lambda_b, \alpha) = \Xi\left[\pi\lambda_b e^{-\frac{\sigma^2}{P_t L}(2^t-1)x^{\alpha/2}}\, e^{-\pi\lambda_b(1+\zeta(t,\alpha))x}\right] \tag{3.15}$$

where $\Xi[f(x,t)] = \int_{t>0}\int_{x>0} f(x,t)\mathrm{d}x\mathrm{d}t$ *and* $\zeta(t,\alpha) = (2^t-1)^{2/\alpha}\int_{(2^t-1)^{-2/\alpha}}^{\infty} \frac{1}{1+u^{\alpha/2}}\,\mathrm{d}u$. *When* $\alpha = 4$, $\zeta(t,4) = \sqrt{2^t-1}\,\mathrm{atan}(\sqrt{2^t-1})$.

Proof. See the proof in [106, Theorem 3]. ∎

Corollary 3.2. *In the interference-limited network, average user rate becomes [106]*

$$\overline{\mathcal{R}}_u(\alpha) = \int_{t>0} \frac{1}{1+\zeta(t,\alpha)}\mathrm{d}t. \tag{3.16}$$

Proof. Substitute $\sigma^2 = 0$ into (3.15) and evaluate the integral. ∎

Similar to coverage probability and for the same reason, average user rate in an interference-limited homogeneous network is also independent of the BS density. However, network densification is still preferred since it reduces the risk of congestion and enhances the average bandwidth per user which improves the average network sum rate.

3.6 HetNet Analysis

HetNets typically combine several tiers of different types of BSs which transmit at different power levels and cover cells of highly varying shapes and sizes. The network design of HetNets is a more challenging task than the design of homogeneous networks. The criteria with which a typical user chooses its parent BS has a huge impact on the overall HetNet performance. For example, a user may connect to any BSs from which it receives the strongest signal, or it may connect to the nearest BS from any tier. Since macro BSs transmit the highest power level, they tend to overwhelm small BSs by covering most of the network area. This may leave most small

BSs idle which is undesirable since their main deployment objective is to offload traffic from congesting macro BSs. This load balancing problem can generally be solved by using *biasing*, in which users are 'artificially' biased to favour small BSs over macro BSs by some degree. If a user is biased to connect to a small BS over a more favorable macro BS, the received power from the macro BS then becomes interference. This increases the aggregate interference suffered by this particular user which generally affects coverage and rate performance. Therefore, although biasing has crucial load balancing benefits, its application presents a tradeoff with coverage probability and average rate performance.

The following investigates existing works on the coverage probability and average rate performance of a general K-tier HetNet using different user association schemes. These results will be used in Chapter 6 to define performance constraints during the optimization of the HetNet deployment configuration that minimizes its APC.

3.6.1 Coverage Probability

Coverage probability of a typical user located at the origin and connected to a k-th tier BS is expressed as [94]

$$\begin{aligned}\mathcal{P}_{c,k} &= \mathbb{E}_x\left[\mathbb{P}(\mathrm{SINR}_k(x) > T)\right] \\ &= \int_0^\infty \mathbb{P}(\mathrm{SINR}_k(x) > T)\, f_{X_k}(x)\,\mathrm{d}x \end{aligned} \tag{3.17}$$

where $f_{X_k}(x)$ is the PDF of the distance between the user and serving BS and depends on the user association scheme. Using $\mathrm{SINR}_k(x)$ in (3.4),

$$\mathbb{P}(\mathrm{SINR}_k(x) > T) = \mathbb{P}\left[h_{x_k} > \frac{Tx^{\alpha_k}}{P_k L}\left(I_j + \frac{\sigma^2}{L}\right)\right] \tag{3.18}$$

$$\overset{(a)}{=} e^{-\left(\frac{T\sigma^2}{P_k L}x^{\alpha_k}\right)} \prod_{j=1}^{K} \mathcal{L}_{I_j}(Tx^{\alpha_k}P_k^{-1}) \tag{3.19}$$

where (a) follows since $h_{x_k} \sim \exp(1)$ [106] and $\mathcal{L}_{I_j}(s_c)$ is the Laplace transform of the interference term evaluated at s_c.

3.6.2 Average User Rate

For the typical user, average rate of a user connected to a k-th tier BS is determined as $\mathcal{R}_{u,k} = \mathbb{E}_x[\mathbb{E}_{\mathrm{SINR}_k}[R_k]]$ where $R_k = \log_2(1 + \mathrm{SINR}_k(x))$. Since R_k is a non-negative random variable, its expectation is determined as $\mathbb{E}_{\mathrm{SINR}_k}[R_k] = \int_{t>0} \mathbb{P}(R_k > t)\,\mathrm{d}t$. Hence,

$$\begin{aligned}\mathcal{R}_{u,k} &= \int_{x>0}\int_{t>0} \mathbb{P}(\log_2(1 + \mathrm{SINR}_k(x)) > t)\, f_{X_k}(x)\,\mathrm{d}t\mathrm{d}x \\ &= \Xi\left[\mathbb{P}(\mathrm{SINR}_k(x) > 2^t - 1)\, f_{X_k}(x)\right] \end{aligned} \tag{3.20}$$

where $\Xi\left[f(x,t)\right] = \int_{t>0}\int_{x>0} f(x,t)\,\mathrm{d}x\,\mathrm{d}t$. The term $\mathbb{P}(\mathrm{SINR}_k(x) > 2^t - 1)$ in (3.20) is of the same form as (3.18)-(3.19) and is therefore solved similarly. Hence,

$$\mathbb{P}(\mathrm{SINR}_k(x) > 2^t - 1) = e^{-\left(\frac{\sigma^2}{P_k L}(2^t-1)x^{\alpha_k}\right)} \prod_{j=1}^{K} \mathcal{L}_{I_j}\left(x^{\alpha_k} P_k^{-1}(2^t-1)\right) \tag{3.21}$$

where $\mathcal{L}_{I_j}(s_r)$ is the Laplace transform of the interference evaluated at s_r.

3.7 Flexible Cell Association

In an open access HetNet, users are allowed to connect to any BS in any tier provided some defined cell association criterion is satisfied. Cell association is an important functionality that impacts the coverage, rate and energy performance of the HetNet. Artificial biasing can be used to manipulate cell association and achieve load balancing objectives and other performance targets. In this thesis, the following three cell association schemes are studied and compared with each other: (i) Maximum ABRP connectivity [94]; (ii) Minimum BTD connectivity [71]; and (iii) Maximum i-SINR connectivity [108], [116], [124].

In maximum ABRP connectivity, a user connects to the BS that provides the strongest average biased received power. Hence a user connects to the nearest k-th tier BS if $P_{r,k} > P_{r,j}$, $\forall j \in K, j \neq k$ where $P_{r,k} = P_k \beta_k r_k^{-\alpha}$ and β_k is the bias value associated with k-th tier BSs. In maximum i-SINR connectivity, a user connects to the BS that provides it with the highest instantaneous SINR [108], [116], [124].

In minimum BTD connectivity, it is assumed that a user knows its relative distances to each of its neighboring BSs in all tiers. If the nearest k-th tier BS is located at a distance r_k from the user, then r_k is multiplied by its respective bias factor,denoted as ν_k. Hence the user associates to the k-th tier if $\nu_k r_k < \nu_j r_j$, $\forall j \in K, j \neq k$. Although this scheme is also discussed in [71], it is analyzed in this thesis (see section 6.2) using an approach that gives results which are directly comparable to those of maximum ABRP and i-SINR schemes.

3.7.1 Maximum ABRP Connectivity

The following analysis describes the tier association probability and the coverage probability and average rate of a typical user in a HetNet using maximum ABRP connectivity scheme.

Lemma 3.1. *The association probability of the k-th tier is expressed as [94, Lemma 1]*

$$\mathcal{A}_k = 2\pi\lambda_k \int_0^\infty r\, e^{-\pi \sum_{j=1}^{K} \lambda_j (\widehat{P}_j \widehat{\beta}_j)^{2/\alpha_j} r^{2/\widehat{\alpha}_j}}\, dr \tag{3.22}$$

where $\widehat{P}_j = \frac{P_j}{P_k}$, $\widehat{\beta}_j = \frac{\beta_j}{\beta_k}$, *and* $\widehat{\alpha}_j = \frac{\alpha_j}{\alpha_k}$. *If all tiers have the same pathloss exponent i.e.*

$\{\alpha_j\} = \alpha$, *(3.22) is simplified as*

$$\mathcal{A}_k = \frac{\lambda_k (P_k \beta_k)^{2/\alpha}}{\sum_{j=1}^{K} \lambda_j (P_j \beta_j)^{2/\alpha}}. \tag{3.23}$$

Proof. See the proof of [94, Lemma 1]. ∎

The result in (3.23) shows that a user tends to connect to a tier with a relatively larger BS density, transmit power and bias factor, which agrees with intuition.

Lemma 3.2. *The PDF $f_X(x)$ of the distance X_k between a typical user and its serving BS is expressed as [94, Lemma 3]*

$$f_{X_k}(x) = \frac{2\pi\lambda_k}{\mathcal{A}_k} x\, e^{-\pi \sum_{j=1}^{K} \lambda_j (\widehat{P}_j \widehat{\beta}_j)^{2/\alpha_j} x^{2/\widehat{\alpha}_j}}. \tag{3.24}$$

Proof. See the proof of [94, Lemma 1]. ∎

Theorem 3.3. *Coverage probability of a typical user in the HetNet is expressed as [94]*

$$\mathcal{P}_{c_P} = \sum_{k=1}^{K} \pi\lambda_k \int_{z>0} e^{-\frac{T\sigma^2}{P_k L} z^{\alpha_k/2}} e^{-\pi \sum_{j=1}^{K} \lambda_j \widehat{P}_j^{2/\alpha_j} \mathcal{C}_j z^{1/\widehat{\alpha}_j}} dz \tag{3.25}$$

where $\mathcal{C}_j = \widehat{\beta}_j^{2/\alpha_j} + \mathcal{Z}(T, \alpha_j, \widehat{\beta}_j)$ *and* $\mathcal{Z}(T, \alpha_j, \widehat{\beta}_j) = T^{2/\alpha_j} \int_{u_j}^{\infty} \frac{1}{1+u^{\alpha_j/2}} du$ *where* $u_j = \left(\frac{\widehat{\beta}_j}{T}\right)^{2/\alpha_j}$.

Proof. See the proof of [94, Theorem 1]. ∎

Corollary 3.3. *If the HetNet is interference-limited and $\{\alpha_j\} = \alpha$, coverage probability becomes*

$$\overline{\mathcal{P}}_{c_P} = \sum_{k=1}^{K} \frac{\lambda_k P_k^{2/\alpha}}{\sum_{j=1}^{K} \lambda_j P_j^{2/\alpha} \mathcal{C}_j}. \tag{3.26}$$

If the HetNet is also unbiased i.e. $\{\beta_j\} = 1$, coverage probability becomes

$$\overline{\mathcal{P}}_{c_P} = [1 + \mathcal{Z}(T, \alpha, 1)]^{-1} = \mathcal{C}^{-1}. \tag{3.27}$$

In the special case of $\alpha = 4$, $\mathcal{Z}(T, 4, 1) = \sqrt{(T)} \cdot \text{atan}\sqrt{T}$. Hence $\overline{\mathcal{P}}_{c_P}$ is in closed form.

Proof. When $\{\alpha_j\} = \alpha$, $\widehat{\alpha}_j = 1$. Substitute $\sigma^2 = 0$ in (3.25) and solve the integral to get the result in (3.26). Since $\mathcal{C}_j$ depends on the variable set $\{k, j\}$, further simplification is only possible when the HetNet is unbiased i.e. $\{\mathcal{C}_j\} = \mathcal{C} = 1 + \mathcal{Z}(T, \alpha, 1)$ for all $j \in K$. ∎

According to (3.27), $\overline{\mathcal{P}}_{c_P}$ of the unbiased HetNet is independent of the tuple $(K, \{\lambda_j\}, \{P_j\})$. This is an interesting result which means that operators can densify the HetNet with any number and types of BSs without affecting coverage probability. The intuitive explanation is that although densification improves the average received signal strength, aggregate interference also increases in equal measure [94], [106]. However if the HetNet is biased, parameters K, $\{\lambda_j\}$ and $\{P_j\}$ influence coverage probability.

Theorem 3.4. *Average user rate in the HetNet is expressed as [94, Theorem 2]*

$$\mathcal{R}_{u_P} = \sum_{k=1}^{K} \Xi \left[\pi \lambda_k e^{-\frac{\sigma^2}{P_k L}(2^t-1) z^{\alpha_k/2}} e^{-\pi \sum_{j=1}^{K} \lambda_j \widehat{P}_j^{2/\alpha_j} \mathcal{D}_j(t)\, z^{1/\widehat{\alpha}_j}} \right] \tag{3.28}$$

where $\mathcal{D}_j(t) = \widehat{\beta}_j^{2/\alpha} + \mathcal{Z}(t, \alpha_k, \widehat{\beta}_j)$, $\mathcal{Z}(t, \alpha_k, \widehat{\beta}_j) = (2^t - 1)^{2/\alpha_j} \int_{u_j}^{\infty} \frac{1}{1+u^{\alpha_j/2}} du$ *and* $u_j = \left(\frac{\widehat{\beta}_j}{2^t-1}\right)^{2/\alpha_j}$.

Proof. See the proof of [94, Theorem 2]. ∎

Corollary 3.4. *If the HetNet is interference-limited and* $\{\alpha_j\} = \alpha$, *average user rate becomes*

$$\overline{\mathcal{R}}_{u_P} = \sum_{k=1}^{K} \int_{t>0} \frac{\lambda_k P_k^{2/\alpha}}{\sum_{j=1}^{K} \lambda_j P_j^{2/\alpha} \mathcal{D}_j(t)} dt. \tag{3.29}$$

Furthermore, if the HetNet is unbiased i.e. $\{\widehat{\beta}_j\} = 1$, *then average user rate is simplified as*

$$\overline{\mathcal{R}}_{u_P} = \int_{t>0} \frac{1}{\mathcal{D}(t)} dt \tag{3.30}$$

where $\mathcal{D}(t) = 1 + \mathcal{Z}(t, \alpha, 1)$ *is a constant. For* $\alpha = 4$, $\mathcal{Z}(t, 4, 1) = \sqrt{2^t - 1}\text{atan}\sqrt{2^t - 1}$.

Proof. Since $\{\widehat{\alpha}_j\} = 1$ in (3.28), let $\sigma^2 = 0$ and solve the resulting integral. If the HetNet is biased, $\mathcal{D}_j(t)$ varies with the set $\{k, j\}$ and further simplification is not possible. In the unbiased HetNet, $\mathcal{D}(t)$ is a constant which allows the simplification shown in (3.30). ∎

Therefore, similar to $\overline{\mathcal{P}}_{c_P}$ in (3.27) and for the same reason, $\overline{\mathcal{R}}_{u_P}$ of the unbiased HetNet is independent of the tuple $(K, \{\lambda_j\}, \{P_j\})$.

3.7.2 Maximum i-SINR Connectivity Scheme

In this scheme, a typical user associates with a BS that provides the highest instantaneous SINR i.e. a user associates with the k-th tier if $\text{SINR}_k(r_k) > \max_{j, j \neq k} \text{SINR}_j(r_j)$ [108].

Theorem 3.5. *Coverage probability in a HetNet is expressed as [108, Theorem 1]*

$$\mathcal{P}_{c_S} = \sum_{k=1}^{K} \pi \lambda_k \int_{z>0} e^{-\frac{T\sigma^2}{P_k L} z^{\alpha/2}} e^{-T^{2/\alpha} \varrho(\alpha) \sum_{j=1}^{K} \lambda_j \widehat{P}_j^{2/\alpha} z}\, dz \tag{3.31}$$

where $\varrho(\alpha) = 2\pi^2 \csc(2\pi/\alpha)\alpha^{-1}$ *is a constant for a given pathloss exponent* α.

Proof. See the proof of [108, Theorem 1]. ∎

Corollary 3.5. *If the HetNet is interference-limited, coverage probability becomes [108]*

$$\overline{\mathcal{P}}_{c_S} = \frac{\pi T^{-2/\alpha}}{\varrho(\alpha)}. \tag{3.32}$$

In the special case of $\alpha = 4$, $\overline{\mathcal{P}}_{c_S} = \frac{2}{\pi\sqrt{T}}$. *These expressions are in closed form.*

Proof. Let $\sigma^2 = 0$ in (3.31) and evaluate the integral. ∎

According to (3.32), $\overline{\mathcal{P}}_{c_S}$ is also independent of the parameter set $\{K, \{\lambda_j\}, \{P_j\}\}$, similar to $\overline{\mathcal{P}}_{c_P}$ of maximum ABRP scheme when the HetNet is unbiased.

Consider a HetNet in which the SINR coverage thresholds are not necessarily similar per tier. Denote the SINR threshold of the k-th tier as T_k. Then coverage probability of the general HetNet depends on the set $\{T_j\}$, $\forall j \in K$. In this case, (3.31) is expressed as [108], [116], [124]

$$\mathcal{P}_{c_S} = \sum_{k=1}^{K} \pi\lambda_k \int_{z>0} e^{-\frac{T_k\sigma^2}{P_k L} z^{\alpha/2}} e^{-\varrho(\alpha)\sum_{j=1}^{K}\lambda_j \widehat{P}_j^{2/\alpha} T_j^{2/\alpha} z}\, \mathrm{d}z. \tag{3.33}$$

When $\sigma^2 = 0$, (3.33) is then expressed as

$$\overline{\mathcal{P}}_{c_S} = \frac{\pi \sum_{j=1}^{K} \lambda_j P_j^{2/\alpha}}{\varrho(\alpha) \sum_{j=1}^{K} \lambda_j P_j^{2/\alpha} T_j^{2/\alpha}}. \tag{3.34}$$

Consider the special case of a homogeneous network such that $K = 1$. Assuming that $\{T_j\} = T$ (3.34) simplifies to (3.32). This shows that under interference-limited conditions, coverage probability of the homogeneous network (shown in (3.12)) is equivalent to that of a multi-tier HetNet with the same SIR thresholds per tier. Fig 3.3 confirms this relationship and shows that coverage probability increases with the pathloss exponent α due to the higher attenuation of interference at higher α. The derivation of coverage probability in (3.32) is based on the assumption that $\{T_j\} > 0\,\mathrm{dB}$, $\forall j \in K$ [108]. This explains the deviation from the accurate result of (3.12) when $\{T_j\} < 0\,\mathrm{dB}$. Generally coverage probability in a homogeneous network is the same as in a multi-tier unbiased HetNet for all connnectivity schemes [106], [108], [115].

Theorem 3.6. *Average user rate in the HetNet is expressed as [108, Theorem 2]*

$$\mathcal{R}_{u_S} = \sum_{k=1}^{K} \Xi\left[\pi\lambda_k e^{-\frac{\sigma^2}{P_k L}(2^t-1)z^{\alpha/2}} e^{-(2^t-1)^{2/\alpha}\varrho(\alpha)\sum_{j=1}^{K}\lambda_j \widehat{P}_j^{2/\alpha} z}\right]. \tag{3.35}$$

Proof. See the proof of [108, Theorem 2]. ∎

Corollary 3.6. *If the HetNet is interference-limited, average user rate becomes*

$$\overline{\mathcal{R}}_{u_S} = \frac{\pi}{\varrho(\alpha)} \int_{t>0} (2^t - 1)^{-2/\alpha}\, dt. \tag{3.36}$$

Proof. Substitute $\sigma^2 = 0$ in (3.35) and evaluate the resulting integral. ∎

The result in (3.36) is simple and only requires the evaluation of one integral. As with coverage probability in (3.32), average user rate in an interference-limited HetNet is independent of the parameter set $\{K, \{\lambda_j\}, \{P_j\}\}$ and is constant for a given α.

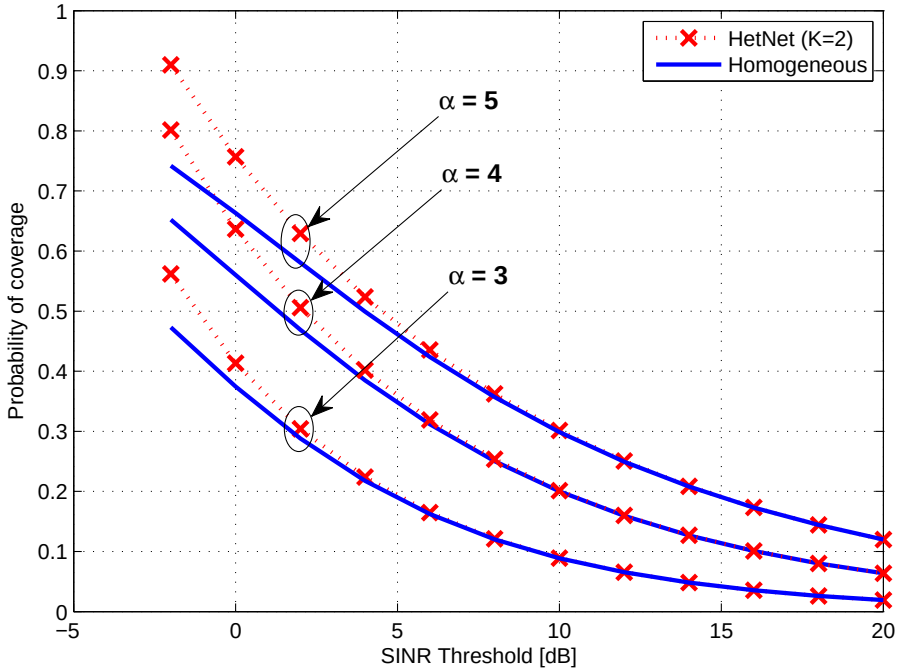


Figure 3.3: Coverage probability of a homogeneous network and a HetNet using maximum i-SINR connectivity, both under interference-limited conditions.

3.8 Summary

This chapter introduced a new analysis approach that uses tools from stochastic geometry to derive simple and tractable network performance results. Important mathematical preliminaries of the PPP model were also introduced. All the main assumptions followed in the analysis of the homogeneous network and HetNet throughout this thesis were discussed. In addition, their justification was presented and various existing works that used similar assumptions were referenced. This chapter also presented all relevant and existing analysis (especially coverage probability and average rate) of homogeneous networks and HetNets.

Chapter 4

Homogeneous Network Deployment

4.1 Introduction

Due to the highly randomized cell sizes in a PV tessellation, the number of users per cell is also highly random. In order to enhance the tractability of network analysis, most existing works make some assumptions: (i) each cell connects at most only one user that is chosen at random from users within its coverage area [106], [71], and (ii) the user density is large enough to guarantee at least one user within each cell [106]. However, both assumptions are unrealistic since BSs have the ability to connect multiple users and the density of users per BS depends on the prevailing BS and user densities which are finite.

In practical networks, the density of active users highly varies in both space and time domains. For example, active users are few at night time and peak some time during the day or early evening. Spatially, user density in cities or business areas peaks during the day and drops off at night as people return to residential areas. In addition, urban areas generally have a higher concentration of users than rural areas. Therefore, the prevailing BS-user density ratio of a practical network changes spatiotemporally throughout the day and this significantly influences the average number of users per BS in the network [49].

After cell association, any idle BSs do not transmit which thins out the aggregate interference experienced by the typical user. Denoting the set of idle BSs as $\{b_{id}\}$, the received SINR at the origin is expressed as

$$\text{SINR} = \frac{hr^{-\alpha}}{\frac{\sigma^2}{P_t L} + \sum_{i \in \Phi_b \setminus \{b_o \cup b_{id}\}} g_i R_i^{-\alpha}} \tag{4.1}$$

4.2 Cell Size and User Distribution

Consider a homogeneous network in which both BSs and users are spatially located according to independent homogeneous PPPs Φ_b and Φ_u and their respective intensities λ_b and λ_u. Each UE is served by the closest BS and the network layout resembles a PV tessellation. The distribution

of cell sizes in the $\mathbb{R}^2$ plane is approximated using the gamma distribution as [107], [125], [115]

$$f_X(A) = \frac{(\mathcal{K}\lambda_b)^{\mathcal{K}}}{\Gamma(\mathcal{K})} A^{\mathcal{K}-1} e^{-\mathcal{K}\lambda_b A} \tag{4.2}$$

where $\mathcal{K} = 3.575$. Since UEs are also located according to a homogeneous PPP, the number of UEs per cell is proportional to its area A and follows a Poisson distribution. Therefore, the distribution of the number of users in typical cell of size A has the following probability density function (see (3.2)):

$$g_A(n) = \frac{(\lambda_u A)^n}{n!} e^{-\lambda_u A}. \tag{4.3}$$

These distributions can be used to find the probability that a typical cell contains N users. In general, the probability that a typical cell of area A contains $N = n$ users is given by

$$\begin{aligned}
\mathbb{P}(N = n) &= \int_0^\infty \mathbb{P}[N = n | X = A]\, f_X(A) \mathrm{d}A \\
&= \frac{\lambda_u^n (\mathcal{K}\lambda_b)^{\mathcal{K}}}{\Gamma(\mathcal{K})\, n!} \int_0^\infty A^{n+\mathcal{K}-1} e^{-(\lambda_u + \mathcal{K}\lambda_b)} \mathrm{d}A \\
&\overset{(a)}{=} \frac{{\lambda_u}^n (\mathcal{K}\lambda_b)^{\mathcal{K}} \Gamma(n+\mathcal{K})}{\Gamma(\mathcal{K}) n! (\lambda_u + \mathcal{K}\lambda_b)^{n+\mathcal{K}}}
\end{aligned} \tag{4.4}$$

where the integral (a) is solved using the identity [126, (3.381.4)]. Therefore, the general probability that a cell has n or less users (i.e. $N \leq n$) is given by

$$\mathbb{P}(N \leq n) = \sum_{m=0}^{n} \frac{{\lambda_u}^m (\mathcal{K}\lambda_b)^K \Gamma(m+\mathcal{K})}{\Gamma(\mathcal{K}) m! (\lambda_u + \mathcal{K}\lambda_b)^{m+\mathcal{K}}}. \tag{4.5}$$

In special cases, the probabilities that a cell has no users or has one user are respectively expressed as

$$\mathbb{P}(N = 0) = \left(1 + \frac{\lambda_u/\lambda_b}{\mathcal{K}}\right)^{-\mathcal{K}} \quad \text{and} \quad \mathbb{P}(N = 1) = \frac{\lambda_u}{\lambda_b} \left(1 + \frac{\lambda_u/\lambda_b}{\mathcal{K}}\right)^{-(\mathcal{K}+1)}. \tag{4.6}$$

Therefore, the probability of an active BS (a BS with at least one user), denoted as p_a, becomes

$$p_a = 1 - \mathbb{P}(N = 0) = 1 - \left(1 + \frac{\lambda_u/\lambda_b}{\mathcal{K}}\right)^{-\mathcal{K}} \tag{4.7}$$

These probabilities clearly depend on the prevailing BS-user density ratio, denoted as $\upsilon = \lambda_b/\lambda_u$. When $\lambda_u \gg \lambda_b$, then $\mathbb{P}(N = 0) \approx 0$ and $p_a \approx 1$. This is the basis of the assumption in [106] which states that the user density is large enough so that all BSs have at least one user.

4.2.1 Single User Connectivity Model

To ease analysis, some works assume that each cell connects only one user, chosen randomly from its constituent users [106]. Denote this as the *single user connectivity model*. If the user density is large enough to ensure that each cell has at least one user, the average number of connected users, denoted as N_c, is equivalent to the average number of BSs. However, due to the

spatiotemporal variability of user density in real networks, there is always a finite probability of a BS being idle as shown in (4.6). Such idle BSs do not transmit and may be put into sleep mode to save energy. In reality therefore, the average number of active BSs is equal to the average number of connected users i.e. $N_c = p_a\lambda_b = \text{ACR} \times \lambda_u$ where ACR is a new measure defined as the average connectivity ratio. ACR represents the ratio of the average number of connected users to the average number of all users in the network. Hence,

$$\text{ACR} = \frac{p_a\lambda_b}{\lambda_u} = \frac{\lambda_b}{\lambda_u}\left[1 - \left(1 + \frac{\lambda_u}{\mathcal{K}\lambda_b}\right)^{-\mathcal{K}}\right] \tag{4.8}$$

which is a function of the BS-user density ratio.

Since each BS connects a single user, any extra users in its coverage area will not get a connection and are therefore in outage due to blocking (or congestion). However, the operator may want to guarantee a certain probability of connection to keep the congestion below a predefined threshold denoted as χ. Given the user density, the operator can determine the average number of BSs required to achieve the blocking constraint. The average number of unconnected (or blocked) users may be expressed as $\text{ABR} \times \lambda_u$ where ABR is a new measure called the average blocking ratio and is simply $\text{ABR} = 1 - \text{ACR}$. Therefore,

$$\text{ABR} = \frac{\lambda_u - p_a\lambda_b}{\lambda_u} \leq \chi \quad \text{s.t.} \quad p_a\lambda_b \geq \lambda_u(1-\chi). \tag{4.9}$$

In practice, χ is always set very low to ensure a good QoS. Considering that there are always some idle BSs, the BS density that achieves such a low blocking rate is very high under the single user connectivity model. It is not possible to determine a closed form expression for the required BS density λ_b in (4.9) but it can easily be determined using numerical methods such as the bisection method [127]. For ease of analysis, consider the first three terms of the binomial series of p_a, expressed as

$$p_a = 1 - \left(1 + \frac{\lambda_u}{\mathcal{K}\lambda_b}\right)^{-\mathcal{K}} \approx \frac{\lambda_u}{\lambda_b}\left[1 - \frac{\lambda_u}{\lambda_b}\left(\frac{\mathcal{K}+1}{2\mathcal{K}}\right)^2\right]. \tag{4.10}$$

Therefore from (4.9), the required BS density is determined using the simple expression

$$\lambda_b > \frac{(\mathcal{K}+1)}{2\mathcal{K}\chi}\lambda_u. \tag{4.11}$$

Hence the required BS density is proportional to the user density and it reduces when the blocking constraint is relaxed.

To illustrate, consider a network with $\chi = 0.1$, $\lambda_u = 10^{-4}\,\text{m}^{-2}$, $A = 2.5\,\text{km} \times 2.5\,\text{km}$ and $\lambda_b \geq 6.4 \times 10^{-4}\,\text{m}^{-2}$. Hence, the average number of users and BSs are $\lambda_u|A| = 625$ and $\lambda_b|A| \geq 4000$ respectively. The approximation gives a 9% error margin but allows a closed form expression. The average number of BSs is much greater than the average number of users which is very unrealistic in practical terms. Therefore, a multi-user connectivity model is necessary to facilitate a realistic study of the network EE under blocking constraints.

4.2.2 Multiple User Connectivity Model

Assume that the bandwidth is divided into a number of channels δ (where $\delta \geq 1$) such that each cell can randomly choose and connect up to a maximum of $N = \delta$ users while avoiding intra-cell interference. The average number of connected users can be expressed in terms of the BS density and the average number of busy channels in the network, i.e.

$$\text{ACR} \cdot \lambda_u = \lambda_b \times \mathbb{E}[\text{number of busy channels per BS}] = \lambda_b \times \mathbb{E}[C_b] \tag{4.12}$$

where $C_b \in \{0, \delta\}$ is the number of busy channels in a BS under the assumption that all channels are busy in an active BS. Since C_b is a random variable that takes on only non-negative integer values, its expectation is expressed as

$$\mathbb{E}[C_b] = \sum_{n=1}^{\delta} n\,\mathbb{P}(C_b = n) \quad \text{where} \quad \mathbb{P}(C_b = n) = \begin{cases} \mathbb{P}(N = n), & n < \delta \\ \mathbb{P}(N \geq \delta), & n = \delta. \end{cases} \tag{4.13}$$

Note that when $\delta = 1$, $\mathbb{E}[C_b] = \mathbb{P}(N \geq 1) = p_a$ and $\text{ACR} = p_a\lambda_b/\lambda_u$ as shown in (4.8). When $\delta = 2$, each cell can connect up to 2 users and $\mathbb{E}[C_b]$ is expressed as

$$\mathbb{E}[C_b] = \mathbb{P}(C_b = 1) + 2\,\mathbb{P}(C_b = 2) = \mathbb{P}(N = 1) + 2\,\mathbb{P}(N \geq 2).$$

Similarly in the case where each BS can connect up to three users ($\delta = 3$),

$$\begin{aligned} \mathbb{E}[C_b] &= \mathbb{P}(C_b = 1) + 2\,\mathbb{P}(C_b = 2) + 3\,\mathbb{P}(C_b = 3) \\ &= \mathbb{P}(N = 1) + 2\,\mathbb{P}(N = 2) + 3\,\mathbb{P}(N \geq 3). \end{aligned}$$

Based on these three cases, the general expression for $\mathbb{E}[C_b]$ can be formulated as

$$\mathbb{E}[C_b] = \begin{cases} \mathbb{P}(N \geq \delta), & \delta = 1 \\ \delta\,\mathbb{P}(N \geq \delta) + \sum_{k=1}^{\delta-1} k\,\mathbb{P}(N = k), & \delta > 1. \end{cases} \tag{4.14}$$

Therefore ACR is obtained from (4.12) as $\text{ACR} = \frac{\lambda_b}{\lambda_u}\mathbb{E}[C_b]$.

To illustrate the importance of this model, consider the simple case of $\delta = 2$. The average number of blocked users is $\lambda_u(1 - \text{ACR})$. Therefore, $\text{ABR} = 1 - \text{ACR} \leq \chi$. From (4.12),

$$\begin{aligned} \text{ABR} &= 1 - \frac{\lambda_b}{\lambda_u}\left[\mathbb{P}(N = 1) + 2\,\mathbb{P}(N \geq 2)\right] \\ &= 1 - \frac{\lambda_b}{\lambda_u}\left[\mathbb{P}(N = 1) + 2(1 - \mathbb{P}(N = 0) - \mathbb{P}(N = 1))\right] \\ &= 1 - \frac{\lambda_b}{\lambda_u}\left[2\,p_a - \mathbb{P}(N = 1)\right]. \end{aligned} \tag{4.15}$$

where $\mathbb{P}(N = 1)$ is shown in (4.6). Using the same parameters χ, λ_u and A, the required BS density then reduces to $\lambda_b \geq 1.35 \times 10^{-4}\,\text{m}^{-2}$ and $\lambda_b|A| \geq 844$ BSs. This is in contrast to the

over 4000 BSs required in the single user connectivity model. Therefore, applying the multi-user connectivity model enables a more realistic analysis of the network's energy performance. The multi-user connectivity model allows flexible allocation of resources to users depending on their capacity or QoS requirements.

The average capacity experienced by a typical user in this multi-channel network depends on how many subchannels are available to the user. Therefore, it is important to first determine the average rate supported on a typical subchannel. The average sum rate is then determined from the average rate per user and the ACR.

4.3 Coverage and Rate Analysis

Existing analysis of coverage probability and average rate discussed in section 3.5 [106] assumes that all BSs are active and implements the single user connectivity model. This is unrealistic because some BSs are inevitably idle as the user density varies in both space and time. Moreover, BSs connect multiple users simultaneously. This section presents coverage and rate analysis of a multi-channel network taking into account the effect of prevailing user density on its performance.

4.3.1 Probability of Coverage

Theorem 4.1. *Coverage probability of a typical user in a homogeneous PPP-based network with idle BSs is expressed as*

$$\mathcal{P}_c(\upsilon, T, \alpha) = \pi\lambda_b \int_{x>0} e^{-\frac{T\sigma^2}{P_t L}x^{\alpha/2}} e^{-\pi\lambda_b(1+\overline{p}_a\rho(T,\alpha))x} dx. \tag{4.16}$$

When the network is interference-limited, (4.16) simplifies to

$$\overline{\mathcal{P}}_c(\upsilon, T, \alpha) = \frac{1}{1+\overline{p}_a\rho(T,\alpha)}. \tag{4.17}$$

Proof. The proof is essentially similar to that of [106, Theorem 1] except that the aggregate interference is thinned by the presence of idle BSs. In this case, the Laplace transform of the interference is expressed as

$$\mathcal{L}_{I_r}(Tr^\alpha) = \exp(-\pi\overline{p}_a\lambda_b r^2\rho(T,\alpha)).$$

When the network is interference-limited, $\sigma^2 = 0$ and the result follows easily. ∎

The results in (4.16)-(4.17) facilitate an investigation of the effect of idle BSs on coverage probability. For instance, as user density increases and more BSs remain active (i.e. $\overline{p}_a$ increases), interference also increases which reduces coverage probability. Therefore, coverage probability generally increases with BS density and reduces with user density. The results in (3.11)-(3.12)

assume that $\lambda_u \gg \lambda_b$ (i.e. $\upsilon \to 0$) such that $p_a \approx 1$, which makes them special cases of (4.16)-(4.17) respectively.

4.3.2 SE and Sum Rate

In a multi-channel network, analysis is performed on a subchannel basis to determine the spectral efficiency of each subchannel. Consider a multi-channel homogeneous network in which the available bandwidth, denoted as $\mathcal{B}$ Hz, is divided into δ subchannels, each of an equal size $\mathcal{B}_\delta$ Hz. The average ergodic rate experienced by a typical user in a typical subchannel, denoted as $\mathcal{R}_{ch}$, is computed as

$$\mathcal{R}_{ch} = \frac{\mathcal{B}}{\delta}\mathbb{E}[\log_2(1+\text{SINR})] \quad \text{[in bps]}. \tag{4.18}$$

Therefore the average rate on a typical subchannel (in bps/Hz) becomes

$$\mathcal{R}_{ch} = \frac{1}{\delta}\int_{r>0} \mathbb{E}[\log_2(1+\text{SINR})]\, f_r(r)\, \mathrm{d}r. \tag{4.19}$$

Theorem 4.2. *The average subchannel rate in a PPP-based homogeneous network with idle BSs is expressed as*

$$\mathcal{R}_{ch}(\upsilon,\alpha) = \frac{1}{\delta}\Xi\left[\pi\lambda_b e^{-\frac{\sigma^2}{P_t L}(2^t-1)x^{\alpha/2}}\, e^{-\pi\lambda_b(1+\overline{p}_a\zeta(t,\alpha))x}\right] \tag{4.20}$$

When the network is interference-limited, average subchannel rate simplifies to

$$\overline{\mathcal{R}}_{ch}(\upsilon,\alpha) = \frac{1}{\delta}\int_{t>0} \frac{1}{1+\overline{p}_a\zeta(t,\alpha)}\mathrm{d}t. \tag{4.21}$$

Proof. The proof is essentially similar to that of [106, Theorem 2] except that the aggregate interference is thinned by the presence of idle BSs. In this case, the Laplace transform of the interference is expressed as

$$\mathcal{L}_{I_r}((2^t-1)r^\alpha) = \exp(-\pi\overline{p}_a\lambda_b r^2\zeta(t,\alpha)).$$

When the network is interference-limited, $\sigma^2 = 0$ and the result follows easily. ∎

Therefore, average rate per subchannel mainly depends on the BS-user density ratio which is consistent with coverage probability. For example, as the user density increases and more BSs remain active, interference increases which reduces average subchannel rate.

Under the full buffer assumption, each active BS sequentially allocates all subchannels to its connected users. This assumption gives the upper bound on bandwidth utilization. Typically, for a BS with N users and δ subchannels, two subchannel allocation scenarios arise:

- If $N < \delta$, each user is initially allocated δ_u subchannels where $\delta_u = \lfloor \delta/N \rfloor$. The remaining subchannels, denoted as $\delta_r = \delta - \delta_u N$, are allocated to any δ_r users chosen at random from the N users. For example, if $N = 4$ and $\delta = 10$, two of the users get 2 subchannels and the other two users get 3 subchannels.

- If $N \geq \delta$, the BS randomly selects δ users and allocates a single subchannel to each (i.e. $\delta_u = 1$). The remaining users, equivalent to $= N - \delta$, are blocked.

Lemma 4.1. *The average number of subchannels per connected user is expressed as*

$$\omega(\upsilon, \delta) = \frac{p_a \delta}{\mathbb{E}[C_b]}. \tag{4.22}$$

In the single user connectivity model where $\delta = 1$, $\mathbb{E}[C_b] = p_a$ such that $\omega = 1$.

Proof. Let $\Omega = p_a \lambda_b \delta$ denote the total number of subchannels in the network. The average number of connected users $N_c = \text{ACR} \times \lambda_u = \lambda_b \mathbb{E}[C_b]$. Hence $\omega = \Omega / N_c$ gives the result. ∎

Corollary 4.1. *Average user rate in a PPP-based multi-channel network is expressed as*

$$\begin{aligned} \mathcal{R}_u(\upsilon, \delta, \alpha) &= \omega(\upsilon, \delta) \times \mathcal{R}_{ch}(\upsilon, \alpha) \\ &= \frac{p_a}{\mathbb{E}[C_b]} \Xi \left[\pi \lambda_b e^{-\frac{\sigma^2}{P_t L}(2^t - 1) x^{\alpha/2}} \, e^{-\pi \lambda_b (1 + \overline{p}_a \zeta(t, \alpha)) x} \right]. \end{aligned} \tag{4.23}$$

In the interference-limited network ($\sigma^2 = 0$), average user rate simplifies to

$$\overline{\mathcal{R}}_u(\upsilon, \delta, \alpha) = \frac{p_a}{\mathbb{E}[C_b]} \int_{t>0} \frac{1}{1 + \overline{p}_a \zeta(t, \alpha)} \mathrm{d}t. \tag{4.24}$$

Corollary 4.2. *Average sum rate in a PPP-based multi-channel network is expressed as*

$$\begin{aligned} \mathcal{T}(\upsilon, \alpha) &= \text{ACR} \times \lambda_u \times \mathcal{R}_u = \lambda_b \mathbb{E}[C_b] \times \mathcal{R}_u \\ &= \pi p_a \lambda_b^2 \Xi \left[e^{-\frac{\sigma^2}{P_t L}(2^t - 1) x^{\alpha/2}} \, e^{-\pi \lambda_b (1 + \overline{p}_a \zeta(t, \alpha)) x} \right]. \end{aligned} \tag{4.25}$$

In the interference-limited network, the average sum rate simplifies to

$$\overline{\mathcal{T}}(\upsilon, \alpha) = p_a \lambda_b \int_{t>0} \frac{1}{1 + \overline{p}_a \zeta(t, \alpha)} \mathrm{d}t. \tag{4.26}$$

If the BS density is increased, the average number of users per BS reduces which enhances both ACR and $\omega(\upsilon, \delta)$. Hence average sum rate is an increasing function of BS density.

The number of subchannels allocated to a typical user depend on the system bandwidth $\mathcal{B}$ and the prevailing BS-user density ratio υ. If $\mathcal{B}$ is fixed, average sum rate is independent of the number of subchannels δ since all available subchannels are occupied. In practice, the subchannel size is normally fixed. Therefore if the desired minimum rate $\mathcal{R}_{min}$ (in bps) of a typical user is known, it is possible to determine the minimum value of $\mathcal{B}$, denoted as $\mathcal{B}_{min}$, that is required to achieve it. Using (4.23), average user rate is $\mathcal{B} \times \mathcal{R}_u$ bps. Therefore $\mathcal{B}_{min}$ must satisfy the inequality

$$\mathcal{B}_{min} \geq \frac{\mathcal{R}_{min} \times \mathbb{E}[C_b]}{p_a \Xi \left[\pi \lambda_b e^{-\frac{\sigma^2}{P_t L}(2^t - 1) x^{\alpha/2}} \, e^{-\pi \lambda_b (1 + \overline{p}_a \zeta(t, \alpha)) x} \right]}. \tag{4.27}$$

Typical QoS requirements for different applications are shown in [128, Table 1.4].

4.4 Optimal Deployment of Homogeneous Networks

Several works have tried to predict the BS density of cellular networks to achieve a target performance. In [129], both coverage and rate constraints are used to determine the optimal transmit power that minimizes the APC of a homogeneous network. However, this work does not perform a joint optimization of the BS density and transmit power to quantify the optimal deployment configuration. In addition, it is assumed that all BSs are always active which ignores the impact of the variable traffic profile. Authors in [117] first derived an expression called *deployment factor* which combines the BS density and transmit power in some mathematical form. The deployment factor was jointly optimized to determine the specific optimal BS density and transmit power configuration subject to a coverage constraint only. This work also assumes that all BSs are always active. Moreover, the deployment factor is not expressed in closed form and can only be determined numerically.

In this thesis, a joint optimization of the BS density and transmit power is performed to obtain the optimal deployment configuration subject to both coverage probability and average user rate constraints. In some special cases, the deployment factor is expressed in closed form. In addition, the effect of the varying user density on the optimal deployment configuration is investigated. Coverage probability and average rate constraints are necessary because satisfying one measure does not automatically satisfy the other. For instance, (i) if SINR is high (good coverage), average user rate may remain low due to insufficient bandwidth, and (ii) if SINR is low (poor coverage), average user rate may still be high due to high bandwidth availability.

Intuitively, additive noise reduces the coverage and average rate performance of the network. Therefore, the interference-limited network defines the upper bound on network performance. Generally as BS density reduces, aggregate interference also reduces which enhances the negative effect of noise on network performance. To manage this, coverage probability and average user rate constraints are expressed in terms of their upper bound values as

$$\mathcal{P}_c \geq \epsilon \overline{\mathcal{P}}_c \quad \text{and} \quad \mathcal{R}_u \geq \kappa \overline{\mathcal{R}}_u \tag{4.28}$$

where $\epsilon \in (0, 1]$ and $\kappa \in (0, 1]$ are the ratios of the coverage probability and average user rate to their upper bound values respectively. The optimization framework is expressed as

$$\begin{cases} \underset{\lambda_b}{\text{minimize}} & \text{APC} = p_a \lambda_b P_{act} + (1 - p_a)\lambda_b P_{sleep} \\ \text{subject to} & \mathcal{P}_c \geq \epsilon\, \overline{\mathcal{P}}_c,\ \mathcal{R}_u \geq \kappa\, \overline{\mathcal{R}}_u,\ P_t \leq \overline{P_t}. \end{cases} \tag{4.29}$$

Due to the complementary relationship between coverage and rate measures (optimization based on one measure improves the other measure as well), this optimization problem can be separated into two individual problems, one constrained by coverage probability and the other constrained by average user rate. The final solution is the maximum value from the two individual optimization problems.

Lemma 4.2. *When $\alpha > 2$ and $\sigma^2 > 0$ which is true in a mobile environment, coverage probability and average user rate increase monotonically with BS density and transmit power.*

Proof. To prove the dependence of coverage probability on transmit power, consider two transmit power values such that $P_{t_2} > P_{t_1}$. Coverage probability corresponding to P_{t_1} is

$$\mathcal{P}_c(P_{t_1}) = \pi\lambda_b \int_{x>0} e^{-\frac{a_1}{P_{t_1}}x^{\alpha/2}} e^{-a_2 x} \mathrm{d}x \tag{4.30}$$

where $a_1 = \frac{T\sigma^2}{L}$ and $a_2 = \pi\lambda_b(1+\overline{p}_a\rho(T,\alpha))$. For P_{t_2}, coverage probability becomes

$$\mathcal{P}_c(P_{t_2}) = \pi\lambda_b \int_{x>0} e^{-\frac{a_1}{P_{t_2}}x^{\alpha/2}} e^{-a_2 x} \mathrm{d}x. \tag{4.31}$$

Since $a_1 > 0$ and $P_{t_2} > P_{t_1}$, then $e^{-\frac{a_1}{P_{t_2}}} > e^{-\frac{a_1}{P_{t_1}}}$ and hence $\mathcal{P}_c(P_{t_2}) > \mathcal{P}_c(P_{t_1})$, which proves that coverage probability increases monotonically with transmit power. Dependence of average user rate on transmit power is proved in the same way.

To prove the dependence of average subchannel rate on BS density, consider two BS densities such that $\lambda_{b_2} > \lambda_{b_1}$. Average subchannel rate corresponding to λ_{b_2} is expressed as

$$\mathcal{R}_{ch}(\lambda_{b_2}) = \frac{\pi}{\delta}\lambda_{b_2}\Xi\left[e^{-a_3(2^t-1)x^{\alpha/2}} e^{-\lambda_{b_2}\pi(1+\zeta(t,\alpha))x}\right] \tag{4.32}$$

where $a_3 = \frac{\sigma}{P_t L}$. Now let $x = y\frac{\lambda_{b_1}}{\lambda_{b_2}}$ and substitute into (4.32). Then,

$$\begin{aligned}\mathcal{R}_{ch}(\lambda_{b_2}) &= \frac{\pi}{\delta}\lambda_{b_1}\Xi\left[e^{-a_3(2^t-1)\left(\frac{\lambda_{b_1}}{\lambda_{b_2}}\right)^{\alpha/2} y^{\alpha/2}} e^{-\lambda_{b_1}\pi(1+\zeta(t,\alpha))y}\right] \\ &\overset{(a)}{>} \frac{\pi}{\delta}\lambda_{b_1}\Xi\left[e^{-a_3(2^t-1)y^{\alpha/2}} e^{-\lambda_{b_1}\pi(1+\zeta(t,\alpha))y}\right] \\ &= \mathcal{R}_{ch}(\lambda_{b_1})\end{aligned} \tag{4.33}$$

where (a) follows since $a_3 > 0$, $\alpha > 2$, and $\left(\frac{\lambda_{b_1}}{\lambda_{b_2}}\right) < 1$. Hence, $\mathcal{R}_{ch}(\lambda_{b_2}) > \mathcal{R}_{ch}(\lambda_{b_1})$ as long as $\lambda_{b_2} > \lambda_{b_1}$. Dependence of coverage probability on BS density is proved in the same way. ∎

4.4.1 Coverage Probability Constraint

Corollary 4.3. *Coverage probability in a PPP-based homogeneous network is approximated as*

$$\mathcal{P}_c \approx \overline{\mathcal{P}}_c\left(1 - \frac{T\sigma^2\psi(\alpha)\overline{\mathcal{P}}_c^{\alpha/2}}{\lambda_b^{\alpha/2}P_t}\right) \tag{4.34}$$

where $\psi(\alpha) = \frac{\Gamma\left(\frac{\alpha}{2}+1\right)}{\pi^{\alpha/2}L}$ *and* $\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}\,\mathrm{d}x$ *is the gamma function.*

Proof. Coverage probability in (4.16) can also be written as

$$\mathcal{P}_c(\mathrm{SNR}, \upsilon, T, \alpha) = 2\pi\lambda_b \int_{s>0} e^{-\frac{T}{\mathrm{SNR}}} e^{-\pi\lambda_b(1+\overline{p}_a\rho(T,\alpha))s^2} s\mathrm{d}s \tag{4.35}$$

where $\mathrm{SNR} = \frac{P_t L r^{-\alpha}}{\sigma^2}$ and $x = s^2$. In a realistic network scenario, target coverage probability is set as high as possible so that $\mathrm{SINR} > T$ in most cases. Furthermore, due to significant

aggregate interference in dense cellular networks, SNR is normally much greater than SINR. Therefore it is reasonable to assume that in a realistic SNR range, SNR $\gg$ SINR or SNR $\gg T$. Using this assumption, the first exponential term in (4.35) can be approximated as

$$e^{-\frac{T}{\text{SNR}}} \approx 1 - \frac{T}{\text{SNR}} \equiv 1 - \frac{T\sigma^2}{P_t L} x^{\alpha/2}. \tag{4.36}$$

Substituting this approximation into (4.16) gives the following integrals:

$$\begin{aligned}\mathcal{P}_c &\overset{(a)}{\approx} \pi\lambda_b \int_{x>0} e^{-\pi\lambda_b(1+\overline{p}_a\rho(T,\alpha))x}\,\mathrm{d}x - \pi\lambda_b \frac{T\sigma^2}{P_t L}\int_{x>0} x^{\alpha/2} e^{-\pi\lambda_b(1+\overline{p}_a\rho(T,\alpha))x}\,\mathrm{d}x \\ &= \overline{\mathcal{P}}_c - \pi\lambda_b \frac{T\sigma^2}{P_t L}\frac{\Gamma\left(\frac{\alpha}{2}+1\right)}{\left[\pi\lambda_b(1+\overline{p}_a\rho(T,\alpha))\right]^{\frac{\alpha}{2}+1}}.\end{aligned} \tag{4.37}$$

The second integral in (a) is solved using the identity [126, (3.381.4)]. Further simplification of (4.37) gives the result. The approximation of coverage probability is verified in Fig. 4.1a which shows increasing accuracy as BS density increases (note that as BS density increases, the BS-user separation distance reduces which enhances received power and consequently SNR). ■

Therefore combining (4.28) and (4.34), the coverage probability constraint can be rewritten as

$$1 - \frac{T\sigma^2\psi(\alpha)\overline{\mathcal{P}}_c^{\alpha/2}}{\lambda_b^{\alpha/2}P_t} \geq \epsilon. \tag{4.38}$$

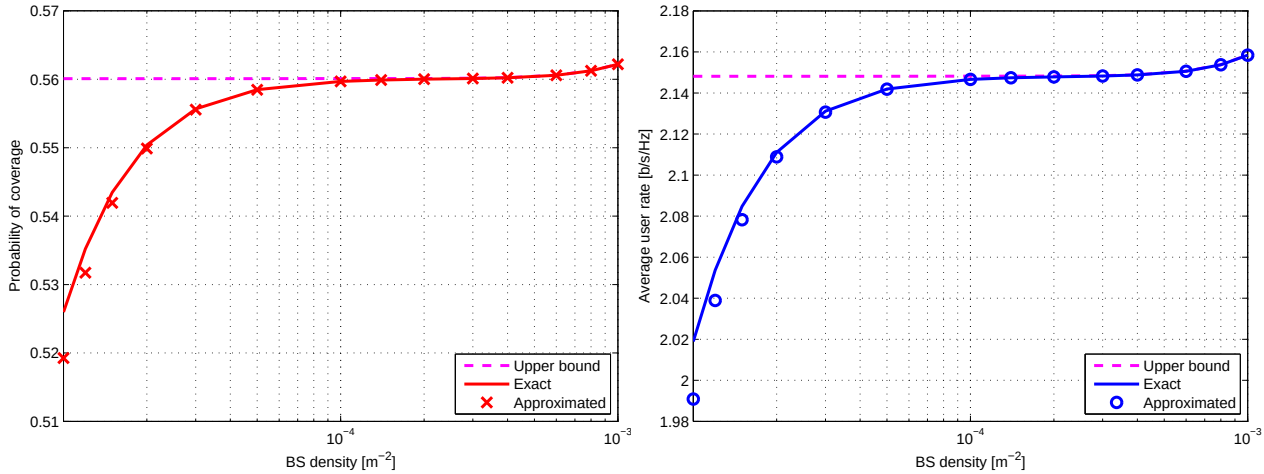
Denote $\mathcal{H}_c = \lambda_b^{\alpha/2}P_t$ as the *deployment factor* of the homogeneous network subject to a coverage probability constraint. Basically, the deployment factor expresses the joint BS density and transmit power in some mathematical form and can therefore be optimized to determine the optimal deployment configuration (optimal BS density and associated optimal transmit power). Since coverage probability and average rate both increase monotonically with the BS density and transmit power (see Lemma 4.2), the optimal deployment factor, denoted as $\mathcal{H}_c^\star$, is obtained when (4.38) is satisfied strictly. Therefore, $\mathcal{H}_c^\star$ satisfies the expression

$$\frac{\lambda_b^{\alpha/2}P_t}{\overline{\mathcal{P}}_c^{\alpha/2}} - \frac{T\sigma^2\psi(\alpha)}{1-\epsilon} = 0. \tag{4.39}$$

Generally, it is not possible to find a closed form expression for $\mathcal{H}_c^\star$ in (4.39) because $\overline{\mathcal{P}}_c$ is also a function of λ_b. To achieve further insight, it is possible to use different values of P_t and for each value, the corresponding BS density λ_b that satisfies the constraint strictly is determined using the bisection method [127]. The optimal deployment configuration ($\lambda_b^\star$, $P_t^\star$) is then the combination of P_t and λ_b that minimizes the APC of the network. This assumption is reasonable since network planning and optimization tasks are generally performed offline. The deployment factor can be expressed in closed form in the following special case.

Special Case 4.1. *When $\lambda_b \ll \lambda_u$, $p_a \approx 1$ and $\overline{\mathcal{P}}_c$ is independent of λ_b (see (3.12)). Hence, the deployment factor in (4.39) can be rewritten in closed form as*

$$\mathcal{H}_c^\star = \frac{T\sigma^2\psi(\alpha)}{(1-\epsilon)}\overline{\mathcal{P}}_c^{\alpha/2}. \tag{4.40}$$



(a) Probability of coverage (b) Average user rate

Figure 4.1: Verification of the approximations of the coverage probability and average user rate with idle BSs in sleep mode, where $\lambda_u = 10^{-2}\,\mathrm{m}^{-2}$, $\alpha = 4$, and $T = 0\,\mathrm{dB}$).

Given the deployment factor, it is optimized as a bivariate problem as follows

$$\begin{cases} \underset{\lambda_b,\, P_t}{\text{minimize}} & \lambda_b(\mathcal{N}_{tx}P_0 + \Delta P_t) \\ \text{subject to} & \lambda_b^{\alpha/2} P_t = \mathcal{H}_c^\star,\ \ P_t \leq \overline{P_t}. \end{cases} \tag{4.41}$$

This problem can easily be converted to a single-variable problem by substituting $\lambda_b = (\mathcal{H}_c^\star P_t^{-1})^{2/\alpha}$. The objective is then to determine the value of P_t that minimizes $F(P_t) = (\mathcal{H}_c^\star P_t^{-1})^{2/\alpha}(\mathcal{N}_{tx}P_0 + \Delta P_t)$. Using the differentiation method, the first derivative is

$$\frac{\mathrm{d}F(P_t)}{\mathrm{d}P_t} = \frac{-2}{\alpha}\mathcal{N}_{tx}P_0\mathcal{H}_c^{\star 2/\alpha}P_t^{-\frac{2+\alpha}{\alpha}} + \Delta\mathcal{H}_c^{\star 2/\alpha}\left(\frac{\alpha-2}{\alpha}\right)P_t^{-2/\alpha} = 0. \tag{4.42}$$

Simplifying (4.42), the optimal solutions of P_t and λ_b are expressed as

$$P_t^\star = \min\left\{\frac{2\mathcal{N}_{tx}P_0}{\Delta(\alpha-2)}, \overline{P_t}\right\} \quad \text{and} \quad \lambda_{b,c}^\star = \left(\frac{\mathcal{H}_c^\star}{P_t^\star}\right)^{2/\alpha}. \tag{4.43}$$

According to (4.43) therefore, the value of $P_t^\star$ does not depend on $\mathcal{H}_c^\star$ and is only a function of the BS power consumption parameters and the pathloss exponent. Hence it can be predetermined if the BS type to be deployed is known.

4.4.2 Average User Rate Constraint

Corollary 4.4. *Average user rate in a PPP-based homogeneous network is approximated as*

$$\mathcal{R}_u \approx \overline{\mathcal{R}}_u - \frac{p_a\sigma^2\psi(\alpha)\psi_r(t,\alpha)}{\mathbb{E}[C_b]\lambda_b^{\alpha/2}P_t} \tag{4.44}$$

where $\psi_r(t,\alpha) = \int_{t>0}\frac{(2^t-1)}{[1+\overline{p}_a\zeta(t,\alpha)]^{\frac{\alpha}{2}+1}}\mathrm{d}t$.

Proof. Average user rate in (4.23) can also be written as

$$\mathcal{R}_u(\text{SNR}, \upsilon, \delta, \alpha) = \frac{p_a}{\mathbb{E}[C_b]} \Xi \left[2\pi\lambda_b s e^{-\frac{2^t-1}{\text{SNR}}} e^{-\pi\lambda_b(1+\overline{p}_a\zeta(t,\alpha))s^2} \right]. \tag{4.45}$$

Using the approach in (4.36), the first exponential term in (4.45) is approximated as

$$e^{-\frac{2^t-1}{\text{SNR}}} \approx 1 - \frac{2^t-1}{\text{SNR}} \equiv 1 - \frac{\sigma^2}{P_t L}(2^t-1)x^{\alpha/2}. \tag{4.46}$$

Substituting (4.46) into (4.23) gives the following integrals:

$$\begin{aligned}\mathcal{R}_u &\overset{(a)}{\approx} \frac{p_a}{\mathbb{E}[C_b]}\pi\lambda_b \Xi\left[e^{-\pi\lambda_b(1+\overline{p}_a\zeta(t,\alpha))x}\right] - \frac{p_a}{\mathbb{E}[C_b]}\pi\lambda_b \frac{\sigma^2}{P_t L}\Xi\left[(2^t-1)x^{\alpha/2}e^{-\pi\lambda_b(1+\overline{p}_a\zeta(t,\alpha))x}\right] \\ &= \frac{p_a}{\mathbb{E}[C_b]}\int_{t>0}\frac{1}{1+p_a\zeta(t,\alpha)}\mathrm{d}t - \frac{p_a}{\mathbb{E}[C_b]}\pi\lambda_b\int_{t>0}\frac{\sigma^2\Gamma\left(\frac{\alpha}{2}+1\right)(2^t-1)}{P_t L(\pi\lambda_b)^{\frac{\alpha}{2}+1}[1+\overline{p}_a\zeta(t,\alpha)]^{\frac{\alpha}{2}+1}}\mathrm{d}t. \end{aligned} \tag{4.47}$$

The second integral in (a) is solved using the identity [126, (3.381.4)]. Further simplification of (4.47) gives the result. This approximation is verified in Fig. 4.1b which also shows increasing accuracy with the BS density. ∎

Therefore using the average user rate constraint in (4.28), (4.44) can be rewritten as

$$1 - \frac{\sigma^2\psi(\alpha)\psi_r(t,\alpha)}{\lambda_b^{\alpha/2}P_t\,\overline{\Re}(\upsilon,\alpha)} \geq \kappa \quad \text{where} \quad \overline{\Re}(\upsilon,\alpha) = \int_{t>0}\frac{1}{1+\overline{p}_a\zeta(t,\alpha)}\mathrm{d}t \tag{4.48}$$

is the SE (in b/s/Hz) in an interference-limited network with $\delta = 1$ [106].

In (4.48), denote $\mathcal{H}_r = \lambda_b^{\alpha/2}P_t$ as the deployment factor of the network subject to the average user rate constraint. Since average user rate increases monotonically with both BS density and transmit power, the optimal deployment configuration, denoted as $\mathcal{H}_r^\star$, is obtained when the constraint in (4.48) is satisfied tightly. Therefore, $\mathcal{H}_r^\star$ satisfies the expression

$$\frac{\lambda_b^{\alpha/2}P_t}{\psi_r(t,\alpha)} - \frac{\sigma^2\psi(\alpha)}{(1-\kappa)\overline{\Re}(\upsilon,\alpha)} \geq 0. \tag{4.49}$$

However, a closed form expression of $\mathcal{H}_r^\star$ is not possible because the terms $\Re(\upsilon,\alpha)$ and $\psi_r(t,\alpha)$ also depend on λ_b. Similar to the coverage constraint, the BS densities that satisfy (4.49) can be obtained for various values of P_t using numerical methods. The optimal network configuration is the $(\lambda_b^\star,\ P_t^\star)$ combination that minimizes the APC of the network. The following special case gives the deployment factor in closed form.

Special Case 4.2. *When $\lambda_b \ll \lambda_u$, $p_a \approx 1$ and both $\psi_r(t,\alpha)$ and $\overline{\Re}(\upsilon,\alpha)$ are independent of λ_b. Hence, the optimal deployment factor in (4.49) is rewritten in closed form as*

$$\mathcal{H}_r^\star = \frac{\sigma^2\psi(\alpha)\psi_r(t,\alpha)}{(1-\kappa)\overline{\Re}(\upsilon,\alpha)}. \tag{4.50}$$

The deployment factor $\mathcal{H}_r^\star$ can be optimized as a bivariate problem to determine the optimal

BS density and transmit power using the following optimization framework:

$$\begin{cases} \underset{\lambda_b,\, P_t}{\text{minimize}} & \lambda_b(\mathcal{N}_{tx}P_0 + \Delta P_t) \\ \text{subject to} & \lambda_b^{\alpha/2} P_t = \mathcal{H}_r^\star, \;\; 0 \le P_t \le \overline{P}_t. \end{cases} \tag{4.51}$$

Following the same procedure as (4.41)-(4.43), the optimal solutions are expressed as

$$P_t^\star = \min\left\{\frac{2\mathcal{N}_{tx}P_0}{\Delta(\alpha-2)}, \overline{P}_t\right\} \quad \text{and} \quad \lambda_{b,r}^\star = \left(\frac{\mathcal{H}_r^\star}{P_t^\star}\right)^{2/\alpha}. \tag{4.52}$$

Therefore according to (4.43) and (4.52), the optimal transmit power is independent of both the coverage probability and average user rate constraints and only depends on the BS power consumption parameters and the pathloss exponent. Therefore, the optimal transmit power can be predetermined if the deployed BS type is known. According to power parameters of different BSs in [19, Table 2], it is optimal to transmit at maximum power which correspondingly minimizes the BS density.

4.4.3 Overall Solution

The coverage and average user rate constraints are complementary to each other because optimization based on one measure also improves the other measure. Therefore, the overall optimal BS density, denoted as $\lambda_b^\star$, is one that satisfies both constraints simultaneously, i.e.

$$\lambda_b^\star = \max\{\lambda_{b,c}^\star, \lambda_{b,r}^\star\}. \tag{4.53}$$

4.5 Numerical Results

This section presents the performance results of a PPP-based dense homogeneous network with the default parameters shown in Table 4.1, unless otherwise stated.

Table 4.1: Simulation Parameters

Parameters	**Values**
Network area	$A = 5\,\text{km} \times 5\,\text{km}$
System bandwidth	$\mathcal{B} = 20\,\text{MHz}$
BS transmit power	$P_t = 21\,\text{dBm}$
Pathloss parameters	$L = -33\,\text{dB}$, $\alpha = 4$
Noise parameters	$F = 10$, $T_a = 300\,\text{K}$
BS power parameters	$\mathcal{N}_b = 2$, $P_0 = 6.8$, $\Delta = 4$, $P_{sl} = 4.3$
Optimization constraints	$\epsilon = \kappa = 0.9$

Fig. 4.2 shows the coverage probability for various SIR threshold values and pathloss exponents. Coverage probability is a reducing function of coverage threshold T because a higher coverage level is more difficult to achieve. In addition as α increases, the interference power, most

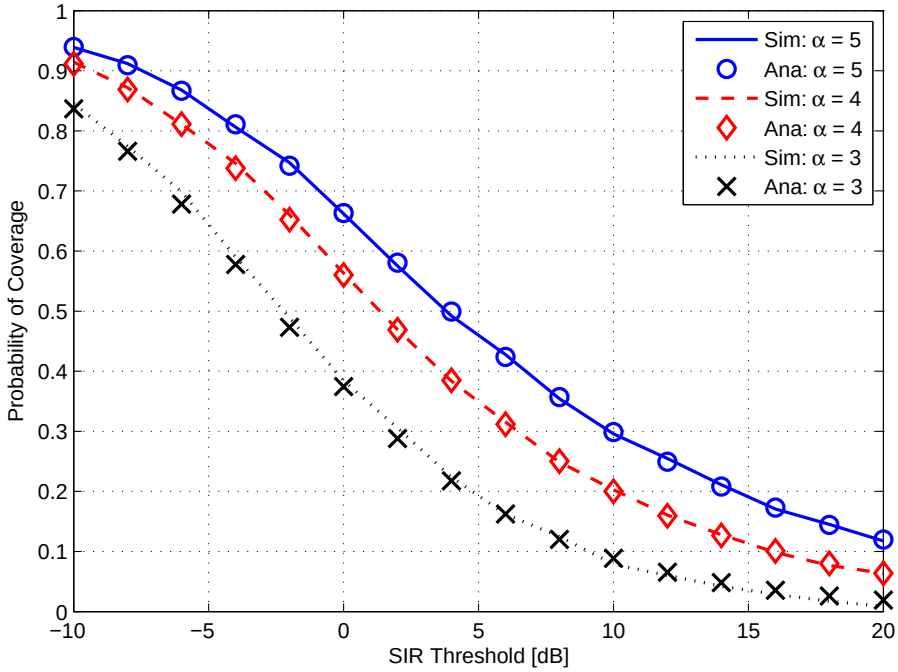


Figure 4.2: Probability of coverage for $\lambda_b = 4 \times 10^{-6}\,\mathrm{m}^{-2}$.

of which comes from distant sources, decays very fast and the SIR generally increases which consequently improves the coverage probability. The analytical results assume that the network extends indefinitely and therefore consider all possible sources of interference. However, Monte Carlo simulations consider a defined section of the network area which includes only a finite number of BSs. Therefore, the simulation results tend to be better than the analytical results. However, the disparity between both results can be minimized by considering a sufficiently large network area for Monte Carlo simulations. In practical deployments, interference from distant sources is often considered negligible so as to facilitate frequency reuse.

Average subchannel rate only depends on the subchannel size and does not vary with system bandwidth since all subchannels are occupied as shown in Fig. 4.3. However, average user rate increases with system bandwidth because the average number of subchannels per user increases.

Coverage probability and average rate are highly influenced by the BS-user density ratio υ as shown in Figs. 4.4-4.5. To explain, consider two distinct υ-regimes:

1. When $\lambda_b \ll \lambda_u$ (low-υ), the coverage probability and average subchannel rate of the interference-limited network are invariant with the BS density since all BSs remain active. In addition, both measures increase with the pathloss exponent α because at higher α, the interference (which originates further away) is attenuated more than the received signal thus increasing the average SNR. However in the presence of noise, the aggregate interference is very small (or negligible) in comparison and the network is essentially noise-limited which makes the coverage probability and average subchannel rate considerably lower than their corresponding upper bound levels in the interference-limited network. In contrast however, both coverage probability and average subchannel rate are now decreasing functions of α because the received signal reduces with α which lowers the average SNR in the noise-limited network.

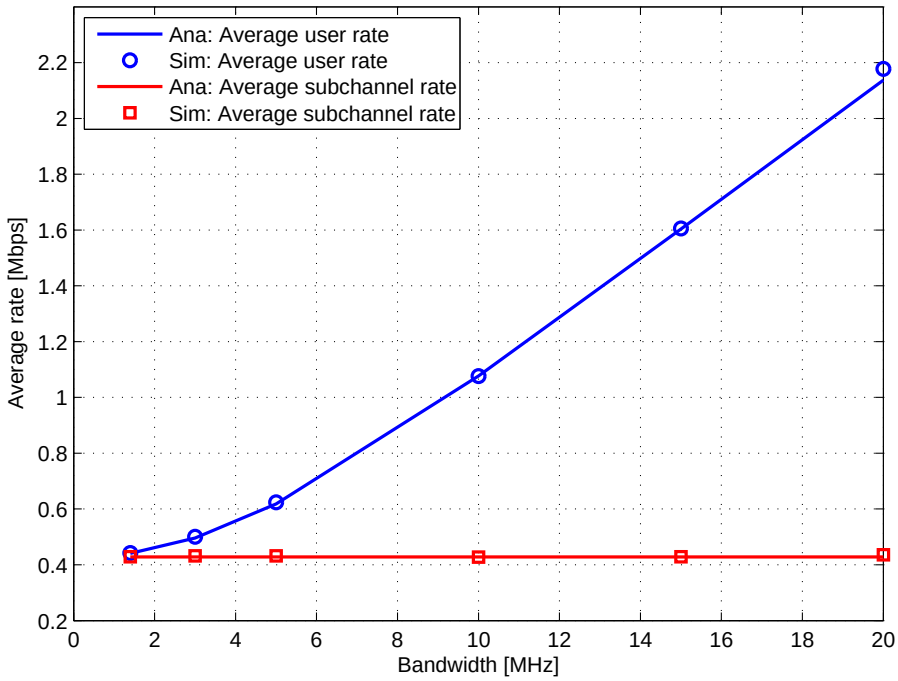


Figure 4.3: Average subchannel and user rates verses system bandwidth for $\lambda_b = 5 \times 10^{-4}\,\mathrm{m}^{-2}$ and $\lambda_u = 100 \times 10^{-4}\,\mathrm{m}^{-2}$.

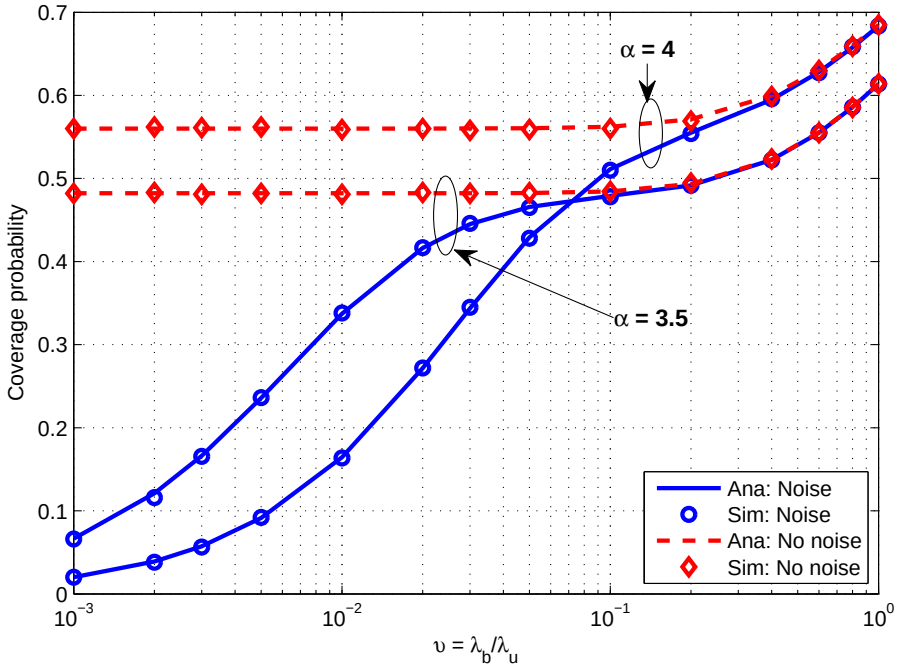


Figure 4.4: Variation of coverage probability with υ for $\lambda_u = 10^{-3}\,\mathrm{m}^{-2}$ and $T = 0\,\mathrm{dB}$.

2. As the BS density increases (υ increases), the gradual increase in the aggregate interference eventually makes the network interference-limited. As a result, both measures approach their corresponding upper bound levels as the effect of noise becomes evermore negligible. In this regime, both coverage probability and average subchannel rate are now increasing functions of α because of the more significant attenuation of interference (which originates far) compared to the received signal. In the very high-υ regime, the density of idle BSs increases and their thinning effect on the aggregate interference begins to enhance coverage probability and average subchannel rate.

Using (4.43) and (4.52), the optimal combination of BS density and transmit power that achieves both coverage probability and average rate targets can be determined. Fig. 4.6

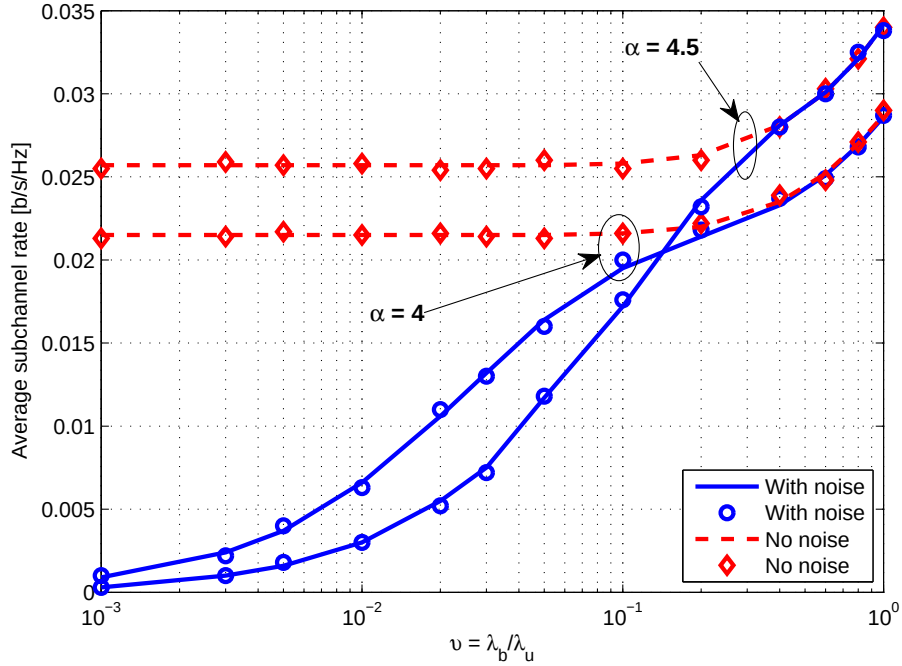


Figure 4.5: Variation of average subchannel rate with υ for $\lambda_u = 10^{-3}\,\mathrm{m}^{-2}$.

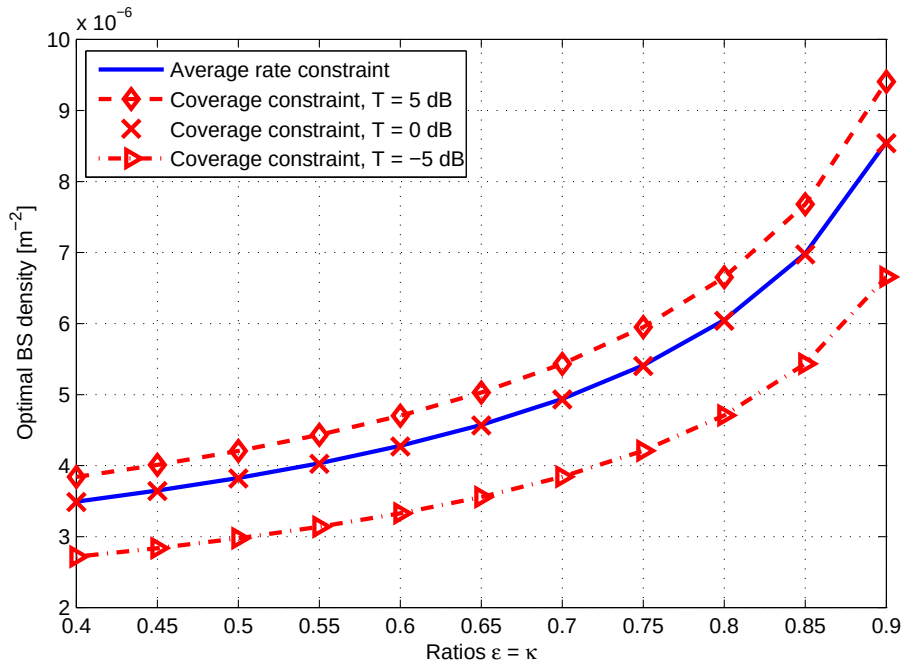


Figure 4.6: Optimal BS density versus optimization constraints where $\lambda_u = 10^{-3}\,\mathrm{m}^{-2}$.

shows that the optimal BS density subject to both constraints increases with the target ratios ϵ and κ because the network requires more BSs to approach its interference-limited scenario. The optimal BS density based on the coverage constraint increases with the SINR coverage threshold T since more BSs are required to enhance SINR level. When $T = 0\,\mathrm{dB}$, the optimal BS densities based on both constraints are approximately equal. However when $T > 0\,\mathrm{dB}$, the coverage constraint requires more BSs (i.e. $\lambda^\star_{b,c} > \lambda^\star_{b,r}$) and therefore it dominates the optimization process. Conversely when $T < 0\,\mathrm{dB}$, $\lambda^\star_{b,c} < \lambda^\star_{b,r}$ and the average rate constraint becomes decisive. This justifies why both coverage and rate constraints are required.

Optimal BS density also depends on the cellular network environment as shown in Fig. 4.7. In general, optimal BS density based on both constraints increases with the pathloss exponent α since wireless signals degrade more rapidly at high α. To further investigate the effect of α,

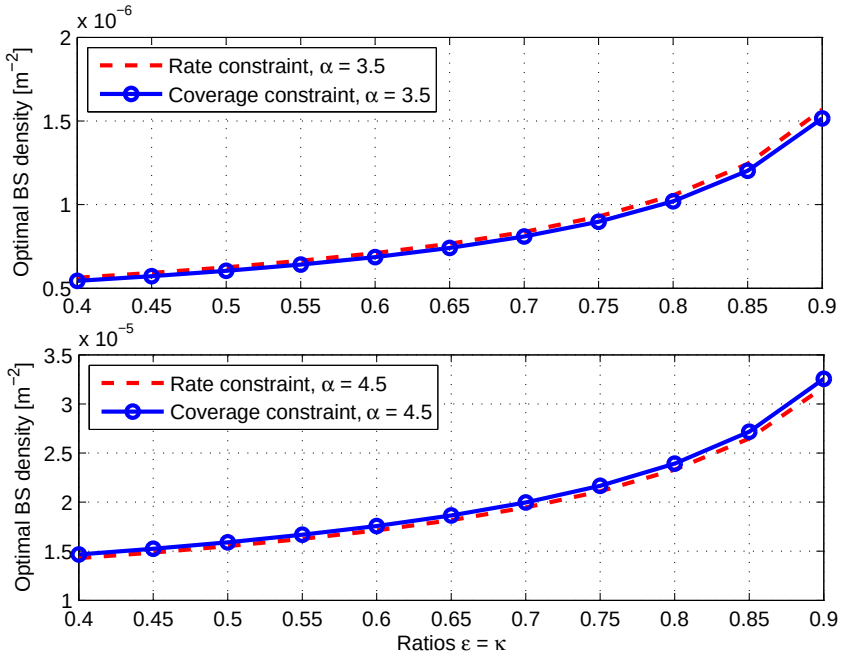


Figure 4.7: Effect of the pathloss exponent α on the optimal BS density subject to coverage and rate constraints ($\lambda_u = 10^{-3}\,\mathrm{m}^{-2},\ T = 0\,\mathrm{dB},\ \alpha = 4$).

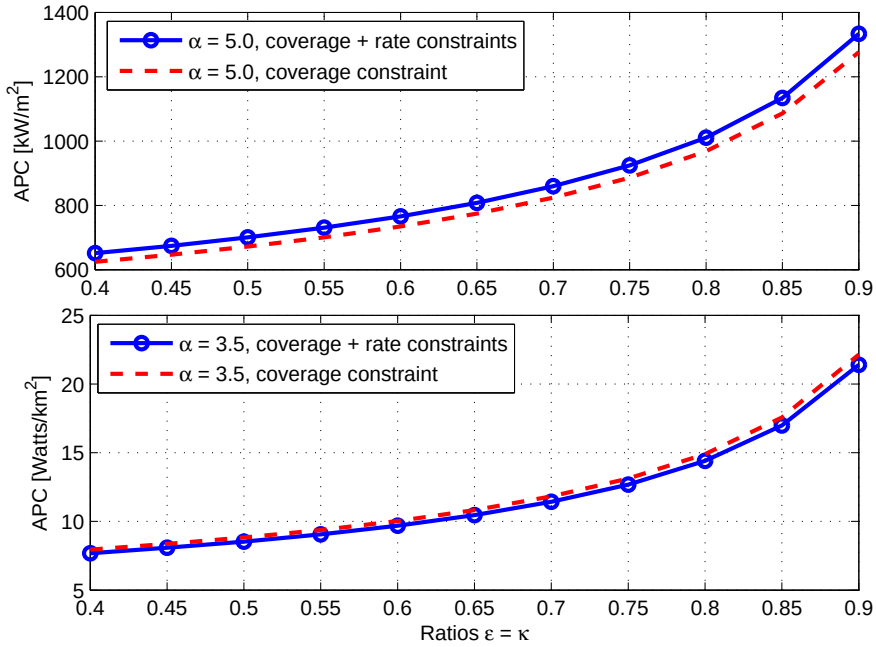


Figure 4.8: Optimal APC versus optimization constraints where $\lambda_u = 10^{-3}\,\mathrm{m}^{-2}$.

assume that $\epsilon = \kappa$ and consider the case of $T = 0\,\mathrm{dB}$ where $\lambda^{\star}_{b,c}$ and $\lambda^{\star}_{b,r}$ are within the same range. When $\alpha < 4$, $\lambda^{\star}_{b,c} > \lambda^{\star}_{b,r}$ but when $\alpha > 4$, $\lambda^{\star}_{b,c} < \lambda^{\star}_{b,r}$. This further emphasizes the necessity of considering both coverage and rate constraints.

The APC, shown in Fig. 4.8, follows a similar trend to the optimal BS density as the optimization constraints change since power consumption per active BS is a constant. The benchmark result considers the approach in [117] which considered only a coverage constraint. When $\alpha = 4$, optimal BS density is similar for either constraint; therefore either constraint is satisfactory in such an environment. However when $\alpha < 4$, considering only a coverage constraint gives a higher APC than when both coverage and rate constraints are used. On the other hand, when $\alpha > 4$, considering only the coverage constraint gives a lower APC but the network will not

achieve the rate target. Hence considering both constraints not only minimizes the APC but also allows the network to achieve all coverage and rate performance targets in all network environments.

4.6 Summary

This chapter presented some of the contributions of this thesis on the performance analysis of PPP-based homogeneous cellular networks. User connectivity models were developed to investigate physical channel allocation and identify idle BSs which impact network performance. Furthermore, an APC-minimization framework was presented to determine the optimal BS density and associated transmit power configuration of the network subject to appropriate coverage and rate constraints. Analysis shows that optimal transmit power is only a function of the BS power consumption parameters. On the other hand, optimal BS density is easily determined using numerical methods. In some special cases, it is expressed in close form.

Chapter 5

Sleep Mode Mechanisms

5.1 Introduction

Future networks will be characterized by a dense deployment of BSs to provide seamless coverage and required capacity enhancements [130]. However, user density is known to vary significantly in space and time and therefore it might be necessary to vary the BS density accordingly to manage the energy performance of the network. A popular technique to adapt the BS density to changes in the user density is to implement sleep mode mechanisms where some BSs are put to sleep mode or completely switched off. In this thesis, two states of sleep mode are considered namely; (i) *sleep mode* where a BS consumes a small amount of power, and (ii) *deep sleep mode* in which a BS is completely switched off and consumes no power. Sleep modes depend on the prevailing BS-UE density ratio and their performance is highly influenced by the criteria used to choose BSs to put into sleep mode.

Several works have studied this idea of managing energy consumption using sleep mode mechanisms [71], [72], [73], [74], [75], [76], [131]. Sleep mode should prioritize BSs with the least activity in the network [71]. In addition, it is beneficial to consider the load profile of the wider geographical coverage area of the network although this makes the implementation of sleep mode more challenging [131]. Energy savings can also be achieved by carefully replacing some macro BSs with micro BSs coupled with appropriate power adjustment [72]. The authors in [73] performed a joint optimization of the BS sleeping control and power matching schemes to achieve flexible tradeoffs between power consumption and user QoS. In [74], the authors investigated the effect of sleep mode on the average user capacity and optimized the transmit power of the remaining BSs to maintain the outage target. The authors in [75] propose a technique of managing power consumption by switching off some BSs and balancing the prevailing traffic load between the remaining active BSs. In [76], the authors proposed a sleep mode algorithm that prioritizes BSs with the least perceived importance to the network as candidates for sleep mode. However, the power consumed by BSs in sleep mode needs to be managed carefully to enhance the benefits of sleep mode. Authors in [132] clearly demonstrated the need for sleep mode power consumption of BSs to be minimized since it can negatively affect the realizable EE especially in dense networks where numerous BSs may be in sleep mode simultaneously.

In this chapter, our contributions include the proposed centralized and distributed strategic sleep mode schemes which enhance both average rate per active BS and network EE compared to existing conventional and random sleep mode schemes. It also investigates the effect of varying user density on network performance. As the user density changes between two extremes, the network is studied over three distinct BS-user density ratios (very low, comparable, and very large). Using appropriate approximations, simpler and tractable analytical expressions of coverage probability, average rate, ACR, etc are derived for each of these classes which allows interesting insights to be made into the performance of the network.

5.2 Conventional Sleep Mode

In this scheme, all idle BSs are put into sleep mode to save energy and no BSs are in deep sleep mode. Idleness refers to a complete lack of users within its coverage area. Hence a typical BS remains active with probability p_a and goes to sleep with probability $1 - p_a$, where

$$p_a = 1 - \mathbb{P}(N = 0) = 1 - \left(1 + \frac{\lambda_u/\lambda_b}{\mathcal{K}}\right)^{-\mathcal{K}}. \tag{5.1}$$

Since all users are already connected to their respective parent BSs, conventional sleep mode does not affect the received power of any user. However, the aggregate interference reduces which enhances the SINR level in the network. Therefore, conventional sleep mode maximizes the network sum rate since it guarantees the maximum possible active BS density. All the previous coverage and average rate analysis of the homogeneous network is based on this scheme. The APC of the network with conventional sleep mode is expressed as

$$\text{APC} = \lambda_b A[p_a(P_{act} - P_{sleep}) + P_{sleep}] \tag{5.2}$$

where A is the network area, and P_{act} and P_{sleep} are the powers consumed by a BS in active and sleep modes respectively. Using average sum rate in (4.25), EE is expressed as

$$\text{EE}_c = \frac{\mathcal{B}\, p_a \Xi[\pi\lambda_b e^{-\frac{\sigma^2}{P_t L}(2^t-1)x^{\alpha/2}}\, e^{-\pi\lambda_b(1+\overline{p}_a\zeta(t,\alpha))x}]}{A[p_a(P_{act} - P_{sleep}) + P_{sleep}]} \text{ [bits/Joule]}. \tag{5.3}$$

The performance tradeoff between EE and SE can therefore be evaluated using (5.3). However, it is not possible to express EE in (5.3) in closed form.

The amount of power consumed by sleep mode BSs impacts and can severely limit EE especially if many BSs are in sleep mode simultaneously [76]. For a given BS type, the maximum sleep mode power consumption that guarantees a target EE, denoted as EE_T, is expressed as

$$P_{sleep} = \left[\frac{\mathcal{B}\, p_a \Xi[\pi\lambda_b e^{-\frac{\sigma^2}{P_t L}(2^t-1)x^{\alpha/2}}\, e^{-\pi\lambda_b(1+p_a\zeta(t,\alpha))x}]}{\text{EE}_T A(1-p_a)} - \frac{p_a P_{act}}{1-p_a}\right]^+ \tag{5.4}$$

where $[\cdot]^+$ guarantees that EE_T is chosen such that $P_{sleep} \geq 0$.

5.3 Random Sleep Mode

Although conventional sleep mode puts all idle BSs in sleep mode, it may be necessary to put extra BSs to sleep so as to further reduce energy consumption. In random sleep mode, the BS density is independently thinned before any cell association and the users are served by the remaining BSs. Therefore, some users may be forced to connect to more distant BSs if their nearest (original or parent) BSs have been selected for deep sleep mode. It is assumed that BSs in deep sleep mode remain in this state for a relatively extended time period (for example, a network may operate with an independently thinned BS density during the low traffic period at night). The multi-user connectivity model facilitates the thinned BS density to take on more users and maintain acceptable blocking rates. This scheme is discussed in [71], [133] and will only be used in this thesis for comparison purposes.

If p_r is the target proportion of active BSs, then the density of BSs in deep sleep mode is $(1-p_r)\lambda_b$. When $p_r = 1$, random sleep mode is equivalent to conventional sleep mode. The analysis of random sleep mode is essentially similar to that of conventional sleep mode, the only difference being that the available BS density is $p_r\lambda_b$ instead of λ_b in the conventional scenario. After cell association, there is a finite probability that some BSs may still remain idle and can also be put to sleep mode. Using analysis similar to the conventional scheme, the probability of an active BS in a network with random sleep mode is expressed as

$$p_{a_r} = 1 - \mathbb{P}(N_r = 0) = 1 - \left(1 + \frac{\lambda_u/p_r\lambda_b}{\mathcal{K}}\right)^{-\mathcal{K}} \tag{5.5}$$

where N_r is the number of users in a cell in a network with random sleep mode. Therefore, the respective densities of BSs in active and sleep modes are $p_{a_r}p_r\lambda_b$ and $(1-p_{a_r})p_r\lambda_b$.

Compared to conventional sleep mode in (4.20), average subchannel rate is now expressed as

$$\mathcal{R}_{ch_r}(\upsilon,\alpha) = \frac{1}{\delta}\Xi\left[\pi p_r\lambda_b e^{-\frac{\sigma^2}{P_t L}(2^t-1)x^{\alpha/2}}\, e^{-\pi p_r\lambda_b(1+\overline{p}_{a_r}\zeta(t,\alpha))x}\right] \tag{5.6}$$

where $\overline{p}_{a_r}\lambda_b = p_{a_r}\lambda_b\backslash\{b_o\}$ and b_o is the serving BS.

ACR, denoted as ACR_r, is $\text{ACR}_r \times \lambda_u = p_r\lambda_b\,\mathbb{E}[C_{br}]$ where $C_{br} \in \{0,\delta\}$ is derived similar to (4.14). The average number of subchannels per user is expressed as

$$\omega_r(\upsilon,\delta) = \frac{\Omega_r(\upsilon,\delta)}{\text{ACR}_r \times \lambda_u} = \frac{p_{a_r}\delta}{p_r\mathbb{E}[C_{br}]} \tag{5.7}$$

where $\Omega_r(\upsilon,\delta) = p_{a_r}\lambda_b\delta$ represents the average number of channels available in the network. The average user rate is expressed as $\mathcal{R}_{u_r}(\upsilon,\delta,\alpha) = \omega_r(\upsilon,\delta) \times \mathcal{R}_{ch_r}(\upsilon,\alpha)$. The average sum rate is then determined as $\mathcal{T}_r(\upsilon,\delta,\alpha) = \text{ACR}_r \times \lambda_u \times \mathcal{R}_{u_r}(\upsilon,\delta,\alpha)$.

Since any idle BSs after cell association are also put to sleep mode to save energy, the APC of a network with random sleep mode is expressed as

$$\text{APC} = p_r\lambda_b A[p_{a_r}(P_{act} - P_{sleep}) + P_{sleep}]. \tag{5.8}$$

Random sleep mode thins the BS density independently without considering user occupancy of affected BSs. However, practical sleep mode schemes are most likely to consider the dynamic spatiotemporal distribution and variation of users in the network. In the following, we propose two strategic sleep mode schemes which exploit this spatiotemporal variation and distribution of users to enhance network energy performance. These schemes may be applied periodically to respond and adapt network energy consumption to variations in user density.

5.4 Centralized Strategic Sleep Mode

In random sleep mode, sleep mode BSs are chosen randomly and hence some BSs with many users may be affected. In centralized strategic sleep mode, a BS remains active or not depending on its perceived degree of importance to the network. The centralized strategic approach considers the whole network and prioritizes BSs with the least number of users for sleep mode [76]. Assuming that the network is initially thinned to $p_s\lambda_b$ where p_s is the probability that a BS remains available for cell association, the density of BSs in *deep* sleep mode is $(1-p_s)\lambda_b$. The value of p_s may be set based on some criteria similar to p_r in random sleep mode. For example, all BSs with $N < n$ users may be put to sleep such that p_s becomes

$$p_s = 1 - \mathbb{P}(N < n) = \sum_{m=0}^{n-1} \frac{\lambda_u^m (\mathcal{K}\lambda_b)^{\mathcal{K}} \Gamma(m+\mathcal{K})}{\Gamma(\mathcal{K}) m! (\lambda_u + \mathcal{K}\lambda_b)^{m+\mathcal{K}}}. \tag{5.9}$$

In general, any given percentage of BSs may be switched off strategically by prioritizing BSs with the fewest users. If any two BSs have the same number of users but only one is to be switched off, the choice is made randomly in this thesis. Since idle BSs are prioritized for sleep mode, the centralized strategic scheme is similar to the conventional scheme over some range of p_s, which creates two regions of interest:

(i) If $p_s \geq p_a$, both centralized strategic and conventional sleep mode schemes are identical since only idle BSs are put to sleep. However, in centralized strategic sleep mode, some BSs are put into deep sleep mode except when $p_s = 1$. The respective density of BSs in active, sleep and deep sleep modes are $p_a\lambda_b$, $(p_s - p_a)\lambda_b$ and $(1-p_s)\lambda_b$.

(ii) If $p_s < p_a$, all available BSs of density $p_s\lambda_b$ remain active. The density of BSs in deep sleep mode is $(1-p_s)\lambda_b$, which consists of all idle BSs of density $(1-p_a)\lambda_b$ and extra BSs of density $(p_a - p_s)\lambda_b$ which contain the fewest users. Hence in this range of p_s, a BS is either in active mode or deep sleep mode only.

Therefore, centralized strategic sleep mode provides two opportunities: (i) it maximizes the number of users that remain connected to their parent BSs, and (ii) it minimizes the total number of affected users which enhances their chance of getting a connection from a neighboring active BS. The network APC with centralized strategic sleep mode is expressed as

$$\text{APC} = \begin{cases} p_s\lambda_b A\, P_{act}, & p_s < p_a \\ \lambda_b A\,[p_a P_{act} + (p_s - p_a) P_{sl}], & p_s \geq p_a. \end{cases} \tag{5.10}$$

5.5 Distributed Strategic Sleep Mode

Although the centralized strategic scheme is optimal, it is difficult to implement in very large networks. In the distributed strategic scheme, the strategic algorithm is implemented in smaller clusters all over the network. The network is subdivided into a grid of N_c equal-sized squares where each square represents a cluster area. Cluster boundaries only determine which cluster each BS belongs to for sleep mode optimization and do not affect cell association since users still connect to their nearest parent BSs. Therefore, BSs within each square form a cluster and together with their associated users define the clusters BS-user density ratio. Although this makes it easier and more manageable to implement, it trades off some of the performance of the centralized strategic scheme. If $N_c = 1$, the centralized and distributed strategic sleep mode schemes are identical. If the i-th cluster has a BS density λ_{b_i} and a user density λ_{u_i}, then the probability of a BS in the cluster being active is expressed as

$$p_{a_i} = 1 - \left(1 + \frac{\lambda_{u_i}/\lambda_{b_i}}{\mathcal{K}}\right)^{-\mathcal{K}}. \tag{5.11}$$

In each cluster, a fraction $1 - p_s$ of all BSs are put into deep sleep mode, prioritizing those with the least number of users. As in the centralized strategic scheme, the distributed strategic scheme in each cluster also gives two regions of interest:

(i) When $p_s \geq p_{a_i}$, the active BS density is $p_{a_i}\lambda_{b_i}$. Only idle BSs are affected and the densities of BSs in sleep and deep sleep modes are $(p_s - p_{a_i})\lambda_{b_i}$ and $(1-p_s)\lambda_{b_i}$ respectively. Denote the number of clusters in this category as $\overline{N}_c$.

(ii) When $p_s < p_{a_i}$, $p_s\lambda_{b_i} < p_{a_i}\lambda_{b_i}$ which means that the active BS density is exactly $p_s\lambda_{b_i}$. All remaining BSs of density $(1 - p_s)\lambda_{b_i}$ are in deep sleep mode. They consist of all idle BSs of density $(1 - p_{a_i})\lambda_{b_i}$ and extra BSs of density $(p_{a_i} - p_s)\lambda_{b_i}$ which contain the least number of users. The number of clusters in this category is $N_c - \overline{N}_c$.

It is important to investigate how the active BS density of a network with distributed strategic scheme compares to that of the centralized strategic scheme. For example, consider the point $p_s = p_a$ in centralized strategic scheme where active BS density, denoted as $\widetilde{B}_c$, is $p_a\lambda_b$ and all idle BSs are prioritized for deep sleep mode. In the distributed strategic scheme, $p_s = p_{a_i}$ for the i-th cluster which means that p_s varies in each cluster depending on its BS-user density ratio. Therefore, in clusters where $p_s < p_{a_i}$, some extra BSs with users are also put in deep sleep mode alongside idle BSs (see (ii) above). Denote the sum of all active BSs in all clusters of the network with distributed strategic scheme as $\widetilde{B}_d$. Then, it is clear that $\widetilde{B}_c \geq \widetilde{B}_d$ i.e.

$$p_s\lambda_b \geq \sum_{i=1}^{\overline{N}_c} p_{a_i}\lambda_{b_i|p_s \geq p_{a_i}} + \sum_{i=\overline{N}_c+1}^{N_c} p_s\lambda_{b_i|p_s < p_{a_i}}. \tag{5.12}$$

This difference in active BS density partly explains the sub-optimal performance of distributed strategic scheme compared to its centralized counterpart. Using (5.12), the APC of the network

with distributed strategic scheme is expressed as

$$\text{APC} = \sum_{i=1}^{\overline{N}_c} \lambda_{b_i} \left[p_{a_i}(P_{act} - P_{sleep}) + p_s P_{sleep}\right]_{|p_s \geq p_{a_i}} + \sum_{i=\overline{N}_c+1}^{N_c} p_s \lambda_{b_i} P_{act|p_s < p_{a_i}}. \tag{5.13}$$

5.6 Effect of Varying User Density

Future networks will consist of a high density deployment of BSs to provide ubiquitous coverage and rate to subscribers. The aggregate power consumption of all BSs is likely to raise energy costs and limit subscriber revenues and profits. Traditionally, cellular networks are planned to support peak traffic but traffic demand is known to vary highly in both space and time domains [49]. Therefore, a large density of BSs may remain idle especially when traffic demand is low. The prevailing user density in the network highly impacts the coverage, average rate, ACR and energy performance of the network. In highly dense networks, extreme variations of the user density may force the network to go through various BS-UE density ratio (υ) regimes. Consider the following three υ regimes [132]: (i) mid-level υ-regime, (ii) very low υ-regime, and (iii) very high υ-regime. The analysis in this section assumes that noise is negligible and that the network has one channel (i.e. $\delta = 1$). Note that in this single-channel network, the average channel and average user rates are equal i.e. $\mathcal{R}_{ch} = \mathcal{R}_u$.

5.6.1 Mid-Level BS-UE Density Ratio

In this regime, $\lambda_b \sim \lambda_u$ and therefore small changes in υ cause significant changes in the active BS probability shown in (4.7). For example when $\upsilon = 1$, $p_a = 0.59$ and when $\upsilon = 2$, $p_a = 0.37$. This makes the analytical approximation of p_a very challenging. In other words, it is difficult to define a scaling function of the average rate or coverage probability with υ. Equally, ACR performance also shows significant variation with υ.

Conventional Sleep Mode: An inspection of the average user rate of the interference-limited network (shown in (4.21)) shows that it scales approximately linearly with υ when $\lambda_b \sim \lambda_u$. Therefore, the average user rate can be approximated as the line [132]

$$\mathcal{R}_u(\upsilon) \approx c_1 \frac{\lambda_b}{\lambda_u} + c_2 = c_1 \upsilon + c_2 \tag{5.14}$$

where c_1 and c_2 are constants that depend on the υ-range considered. For example, considering the range $0.25 \leq \upsilon \leq 4.0$ which dimensions a network in which the BS density is up to four times smaller or bigger than the user density, the line of best fit[1] gives the constants as $c_1 = 0.6274$ and $c_2 = 2.2293$ as shown in Fig. 5.11.

According to (5.14), average subchannel rate is directly proportional to the BS density but inversely proportional to the user density. This is intuitive because a higher BS density improves the received signal power but reduces p_a (the consequence is lower aggregate interference) which

[1]It is obtained using the MATLAB functions `polyfit(x,y,n)` and `polyval(p,x)`.

enhances SINR. In contrast, a higher user density increases the density of active BSs which increases aggregate interference and lowers SINR. Average network sum rate becomes

$$\mathcal{T}_c(\upsilon) \approx p_a \lambda_b \mathcal{R}_u(\upsilon) = p_a \lambda_b (c_1 \upsilon + c_2). \tag{5.15}$$

Therefore, the average sum rate highly depends on the BS density and significant improvements can be derived from deploying more BSs. Due to this dependence on the BS density, EE performance is also very sensitive to the υ ratio. The EE is expressed as

$$\mathrm{EE}_c = \frac{\mathcal{T}_c}{E_c} = \frac{p_a(c_1 \upsilon + c_2)}{p_a P_{act} + (1 - p_a) P_{sleep}}. \tag{5.16}$$

Therefore, the power consumed by BSs in sleep mode has a big impact on network EE. If P_{sleep} is low, the operator can deploy a dense network with sleep mode to enhance average sum rate without worsening network EE. In general for a given υ-range, EE is maximized at a given optimal υ which can be obtained using the differentiation method. Using the original expression of $\mathcal{T}_c$, the first derivative of EE, denoted as $\mathcal{D}_c$, is determined as

$$\mathcal{D}_c = \frac{\mathrm{dEE}_c}{\mathrm{d}p_a} = \frac{\mathrm{EE}_c \widetilde{\mathcal{T}}_c - \mathcal{T}_c \widetilde{\mathrm{EE}}_c}{\mathrm{EE}_c^2} \tag{5.17}$$

where $\widetilde{\mathrm{EE}}_c = \lambda_b (P_{act} - P_{sleep})$ and $\widetilde{\mathcal{T}}_c = \int_{t>0} [1 + p_a \zeta(t, \alpha)]^{-2} \mathrm{d}t$ are the derivatives of EE_c and $\mathcal{T}_c$ respectively. The derivative $\widetilde{\mathcal{T}}_c$ of $\mathcal{T}_c$ is derived using the product rule as follows

$$\begin{aligned} \widetilde{\mathcal{T}}_c &= \lambda_b \int_{t>0} [1 + p_a \zeta(t, \alpha)]^{-1} \mathrm{d}t - p_a \lambda_b \int_{t>0} [1 + p_a \zeta(t, \alpha)]^{-2} \zeta(t, \alpha) \mathrm{d}t \\ &= \lambda_b \int_{t>0} \frac{1}{[1 + p_a \zeta(t, \alpha)]^2} \mathrm{d}t. \end{aligned}$$

To determine the optimal p_a, denoted as $p_a^\star$, then $\mathcal{D}_c = 0$. Hence

$$\begin{aligned} \mathcal{D}_c &= \int_{t>0} \left[\frac{p_a P_{act} + (1 - p_a) P_{sleep}}{[1 + p_a \zeta(t, \alpha)]^2} - \frac{p_a (P_{act} - P_{sleep})}{[1 + p_a \zeta(t, \alpha)]} \right] \mathrm{d}t = 0 \\ &\Rightarrow \int_{t>0} \frac{P_{sleep} + p_a^2 (P_{act} + P_{sleep})}{[1 + p_a \zeta(t, \alpha)]^2} \mathrm{d}t = 0. \end{aligned} \tag{5.18}$$

Using (5.18), $p_a^\star$ can be determined numerically. Then $\upsilon^\star$ is obtained at the point where $\mathcal{D}_c(p_a^\star) = 0$. According to (4.7), $\upsilon^\star$ is obtained from $p_a^\star$ in closed form as

$$\upsilon^\star = \frac{1}{\mathcal{K}} \left[(1 - p_a^\star)^{-1/\mathcal{K}} - 1 \right]^{-1}. \tag{5.19}$$

Random Sleep Mode: In this regime, the node densities are such that $p_r \lambda_b \sim \lambda_u$. In a network with random sleep mode, the BS-UE density ratio is expressed as $\upsilon_r = p_r \lambda_b / \lambda_u = p_r \upsilon$. Similar to the conventional scheme, average user rate can be approximated using the line [132]

$$\mathcal{R}_{u_r}(\upsilon) = r_1 \frac{p_r \lambda_b}{\lambda_u} + r_2 = r_1 p_r \upsilon + r_2 \tag{5.20}$$

where constants r_1 and r_2 are determined from the line of best fit and depend on the considered range of υ_r. For the υ_r-range $0.25 \leq \upsilon \leq 4.0$ in Fig. 5.11, the line constants are determined as $r_1 = 0.5470$ and $r_2 = 2.1541$. Therefore, the average channel rate is directly proportional to the probability p_r. The average network sum rate is approximated as

$$\mathcal{T}_r(\upsilon) \approx p_{a_r}\lambda_b(r_1 p_r \upsilon + r_2). \tag{5.21}$$

Similar to conventional sleep mode, there is an optimal υ-ratio at which network EE is maximized. Using the same differentiation method, the optimal point can also be obtained.

5.6.2 Very High BS-UE Density Ratio

In this regime, $\lambda_b \gg \lambda_u$ such that $\upsilon \to \infty$. Due to the relatively few users, most BSs are likely to remain idle. Furthermore, active BSs are likely to individually have very few users due to their relatively small coverage areas. This υ-regime is also studied in [134] which only analyzes the outage performance of the homogeneous network with conventional sleep mode.

Conventional Sleep Mode: When $\upsilon \to \infty$, the probability of an active BS shown in (5.1) can be approximated using the first two terms of its binomial series as

$$p_a \approx 1 - \left(1 - \mathcal{K}\frac{\lambda_u}{\mathcal{K}\lambda_b}\right) = \frac{\lambda_u}{\lambda_b} = \frac{1}{\upsilon}. \tag{5.22}$$

In other words, most BSs will be idle as $\upsilon \to \infty$. Using (5.22), average user rate becomes

$$\overline{\mathcal{R}}_u(\upsilon, \alpha) \approx \int_{t>0} \frac{1}{1 + (\lambda_u/\lambda_b)\zeta(t, \alpha)} \mathrm{d}t. \tag{5.23}$$

Similarly, the average sum rate of the network is simplified as

$$\overline{\mathcal{T}}(\upsilon, \alpha) \approx \lambda_u \int_{t>0} \frac{1}{1 + (\lambda_u/\lambda_b)\zeta(t, \alpha)} \mathrm{d}t = \lambda_u \overline{\mathcal{R}}_u(\upsilon, \alpha). \tag{5.24}$$

Therefore, as λ_b increases (or λ_u reduces), λ_u/λ_b reduces and $\overline{\mathcal{R}}_u(\upsilon, \alpha)$ increases further. In addition, the average sum rate is directly proportional to the user density i.e. $\overline{\mathcal{T}}(\upsilon, \alpha) = \lambda_u \overline{\mathcal{R}}_u(\upsilon, \alpha)$. To explain the direct proportionality, consider the ACR performance of the network in this regime which is approximated as

$$\text{ACR} = p_a \frac{\lambda_b}{\lambda_u} \approx 1. \tag{5.25}$$

In other words, each added user is highly likely to activate a previously idle BS because the coverage cells are very small. Therefore any increase in user density is matched by a comparable increase in active BS density which gives the linear relationship in average sum rate. Fig. 5.16 verifies that the approximation of $\overline{\mathcal{T}}(\upsilon, \alpha)$ is very accurate and becomes tighter as $\upsilon \to \infty$.

Coverage probability of an interference-limited network with conventional sleep mode becomes

$$\mathcal{P}_c(\upsilon, T, \alpha) \approx \frac{1}{1 + (\lambda_u/\lambda_b)\rho(T, \alpha)} \overset{(a)}{\approx} 1 - \frac{\lambda_u}{\lambda_b}\rho(T, \alpha), \tag{5.26}$$

where (a) uses the first two terms of its binomial series. Hence as $\upsilon \to \infty$, $\mathcal{P}_c \to 1$.

Random Sleep Mode: In the very high υ-regime (when $\lambda_b \gg \lambda_u$), an operator may decide to initially reduce the available user density by some degree. Active BS probability in a network with random sleep mode, shown in (5.5), is approximated using its binomial series as

$$p_{ar} \approx \frac{\lambda_u}{p_r \lambda_b}. \tag{5.27}$$

Hence, the average user rate and average sum rate of the network are approximated as

$$\overline{\mathcal{R}}_{u_r}(\upsilon, \alpha) = \int_{t>0} \frac{1}{1 + (\lambda_u/p_r\lambda_b)\zeta(t, \alpha)} \mathrm{d}t, \tag{5.28a}$$

$$\overline{\mathcal{T}}(\upsilon, \alpha) = \lambda_u \int_{t>0} \frac{1}{1 + (\lambda_u/p_r\lambda_b)\zeta(t, \alpha)} \mathrm{d}t. \tag{5.28b}$$

The average user rate with random sleep mode also increases as the υ-ratio increases and it also an increasing function of probability p_r. Similar to the conventional sleep mode scheme and for the same reason, the average sum rate also has a linear relationship with the user density.

Centralized and Distributed Strategic Sleep Mode: Since most BSs are idle, the strategic sleep mode algorithm prioritizes them for sleep mode. Therefore, unless p_s is very small, the performance of both strategic schemes resembles that of conventional sleep mode since they both prioritize idle BSs for sleep mode. Hence, both strategic schemes have no particular advantage over conventional and random sleep modes.

5.6.3 Very Low BS-UE Density Ratio

When $\lambda_b \ll \lambda_u$ (or $\upsilon \ll 1$) in a network with conventional sleep mode, the active BS probability shown in (5.1) is approximated as $p_a \approx 1$ i.e. every BS is likely to have at least one user. The ACR is approximated as

$$\text{ACR} = \frac{p_a \lambda_b}{\lambda_u} \approx \frac{\lambda_b}{\lambda_u} = \upsilon \ll 1. \tag{5.29}$$

Hence most users remain unconnected since ACR is very low. Average user rate becomes

$$\overline{\mathcal{R}}_u(\alpha) \approx \int_{t>0} \frac{1}{1 + \zeta(t, \alpha)} \mathrm{d}t \tag{5.30}$$

which gives a unique value that is independent of υ. For example, $\overline{\mathcal{R}}_u(\alpha = 4) = 2.15\,\text{b/s/Hz}$. The average sum rate of the network is approximated as

$$\mathcal{T}(\lambda_b, \alpha) \approx \lambda_b \int_{t>0} \frac{1}{1 + \zeta(t, \alpha)} \mathrm{d}t. \tag{5.31}$$

Therefore the average sum rate is directly proportional to the BS density. For example, $\mathcal{T}(\lambda_b, \alpha = 4) = 2.15\lambda_b$ b/s/Hz. This direct proportionality is a consequence of the fact that each additional BS is likely to associate with at least one user since cell sizes are relatively large. Hence, significant gains in average sum rate are possible by further densification of the network with more BSs. Furthermore, increasing the user density has no effect on average sum rate since all the BSs are already active.

Similarly, coverage probability does not depend on υ and is therefore a constant value for a given α. Coverage probability in this regime is shown in (3.12); for example $\mathcal{P}_c(T = 0\,\text{dB}, \alpha = 4) = 0.56$. This special network regime is studied in [106] which also assumes a single user connectivity model. Since $\lambda_b \ll \lambda_u$, random and strategic schemes are ignored because it is very unrealistic to further reduce the available BS density under such low ACR conditions.

5.7 Numerical Results

To investigate the performance of different *sleep mode mechanisms*, consider a PPP-based dense homogeneous network with the default parameters shown in Table 5.1, unless otherwise stated.

Table 5.1: Simulation Parameters

Parameters	Values
Network size	$A = 5\,\text{km} \times 5\,\text{km}$
Total BS bandwidth	$\mathcal{B} = 20\,\text{MHz}$
BS and user densities	$\lambda_b = 1.6 \times 10^{-5}$ m^2, $\lambda_u = 2\lambda_b$
Transmit power	$P_t = 21\,\text{dBm}$
Pathloss parameters	$L = -33\,\text{dB}$, $\alpha = 4$
Power parameters	$\mathcal{N}_b = 2$, $P_0 = 6.8\,\text{W}$, $\Delta = 4$, $P_{sl} = 4.3$
Additive noise parameters	$F = 10$, $T_a = 300\,\text{K}$

Coverage probability, average rate and energy consumption of the network depend on the sleep mode scheme implemented. Fig. 5.1 shows the density of BSs that remain active in the considered network area for each sleep mode scheme. The active BS density partly explains the performance capabilities of different sleep mode schemes. When $p_r = p_s = 1$, all schemes are identical and active BS density is maximized as $p_a\lambda_b$. In random sleep mode, active BS density varies proportionally with p_r. In contrast, centralized strategic sleep mode improves active BS density which approaches its upper bound ($p_a\lambda_b$) much faster. The active BS density of distributed strategic sleep mode is slightly lower than that of its centralized counterpart, which verifies the analysis in (5.12). The difference in active BS density between the two schemes increases with the cluster density but in a concave manner. To illustrate, consider the point $p_s = p_a \approx 0.8$ where the differences in active BS density with $N_c = 8$ and $N_c = 16$ are 3.5% and 5% respectively. This reducing effect of the cluster density on active BS density is important where operators prefer smaller and more manageable cluster sizes.

The performance of a sleep mode scheme can also be traced in the way it distributes users

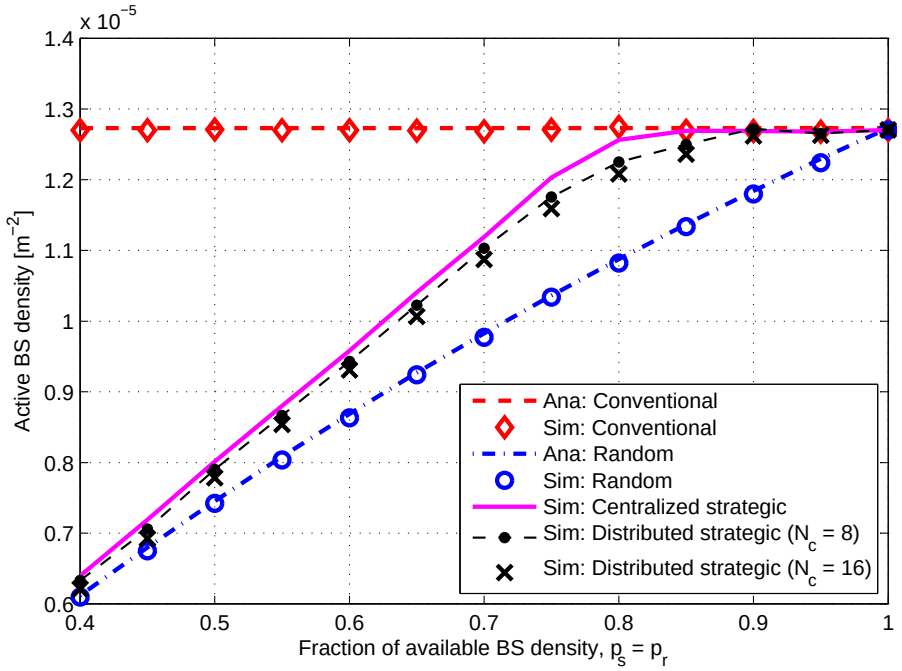


Figure 5.1: Density of active BSs for the various sleep mode schemes.

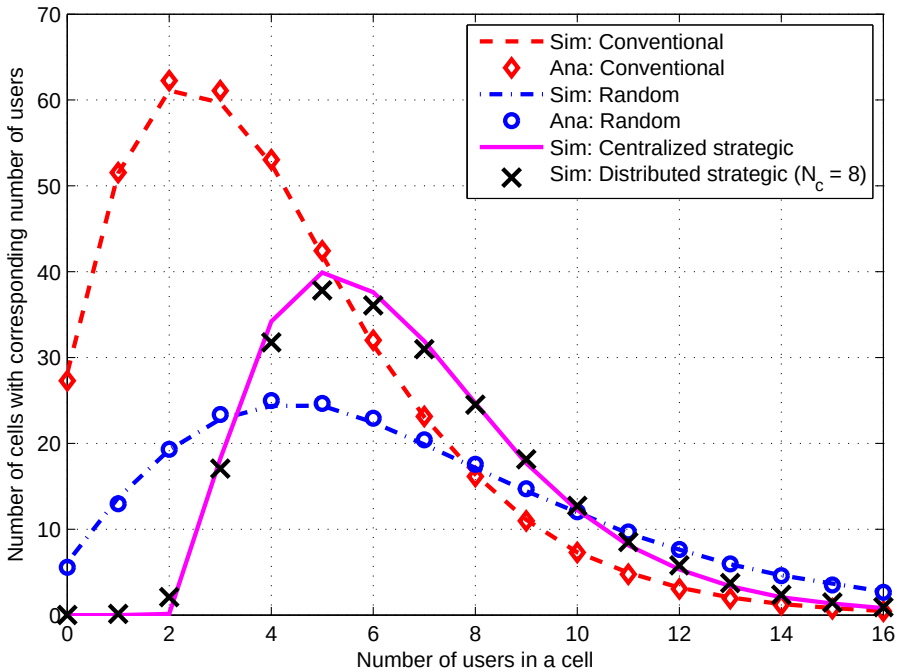


Figure 5.2: The total number of BSs with corresponding number of users, where $\lambda_u = 4\lambda_b$ and $p_r = p_s = 0.6$.

between remaining BSs. To illustrate, consider a network in which $p_s < p_a$. Fig. 5.2 shows the sum total of BSs with a corresponding number of users after cell association. Although the BS density is significantly lower than user density ($\lambda_u = 4\lambda_b$), many BSs still remain idle (about 35 BSs in this case) and the majority of active BSs generally have very few users. Conventional sleep mode may lead to underutilization of bandwidth resources in these cells especially where users have low traffic requirements. Both random and the strategic schemes reduce the available BS density and redistribute users among remaining BSs. After user re-association, random sleep mode still has some idle BSs (about six BSs) but the strategic schemes ensure that all available BSs remain active which enhances average sum rate. To understand the advantage of the strategic schemes, note that on the left hand side of Fig. 5.2, both conventional and random

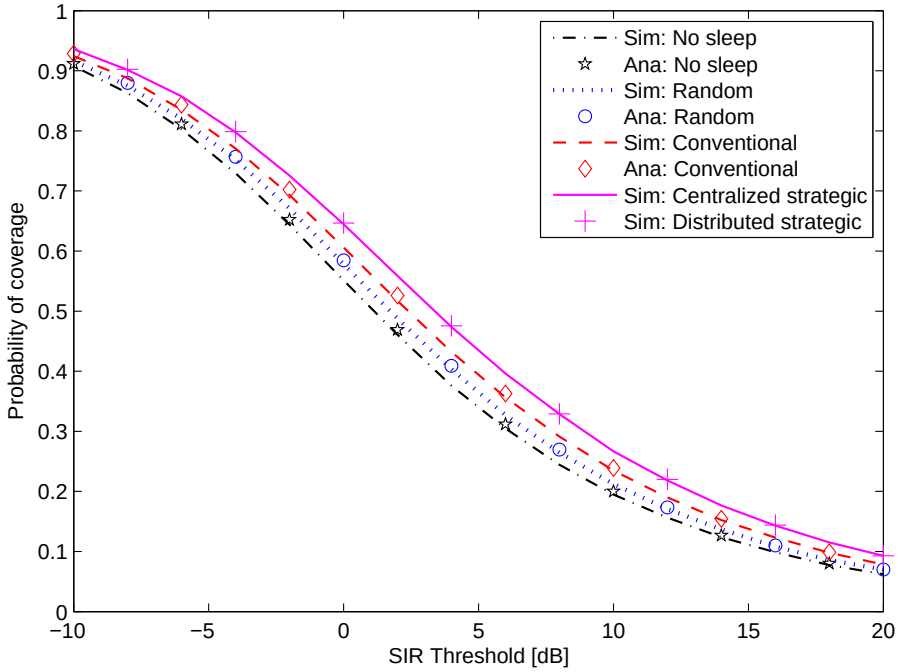


Figure 5.3: Coverage probability of the interference-limited homogeneous network under the various sleep mode schemes.

sleep modes are more likely to have idle BSs which consume some energy even in sleep mode. Furthermore on the right hand side, random sleep mode is more likely to cause congestion than the strategic schemes. Hence the strategic schemes ensure a fairer distribution of bandwidth resources among users compared to conventional and random schemes.

Sleep mode mechanisms generally improve the SINR coverage probability of a typical user as shown in Fig. 5.3. The 'no sleep' scenario is a special case which assumes that $\lambda_u \gg \lambda_b$ such that $p_a \approx 1$ i.e. all BSs always transmit which maximizes the aggregate interference and gives the lower bound on coverage probability [106]. Centralized strategic sleep mode guarantees the best coverage probability because of two reasons: (i) it minimizes aggregate interference compared to conventional sleep mode; and (ii) it minimizes the effect of sleep mode on average received signal strength compared to random sleep mode. Note however that its performance generally depends on the choice of p_s. For example it has the same performance as conventional sleep mode when $p_s \geq p_a$. Coverage probability with distributed strategic scheme is very close to that of its centralized counterpart due to their comparable active BS densities. Although conventional sleep mode puts all idle BSs to sleep and maintains the received signal of every connected user, its aggregate interference is still higher compared to the strategic schemes. This reduces the average SINR and consequently gives a lower coverage probability in comparison. Although random sleep mode has the lowest aggregate interference, it has the worst SINR coverage because of its significant negative effect on the average received signal of users.

The benefit of sleep mode is to enhance network EE but this is normally achieved at the expense of other performance measures. The EE of a given scheme depends on its ability to maximize average sum rate from the remaining active BSs. In this regard, the strategic schemes maximize average rate per active BS compared to conventional and random schemes as shown in Fig. 5.4. Hence the strategic schemes give a better network EE compared to conventional and random sleep modes as shown in Fig. 5.5. In the range $p_s \geq p_a$ where conventional and both strategic

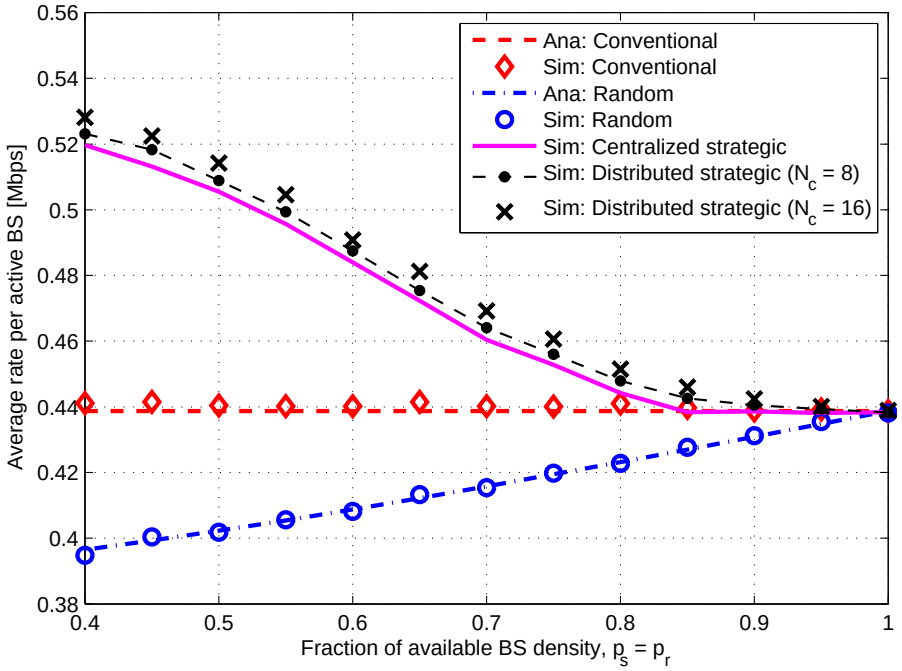


Figure 5.4: Average rate per active BS for different sleep mode schemes.

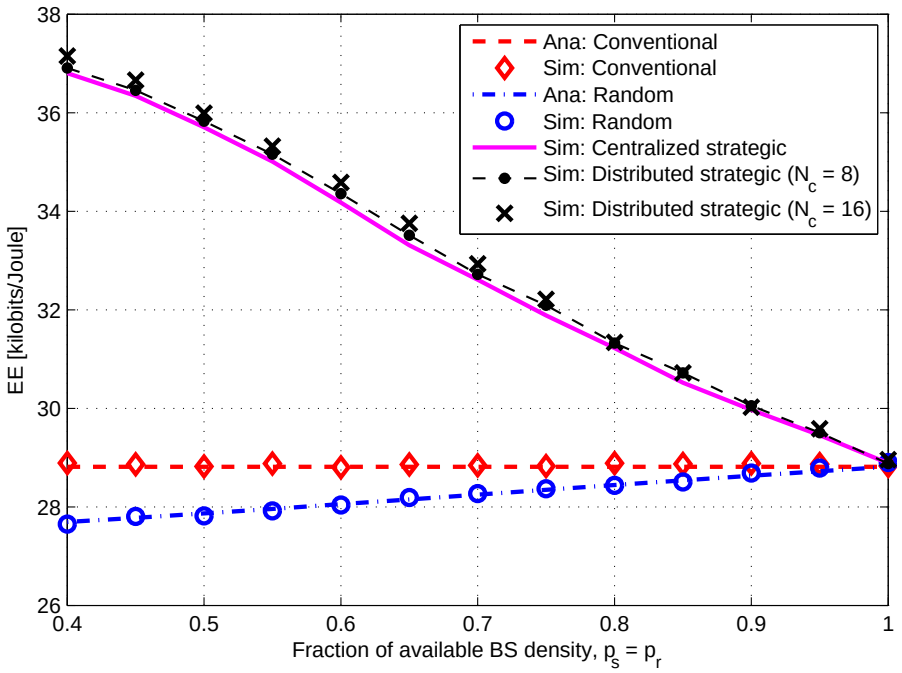


Figure 5.5: Network EE performance of different sleep mode schemes.

schemes are identical, the strategic schemes still have better EE because they put some BSs in deep sleep mode (compare (5.2) and (5.10)). Random sleep mode has the worst EE performance due to its significant effect on SINR which results in low average sum rate.

Sleep mode schemes have a big impact on the realizable average rate performance of the network as shown in Figs. 5.6-5.7. Conventional sleep mode has constant average user rate and average sum rate because it maintains the same average active BS density of $p_a\lambda_b$. In addition, it defines the upper bound on both measures because of its superior active BS density that maximizes both the connected users and the bandwidth per connected user (see Fig. 5.1). With random sleep mode however, both average user and average sum rate increase proportionally with p_r until they reach their respective upper bound levels at $p_r = 1$. In general, random sleep mode

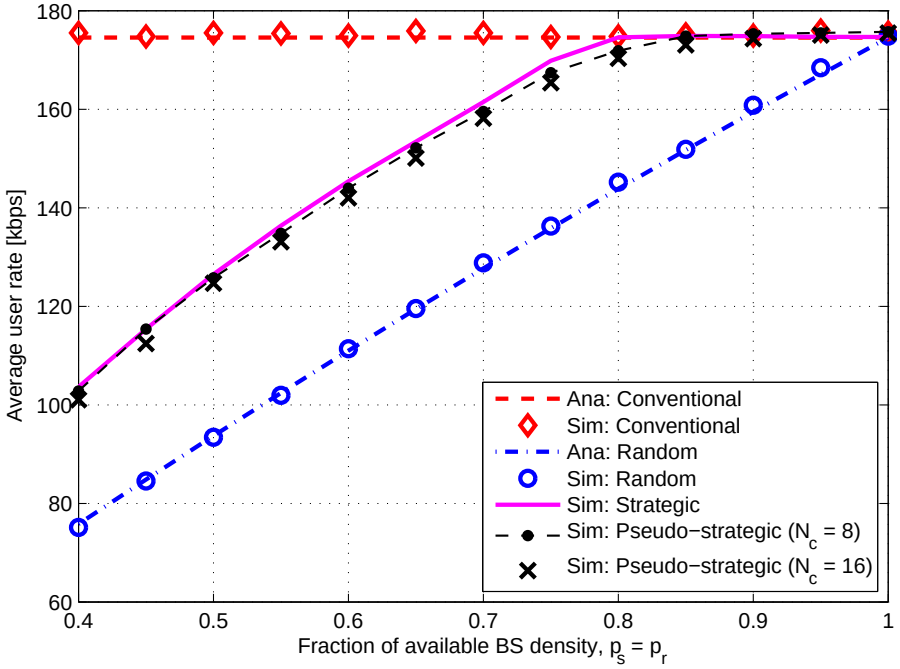


Figure 5.6: Average user rates for the different sleep mode schemes.

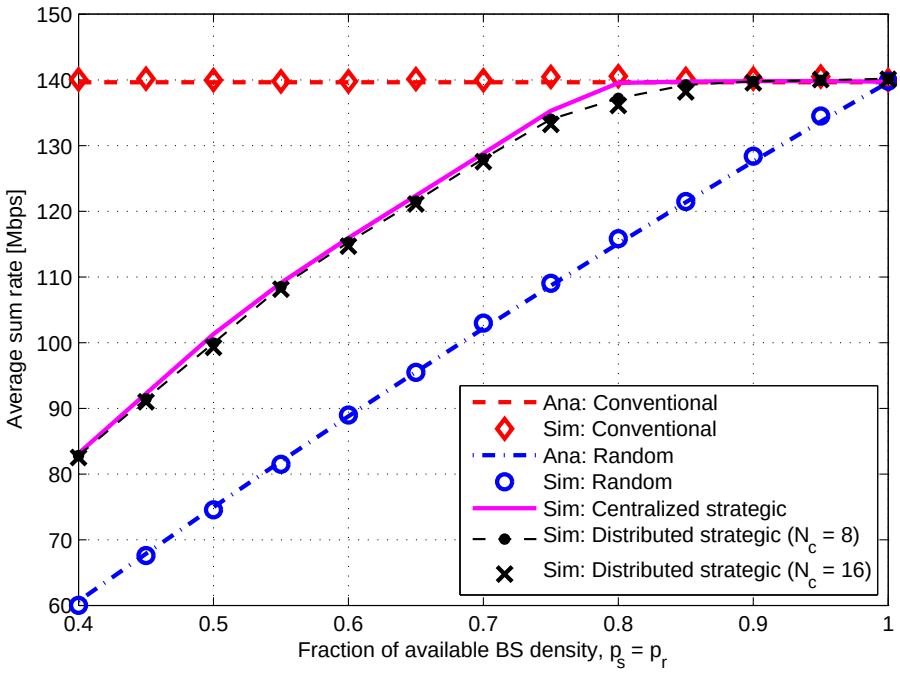


Figure 5.7: Average sum rates with the different sleep mode schemes.

has the worst average rate performance because of its random selection of sleep mode BSs. In contrast, centralized strategic sleep mode optimizes the selection process of sleep mode BSs to enhance average rate performance and as a result, both average user and sum rate performances approach their upper bounds much faster than random sleep mode. In the $p_s \geq p_a$ range, average user and sum rates of centralized strategic scheme are equivalent to their respective upper bound levels because the scheme is identical to the conventional scheme in this range. Furthermore, centralized strategic scheme gives slightly better average user rate and average sum rate performances than its distributed counterpart due to its superior active BS density. Hence although conventional sleep mode maximizes average sum rate, it does so at the expense of significantly more energy consumption which affects EE as shown in Fig. 5.5.

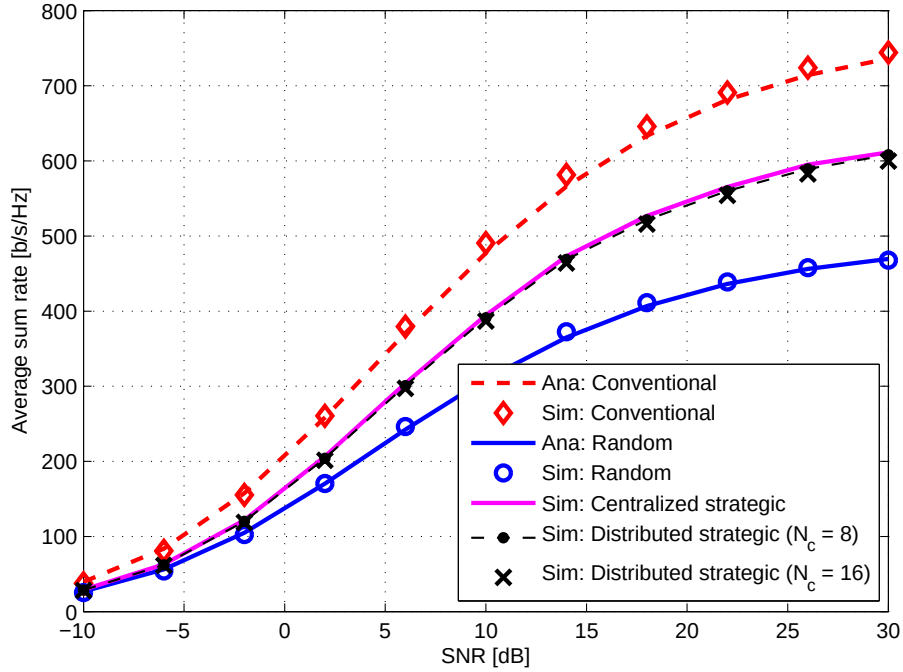


Figure 5.8: Average sum rate versus SNR for the sleep mode schemes ($p_s = p_r = 0.6$ and $\sigma^2 = 0.01\,\mathrm{W}$).

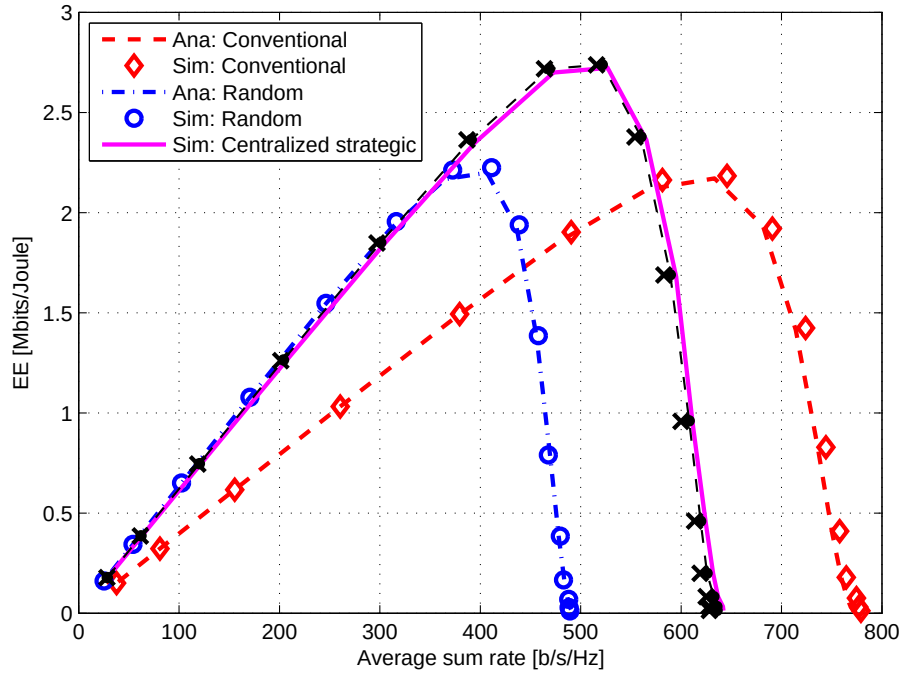


Figure 5.9: EE versus average sum rate for the sleep mode schemes ($p_s = p_r = 0.6$).

Fig. 5.8 illustrates the variation of sum SE of the network with SNR for different sleep mode schemes. At very low SNR where the network is noise-limited, sum SE is very low and comparable for all sleep mode schemes. As SNR increases, sum SE increases accordingly and differences in sleep mode performance become more apparent. In general, conventional sleep mode has the best sum SE performance over the whole SNR range due it its superior active BS density. The strategic schemes also outperform random sleep mode due to their ability to maintain a considerably better active BS density. At very high SNR where the network is interference-limited, sum SE saturates at different levels for each sleep mode scheme.

The tradeoff analysis of network EE and sum SE, illustrated in Fig. 5.9, shows that EE generally increases and then drops as sum SE increases towards its saturation level. Maximum EE

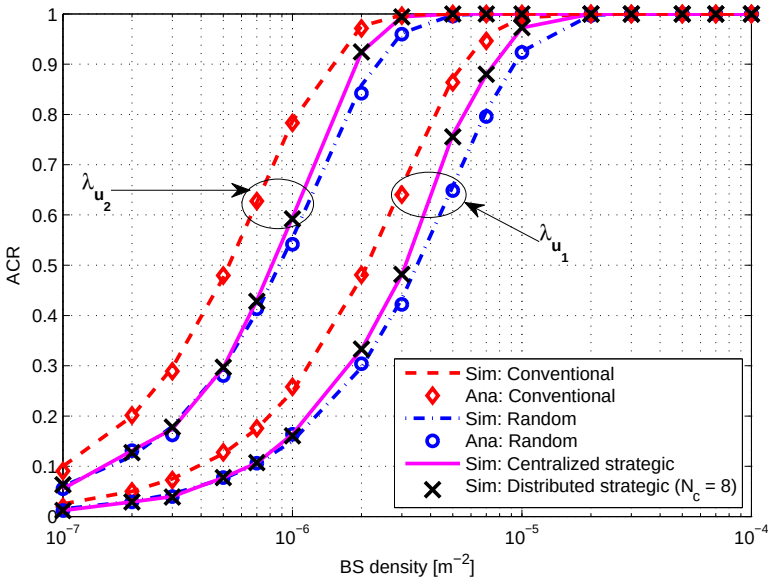


Figure 5.10: ACR verses BS density for the sleep mode schemes, where $p_r = p_s = 0.6$, $\mathcal{B} = 10\,\text{MHz}$, $\lambda_{u_1} = 5 \times 10^{-5}\,\text{m}^{-2}$ and $\lambda_{u_2} = 2 \times 10^{-4}\,\text{m}^{-2}$.

occurs at different sum SE values for different sleep mode schemes. With random sleep mode, maximum EE is achieved at the lowest sum SE in comparison to the other schemes. In contrast, conventional sleep mode achieves maximum EE at the highest sum SE of all the schemes. The improved EE-SE tradeoff performance of random sleep mode over conventional sleep mode is a consequence of the significantly more aggregate fixed (or circuit) power consumption in a network with conventional sleep mode. In comparison, both strategic schemes guarantee the highest maximum EE performance although this is achieved at a lower sum SE than conventional scheme. Distributed strategic scheme fairly matches its centralized counterpart. Hence the strategic schemes give useful flexibility between conventional and random schemes towards achieving a very high EE at a relatively good sum SE.

Another consequence of sleep mode is how it affects user connectivity in the network. ACR mainly depends on the prevailing BS-user density ratio υ but it is also influenced by sleep mode schemes as shown in Fig. 5.10. Generally for a given BS density, ACR reduces as the user density increases since a typical user is more likely to remain unconnected. For a given user density λ_u, ACR increases with the BS density since the average number of users per cell reduces. Conventional sleep mode defines the upper bound of ACR performance because it guarantees the highest density of active BSs. At very low BS density where $\lambda_b \ll \lambda_u$, most BSs cover many more users than they can provide a connection, resulting into a low ACR for all sleep mode schemes. In this range, both strategic schemes have no tangible gain over random scheme because BSs with many associated users are forced to sleep which affects many users. In general, sleep mode should not be implemented in this υ-regime because although energy saving can be achieved, congestion is too severe to achieve good QoS. However as BS density increases and average cell size reduces, the average number of users per cell reduces and ACR generally increases in all schemes. In addition, some cells become idle or cover very few users and the strategic schemes begin to outperform the random scheme by prioritizing such BSs for sleep mode. As the BS density increases further, eventually all BSs contain few users who all get connected irrespective of the sleep mode scheme. Beyond this point, additional BS deployment does not yield any ACR gain although it may improve other measures.

To investigate the effect of *varying user density*, consider a dense PPP-based homogeneous network of small BSs with the default parameters shown in Table 5.2, unless otherwise stated.

Table 5.2: Parameters used to obtain results

Parameter	Value	Parameter	Value
Very low υ: λ_u/m^2	4×10^{-4}	$\mathcal{N}_{tx}$, $P_{0,b}$, P_{sl}, Δ_b	1, 6.8 W, 1.5 W, 4.0
Mid-level υ: λ_u/m^2	4×10^{-4}	Network area A	$2.5\,\mathrm{km} \times 2.5\,\mathrm{km}$
Mid-level υ: λ_b/m^2	$\{1, 16\} \times 10^{-4}$	P_t	0.13 W
Very high υ: λ_u/m^2	$\{4, 40\} \times 10^{-6}$	α	4
Very high υ: λ_b/m^2	$\{4, 5\} \times 10^{-4}$	σ^2	0

In the mid-level υ-regime, consider a network in which the BS density is up to four times less than or greater than the user density i.e. $0.25 \leq \upsilon \leq 4.0$. Fig. 5.11 verifies the accuracy of the approximation of the average user rate as a linear function of υ. It also shows the exact analytical results verified by Monte Carlo simulation. For conventional sleep mode, the line constants are $c_1 = 0.6274$ and $c_2 = 2.2243$ over this υ-range. Similarly, the constants in the random sleep mode case are $r_1 = 0.7445$ and $r_2 = 2.0912$.

The difference in average user rate performance between conventional and random sleep modes increases with υ (or as the user density reduces). When the υ-ratio is sufficiently small, all BSs remain active and the average user rate is the same with both schemes. As υ increases, the average number of idle BSs increases which thins out the aggregate interference. On one hand, conventional sleep mode ensures that users remain connected to their parent BSs which maintains their received signal and enhances their SINR level. On the other hand, random sleep mode increases the average transmitter-receiver separation distance which reduces the average received signal strength and gives a lower SINR level compared to conventional sleep mode scheme. This difference in SINR level between conventional and random sleep modes increases with the υ-ratio.

Fig. 5.12 shows the approximated average network sum rate, verified by both its exact analytical and Monte Carlo simulation over the mid-level υ-range. Conventional sleep mode gives a higher average sum rate since it connects more users and has a better average rate per user. In the PPP model, it is assumed that each active BS randomly selects and connects a single user from all the users within its coverage area. If each BS were to prioritize its highest SINR user, the average sum rate would be enhanced as shown in the *special case* characteristic. Generally for a given user density, network densification enhances the average sum rate since the average user rate increases and more users get connected to the network.

The mid-level regime shows an optimal υ-ratio, denoted as $\upsilon^\star$, at which network EE is maximized as shown in Fig. 5.13. Even though random sleep mode reduces the APC of the network, this is achieved by significantly sacrificing the average sum rate. Therefore, conventional sleep mode always gives a better average EE than random sleep mode. Prioritizing channel allocation to the highest SINR users significantly enhances the average EE performance as shown in the *special case* characteristic.

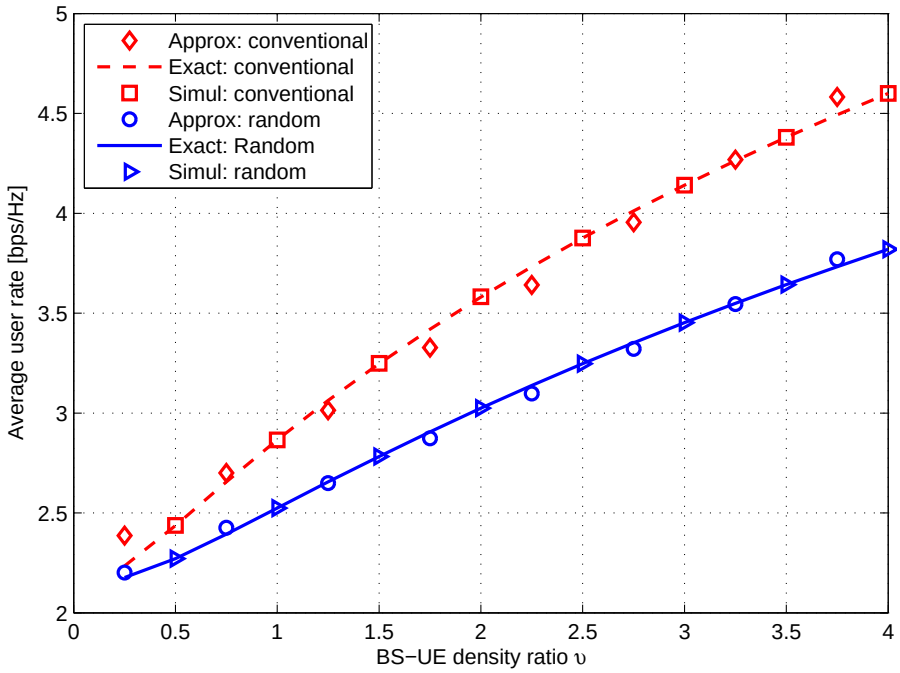


Figure 5.11: Average user rate in the mid-level υ-regime (for $p_r = 0.6$).

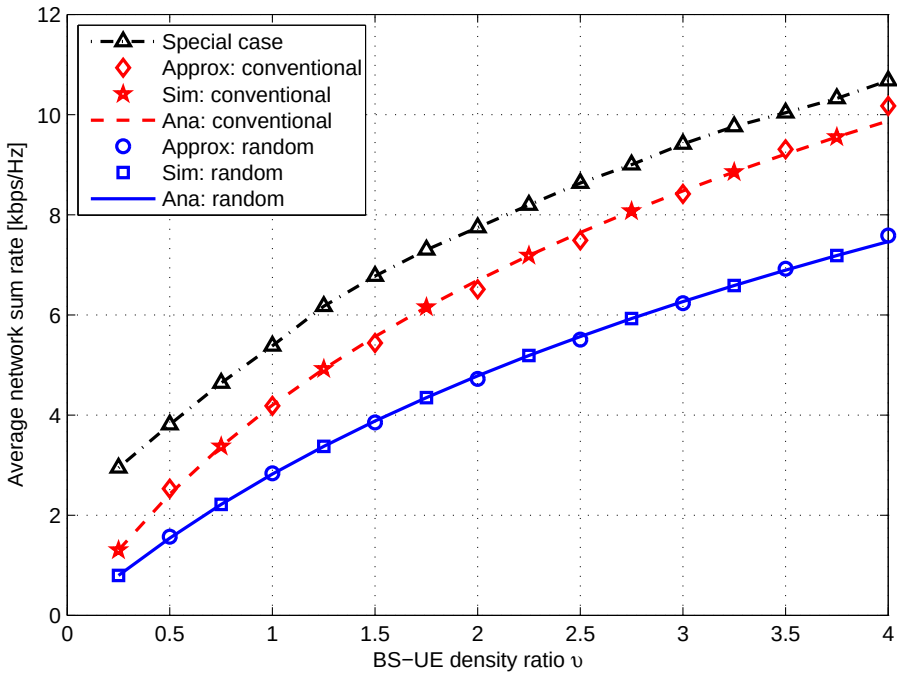


Figure 5.12: Variation of average network sum rate with υ in the '$\lambda_b \sim \lambda_u$' regime.

The power consumption of BSs in sleep mode has a significant impact on the average EE especially in dense networks where numerous BSs may be in sleep mode simultaneously. When P_{sl} is very small or negligible, average EE increases monotonically over this υ-regime. Conversely, average EE decreases monotonically when P_{sl} is higher than a certain level. Therefore, two P_{sl} threshold values exist between which average EE has a maximum value at a given υ value. Denote the lower and upper thresholds as P_{sl_1} and P_{sl_2} respectively. If $P_{sl}^{\star}$ is the actual value at which average EE is maximized, denoted as $\text{EE}_{\max}$, then $P_{sl_1} \leq P_{sl}^{\star} \leq P_{sl_2}$. $\text{EE}_{\max}$ occurs at different υ-ratios (denoted $\xi^{\star}$) as P_{sl} varies. The $\text{EE}_{\max}$ values and the corresponding $\xi^{\star}$ are determined by analyzing the derivatives of average EE, as shown in (5.16)-(5.19).

Fig. 5.14 shows $\text{EE}_{\max}$ and the corresponding optimal υ-ratios ($\xi^{\star}$) at which it occurs for both

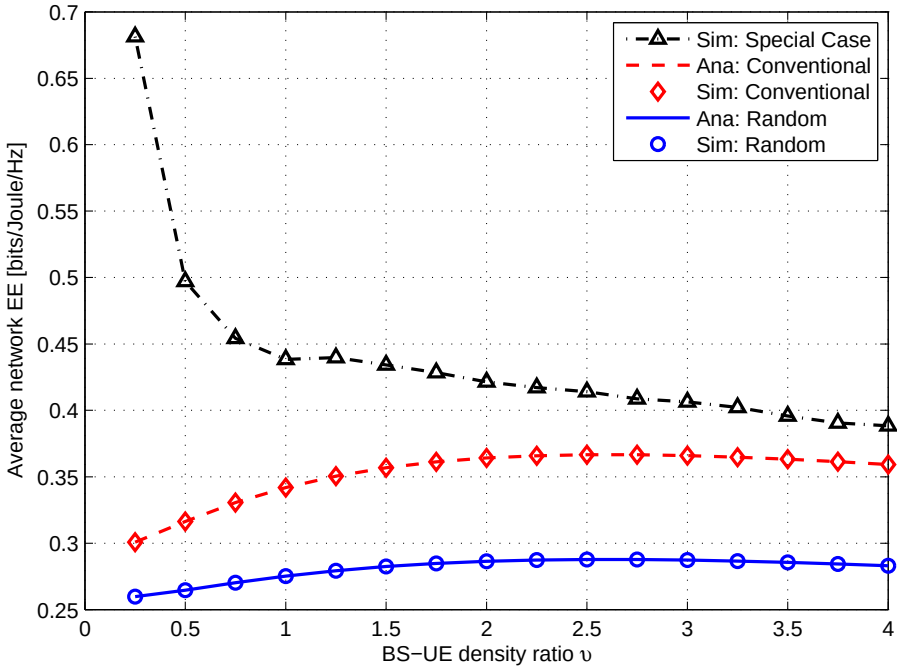


Figure 5.13: Variation of average EE with υ in the '$\lambda_b \sim \lambda_u$' regime.

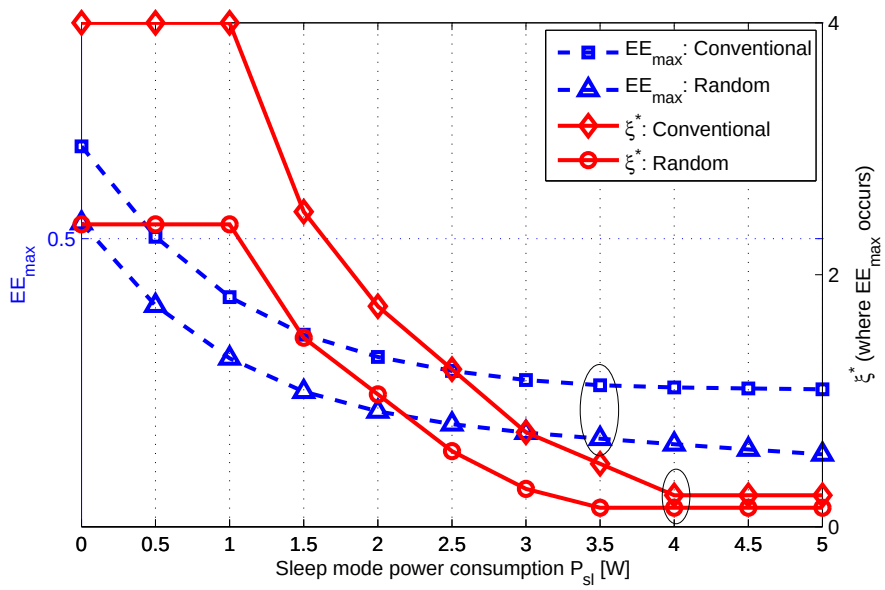


Figure 5.14: Variation of $\text{EE}_{\max}$ and $\upsilon^\star$ with sleep mode power consumption in the '$\lambda_b \sim \lambda_u$' regime ($p_r = 0.6$). Dashed lines represent $\text{EE}_{\max}$ (left y-axis) and solid lines represent $\xi^\star$ (right y-axis).

conventional and random sleep modes. As expected, a lower P_{sl} guarantees a higher $\text{EE}_{\max}$ in both sleep mode schemes. With conventional sleep mode, average EE increases monotonically when $P_{sl} < 1.0\,\text{W}$ such that $\text{EE}_{\max}$ always occurs at $\upsilon = 4.0$ (hence, the curve is flat in this range). Similarly, average EE decreases monotonically when $P_{sl} > 3.8\,\text{W}$ such that $\text{EE}_{\max}$ always occurs when $\upsilon = 0.25$ (the curve is flat in this range as well). With random sleep mode, these thresholds are 1.0 W and 3.5 W respectively but the general behavior is consistent.

Furthermore, a lower P_{sl} value ensures that the average EE is maximized at a higher υ-ratio with both sleep modes. This is desirable because it allows more BSs to remain active which increases the average bandwidth per user, reduces the probability of congestion, reduces the risk of coverage holes and improves the average sum rate of the network. Therefore it is possible to know the optimal $\upsilon^\star$-ratio at which average EE is maximized if the sleep mode

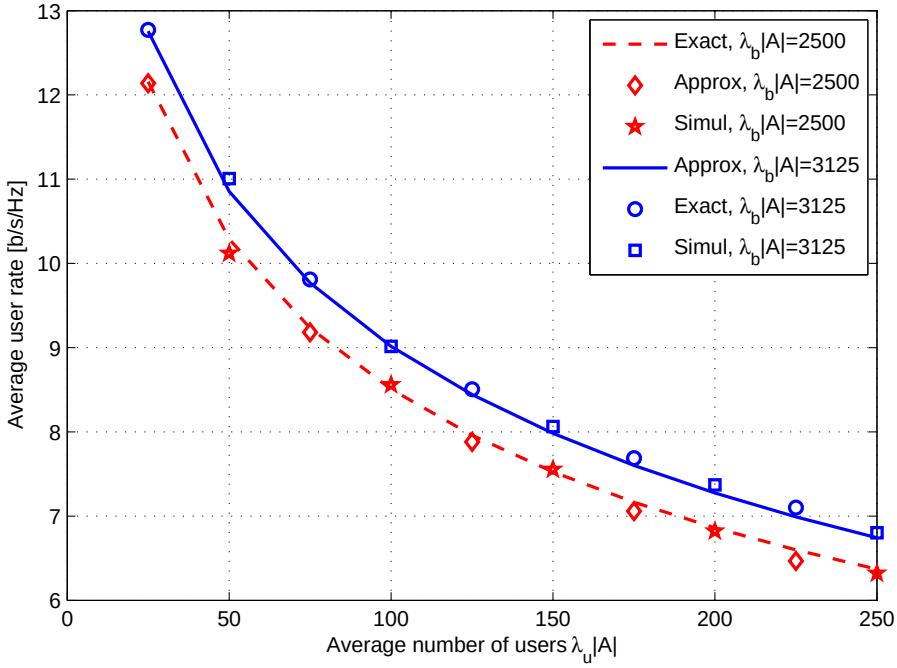


Figure 5.15: Average user rate in 'very high υ-regime'

power consumption of deployed BSs is known. To maintain a desired $\mathrm{EE}_{\max}$, the BS density can be varied as the user density changes.

In the 'very high υ-regime', consider a network with a very small user density (or very large BS density) such that $10 \leq \upsilon \leq 100$. As a result, many BSs are idle which thins the interference and gives a large average user rate. Fig. 5.15 shows the variation of the average user rate with user density for two different BS densities. The approximation of the average user rate is verified by both its exact and Monte Carlo simulations. As user density increases, more BSs become active which increases the interference and reduces average user rate.

Fig. 5.16 shows that average network sum rate increases almost linearly with user density which verifies the analysis shown in (5.24). Although the average user rate reduces with increasing user density $\lambda_u|A|$, more users actually get connected which increases the average sum rate. In this regime, large gains in the sum rate are possible even when the user density increases marginally. However, further BS densification of the network only gives marginal gains in the average sum rate since most of them remain idle. Note that due to the large number of idle BSs, the sleep mode power consumption P_{sl} severely impacts average network EE performance. Therefore, deep sleep strategies to completely switch off BSs should be devised.

In the 'very low υ-regime, consider the υ-range $0.02 \leq \upsilon \leq 0.2$. Since all BSs remain active, average user rate is approximately equal to 2.15 b/s/Hz. Fig. 5.17 verifies the approximation of average sum rate using both its exact and simulation results. The average network sum rate varies linearly with υ-ratio which verifies the analysis shown in (5.31). Therefore, further BS densification of the network results into a linear increase in the average sum rate. In this regime, sleep mode power consumption is not of major concern since the number of BSs in sleep mode is negligible. However, such a network is characterized by low user connectivity (high congestion) and low average sum rate.

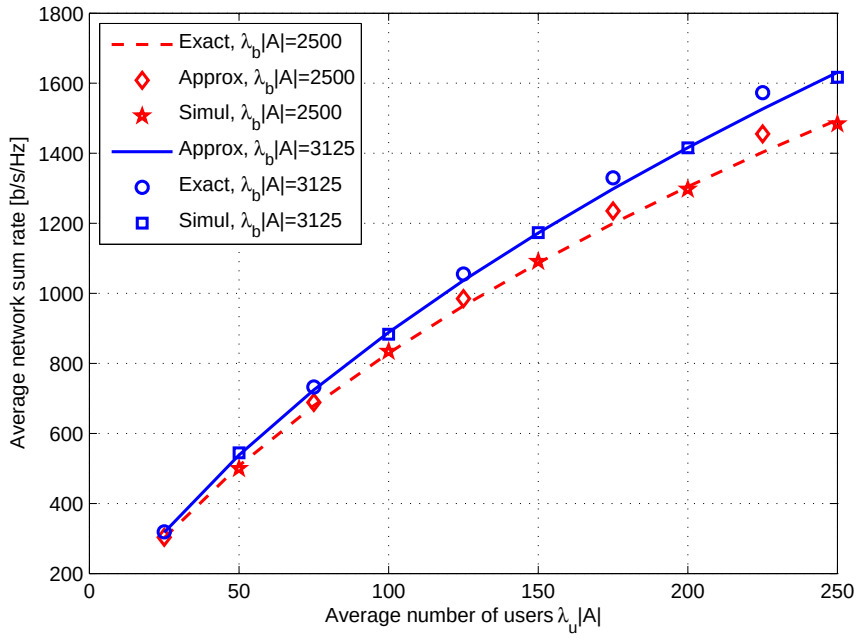


Figure 5.16: Average network sum rate in the 'very high υ-regime'.

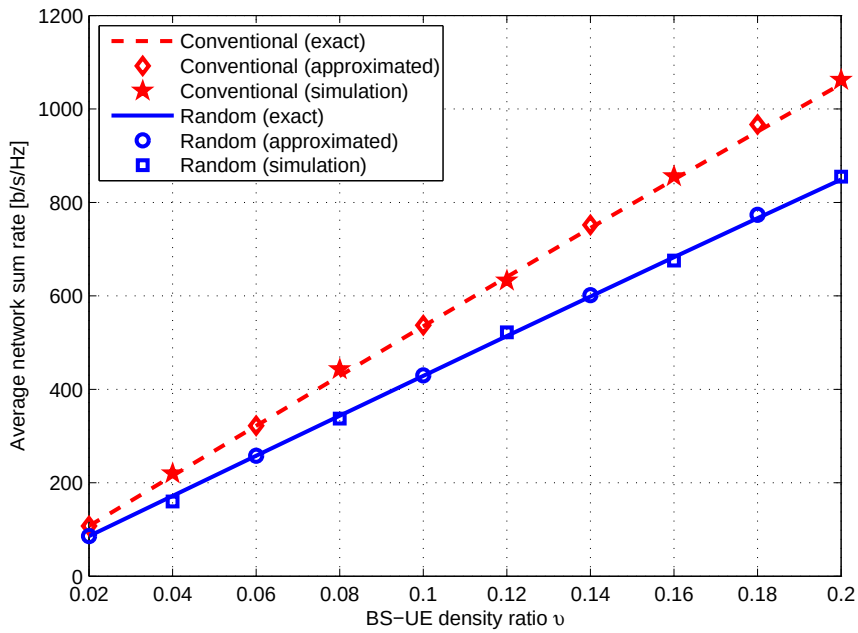


Figure 5.17: Average network sum rate in the 'very low υ-regime'.

5.8 Discussion Points

The approximations of coverage probability and average rate under different υ-ratios have allowed interesting insights to be made on how variable user density affects network performance. In some cases, the performance measures are expressed in closed-form while others only require the evaluation of a simple integral. The following key insights can be drawn from the analysis.

(i) In the mid-level υ-regime where $\lambda_b \sim \lambda_u$, there is an optimal υ-ratio at which the average EE performance is maximized. Therefore, if changes in user density are known, the BS density can dynamically change to maintain the optimal υ-ratio.

(ii) The power consumption of BSs in sleep mode has a major impact on the average EE especially in dense networks where numerous BSs may be in sleep mode during low traffic

periods. If the sleep mode power consumption of BSs is low, more BSs can be deployed to enhance connectivity, coverage probability and average rate performance without worsening average EE. Therefore, it is imperative that equipment manufacturers design equipment that consume minimal power in sleep mode. In addition, deep sleep mechanisms which completely switch off BSs are beneficial.

(iii) Conventional sleep mode maximizes the density of active BSs and consequently the average sum rate of the network. Random sleep mode has no advantage over conventional sleep mode in all three v-regimes. However, strategic selection of candidate BSs for sleep mode enhances the probability of coverage and average network EE. Distributing the strategic algorithm in clusters across the network gives a good fit of the performance of its optimal centralized counterpart.

5.9 Summary

This chapter presented sleep mode mechanisms that can help to adapt the energy consumption of homogeneous networks to the prevailing traffic or user intensity. Both centralized and distributed strategic algorithms improve the coverage, rate and energy performance of the network. If the prevailing user density is low compared to the BS density, most BSs remain idle and the network only needs to retain only enough active BSs to provide seamless coverage. However, if the user density is much larger than the BS density, all BSs remain active to meet the high traffic demand.

Chapter 6

HetNet Deployment Optimization

6.1 Introduction

In the deployment of multi-tier HetNets, the choice of BS densities and transmit powers per tier is an interesting problem that affects the SE and EE performance of the HetNet. Some researchers have attempted to predict the optimal deployment configuration of HetNets subject to appropriate performance measures. In [135], a two-tier unbiased HetNet is investigated to determine the optimal densities of macro BSs and small BSs that maximize its EE subject to a blocking probability constraint. In [115], an energy cost minimization framework is formulated to determine the optimal combination of BS densities in a two-tier unbiased HetNet subject to a service outage constraint. This work defines the circumstances under which it is preferable to densify the existing HetNet with either a macro BS or a micro BS. However this work does not optimize the associated transmit power in each tier. In addition, it assumes that the HetNet is unbiased yet biasing is very important for load balancing in HetNets.

The authors in [117] implement an APC minimization framework to determine the optimal BS densities and their associated transmit powers in a two-tier unbiased HetNet subject to a coverage probability constraint. The authors define an optimal network *deployment factor* $\mathcal{H}^\star$ which is expressed as $\mathcal{H}^\star = \lambda_M P_M^{2/\alpha} + \lambda_m P_m^{2/\alpha}$ where λ_M and λ_m are the macro BS and micro BS densities while P_M and P_m are the respective macro BS and micro BS transmit powers. Using $\mathcal{H}^\star$, they analyze the performance scenarios under which the network maybe densified using either a macro BS or micro BS. However, this work also ignores biasing and considers a single performance constraint. Moreover, the deployment factor is not expressed in closed form and must be determined numerically, for example using the bisection method [127].

In this chapter, a closed form expression for the deployment factor of a general K-tier unbiased HetNet using different user association mechanisms is derived. In addition, the deployment factor of a two-tier biased HetNet is determined to investigate the effect of biasing on coverage probability, average rate and energy performance of HetNets. Both coverage probability and average user rate constraints are applied in the optimization of the HetNet to ensure that both performance targets are achieved in all network environments.

It is important to investigate how coverage probability and average user rate vary with changes in BS density and transmit power per tier. This behaviour will be used later to optimize the HetNet deployment configuration subject to both performance measures.

Lemma 6.1. *In a mobile environment where $\alpha > 2$ and $\sigma^2 > 0$, both coverage probability and average user rate increase monotonically with the BS density per tier.*

Proof. This is proved using the maximum ABRP connectivity-based expression of the coverage probability. The proof can easily be extended to the other association schemes.

Assume that a typical user is associated with the k-th tier and consider two k-th tier BS densities λ_{k_1} and λ_{k_2}, where $\lambda_{k_2} > \lambda_{k_1}$. Using (3.25), coverage probability corresponding to the BS density λ_{k_2} is expressed as

$$\mathcal{P}_{c_P}(\lambda_{k_2}) = \sum_{k=1}^{K} \pi \lambda_{k_2} \int_0^\infty e^{-a_1 z^{\alpha/2}} e^{-\pi(\lambda_{k_2}\mathcal{C}+a_2)z} \, \mathrm{d}z \tag{6.1}$$

where $a_1 = \frac{T\sigma^2}{P_k L}$ and $a_2 = \mathcal{C} \sum_{j=1, j\neq k}^{K} \lambda_j \widehat{P}_j^{2/\alpha}$ where $\mathcal{C} = 1 + \mathcal{Z}(T, \alpha, 1)$ is a constant. Substituting $z = (\lambda_{k_1}/\lambda_{k_2})x$ into (6.1) gives

$$\begin{aligned}
\mathcal{P}_{c_P}(\lambda_{k_2}) &= \sum_{k=1}^{K} \pi \lambda_{k_1} \int_0^\infty e^{-a_1 \left(\frac{\lambda_{k_1}}{\lambda_{k_2}}\right)^{\alpha/2} x^{\alpha/2}} \times e^{-\pi\left(\lambda_{k_1}\mathcal{C}+a_2\frac{\lambda_{k_1}}{\lambda_{k_2}}\right)x} \mathrm{d}x \\
&\overset{(a)}{>} \sum_{k=1}^{K} \pi \lambda_{k_1} \int_0^\infty e^{-a_1 x^{\alpha/2}} e^{-\pi(\lambda_{k_1}\mathcal{C}+a_2)x} \mathrm{d}x \\
&= \mathcal{P}_{c_P}(\lambda_{k_1})
\end{aligned}$$

where (a) follows because $a_1 > 0$, $a_2 > 0$, $\mathcal{C} > 0$, $\alpha > 2$ and $(\lambda_{k_1}/\lambda_{k_2}) < 1$. Therefore, coverage probability always increases if the BS density is increased from λ_{k_1} to λ_{k_2}. Note that $a_2 > 0$ in biased HetNets and therefore the proof holds. Monotonicity of average user rate with BS density is also proved in the same way. ∎

Lemma 6.2. *In a mobile environment where $\alpha > 2$ and $\sigma^2 > 0$, both coverage probability and average user rate increase monotonically with the BS transmit power per tier.*

Proof. This is proved using the maximum ABRP connectivity-based expression of the average user rate. The proof can easily be extended to the other association schemes.

Consider two transmit power values P_{k_1} and P_{k_2}, where $P_{k_2} > P_{k_1}$. Using (3.28), the average user rate corresponding to P_{k_2} is expressed as

$$\mathcal{R}_{u_P}(P_{k_2}) = \sum_{k=1}^{K} \Xi \left[\pi \lambda_k \, e^{-(b_1/P_{k_2})z^{\alpha/2}} e^{-(b_2/P_{k_2}^{2/\alpha})z} \right] \tag{6.2}$$

where $b_1 = \frac{\sigma^2}{L}(2^t - 1)$ and $b_2 = \pi \sum_{j=1}^{K} \lambda_j P_j^{2/\alpha} \mathcal{D}_j(t)$. Using P_{k_1}, the average user rate becomes

$$\mathcal{R}_{u_P}(P_{k_1}) = \sum_{k=1}^{K} \Xi \left[\pi \lambda_k \, e^{-(b_1/P_{k_1})z^{\alpha/2}} e^{-(b_2/P_{k_1}^{2/\alpha})z} \right] \tag{6.3}$$

Since $b_1 > 0$, $b_2 > 0$, $\alpha > 2$ and $P_{k_2} > P_{k_1}$, both respective exponential terms in (6.2) are larger than those in (6.3). Hence $\mathcal{R}_{u_P}(P_{k_2}) > \mathcal{R}_{u_P}(P_{k_1})$ which proves Lemma 6.2. The monotonicity of coverage probability with transmit power is proved in the same way. ∎

Lemmas 6.1-6.2 are used later to determine the optimal BS density and transmit power per tier that minimize the APC of the HetNet subject to coverage and average user rate constraints.

6.2 Minimum BTD Connectivity

Without loss of generality, assume that a typical user is located at the origin and associates with a k-th tier BS located a distance r_k away.

Lemma 6.3. *The probability that the typical user in a HetNet with minimum BTD connectivity is associated to the k-th tier is expressed as*

$$\mathcal{A}_k = \frac{\lambda_k}{\sum_{j=1}^{K} \lambda_j \left(\frac{\nu_k}{\nu_j}\right)^2}. \tag{6.4}$$

Proof. Denote the tier of association as tier m such that $\mathcal{A}_k = \mathbb{P}[m = k]$. Then,

$$\begin{aligned}
\mathcal{A}_k = \mathbb{P}[m = k] &= \mathbb{E}_{r_k}\left[\mathbb{P}\left[\nu_k r_k < \nu_j r_j, \forall j \in K, j \neq k\right]\right] \\
&= \mathbb{E}_{r_k}\left[\prod_{j=1, j\neq k}^{K} \mathbb{P}\left[r_j > \frac{\nu_k}{\nu_j} r_k\right]\right] \\
&= \int_0^\infty \prod_{j=1, j\neq k}^{K} \mathbb{P}\left[r_j > \frac{\nu_k}{\nu_j} r\right] f_{r_k}(r)\, \mathrm{d}r.
\end{aligned} \tag{6.5}$$

Similar to (3.10), $f_{r_k}(r)$ is expressed as

$$f_{r_k}(r) = 2\pi\lambda_k r\, e^{-\pi\lambda_k r^2}. \tag{6.6}$$

In addition, the product term in (6.5) is evaluated as

$$\begin{aligned}
\prod_{j=1, j\neq k}^{K} \mathbb{P}\left[r_j > \frac{\nu_k}{\nu_j} r\right] &= \mathbb{P}\left[\text{No } j\text{-th tier BS closer than } \frac{\nu_k}{\nu_j} r\right] \\
&= \prod_{j=1, j\neq k}^{K} e^{-\pi\lambda_j \left(\frac{\nu_k}{\nu_j} r\right)^2} = e^{-\pi \sum\limits_{j=1, j\neq k}^{K} \lambda_j \left(\frac{\nu_k}{\nu_j} r\right)^2}.
\end{aligned} \tag{6.7}$$

Substituting (6.6) and (6.7) into (6.5) gives

$$\begin{aligned}
\mathcal{A}_k &= 2\pi\lambda_k \int_0^\infty r\, e^{-\pi\lambda_k r^2} e^{-\pi \sum_{j=1, j\neq k}^{K} \lambda_j (\nu_k/\nu_j)^2 r^2}\, \mathrm{d}r \\
&= 2\pi\lambda_k \int_0^\infty r\, e^{-\pi \sum_{j=1}^{K} \lambda_j (\nu_k/\nu_j)^2 r^2}\, \mathrm{d}r.
\end{aligned} \tag{6.8}$$

Evaluating the integral in (6.8) by substituting $y = r^2$ gives the result. ∎

According to Lemma 6.3, more users connect to a tier with a higher BS density and a smaller bias value. This is intuitive because a smaller bias value makes BSs of that tier appear to be closer than those of a tier with a larger bias value.

Lemma 6.4. *The distance between a typical user and its serving k-th tier BS, denoted as X_k, is a random variable whose pdf $f_{X_k}(x)$ is expressed as*

$$f_{X_k}(x) = \frac{2\pi\lambda_k}{\mathcal{A}_k} x\, e^{-\pi \sum_{j=1}^{K} \lambda_j \left(\frac{\nu_k}{\nu_j}\right)^2 x^2}. \tag{6.9}$$

Proof. The probability that $X_k > x$ is expressed as

$$\mathbb{P}[X_k > x] = \mathbb{P}[r_k > x | m = k] = \frac{\mathbb{P}[r_k > x, m = k]}{\mathbb{P}[m = k]}. \tag{6.10}$$

In this case, $\mathbb{P}[m = k] = \mathcal{A}_k$. In addition, $\mathbb{P}[r_k > x, m = k]$ is evaluated as

$$\begin{aligned}
\mathbb{P}[r_k > x, m = k] &= \mathbb{P}[r_k > x, \nu_k r_k < \nu_j r_j, \forall j \in K, j \neq k] \\
&= \int_x^\infty \prod_{j=1, j\neq k}^{K} \mathbb{P}\left[r_j > \frac{\nu_k}{\nu_j} r\right] f_{r_k}(r)\, \mathrm{d}r \\
&\overset{(a)}{=} 2\pi\lambda_k \int_x^\infty r\, e^{-\pi \sum_{j=1}^{K} \lambda_j \left(\frac{\nu_k}{\nu_j}\right)^2 r^2}\, \mathrm{d}r
\end{aligned} \tag{6.11}$$

where (a) follows from (6.6) and (6.7). Substituting (6.4) and (6.11) into (6.10) gives

$$\mathbb{P}[X_k > x] = \frac{2\pi\lambda_k}{\mathcal{A}_k} \int_x^\infty r\, e^{-\pi \sum_{j=1}^{K} \lambda_j \left(\frac{\nu_k}{\nu_j}\right)^2 r^2}\, \mathrm{d}r. \tag{6.12}$$

However, the CDF $F_{X_k}(x) = 1 - \mathbb{P}[X_k > x]$. Therefore, the PDF $f_{X_k}(x) = \frac{\mathrm{d}F_{X_k}(x)}{\mathrm{d}x}$. ∎

6.2.1 Coverage Probability

Theorem 6.1. *Coverage probability of a typical user in the HetNet is expressed as*

$$\mathcal{P}_{c_D} = \sum_{k=1}^{K} \pi\lambda_k \int_{z>0} e^{-\frac{T\sigma^2}{P_k L} z^{\alpha/2}}\, e^{-\pi \sum_{j=1}^{K} \lambda_j \widehat{P}_j^{2/\alpha} \mathcal{E}_j z}\, dz \tag{6.13}$$

where $\mathcal{E}_j = \left[\mathcal{S}_j + \widehat{P}_j^{-2/\alpha} \widehat{\nu}_j^{-2}\right]$, $\mathcal{S}_j = T^{2/\alpha} \int_{u_j}^{\infty} \frac{1}{1+u^{\alpha/2}} du$, $u_j = (\widehat{P}_j T)^{-2/\alpha} \widehat{\nu}_j^{-2}$ and $\widehat{\nu}_j = \frac{\nu_j}{\nu_k}$. In the special case of $\alpha = 4$, $\mathcal{S}_j = \sqrt{T}\left[\frac{\pi}{2} - \arctan\left(\frac{\widehat{\nu}_j^{-2}}{\sqrt{\widehat{P}_j T}}\right)\right]$.

Proof. The Laplace transform of the aggregate interference is expressed as

$$\begin{aligned}
\mathcal{L}_{I_j}(s_c) = \mathbb{E}[e^{-s_c I_j}] &= \mathbb{E}_{\Phi_j}\left[e^{-T x^{\alpha_k} P_k^{-1} \sum_{j\in\Phi_j} P_j h_{x_j} x_j^{-\alpha_j}}\right] \\
&= \mathbb{E}_{\Phi_j}\left[e^{-T x^{\alpha_k} \sum_{j\in\Phi_j} \widehat{P}_j h_{x_j} x_j^{-\alpha_j}}\right].
\end{aligned} \tag{6.14}$$

Using the probability generating functional (PGFL) of the PPP [106], [107], and the fact that channel gain is i.i.d exponential with $h \sim \exp(1)$, (6.14) is simplified as

$$\begin{aligned}\mathcal{L}_{I_j}(s_c) &= \exp\left\{-2\pi\lambda_j \int_{q_j}^{\infty} \left(1 - \frac{1}{1 + Tx^{\alpha_k}\widehat{P}_j q^{-\alpha_j}}\right) q\,\mathrm{d}q\right\} \\ &= \exp\left\{-2\pi\lambda_j \int_{q_j}^{\infty} \frac{q}{1 + (Tx^{\alpha_k}\widehat{P}_j)^{-1} q^{\alpha_j}}\,\mathrm{d}q\right\}. \end{aligned} \tag{6.15}$$

The limit q_j, which refers to the distance to the nearest interferer, is given as $q_j = (\nu_k/\nu_j)\,x$. Let $u = (Tx^{\alpha_k}\widehat{P}_j)^{-2/\alpha_j} q^2$ such that $q\mathrm{d}q = (Tx^{\alpha_k}\widehat{P}_j)^{2/\alpha_j}\mathrm{d}u/2$. The lower limit u_j becomes $u_j = (Tx^{\alpha_k}\widehat{P}_j)^{-2/\alpha_j}(\nu_k/\nu_j)^2 x^2 = (T\widehat{P}_j)^{-2/\alpha_j}\widehat{\nu}_j^{-2} x^{-2/\widehat{\alpha}_j} x^2$. At this point, the only way to eliminate the term $x^{-2/\widehat{\alpha}_j} x^2$ is by assuming that $\{\alpha_j\} = \alpha$ s.t. $\widehat{\alpha}_j = 1$. Then $u_j = (T\widehat{P}_j)^{-2/\alpha}\widehat{\nu}_j^{-2}$ and hence,

$$\begin{aligned}\mathcal{L}_{I_j}(s_c) &= \exp\left\{-\pi\lambda_j \widehat{P}_j^{2/\alpha} x^2 T 2/\alpha \int_{u_j}^{\infty} \frac{1}{1 + u^{\alpha/2}}\,\mathrm{d}u\right\} \\ &= \exp\{-\pi\lambda_j \widehat{P}_j^{2/\alpha} \mathcal{S}_j\, x^2\}. \end{aligned} \tag{6.16}$$

Coverage probability of the k-th tier is then determined according to (3.17) as

$$\begin{aligned}\mathcal{P}_{c,k} &= \int_{x>0} e^{-\frac{T\sigma^2}{P_k L}x^\alpha} \prod_{j=1}^{K} \mathcal{L}_{I_j}(s_c)\, f_{X_k}(x)\,\mathrm{d}x \\ &= \frac{2\pi\lambda_k}{\mathcal{A}_k} \int_{x>0} e^{-\frac{T\sigma^2}{P_k L}x^\alpha} e^{-\pi\sum_{j=1, j\neq k}^{K}\lambda_j \widehat{P}_j^{2/\alpha}\mathcal{S}_j\, x^2} \cdot x\, e^{-\pi\sum_{j=1}^{K}\lambda_j\left(\frac{\nu_k}{\nu_j}\right)^2 x^2}\,\mathrm{d}x \\ &= \frac{2\pi\lambda_k}{\mathcal{A}_k} \int_{x>0} e^{-\frac{T\sigma^2}{P_k L}x^\alpha} e^{-\pi\sum_{j=1}^{K}\lambda_j \widehat{P}_j^{2/\alpha}\left[\mathcal{S}_j + \widehat{P}_j^{-2/\alpha}\widehat{\nu}_j^{-2}\right] x^2}\,\mathrm{d}x. \end{aligned} \tag{6.17}$$

Combining (3.7) and (6.17) and substituting $z = x^2$ gives overall coverage probability. ∎

Corollary 6.1. *Coverage probability of a user in an interference-limited HetNet with minimum BTD connectivity is expressed as*

$$\overline{\mathcal{P}}_{c_D} = \sum_{k=1}^{K} \frac{\lambda_k P_k^{2/\alpha}}{\sum_{j=1}^{K}\lambda_j P_j^{2/\alpha}\mathcal{E}_j}. \tag{6.18}$$

Proof. Let $\sigma^2 = 0$ in (6.13) and solve the integral. ∎

In the unbiased HetNet (i.e. $\{\widehat{\nu}_j\} = 1,\ \forall j \in K$), the term $\mathcal{E}_j$ is still not a constant since it also depends on $\{P_j\}$. Note that in a typical HetNet, $\{\widehat{P}_j\} \neq 1$. Therefore, $\overline{\mathcal{P}}_{c_D}$ of the unbiased interference-limited HetNet always depends on the parameter set $\{K, \{\lambda_j\}, \{P_j\}\}$ which is in contrast to maximum ABPR connectivity in (3.26). This dependence is a consequence of forcing users to associate with their closest BSs instead of BSs that provide the highest received signal strength. Hence, even without artificial biasing, user association in an interference-limited HetNet with minimum BTD connectivity is already *biased* by separation distances.

6.2.2 Average User Rate

Theorem 6.2. *Average user rate in a HetNet with minimum BTD connectivity is expressed as*

$$\mathcal{R}_{u_D} = \sum_{k=1}^{K} \Xi \left[\pi \lambda_k \, e^{\frac{\sigma^2}{P_k L}(2^t-1)z^{\alpha/2}} \, e^{-\pi \sum_{j=1}^{K} \lambda_j \widehat{P}_j^{2/\alpha} \mathcal{G}_j(t) z} \right] \tag{6.19}$$

where $\mathcal{G}_j(t) = \mathcal{U}_j(t) + \frac{1}{\widehat{P}_j^{2/\alpha}\widehat{\nu}_j^2}$, $\mathcal{U}_j(t) = (2^t-1)^{2/\alpha} \int_{\varrho_j(t)}^{\infty} \frac{1}{1+u^{\alpha/2}} du$ *and* $\varrho_j(t) = \frac{1}{\widehat{\nu}_j^2 (\widehat{P}_j (2^t-1))^{2/\alpha}}$. *In the special case of* $\alpha = 4$, $\mathcal{U}_j = \sqrt{2^t-1} \left[\frac{\pi}{2} - \operatorname{atan}\left(\frac{1}{\widehat{\nu}_j^2 \sqrt{\widehat{P}_j (2^t-1)}} \right) \right]$.

Proof. Assuming $\{\alpha_j\} = \alpha$ and using (3.20)-(3.21), average user rate of the k-th tier becomes

$$\mathcal{R}_{u,k} \int_{t>0} \int_{x>0} e^{-\frac{\sigma^2}{P_k L}(2^t-1)x^\alpha} \prod_{j=1}^{K} \mathcal{L}_{I_j}(s_r) \, f_{X_k}(x) \, \mathrm{d}x \, \mathrm{d}t \tag{6.20}$$

where $s_r = (2^t-1)x^\alpha P_k^{-1}$. Following the proof of Theorem 6.1, $\mathcal{L}_{I_j}(s_r)$ is expressed as

$$\mathcal{L}_{I_j}(s_r) = \exp\left\{ -\pi \lambda_j \widehat{P}_j^{2/\alpha} \mathcal{U}_j(t) x^2 \right\}. \tag{6.21}$$

$\mathcal{R}_{u,k}$ is then obtained by substituting (6.9) and (6.21) into (6.20). Combining the resulting expression with (3.7) and (6.4) and substituting $z = x^2$ gives the result in (6.19). ∎

Corollary 6.2. *Average user rate in an interference-limited HetNet with minimum BTD connectivity is expressed as*

$$\overline{\mathcal{R}}_{u_D} = \sum_{k=1}^{K} \int_{t>0} \frac{\lambda_k P_k^{2/\alpha}}{\sum_{j=1}^{K} \lambda_j P_j^{2/\alpha} \mathcal{G}_j(t)} \mathrm{d}t. \tag{6.22}$$

Proof. Let $\sigma^2 = 0$ in (6.19) and solve the integral. ∎

Similar to $\overline{\mathcal{P}}_{c_D}$, when $\{\widehat{\nu}_j\} = 1$, $\mathcal{G}_j(t)$ is not a constant since it depends on $\{P_j\}$ where $\{\widehat{P}_j\} \neq 1$. As a result, $\overline{\mathcal{R}}_{u_D}$ is dependent on the parameter set $\{K, \{\lambda_j\}, \{P_j\}\}$. Again, this is contrary to maximum ABRP connectivity which gives a constant $\overline{\mathcal{R}}_{u_P}$ when the HetNet is unbiased.

6.3 Optimization Constraints

Intuitively, noise reduces the coverage probability and average user rate in the HetNet and therefore the interference-limited HetNet defines the upper bounds of both measures. Biasing generally reduces the coverage probability and average rate performance compared to an unbiased HetNet [94]. However, biasing is an important operational technique that enhances load balancing and average sum rate in HetNets [93], [96], [97]. Therefore to facilitate its implementation while managing its negative effect on coverage probability and average rate, the coverage probability and average user rate constraints are defined as [97]

$$\mathcal{P}_c \geq \epsilon \, \overline{\mathcal{P}}_c \quad \text{and} \quad \mathcal{R}_u \geq \kappa \, \overline{\mathcal{R}}_u \tag{6.23}$$

respectively, where $\epsilon \in (0,1]$ and $\kappa \in (0,1]$ are ratios of the coverage probability and average user rate to their respective upper bound values.

In this chapter, the deployment configuration of the HetNet is optimized to minimize its APC subject to coverage probability and average user rate constraints. Since coverage probability and average user rate are complementary performance measures (i.e. optimization based on one measure improves the other measure as well), the optimal solution is one that satisfies both measures. Therefore the optimization problem is divided into two separate problems, one subject to a coverage probability constraint and the other subject to an average user rate constraint. The final solution is the maximum of the two individual solutions.

The purpose of HetNet optimization is to devise an optimal deployment strategy in terms of its deployment factor. The deployment factor, expressed as $\mathcal{H} = \sum_{j=1}^{K} \lambda_j P_j^{2/\alpha}$, is essentially a combination of BS densities and their associated transmit powers per tier that jointly achieve a given performance constraint. It can easily be optimized to determine the specific optimal BS densities and associated transmit powers per tier that minimize the APC of the HetNet.

6.4 Maximum ABRP Connectivity

In this user association, a typical user connects to the BS from any tier that provides the best long-term average biased received power. The coverage probability and average user rate analysis of a HetNet using this association scheme is discussed in subsection 3.7.1.

6.4.1 Coverage Probability Constraint

Theorem 6.3. *The coverage probability of a typical user in the HetNet is approximated as*

$$\mathcal{P}_{c_P} \approx \overline{\mathcal{P}}_{c_P} - T\sigma^2\psi(\alpha) \sum_{k=1}^{K} \frac{\lambda_k P_k^{2/\alpha}}{\left[\sum_{j=1}^{K} \lambda_j P_j^{2/\alpha} \mathcal{C}_j\right]^{\frac{\alpha}{2}+1}} \tag{6.24}$$

where $\overline{\mathcal{P}}_{c_P}$ is shown in (3.26)-(3.27) and $\psi(\alpha) = \frac{\Gamma\left(\frac{\alpha}{2}+1\right)}{\pi^{\alpha/2}L}$. If the HetNet is unbiased, $\{\mathcal{C}_j\} = \mathcal{C}$ which is a constant. Hence (6.24) can be simplified as

$$\mathcal{P}_{c_P} \approx \overline{\mathcal{P}}_{c_P} - \frac{T\sigma^2\psi(\alpha)}{\mathcal{C}^{\frac{\alpha}{2}+1}\left[\sum_{j=1}^{K} \lambda_j P_j^{2/\alpha}\right]^{\alpha/2}}. \tag{6.25}$$

Proof. Similar to (4.36), consider the following approximation based on SNR $\gg T$:

$$e^{-\frac{T}{\mathrm{SNR}_k}} \equiv e^{-\frac{T\sigma^2}{P_k L} z^{\alpha/2}} \approx 1 - \frac{T\sigma^2}{P_k L} z^{\alpha/2}. \tag{6.26}$$

Substituting (6.26) into (3.25) gives the following two integrals:

$$\begin{aligned}\mathcal{P}_{c_P} &\approx \sum_{k=1}^{K} \pi\lambda_k \left[\int_{z>0} e^{-\pi\sum_{j=1}^{K}\lambda_j \widehat{P}_j^{2/\alpha}\mathcal{C}_j z}\mathrm{d}z - \frac{T\sigma^2}{P_k L}\int_{z>0} z^{\alpha/2} e^{-\pi\sum_{j=1}^{K}\lambda_j \widehat{P}_j^{2/\alpha}\mathcal{C}_j z}\mathrm{d}z\right] \\ &\overset{(a)}{=} \sum_{k=1}^{K} \frac{\lambda_k}{\sum_{j=1}^{K}\lambda_j \widehat{P}_j^{2/\alpha}\mathcal{C}_j} - \sum_{k=1}^{K}\pi\lambda_k \frac{T\sigma^2\,\Gamma\left(\frac{\alpha}{2}+1\right)}{P_k L\left[\pi\sum_{j=1}^{K}\lambda_j \widehat{P}_j^{2/\alpha}\mathcal{C}_j\right]^{\frac{\alpha}{2}+1}} \\ &= \overline{\mathcal{P}}_{c_P} - \frac{T\sigma^2\,\Gamma\left(\frac{\alpha}{2}+1\right)}{\pi^{\alpha/2}L}\sum_{k=1}^{K}\frac{\lambda_k P_k^{2/\alpha}}{\left[\sum_{j=1}^{K}\lambda_j P_j^{2/\alpha}\mathcal{C}_j\right]^{\frac{\alpha}{2}+1}} \end{aligned} \tag{6.27}$$

where (a) follows using the identity in [126, (3.381.4)]. ■

The approximated coverage probability in (6.24) and (6.25) can be used to determine the required deployment factor of the HetNet subject to this measure. However, the analysis is different for biased and unbiased HetNets as discussed next.

Unbiased HetNet: Combining the coverage probability constraint $\mathcal{P}_{c_P} \geq \epsilon\,\overline{\mathcal{P}}_{c_P}$ in (6.23) with (3.27), (6.25) can be rewritten as

$$\mathcal{H}_c \equiv \sum_{j=1}^{K}\lambda_j P_j^{2/\alpha} \geq \frac{1}{\mathcal{C}}\left(\frac{T\sigma^2\psi(\alpha)}{1-\epsilon}\right)^{2/\alpha} \tag{6.28}$$

where $\mathcal{H}_c$ is defined as the deployment factor of a general K-tier HetNet that satisfies the coverage probability constraint. Hence, the deployment factor of the K-tier unbiased HetNet is expressed in closed form. According to Lemmas 6.1-6.2, the optimal deployment factor is expressed as $\mathcal{H}_c^\star = \frac{1}{\mathcal{C}}\left(\frac{T\sigma^2\psi(\alpha)}{1-\epsilon}\right)^{2/\alpha}$ and it is an increasing function of both σ^2 and ϵ.

Consider a typical K-tier unbiased HetNet defined by the parameters $K = 2$, $\lambda_j \in \{\lambda_b, \lambda_s\}$ and $P_j \in \{P_b, P_s\}$. The optimal deployment factor is expressed as $\lambda_b P_b^{2/\alpha} + \lambda_s P_s^{2/\alpha} = \mathcal{H}_c^\star$. This special case of the two-tier unbiased HetNet using maximum ABRP connectivity is considered in [117] but a closed form expression for the deployment factor is not provided. A network APC minimization framework is formulated to determine the optimal macro and micro BS densities and their optimal transmit powers subject to a coverage probability constraint.

Biased HetNet: In this case, $\{\mathcal{C}_j\} \neq \mathcal{C}$ and further simplification of (6.24) is not possible. Using $\mathcal{P}_{c_P} \geq \epsilon\overline{\mathcal{P}}_{c_P}$, (6.24) is instead rewritten as

$$(1-\epsilon)\overline{\mathcal{P}}_{c_P} - T\sigma^2\psi(\alpha)\sum_{k=1}^{K}\frac{\lambda_k P_k^{2/\alpha}}{\mathcal{Q}_c^{\frac{\alpha}{2}+1}} \geq 0 \tag{6.29}$$

where $\mathcal{Q}_c = \sum_{j=1}^{K}\lambda_j P_j^{2/\alpha}\mathcal{C}_j$. Due to the variation of $\mathcal{C}_j$ with the $\{k, j\}$ pair, the deployment factor cannot be isolated and expressed in closed form. In order to investigate the impact of biasing on the deployment factor and energy performance of the HetNet, consider a conventional two-tier biased HetNet of macro BSs and small BSs in which the macrocell tier deployment factor $\mathcal{H}_b = \lambda_b P_b^{2/\alpha}$ is known. This is often the case where an operator has a macrocell network

providing outdoor coverage and wishes to densify it with small BSs for targeted coverage and capacity enhancements in indoor and other environments. For such a HetNet, $\mathcal{Q}_c = \mathcal{H}_b\mathcal{C}_b + \mathcal{H}_{s,c}\mathcal{C}_s$ where $\mathcal{H}_{s,c} = \lambda_{s,c}P_s^{2/\alpha}$ is the required deployment factor of the small cell tier that jointly satisfies the coverage probability constraint.

The optimal value of $\mathcal{H}_{s,c}$, denoted as $\mathcal{H}^\star_{s,c}$, is easily determined using numerical methods such as the bisection method [127]. To determine the optimal deployment configuration of the small cell tier, consider the following APC minimization framework that defines this bivariate optimization problem:

$$\begin{cases} \underset{\lambda_{s,c},\, P_s}{\text{minimize}} & \lambda_{s,c}(\mathcal{N}_s P_{0,s} + \Delta_s P_s) \\ \text{subject to} & \lambda_{s,c}P_s^{2/\alpha} = \mathcal{H}^\star_{s,c}, \;\; 0 \le P_s \le \overline{P}_s \end{cases} \tag{6.30}$$

where $\overline{P}_s$ is the maximum small BS transmit power. This problem can easily be converted into a single variable problem by using the substitution $\lambda_{s,c} = \mathcal{H}^\star_{s,c}P_s^{-2/\alpha}$. The optimization framework of the univariate problem is

$$\begin{cases} \underset{P_s}{\text{minimize}} & \mathcal{H}^\star_{s,c}P_s^{-2/\alpha}(\mathcal{N}_s P_{0,s} + \Delta_s P_s) \\ \text{subject to} & 0 \le P_s \le \overline{P}_s. \end{cases} \tag{6.31}$$

Therefore, $F(P_s) = \mathcal{H}^\star_{s,c}P_s^{-2/\alpha}(\mathcal{N}_s P_{0,s} + \Delta_s P_s)$. Using differentiation to determine its minimum,

$$\frac{\mathrm{d}F(P_s)}{\mathrm{d}P_s} = \frac{\alpha-2}{\alpha}P_s^{-2/\alpha}\Delta_s\mathcal{H}^\star_{s,c} - \frac{2}{\alpha}P_s^{\frac{-2-\alpha}{\alpha}}\mathcal{N}_s P_{0,s}\mathcal{H}^\star_{s,c} = 0.$$

$$\Rightarrow P_s^\star = \frac{2\mathcal{N}_s P_{0,s}}{\Delta_s(\alpha-2)}.$$

Hence, the resulting optimal solutions, denoted as $P_s^\star$ and $\lambda^\star_{s,c}$, are expressed as

$$P_s^\star = \min\left\{\frac{2\mathcal{N}_s P_{0,s}}{\Delta_s(\alpha-2)}, \overline{P}_s\right\} \quad \text{and} \quad \lambda^\star_{s,c} = \mathcal{H}^\star_{s,c}P_s^{\star\,-2/\alpha}. \tag{6.32}$$

Hence $P_s^\star$ only depends on the small BS power consumption parameters. Furthermore, the value of $\mathcal{H}^\star_{s,c}$ depends on the bias ratio, denoted as $\beta = \beta_s/\beta_b$. Therefore β can be varied to determine the optimal set $\{\mathcal{H}^\star_{s,c}, \beta^\star\}$ that minimizes the APC.

6.4.2 Average User Rate Constraint

Theorem 6.4. *The average rate of a typical user in a K-tier HetNet is approximated as*

$$\mathcal{R}_{u_P} \approx \overline{\mathcal{R}}_{u_P} - \sigma^2\psi(\alpha)\sum_{k=1}^{K}\int_{t>0}\frac{\lambda_k P_k^{2/\alpha}(2^t-1)}{\left[\sum_{j=1}^{K}\lambda_j P_j^{2/\alpha}\mathcal{D}_j(t)\right]^{\frac{\alpha}{2}+1}}\,dt \tag{6.33}$$

where $\overline{\mathcal{R}}_{u_P}$ is shown in (3.29)-(3.30). If the HetNet is unbiased, $\{\mathcal{D}_j(t)\} = \mathcal{D}(t) = 1+\mathcal{Z}(t,\alpha,1)$ which is a constant. Hence, (6.33) can be simplified to

$$\mathcal{R}_{u_P} \approx \overline{\mathcal{R}}_{u_P} - \frac{\sigma^2\psi(\alpha)}{\left[\sum_{j=1}^{K}\lambda_j P_j^{2/\alpha}\right]^{\alpha/2}} \int_{t>0} \frac{(2^t-1)}{\mathcal{D}(t)^{\frac{\alpha}{2}+1}}\, dt. \tag{6.34}$$

Proof. Over a realistic SNR range, the first exponential term in (3.28) is approximated as

$$e^{\frac{-\sigma^2}{P_k L}(2^t-1)z^{\alpha/2}} \approx 1 - \frac{\sigma^2}{P_k L}(2^t-1)z^{\alpha/2}. \tag{6.35}$$

Substituting (6.35) into (3.28) gives the following integrals:

$$\begin{aligned}\mathcal{R}_{u_P} &\approx \sum_{k=1}^{K}\pi\lambda_k\left\{\Xi\left[e^{-\pi\sum_{j=1}^{K}\lambda_j\widehat{P}_j^{2/\alpha}\mathcal{D}_j(t)z}\right] - \Xi\left[\frac{\sigma^2}{P_k L}(2^t-1)z^{\alpha/2}\, e^{-\pi\sum_{j=1}^{K}\lambda_j\widehat{P}_j^{2/\alpha}\mathcal{D}_j(t)z}\right]\right\} \\ &= \overline{\mathcal{R}}_{u_P} - \sum_{k=1}^{K}\frac{\sigma^2}{P_k L}\pi\lambda_k\int_{t>0}\frac{\Gamma\left(\frac{\alpha}{2}+1\right)(2^t-1)}{\left[\pi\sum_{j=1}^{K}\lambda_j\widehat{P}_j^{2/\alpha}\mathcal{D}_j(t)\right]^{\frac{\alpha}{2}+1}}\,\mathrm{d}t.\end{aligned} \tag{6.36}$$

Simplifying (6.36) gives the result in (6.33). ∎

Unbiased HetNet: Combining the average user rate constraint $\mathcal{R}_{u_P} \geq \kappa\overline{\mathcal{R}}_{u_P}$ in (6.23) and (3.30), (6.34) can be rewritten in terms of the deployment factor as

$$\mathcal{H}_r \equiv \sum_{j=1}^{K}\lambda_j P_j^{2/\alpha} \geq \left(\frac{\sigma^2\psi(\alpha)g(t,\alpha)}{1-\kappa}\right)^{2/\alpha} \tag{6.37}$$

where $\mathcal{H}_r$ is the deployment factor and $g(t,\alpha) = \int_{t>0}\frac{(2^t-1)}{\mathcal{D}(t)^{\frac{\alpha}{2}+1}}\,\mathrm{d}t \times \left(\int_{t>0}\frac{1}{\mathcal{D}(t)}\right)^{-1}$ is a constant for a given α. Hence, the deployment factor of a general K-tier unbiased HetNet that satisfies the average user rate constraint is also expressed in closed form. According to Lemmas 6.1-6.2, the optimal deployment factor is expressed as $\mathcal{H}_r^\star = \left(\frac{\sigma^2\psi(\alpha)g(t,\alpha)}{1-\kappa}\right)^{2/\alpha}$ and it is an increasing function of σ^2 and κ.

Biased HetNet: In this case, $\{\mathcal{D}_j(t)\} \neq \mathcal{D}(t)$ and further simplification of (6.33) is not possible. Using $\mathcal{R}_{u_P} \geq \kappa\overline{\mathcal{R}}_{u_P}$, (6.33) is instead rewritten as

$$(1-\kappa)\overline{\mathcal{R}}_{u_P} - \sigma^2\psi(\alpha)\sum_{k=1}^{K}\int_{t>0}\frac{\lambda_k P_k^{2/\alpha}(2^t-1)}{\mathcal{Q}_r^{\frac{\alpha}{2}+1}}\,\mathrm{d}t \geq 0 \tag{6.38}$$

where $\mathcal{Q}_r = \sum_{j=1}^{K}\lambda_j P_j^{2/\alpha}\mathcal{D}_j(t)$. Due to the dependence of $\mathcal{D}_j(t)$ on the $\{k,j\}$ pair, it is not possible to isolate and express the deployment factor $\mathcal{H}_{s,r} = \sum_{j=1}^{K}\lambda_j P_j^{2/\alpha}$ in closed form. Similar to the coverage probability constraint scenario, assume a conventional two tier HetNet with the same macrocell tier deployment factor $\mathcal{H}_b = \lambda_b P_b^{2/\alpha}$. Then, $\mathcal{Q}_r = \mathcal{H}_b\mathcal{D}_b(t) + \mathcal{H}_{s,r}\mathcal{D}_s(t)$ where $\mathcal{H}_{s,r} = \lambda_s P_s^{2/\alpha}$ is the required small cell tier deployment factor that jointly satisfies the average user rate constraint. Using numerical methods, the optimal value of $\mathcal{H}_{s,r}$, denoted as $\mathcal{H}_{s,r}^\star$, can easily be determined.

The optimization of $\mathcal{H}^{\star}_{s,r}$ to determine the values of $\lambda^{\star}_{s,r}$ and $P^{\star}_{s}$ that minimize the APC of the HetNet follows the same procedure shown in (6.30)-(6.32). The solutions are

$$P^{\star}_{s} = \min\left\{\frac{2\mathcal{N}_s P_{0,s}}{\Delta_s(\alpha-2)}, \overline{P}_s\right\} \quad \text{and} \quad \lambda^{\star}_{s,r} = \mathcal{H}^{\star}_{s,r} P^{\star -2/\alpha}_{s}. \tag{6.39}$$

The optimal transmit power $P^{\star}_{s}$ subject to both coverage probability and average user rate constraints is similar as shown in (6.32) and (6.39). Therefore, if the type of deployed small BSs is known, the optimal transmit power can be predetermined. Since β also influences $\mathcal{H}^{\star}_{s,r}$, it can be varied to determine the optimal set $\{\mathcal{H}^{\star}_{s,r}, \beta^{\star}\}$ at which APC is minimized.

6.4.3 Overall Solution

Since coverage probability and average rate constraints are complementary to each other (i.e. optimization based on one measure also improves the other), the optimal solution is one that satisfies both constraints. The overall optimal solution will be the maximum of the two solutions i.e. the optimal small BS density is $\lambda^{\star}_{s} = \max\{\lambda^{\star}_{s,c}, \lambda^{\star}_{s,r}\}$ and the optimal transmit power is $P^{\star}_{s}$. In addition, both $\lambda^{\star}_{s,c}$ and $\lambda^{\star}_{s,r}$ are influenced by the bias ratio in the HetNet. The optimal bias ratio at which $\lambda^{\star}_{s}$ (or, $\max\{\mathcal{H}^{\star}_{s,c}, \mathcal{H}^{\star}_{s,r}\}$) is minimized is determined as

$$\begin{cases} \underset{\beta}{\text{minimize}} & \max\{\lambda^{\star}_{s,c}, \lambda^{\star}_{s,r}\} \\ \text{subject to} & \beta > 0,\ \lambda^{\star}_{s,c} \geq 0,\ \lambda^{\star}_{s,r} \geq 0, P^{\star}_{s} > 0. \end{cases} \tag{6.40}$$

The resulting APC obtained by operating the HetNet at its optimal bias value $\beta^{\star}$ becomes

$$\text{APC} = \lambda_b(\mathcal{N}_b P_{0,b} + \Delta_b P_b) + \lambda^{\star}_{s}(\mathcal{N}_s P_{0,s} + \Delta_s P^{\star}_{s}). \tag{6.41}$$

6.5 Minimum BTD Connectivity

6.5.1 Coverage Probability Constraint

Theorem 6.5. *Coverage probability of a typical user can be approximated as*

$$\mathcal{P}_{c_D} \approx \overline{\mathcal{P}}_{c_D} - T\sigma^2\psi(\alpha)\sum_{k=1}^{K}\frac{\lambda_k P_k^{2/\alpha}}{\left[\sum_{j=1}^{K}\lambda_j P_j^{2/\alpha}\mathcal{E}_j\right]^{\frac{\alpha}{2}+1}}. \tag{6.42}$$

Proof. Substituting (6.26) into (6.13) gives the following integrals:

$$\mathcal{P}_c \approx \sum_{k=1}^{K}\pi\lambda_k\left\{\int_0^{\infty} e^{-\pi\sum_{j=1}^{K}\lambda_j\widehat{P}_j^{2/\alpha}\mathcal{E}_j z}\,\mathrm{d}z - \frac{T\sigma^2}{P_k L}\int_0^{\infty} z^{\alpha/2}e^{-\pi\sum_{j=1}^{K}\lambda_j\widehat{P}_j^{2/\alpha}\mathcal{E}_j z}\,\mathrm{d}z\right\}. \tag{6.43}$$

Solving both integrals gives the result. ∎

In this scheme, $\mathcal{E}_j$ is a function of $\{\widehat{\nu}_j\}$ and $\{\widehat{P}_j\}$ both of which vary with the $\{k, j\}$ pair. As a result, even when the HetNet is unbiased (i.e. $\widehat{\nu}_j = 1$), the term $\mathcal{E}_j$ is still not a constant since $\{\widehat{P}_j\} \neq 1$ in a typical HetNet. Therefore, the deployment factor cannot be isolated and expressed in closed form. Instead, (6.42) is rewritten using $\mathcal{P}_{c_D} \geq \epsilon \overline{\mathcal{P}}_{c_D}$ as

$$(1-\epsilon)\overline{\mathcal{P}}_{c_D} - T\sigma^2\psi(\alpha)\sum_{k=1}^{K}\frac{\lambda_k P_k^{2/\alpha}}{\mathcal{Q}_c^{\frac{\alpha}{2}+1}} \geq 0 \tag{6.44}$$

where $\mathcal{Q}_c = \sum_{j=1}^{K}\lambda_j P_j^{2/\alpha}\mathcal{E}_j$. Similar to the case of the biased HetNet using maximum ABRP connectivity, it is still possible to investigate the optimal performance of the HetNet by assuming a conventional two-tier HetNet whose macrocell tier deployment factor $\mathcal{H}_b = \lambda_b P_b^{2/\alpha}$ is known. Then, $\mathcal{Q}_c = \mathcal{H}_b\mathcal{E}_b + \mathcal{H}_{s,c}\mathcal{E}_s$ where $\mathcal{H}_{s,c} = \lambda_s P_s^{2/\alpha}$ is the small cell tier deployment factor. The optimal value $\mathcal{H}_{s,c}^{\star}$ can be determined using numerical methods. The optimization to determine $P_s^{\star}$ and $\lambda_{s,c}^{\star}$ follows the procedure shown in (6.30)-(6.32). The expressions of the optimal solutions (transmit power and BS density) are otherwise similar to those shown in (6.32) except that the deployment factor $\mathcal{H}_{s,c}^{\star}$ is determined from (6.44).

6.5.2 Average User Rate Constraint

Theorem 6.6. *Average user rate in the HetNet can be approximated as*

$$\mathcal{R}_{u_D} \approx \overline{\mathcal{R}}_{u_D} - \sigma^2\psi(\alpha)\sum_{k=1}^{K}\int_{t>0}\frac{\lambda_k P_k^{2/\alpha}(2^t-1)}{\left[\sum_{j=1}^{K}\lambda_j P_j^{2/\alpha}\mathcal{G}_j(t)\right]^{\frac{\alpha}{2}+1}}\,dt. \tag{6.45}$$

Proof. Substituting (6.35) into (6.19) gives the following integrals:

$$\mathcal{R}_{u_D} \approx \sum_{k=1}^{K}\pi\lambda_k\left\{\Xi\left[e^{-\pi\sum_{j=1}^{K}\lambda_j\widehat{P}_j^{2/\alpha}\mathcal{G}_j(t)z}\right] - \Xi\left[\frac{\sigma^2}{P_k L}(2^t-1)z^{\alpha/2}e^{-\pi\sum_{j=1}^{K}\lambda_j\widehat{P}_j^{2/\alpha}\mathcal{G}_j(t)z}\right]\right\}. \tag{6.46}$$

Evaluating both integrals and simplifying gives the result. ∎

Similar to the coverage probability constraint scenario, the term $\mathcal{G}_j(t)$ also depends on both $\{\widehat{\nu}_j\}$ and $\{\widehat{P}_j\}$ and therefore continues to vary with the $\{k, j\}$ pair even when $\{\widehat{\nu}_j\} = 1$. Hence, further simplification of (6.45) using the constraint term $\mathcal{R}_{u_D} \geq \kappa\,\overline{\mathcal{R}}_{u_D}$ is not possible and the deployment factor cannot be expressed in closed form. Instead, (6.45) is rewritten as

$$(1-\kappa)\overline{\mathcal{R}}_{u_D} - \sigma^2\psi(\alpha)\sum_{k=1}^{K}\int_{t>0}\frac{\lambda_k P_k^{2/\alpha}(2^t-1)}{\mathcal{Q}_r^{\frac{\alpha}{2}+1}}\,\mathrm{d}t \geq 0 \tag{6.47}$$

where $\mathcal{Q}_r = \sum_{j=1}^{K}\lambda_j P_j^{2/\alpha}\mathcal{G}_j(t)$. Assuming a two-tier HetNet with the same $\mathcal{H}_b$, then $\mathcal{Q}_r = \mathcal{H}_b\mathcal{G}_b(t) + \mathcal{H}_{s,r}\mathcal{G}_s(t)$ where $\mathcal{H}_{s,r} = \lambda_s P_s^{2/\alpha}$ is the required small cell tier deployment factor subject to the average user rate constraint. The optimal value $\mathcal{H}_{s,r}^{\star}$ can easily be determined using numerical methods. The optimization to determine $P_s^{\star}$ and $\lambda_{s,c}^{\star}$ follows the procedure shown in (6.30)-(6.32), with similar solutions.

The overall optimal BS density is determined as $\lambda_s^\star = \max\{\lambda_{s,c}^\star, \lambda_{s,r}^\star\}$ and the optimal transmit power $P_s^\star$. The optimal bias factor, denoted as $\upsilon^\star$, at which the APC is minimized can also be obtained using the same procedure shown in (6.40).

6.6 Maximum i-SINR Connectivity

In this scheme, a typical user connects to the BS from any tier that provides the best instantaneous received SINR. This scheme has been discussed in subsection 3.7.2 [108].

6.6.1 Coverage Probability Constraint

Theorem 6.7. *Coverage probability of a typical user can be approximated as*

$$\mathcal{P}_{c_S} \approx \overline{\mathcal{P}}_{c_S} - \frac{\pi\sigma^2\xi(\alpha)T^{-2/\alpha}}{\varrho(\alpha)\left[\sum_{j=1}^K \lambda_j P_j^{2/\alpha}\right]^{\alpha/2}} \tag{6.48}$$

where $\xi(\alpha) = \frac{\Gamma\left(\frac{\alpha}{2}+1\right)}{L\varrho(\alpha)^{\alpha/2}}$ *is a constant for a given* α.

Proof. Substituting (6.26) into (3.31) gives the following integrals:

$$\begin{aligned}\mathcal{P}_{c_S} &\approx \sum_{k=1}^K \pi\lambda_k \left[\int_{z>0} e^{-T^{2/\alpha}\varrho(\alpha)\sum_{j=1}^K \lambda_j \widehat{P}_j^{2/\alpha} z}\,\mathrm{d}z - \frac{T\sigma^2}{P_k L}\int_{z>0} z^{\alpha/2} e^{-T^{2/\alpha}\varrho(\alpha)\sum_{j=1}^K \lambda_j \widehat{P}_j^{2/\alpha} z}\,\mathrm{d}z\right] \\ &\overset{(a)}{=} \overline{\mathcal{P}}_{c_S} - \sum_{k=1}^K \frac{\pi\lambda_k P_k^{\frac{\alpha+2}{\alpha}}\, T\sigma^2\,\Gamma\left(\frac{\alpha}{2}+1\right)}{P_k L\left[T^{2/\alpha}\varrho(\alpha)\sum_{j=1}^K \lambda_j P_j^{2/\alpha}\right]^{\frac{\alpha}{2}+1}}\end{aligned} \tag{6.49}$$

where (a) follows using [126, (3.381.4)]. Simplification of (6.49) gives the result. ∎

Using $\mathcal{P}_c \geq \epsilon\,\overline{\mathcal{P}}_c$, (6.48) can be rewritten in terms of deployment factor $\mathcal{H}_c$ as

$$\mathcal{H}_c \equiv \sum_{j=1}^K \lambda_j P_j^{2/\alpha} \geq \left(\frac{\sigma^2\xi(\alpha)}{1-\epsilon}\right)^{2/\alpha}. \tag{6.50}$$

Hence the deployment factor is expressed in closed form. According to Lemmas 6.1-6.2, the optimal deployment factor is $\mathcal{H}_c^\star = \left(\frac{\sigma^2\xi(\alpha)}{1-\epsilon}\right)^{2/\alpha}$ and it is an increasing function of σ^2 and ϵ.

6.6.2 Average User Rate Constraint

Theorem 6.8. *Average user rate in the HetNet can be approximated as*

$$\mathcal{R}_{u_S} \approx \overline{\mathcal{R}}_{u_S} - \frac{\pi\sigma^2\xi(\alpha)}{\varrho(\alpha)\left[\sum_{j=1}^K \lambda_j P_j^{2/\alpha}\right]^{\alpha/2}} \int_{t>0} (2^t-1)^{-2/\alpha}\,dt. \tag{6.51}$$

Proof. Substituting (6.35) into (3.35) gives the following integrals

$$\mathcal{R}_{u_S} \approx \sum_{k=1}^{K} \pi\lambda_k \Xi \left[e^{-(2^t-1)^{2/\alpha}\varrho(\alpha)\sum_{j=1}^{K}\lambda_j \widehat{P}_j^{2/\alpha} z} - \frac{\sigma^2}{P_k L}(2^t-1) z^{\alpha/2} e^{-(2^t-1)^{2/\alpha}\varrho(\alpha)\sum_{j=1}^{K}\lambda_j \widehat{P}_j^{2/\alpha} z} \right]$$

$$= \overline{\mathcal{R}}_{u_S} - \pi \sum_{k=1}^{K} \lambda_k \int_{t>0} \frac{\sigma^2(2^t-1)}{P_k L} \frac{\Gamma\left(\frac{\alpha}{2}+1\right) P_k^{\frac{\alpha+2}{\alpha}}}{\left[(2^t-1)^{2/\alpha}\varrho(\alpha)\right]^{\frac{\alpha}{2}+1} \left[\sum_{j=1}^{K}\lambda_j P_j^{2/\alpha}\right]^{\frac{\alpha}{2}+1}} \, \mathrm{d}t. \tag{6.52}$$

Further simplification of (6.52) gives the result. ∎

Using $\mathcal{R}_u \geq \kappa \overline{\mathcal{R}}_u$, (6.51) can be rewritten in terms of its deployment factor $\mathcal{H}_r$ as

$$\mathcal{H}_r \equiv \sum_{j=1}^{K} \lambda_j P_j^{2/\alpha} \geq \left(\frac{\sigma^2 \xi(\alpha)}{1-\kappa} \right)^{2/\alpha}. \tag{6.53}$$

Therefore according to Lemmas 6.1-6.2, the optimal deployment factor is $\mathcal{H}_r^\star = \left(\frac{\sigma^2\xi(\alpha)}{1-\epsilon}\right)^{2/\alpha}$ and it is an increasing function of the parameters σ^2 and κ.

Hence according to (6.50) and (6.53), the individual optimal deployment factors of the HetNet based on coverage probability and average user rate constraints are similar when $\epsilon = \kappa$.

6.7 Numerical Results

Consider a HetNet with the default parameters shown in Table 6.1, unless otherwise stated.

Table 6.1: Parameters used to obtain results

Parameters	**Value**
Network size	$A = 10\,\text{km} \times 10\,\text{km}$
System bandwidth	$\mathcal{B} = 20\,\text{MHz}$
Pathloss parameters	$L = -33\,\text{dB}$, $\alpha = 4$
Additive noise parameters	$F = 10$, $T_a = 300\,\text{K}$
Coverage threshold	$T = 0\,\text{dB}$
Macrocell tier parameters	$\mathcal{H}_b = 4 \times 10^{-5}\,\text{m}^2$, $\beta_b = 0\,\text{dB}$
User density	$\lambda_u = 10^{-3}\,\text{m}^2$
Optimization constraints	$\epsilon = 0.9$, $\kappa = 0.9$
Macro BS power parameters	$\mathcal{N}_b = 6$, $P_{0,b} = 130$, $\Delta_b = 4.7$, $\overline{P}_b = 20\,\text{W}$
Small BS power parameters	$\mathcal{N}_s = 2$, $P_{0,s} = 6.8$, $\Delta_s = 4.0$, $\overline{P}_s = 0.13\,\text{W}$

Intuitively, additive noise reduces the coverage probability and average rate of a typical user in the HetNet. However, the amount of this reduction depends on the prevailing aggregate interference in the HetNet which varies with the BS density and transmit power. As Figs. 6.1-6.6 show, when the small BS density increases, both coverage probability and average user rate approach their respective upper bound values $\overline{\mathcal{P}}_c$ and $\overline{\mathcal{R}}_u$. This is because the HetNet gradually becomes interference-limited and the effect of noise becomes ever more negligible.

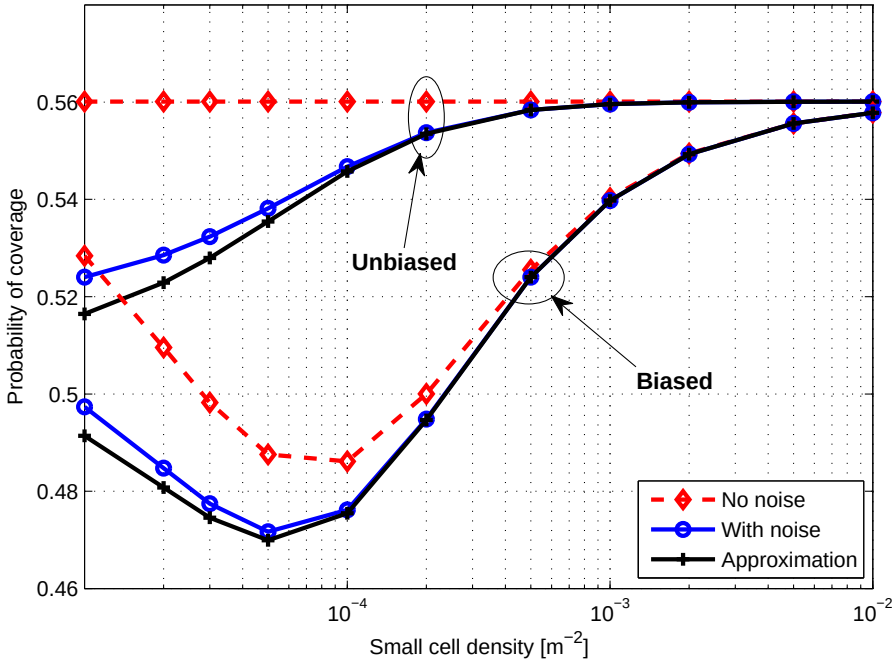


Figure 6.1: Verification of coverage probability approximation in the biased and unbiased HetNet using maximum ABRP connectivity, where $\beta = 10\,\text{dB}$.

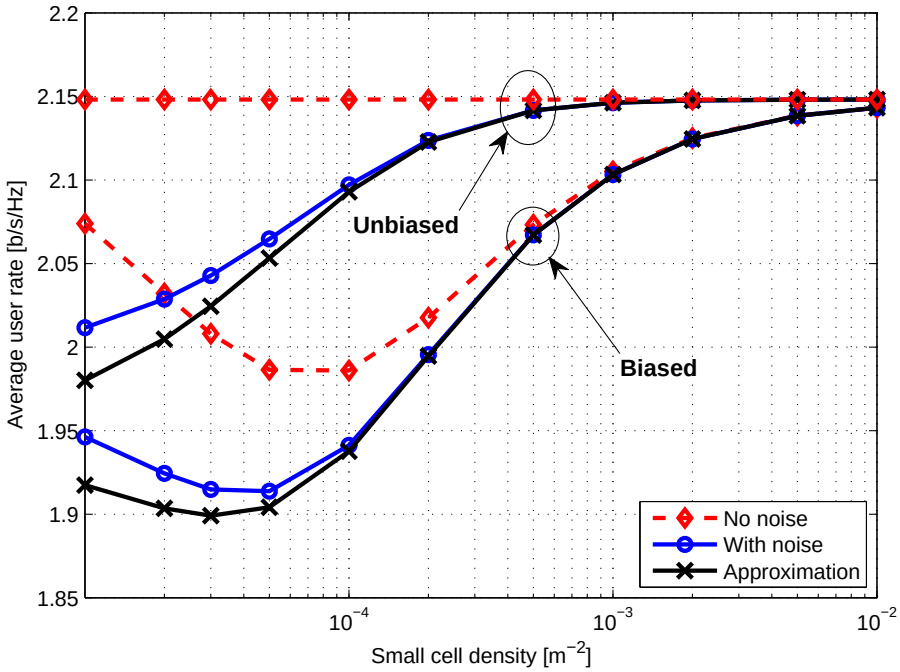


Figure 6.2: Verification of average user rate approximation in the biased and unbiased HetNet using maximum ABRP connectivity, where $\beta = 10\,\text{dB}$.

Beyond this point, deploying extra small BSs does not give any gain in terms of coverage probability or average user rate although they consume energy. Note however that network densification enhances the average network sum rate if these added small BSs cover users.

Figs. 6.1-6.6 also verify the approximations of coverage probability and average user rate for maximum ABRP, maximum i-SINR and minimum BTD connectivity schemes. In all cases, the approximations are very tight and the accuracy increases as the BS density increases since the effect of noise on network performance becomes ever more negligible. Therefore, the approach in this thesis can be used to determine very accurate deployment configurations in dense HetNets.

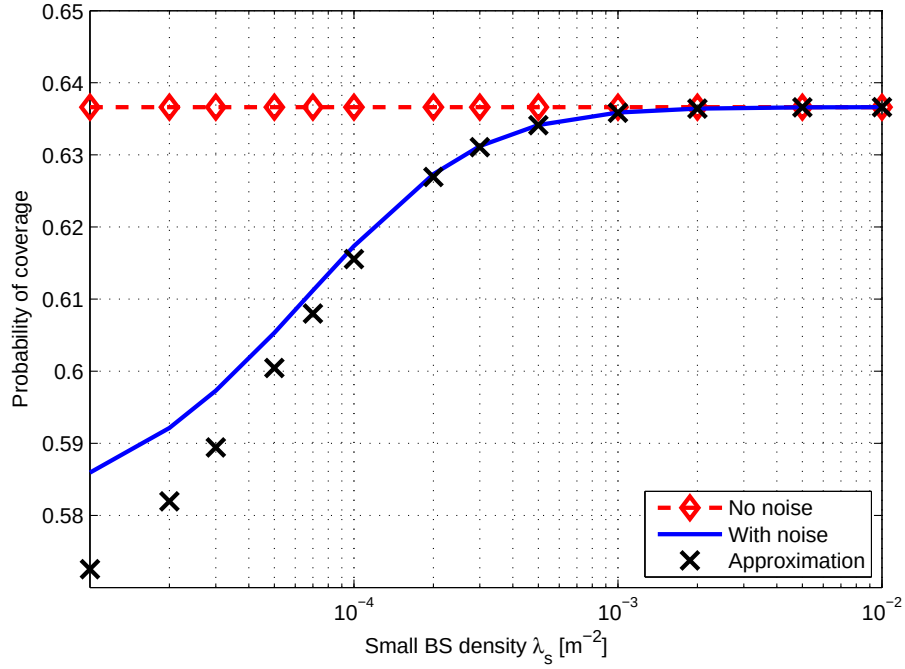


Figure 6.3: Verification of coverage probability approximation in a HetNet using maximum i-SINR connectivity, where ($T = 0\,\text{dB}$).

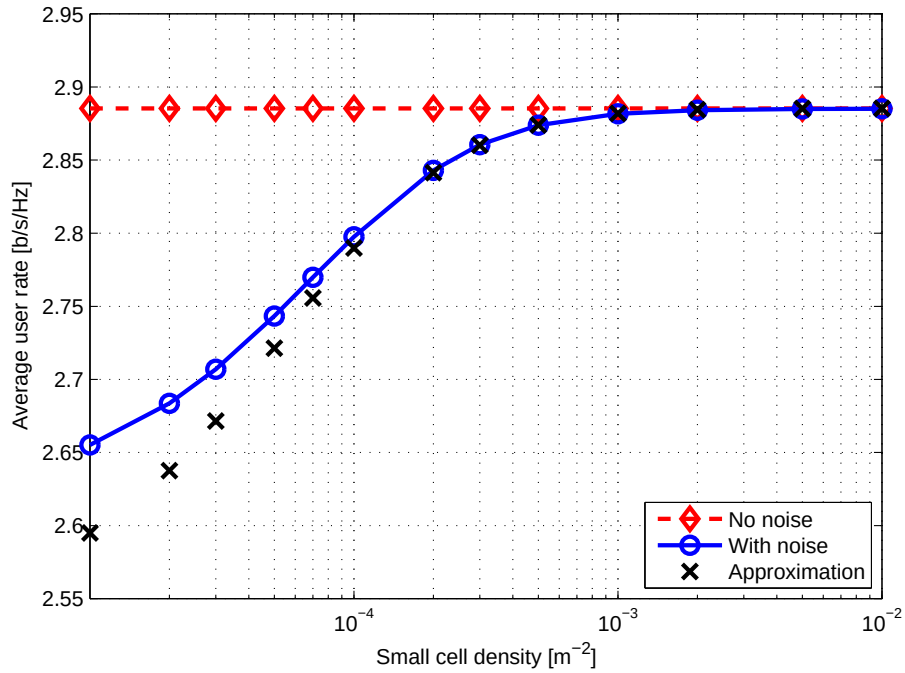


Figure 6.4: Verification of average user rate approximation in a HetNet using maximum i-SINR connectivity.

Consider a HetNet using maximum ABRP connectivity and illustrated by Figs. 6.1-6.2. In both biased and unbiased scenarios, coverage probability and average rate approach their upper bounds $\overline{\mathcal{P}}_c$ and $\overline{\mathcal{R}}_u$ respectively as the HetNet becomes more interference-limited. In the biased HetNet, both coverage probability and average user rate are lower than in the unbiased HetNet because some users are forced to associate with less favorable small BSs which provide lower signal strength than the nearest macro BSs. When the small BS density λ_s is low, increasing λ_s allows more users to be offloaded to the small cell tier. However, since λ_s is still low, the aggregate interference increases while the received signal strength is still weak due to large distances between users and their serving small BSs. Hence, depending on the bias ratio, both

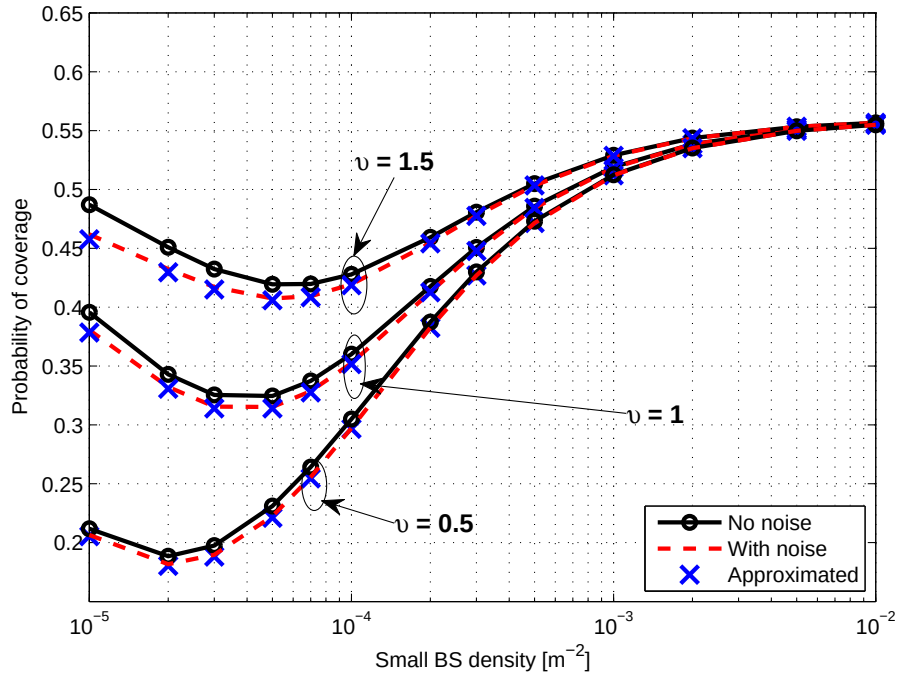


Figure 6.5: Verification of coverage probability approximation in the biased and unbiased HetNet using minimum BTD connectivity.

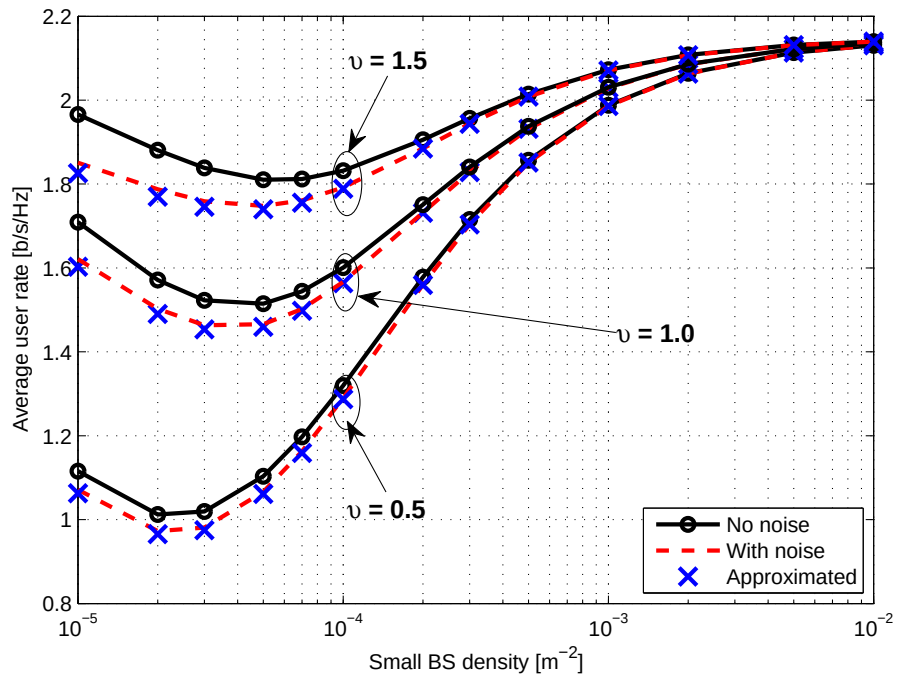


Figure 6.6: Verification of average user rate approximation in the biased and unbiased HetNet using minimum BTD connectivity.

coverage probability and average rate may initially drop. When the small BS density is further increased however, the separation distances between small cell users and their serving BSs reduces to a point where there is a performance gain.

Furthermore, when $\lambda_s \gg \lambda_b$, both coverage probability and average user rate in the biased HetNet approach their respective counterparts in the unbiased HetNet because the small cell tier dominates the macrocell tier so much that it effectively covers most of the network area (i.e. from (3.23), $\mathcal{A}_s \approx 1$). As a result, the network essentially resembles a 'homogenous' network of small BSs and therefore biasing has no effect on its performance.

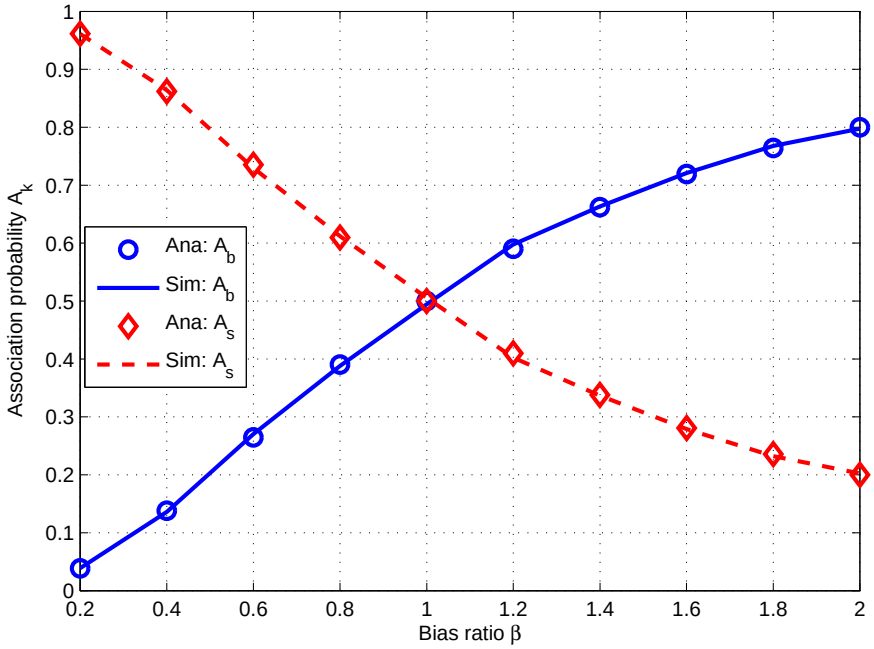


Figure 6.7: Cell association probability in a two-tier biased HetNet using maximum BTD connectivity, where $\lambda_b = \lambda_s$.

Consider a two-tier HetNet using minimum BTD connectivity and whose performance is illustrated by Figs. 6.5-6.6. As opposed to maximum ABRP, both coverage probability and average user rate of the unbiased interference-limited HetNet vary with the number of tiers K, BS densities $\{\lambda_j\}$ and their associated transmit powers $\{P_j\}$. In general, both measures are lower than in maximum ABRP connectivity as a consequence of forcing some users to associate with less favorable small BSs. In other words, even in the absence of biasing, an interference-limited HetNet with minimum BTD connectivity is already *biased* by separation distance. This can further be illustrated by the association probability in Fig. 6.7 which shows that when $\lambda_b = \lambda_s$ in the unbiased scenario, both tiers have equal association probability even though the macrocell tier surely dominates the small cell tier in terms of SINR coverage.

When $\nu > 1$ in Figs. 6.5-6.6, it means that the HetNet is actually biased to favour macro BSs instead of small BSs which explains the performance gain over the unbiased HetNet. However, biasing should always favour small BSs over macro BSs to improve load balancing and enhance the average sum rate of the HetNet. Hence, biasing in the practical range of $\nu < 1$ reduces the coverage probability and average rate of a typical user in the HetNet. However, similar to maximum ABRP connectivity in the range $\lambda_s \gg \lambda_b$, the performance of the biased HetNet approaches that of the unbiased HetNet due to the overwhelming dominance of the small cell tier over the macro tier which nullifies biasing.

Another important aspect of HetNet analysis is its performance in different cellular environments. Although this behavior is the same with all schemes, it is illustrated here using maximum i-SINR connectivity as shown in Figs. 6.8-6.9. In the low λ_s-regime, the aggregate interference is generally low and additive noise has a significant impact on performance. When $\alpha = 4.5$, the aggregate interference is less than when $\alpha = 4$ due to the greater attenuation rate. Hence, compared to the interference-limited HetNet, additive noise has a larger impact on the HetNet performance at $\alpha = 4.5$ than at $\alpha = 4$. Therefore when λ_s is sufficiently low, the HetNet

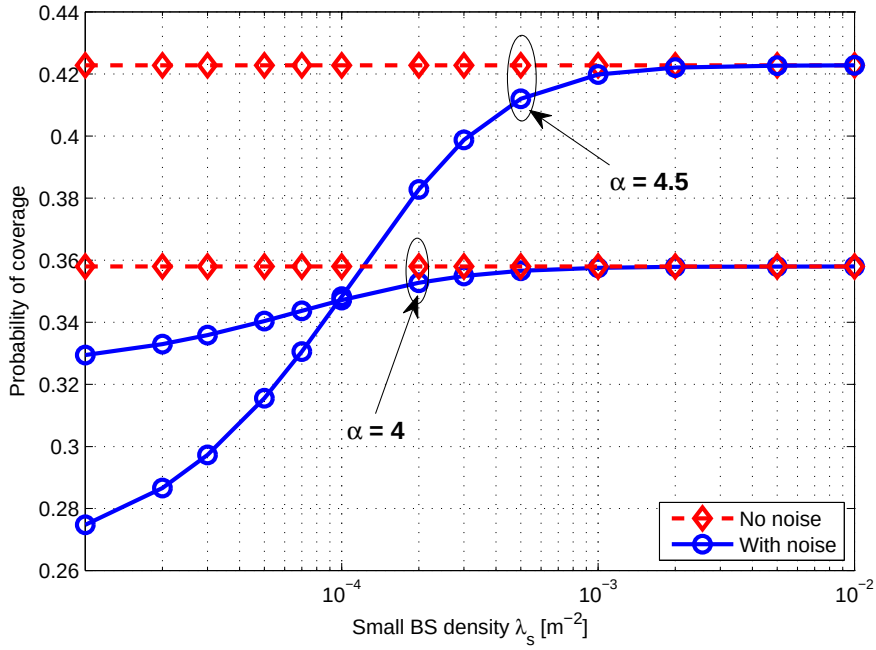


Figure 6.8: Coverage probability at various $\{\lambda_s, \alpha\}$ combinations in a HetNet using maximum i-SINR connectivity, where $T = 5\,\text{dB}$.

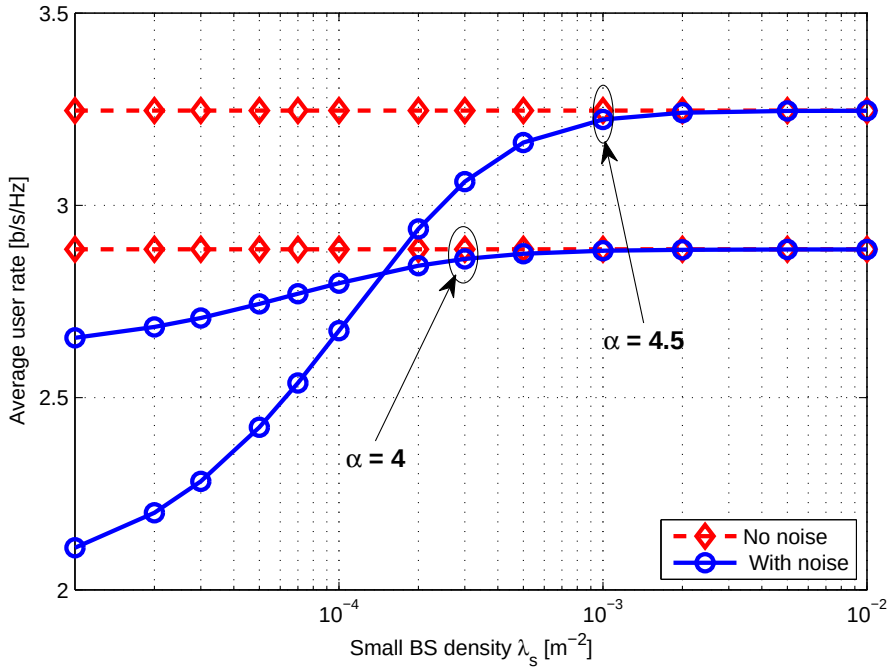


Figure 6.9: Variation of average user rate with small BS density in a HetNet using maximum i-SINR connectivity.

performance is worse off at $\alpha = 4.5$ than at $\alpha = 4$. This basically means that both performance measures reduce as α increases, which is not intuitive. However as λ_s increases, the additive noise gradually loses its impact on SINR and eventually the HetNet performance begins to increase with α. This is an interesting result that further emphasizes the importance of network densification to combat the effect of noise and enhance performance in high pathloss environments such as urban areas.

Average sum rate is highly influenced by the prevailing user density in the HetNet. Fig. 6.10 shows the variation of average sum rate with small BS density at different small BS-user density ratios in a HetNet using maximum ABRP connectivity. Generally for a given user density λ_u,

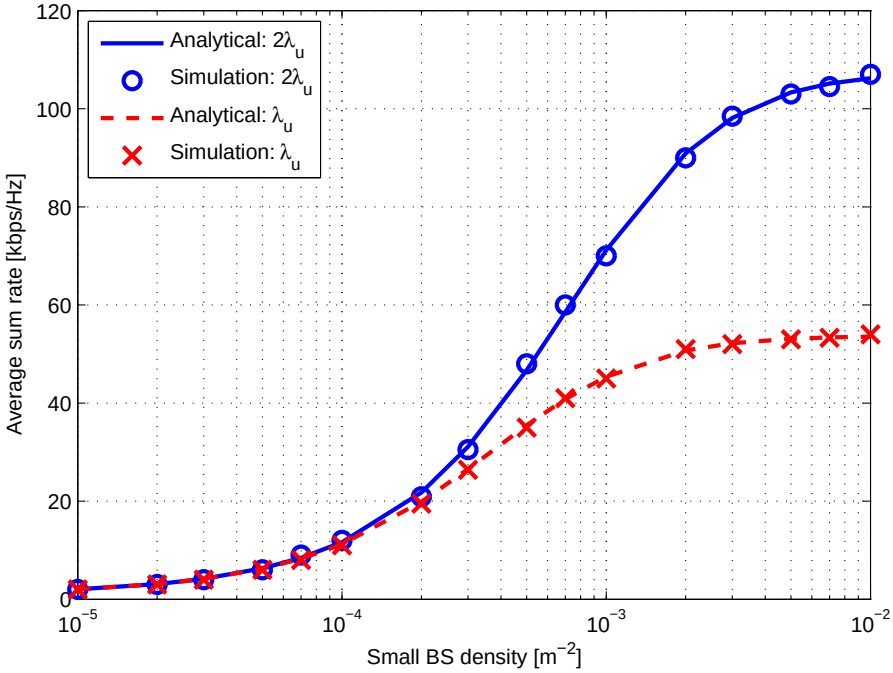


Figure 6.10: Average sum rate of the unbiased HetNet using maximum ABRP scheme.

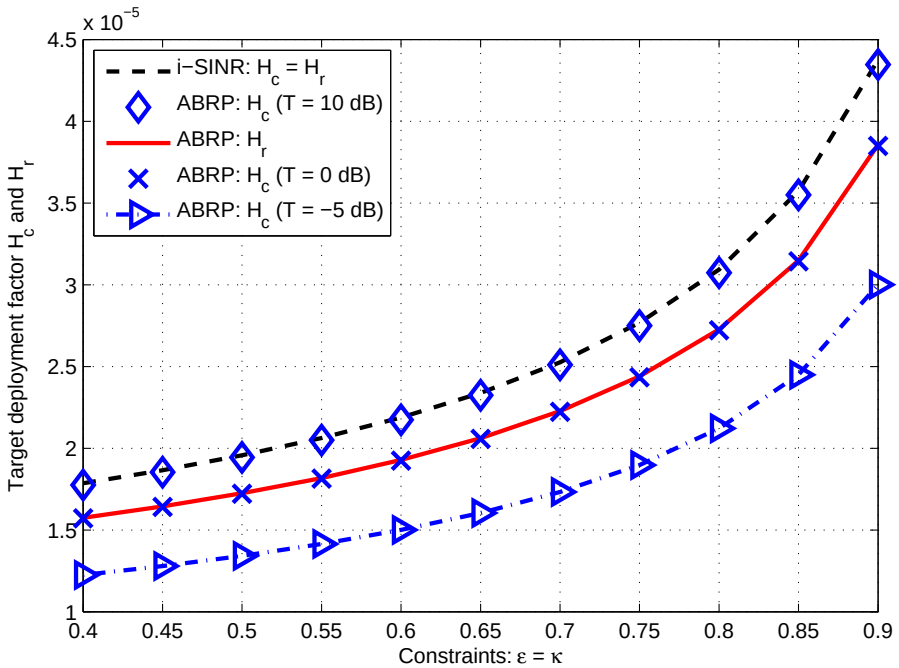


Figure 6.11: Deployment factors of the unbiased HetNet using maximum i-SINR and ABRP connectivity schemes.

average sum rate increases with the small BS density because as more small BSs are deployed, the average number of users per BS in both tiers reduces which enhances the average bandwidth per user. In the low λ_s-regime (where $\lambda_u \gg \lambda_s$), all BSs are likely to contain users and the average sum rate increases almost linearly with the small BS density λ_s. Increasing the user density does not improve the average sum rate in this regime since additional users simply share the same bandwidth with existing users. As λ_s increases further, the average bandwidth per user continues to increase which enhances the average sum rate. In addition, idle BSs begin to emerge as more small BSs are deployed. Increasing the user density reduces the density of these idle BSs which further enhances the average sum rate. However, the increasing interference begins to limit the average sum rate which eventually saturates when $\lambda_s \gg \lambda_u$.

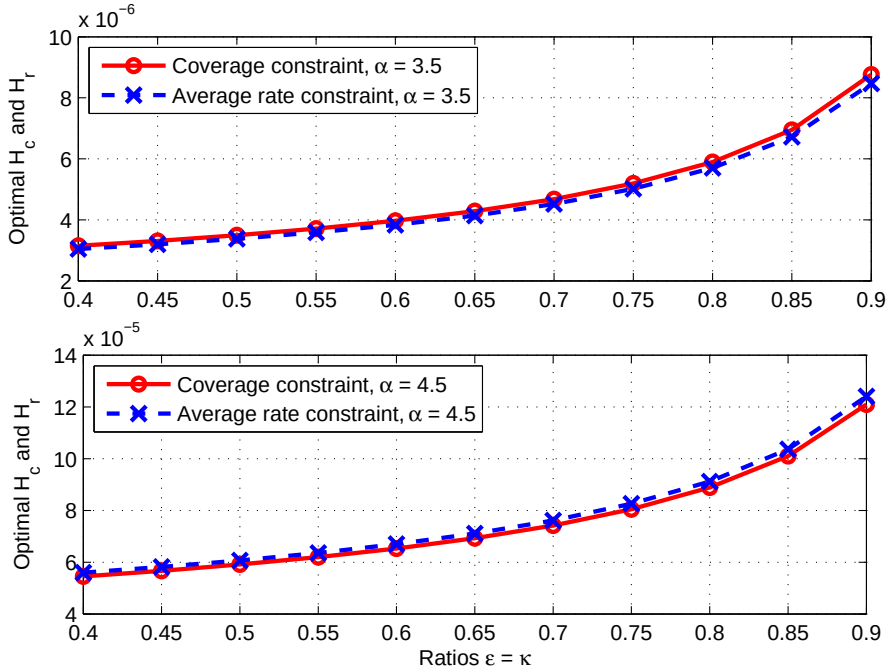


Figure 6.12: Deployment factors $\mathcal{H}_c$ and $\mathcal{H}_r$ of the unbiased K-tier HetNet using maximum ABRP connectivity, for $T = 0\,\text{dB}$.

In general, the HetNet deployment factors $\mathcal{H}_c$ and $\mathcal{H}_r$ increase with their respective constraints ϵ and κ as shown in Fig. 6.11 which considers maximum ABRP and maximum i-SINR schemes. This is intuitive because as ϵ and κ increase, the HetNet is basically required to approach its interference-limited state which requires a higher deployment factor. For maximum i-SINR connectivity, the deployment factors $\mathcal{H}_c$ and $\mathcal{H}_r$ are both independent of T, and are equal for $\epsilon = \kappa$ (see (6.50) and (6.53)). For maximum ABRP connectivity however, $\mathcal{H}_c$ depends on T but $\mathcal{H}_r$ is independent of T. Specifically, $\mathcal{H}_c$ reduces with T since a lower SINR target can be achieved with a lower combination of BS densities and transmit powers. When $T = 0\,\text{dB}$ and $\alpha = 4$, $\mathcal{H}_c$ and $\mathcal{H}_r$ are approximately equal. However, when $T > 0\,\text{dB}$, $\mathcal{H}_c > \mathcal{H}_r$ and when $T < 0\,\text{dB}$, $\mathcal{H}_c < \mathcal{H}_r$. Furthermore, Fig. 6.12 shows that when $\alpha < 4$, $\mathcal{H}_c > \mathcal{H}_r$ and when $\alpha > 4$, $\mathcal{H}_r > \mathcal{H}_c$. This dependence of $\max\{\mathcal{H}_c, \mathcal{H}_r\}$ on the set $\{T, \alpha\}$ also justifies the necessity of using both coverage and average rate constraints during HetNet optimization.

In addition, maximum ABRP connectivity requires a lower deployment factor than maximum i-SINR connectivity at reasonably low of T. For example, the deployment factors $\mathcal{H}_c$ of both schemes are approximately equal when $T = 10\,\text{dB}$ as shown in Fig. 6.11. However, typical values of T are normally set much lower than $10\,\text{dB}$ since reliable QoS can be provided at lower SINR levels. Therefore at practical values of T, maximum ABRP connectivity is the better user association strategy since it minimizes the required deployment factor. In other words, a HetNet with maximum ABRP connectivity requires a lower combination of optimal BS density and transmit power which minimizes the resulting APC.

For minimum BTD connectivity, the range of bias ratio is different and for this reason, its results are presented separately and comparisons are made where possible. Consider a two-tier HetNet with a known macrocell tier deployment factor $\mathcal{H}_b$. Fig. 6.13 shows the variation of the required $\mathcal{H}_{s,c}$ and $\mathcal{H}_{s,r}$ with their respective constraint ratios ϵ and κ. In general, both $\mathcal{H}_{s,c}$ and $\mathcal{H}_{s,r}$ also increase with their respective constraints. When $T = 0\,\text{dB}$ and $\epsilon = \kappa$, $\mathcal{H}_{s,c}$ and

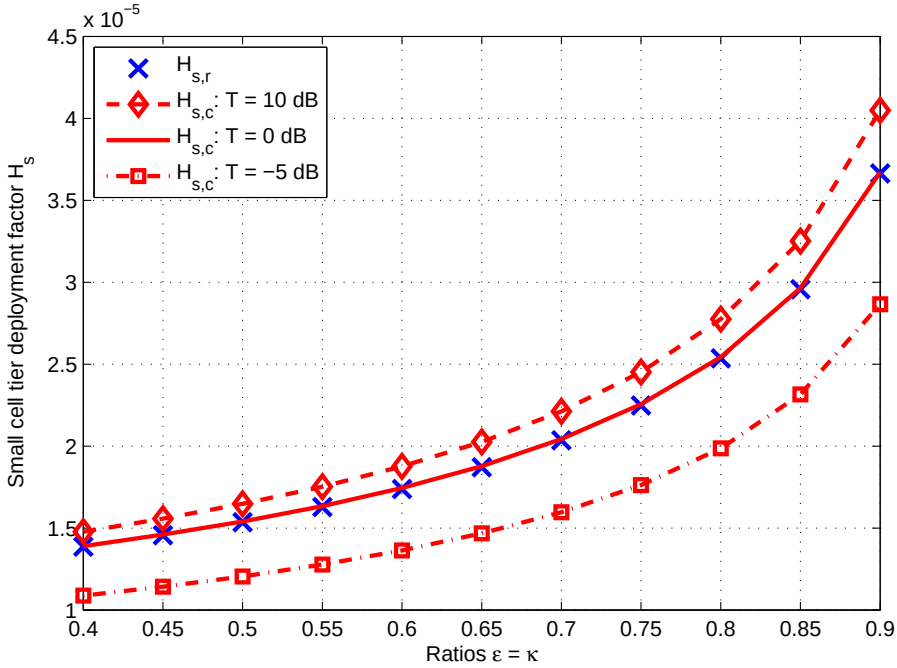


Figure 6.13: Variation of the deployment factors $\mathcal{H}_{s,c}$ and $\mathcal{H}_{s,r}$ with ratios ϵ and κ in the unbiased HetNet using minimum BTD connectivity ($\mathcal{H}_b = 10^{-6}\,\mathrm{m}^{-2}$ and $\alpha = 4$).

$\mathcal{H}_{s,r}$ are equal. However when $T < 0\,\mathrm{dB}$, $\mathcal{H}_{s,r} > \mathcal{H}_{s,c}$ and when $T > 0\,\mathrm{dB}$, $\mathcal{H}_{s,c} > \mathcal{H}_{s,r}$. This response is similar to the case of maximum ABRP connectivity in Fig. 6.11, and also justifies why both constraints are necessary in the optimization.

Given $\mathcal{H}_b$, the effect of biasing on the required small cell tier deployment factor $\mathcal{H}_s$ can be investigated. Fig. 6.14 shows the variation of $\mathcal{H}_{s,c}$ and $\mathcal{H}_{s,r}$ with the bias ratio in a HetNet using maximum ABRP connectivity, where $\epsilon = \kappa$ and $T = \{-3, 0, 5\}\,\mathrm{dB}$. Whereas $\mathcal{H}_{s,r}$ is independent of T, $\mathcal{H}_{s,c}$ is an increasing function of T. It is clear that if T is very large, the overall deployment factor $\mathcal{H}_s = \mathcal{H}_{s,c}$ over the β-range shown in Fig. 6.14. Conversely, if T is small enough, the deployment factor $\mathcal{H}_s = \mathcal{H}_{s,r}$ over the same β-range. In general, $\mathcal{H}_{s,c}$ and $\mathcal{H}_{s,r}$ vary differently with the bias ratio but both show a minimum point over the β-range considered. Their exact minima are influenced by the value of $\mathcal{H}_b$ and the ratios ϵ and κ. As the bias ratio increases, both $\overline{\mathcal{P}}_c$ and $\overline{\mathcal{R}}_u$ reduce accordingly as shown in Fig. 6.15. This initially makes it easier to achieve the performance constraints and both $\mathcal{H}_{s,c}$ and $\mathcal{H}_{s,r}$ reduce as shown in Fig. 6.14. As the bias ratio increases further, the small cell tier begins to dominate the macrocell tier. Eventually a point is reached beyond which both $\overline{\mathcal{P}}_c$ and $\overline{\mathcal{R}}_u$ become approximately invariant with β. Since the HetNet now resembles and behaves like a homogeneous network of small BSs, it requires even more densification to achieve both performance targets. This explains why both deployment factors $\mathcal{H}_{s,c}$ and $\mathcal{H}_{s,r}$ increase in this high β-regime.

For minimum BTD connectivity, Fig. 6.16 shows the variation of $\mathcal{H}_{s,c}$ and $\mathcal{H}_{s,r}$ with the bias ratio $\nu = \nu_s/\nu_b$. Although $\mathcal{H}_{s,c}$ has an increasing relationship with T, $\mathcal{H}_{s,r}$ is independent of T. Therefore, there is a range of T values at which $\mathcal{H}_{s,c}$ and $\mathcal{H}_{s,r}$ are comparable, for instance $T = 0\,\mathrm{dB}$. In general, both $\mathcal{H}_{s,c}$ and $\mathcal{H}_{s,r}$ are minimized at some optimal bias ratios $\nu_c^\star$ and $\nu_r^\star$ respectively. However both $\nu_c^\star$ and $\nu_r^\star$ correspond to the range $\nu_s > 0\,\mathrm{dB}$ which basically means that user association is biased towards the macrocell tier. In the practical biasing range of $\nu < 0\,\mathrm{dB}$, both deployment factors $\mathcal{H}_{s,c}$ and $\mathcal{H}_{s,r}$ generally increase as ν reduces.

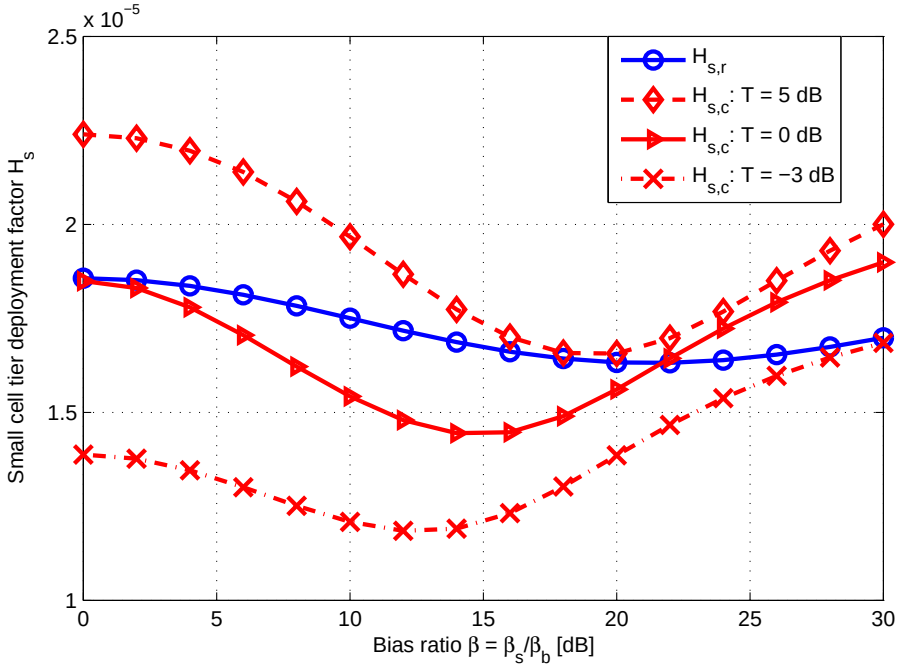


Figure 6.14: Variation of small cell deployment factor $\mathcal{H}_s$ in a biased HetNet (for $\mathcal{H}_b = 2 \times 10^{-5}\,\mathrm{m}^{-2}$ and $\epsilon = \kappa = 0.9$).

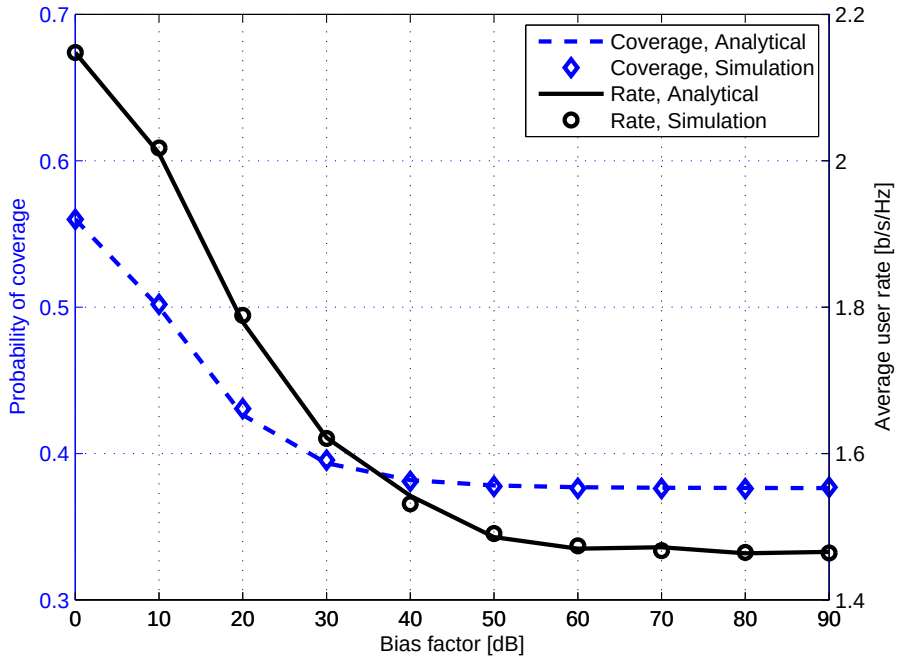


Figure 6.15: $\overline{\mathcal{P}}_c$ and $\overline{\mathcal{R}}_u$ versus bias ratio in a HetNet using maximum ABRP connectivity (for $\mathcal{H}_b = 2 \times 10^{-5}\,\mathrm{m}^{-2}$ and $\lambda_s = 10^{-4}\,\mathrm{m}^{-2}$).

If $\beta^\star$ is the optimal bias ratio at which the overall deployment factor $\mathcal{H}_s = \max\{\mathcal{H}_{s,c}, \mathcal{H}_{s,r}\}$ is minimized, then the APC of the HetNet is also minimized at $\beta^\star$. Fig. 6.17 shows the APC of the HetNet with maximum ABRP connectivity, clearly demonstrating that significant energy savings are possible if the HetNet is operated at its optimal bias point. For instance, at the optimal bias factor $\beta^\star = 20\,\mathrm{dB}$, a power saving of approximately $8.4\,\mathrm{kW}$ is realized in the simulation area compared to the unbiased HetNet. Therefore, in addition to its load balancing potential, biasing can potentially enhance the energy performance of the HetNet by minimizing its required BS density and transmit power configuration per tier.

For minimum BTD connectivity, Fig. 6.18 shows that the optimal bias ratio $\nu^\star$ at which the

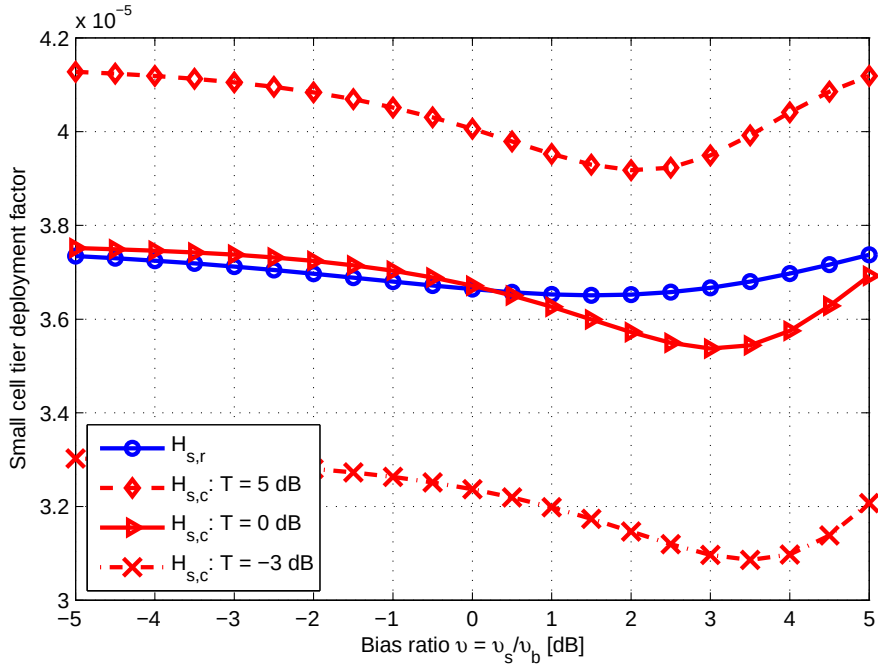


Figure 6.16: Variation of $\mathcal{H}_{s,c}$ and $\mathcal{H}_{s,r}$ with bias ratio in a HetNet using minimum BTD connectivity (for $\mathcal{H}_b = 10^{-6}\,\mathrm{m}^{-2}$, $\epsilon = \kappa = 0.9$ and $\alpha = 4$).

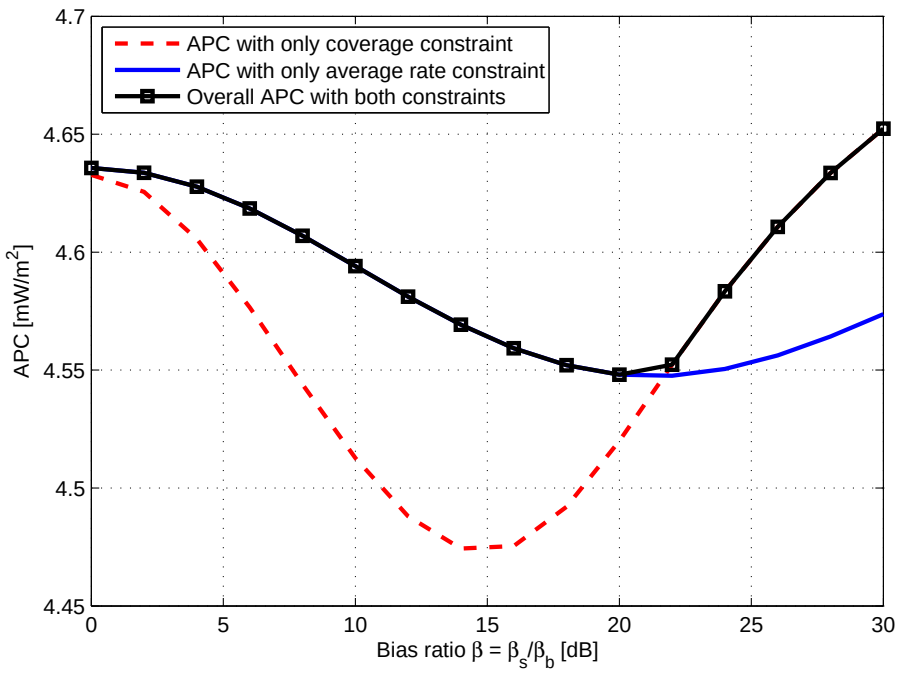


Figure 6.17: Variation of the APC of the biased HetNet with the bias factor (for $\mathcal{H}_b = 2 \times 10^{-5}\,\mathrm{m}^{-2}$, $\epsilon = \kappa = 0.9$ and $T = 0\,\mathrm{dB}$).

APC is minimized is $\nu^\star = 1.5\,\mathrm{dB}$. Considering the practical range of bias values $\nu < 0\,\mathrm{dB}$, the APC of the biased HetNet will generally increase compared to the unbiased HetNet. Therefore, any biasing for load balancing and other purposes has to be traded off for energy consumption. This makes the maximum ABRP connectivity scheme more desirable for biased HetNets.

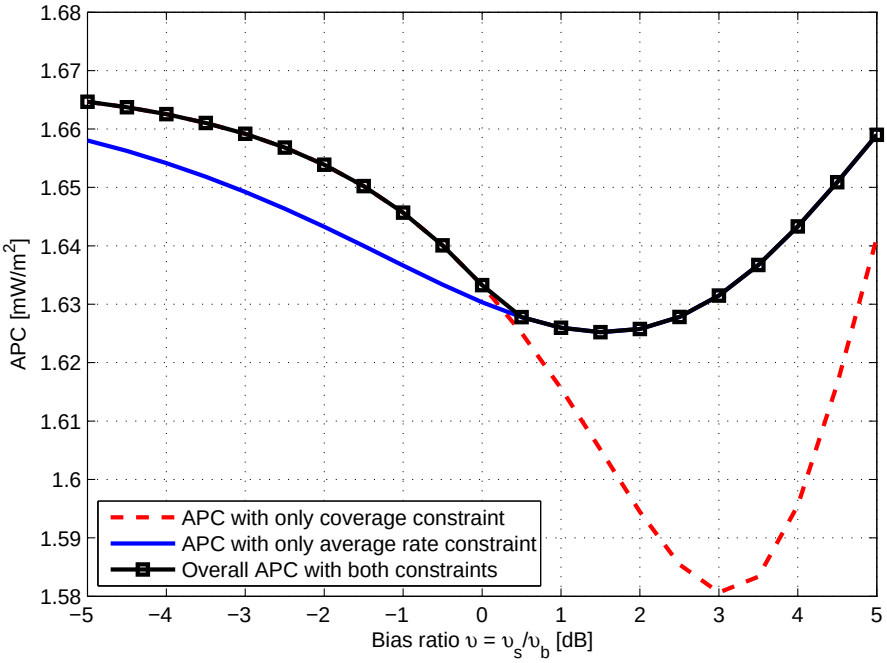


Figure 6.18: APC versus bias factor in a HetNet using minimum BTD connectivity.

6.8 Summary

This chapter analysed the minimum BTD connectivity scheme for general K-tier HetNets. The effect of biasing on the coverage probability and average rate performance of the HetNet using maximum ABRP, maximum i-SINR and minimum BTD connectivity schemes is analyzed and illustrated. This analysis is followed by optimization of the general biased and unbiased HetNet using all three schemes to determine its optimal deployment configuration of BS density and transmit power that minimizes its APC subject to both coverage and average rate constraints. The optimization results show that optimal transmit power is the same irrespective of the constraint applied and can be predetermined if the type of deployed BSs per tier is known. In unbiased HetNets using maximum ABRP and maximum i-SINR connectivity schemes, the deployment factor is expressed in closed form but in biased HetNets, the deployment factor is determined using numerical methods. For minimum BTD connectivity, the HetNet is biased even in the absence of artificial biasing and as a result, its deployment factor is also determined using numerical methods. Although biasing reduces the coverage probability and average rate of users, it can potentially enhance coverage, rate and energy performance if applied appropriately.For instance in a HetNet using maximum ABRP connectivity, analysis shows that appropriate biasing can enable the operator to make significant energy savings by minimizing the deployment factor required to achieve performance targets in all cellular environments. For a HetNet using minimum BTD connectivity scheme however, biasing in the practical range deteriorates energy consumption performance compared to the unbiased HetNet.

Chapter 7

Biased HetNets with Sleep Mode

7.1 Introduction

In a typical HetNet, BSs in different tiers transmit at different power levels and may have different bias values. The average cell size in the k-th tier is influenced by its BS density λ_k, transmit power P_k, and associated bias value β_k relative to other tiers. If the statistical distributions of cell sizes in each tier are known, it is possible to analyze the probabilities of active and idle BSs. This facilitates an investigation of the coverage probability, average rate and EE performance of the HetNet which considers thinned interference due to idle BSs. If idle BSs are put into sleep mode, this conventional scheme ensures that all users remain connected to their parent BSs but the aggregate interference reduces. The application of conventional scheme in homogeneous networks has been discussed in section 5.2.

Cell size distributions in HetNets are more difficult to formulate compared to homogeneous networks. In [115], the authors approximate cell size distributions of a conventional two-tier unbiased HetNet of macro and micro BSs using the gamma distribution. Our main contributions in this chapter are to extend this work and formulate cell size distributions for a general biased two-tier HetNet of macro BSs and small BSs. These distributions facilitate an investigation of the effect of biasing on coverage probability, average rate and EE performance of the HetNet with conventional sleep mode. Furthermore, the multiple user connectivity model is applied to the HetNet which enables an investigation of the ACR performance of the HetNet. The analysis of multiple user connectivity in homogeneous networks has been discussed in subsection 4.2.2.

7.2 User Connectivity in HetNets

The cell sizes and shapes vary greatly since different tiers have different BS densities, transmit powers and bias values. This variability increases with the number of tiers in the HetNet. In general, user association is biased towards low-power and small-coverage BSs which offload traffic from larger BSs in order to achieve load balancing and other performance objectives. Since idle BSs do not transmit, the aggregate interference in the HetNet thins accordingly. For

example, the density of active BSs in the k-th tier is $p_k\lambda_k$ where p_k is the probability of a k-th BS being active. Assume that each BS in the HetNet has access to the same frequency band with a bandwidth of $\mathcal{B}$ Hz. Then each BS has the same number of subchannels, denoted as δ, where each subchannel has a size of $\mathcal{B}_\delta = \mathcal{B}/\delta$. If $\delta > 1$, the HetNet implements the multiple user connectivity model where its resources are shared between all of its users. The single user connectivity model is a special case where each active BS chooses and connects only one user. Therefore, each BS can connect up to a maximum of $N = \delta$ users to avoid intra-cell interference. Denote N_k as the number of users in a typical k-th tier BS. Then, the average connectivity ratio of the k-th tier, denoted as ACR_k, represents the average proportion of all users that get connected by BSs in the k-th tier.

Theorem 7.1. *The overall* ACR *of the HetNet is expressed as*

$$\mathrm{ACR} = \sum_{k=1}^{K} \mathrm{ACR}_k = \sum_{k=1}^{K} \frac{\lambda_k}{\lambda_u}\mathbb{E}[C_k] \tag{7.1}$$

where $C_k \in \{0, \delta\}$ is the number of busy subchannels in a k-th tier BS, and

$$\mathbb{E}[C_k] = \begin{cases} \mathbb{P}(N_k \geq 1), & \delta = 1 \\ \delta\,\mathbb{P}(N_k \geq \delta) + \sum\limits_{k=0}^{\delta-1} k\,\mathbb{P}(N_k = k), & \delta > 1. \end{cases} \tag{7.2}$$

Proof. The average number of connected users in the k-th tier is equivalent to the number of busy subchannels in the tier i.e. $\mathrm{ACR}_k \times \lambda_u = \lambda_k \times \mathbb{E}[C_k]$. Since C_k is a non-negative random variable, its expectation is expressed as

$$\mathbb{E}[C_k] = \sum_{n=1}^{\delta} n\,\mathbb{P}(C_k = n) \quad \text{where} \quad \mathbb{P}(C_k = n) = \begin{cases} \mathbb{P}(N_k = n), & n < \delta \\ \mathbb{P}(N_k \geq \delta), & n = \delta. \end{cases} \tag{7.3}$$

Substituting different values of n in (7.3) shows that $\mathbb{E}[C_k]$ can generally be expressed as shown in (7.2). Overall ACR is a linear sum of the ACRs per tier. ∎

Special Case 7.1. *If a HetNet uses the single user connectivity model (i.e. $\delta = 1$), then $\mathbb{E}[C_k] = \mathbb{P}(N_k \geq 1) = 1 - \mathbb{P}(N_k = 0) = p_k$. Hence,*

$$\mathrm{ACR}_k = \frac{p_k\lambda_k}{\lambda_u}. \tag{7.4}$$

Special Case 7.1 shows that if each k-th tier BS connects a single user, the average number of users connected to the k-th tier is equivalent to its average number of active BSs.

Special Case 7.2. *In a homogeneous network with a BS density of λ_b,* ACR *is expressed as*

$$\mathrm{ACR} = \frac{\lambda_b}{\lambda_u}\mathbb{E}[C_b] \tag{7.5}$$

where $C_b \in \{0, \delta\}$. If $\delta = 1$, then $\mathbb{E}[C_b] = p_a$ where p_a is the probability of a BS being active. Connectivity in a homogeneous network was discussed in subsection 4.2.2.

Intuitively, if a connectivity constraint is imposed on the network, any increase in the user density triggers an appropriate increase in the BS density per tier to maintain the performance target. Denoting ξ as the ACR constraint, the required BS density configuration of the HetNet must satisfy the expression

$$\sum_{k=1}^{K} \lambda_k \mathbb{E}[C_k] \geq \lambda_u \xi. \tag{7.6}$$

Generally, such a connectivity constraint may also be applied per tier to enhance load balancing. If the ACR_k constraint is ξ_k, the required BS density of the k-th tier must satisfy

$$\lambda_k \mathbb{E}[C_k] \geq \lambda_u \xi_k. \tag{7.7}$$

Since ACR is an increasing function of BS density, optimal BS density satisfies the constraint tightly. The optimal BS density per tier is easily obtained using numerical methods.

7.3 Effect of User Density on HetNet Performance

Consider a typical user associated to the nearest k-th tier BS which is located a distance of r away. Considering interference thinning, its received SINR is expressed as

$$\text{SINR}_k(r) = \frac{P_k h_r r^{-\alpha}}{\frac{\sigma^2}{L} + \sum_{j=1}^{K} \sum_{r_j \in \Phi_j \setminus \{B_{k,0} \cup B_{id}\}} P_j h_{r_j} \|R_j\|^{-\alpha}} \tag{7.8}$$

where $B_{k,0}$ is the serving BS and $\{B_{id}\}$ is the set of idle BSs in all tiers of the HetNet. Assume that the HetNet uses maximum ABRP cell association scheme discussed in subsection 3.7.1.

7.3.1 Average Rate

Theorem 7.2. *The average subchannel rate in a typical k-th tier BS is expressed as*

$$\mathcal{R}_{ch,k} = \frac{\mathcal{B}}{\delta \mathcal{A}_k} \pi \lambda_k \, \Xi \left[e^{-\frac{\sigma^2 (2^t - 1)}{P_k L} z^{\alpha/2}} \, e^{-\pi \sum_{j=1}^{K} \lambda_j \widehat{P}_j^{2/\alpha} \mathcal{D}_j(t) \, z} \right] \quad \text{[bits/second]} \tag{7.9}$$

where $\mathcal{D}_j(t) = \widehat{\beta}_j^{2/\alpha} + p_j \mathcal{Z}(t, \alpha, \widehat{\beta}_j)$, $\mathcal{A}_k$ is the k-th tier association probability shown in (3.23) and p_j is the probability of a j-th tier BS being active.

The overall average subchannel rate in the HetNet is expressed as

$$\mathcal{R}_{ch} = \sum_{k=1}^{K} \mathcal{R}_{ch,k} \times \mathcal{A}_k = \sum_{k=1}^{K} \frac{\mathcal{B}}{\delta} \pi \lambda_k \, \Xi \left[e^{-\frac{\sigma^2 (2^t - 1)}{P_k L} z^{\alpha/2}} \, e^{-\pi \sum_{j=1}^{K} \lambda_j \widehat{P}_j^{2/\alpha} \mathcal{D}_j(t) \, z} \right] \quad \text{[bits/second]}. \tag{7.10}$$

In an interference-limited environment, the average subchannel rate simplifies to

$$\overline{\mathcal{R}}_{ch} = \frac{\mathcal{B}}{\delta} \sum_{k=1}^{K} \int_{t>0} \frac{\lambda_k P_k^{2/\alpha}}{\sum_{j=1}^{K} \lambda_j P_j^{2/\alpha} \mathcal{D}_j(t)} \mathrm{d}t. \tag{7.11}$$

Proof. Following the analysis of (3.20)-(3.21), the average rate per subchannel is determined as

$$\mathcal{R}_{ch,k} = \mathcal{B}_\delta \, \mathbb{E}_x[\mathbb{E}_{\mathrm{SINR}_k}[\log_2(1 + \mathrm{SINR}_k(x))]]. \tag{7.12}$$

Other than incorporating interference thinning, the proof is otherwise similar to the proof of Theorem 3.4. The Laplace transform of the thinned interference in this case is expressed as

$$\mathcal{L}_{I_j}(s_r) = \exp\left(-\pi p_j \lambda_j \widehat{P}_j^{2/\alpha} \mathcal{Z}(t, \alpha, \widehat{\beta}_j)\, r^2\right) \tag{7.13}$$

For the interference-limited HetNet, substitute $\sigma^2 = 0$ and solve the resulting integrals. ∎

Assuming full buffer traffic, a typical k-th tier BS with N_k users in its area of coverage will allocate all the subchannels to all users sequentially until they are finished. Two scenarios arise:

1. If $N_k < \delta$, all users are connected and each user is initially allocated $\delta_u = \lfloor \delta / N_k \rfloor$ subchannels. The remaining subchannels $\delta_r = \delta - \delta_u N_k$ are allocated to any δ_r users chosen randomly from the N_k users.

2. If $N_k \geq \delta$, the BS randomly selects δ users and allocates a single subchannel to each i.e. $\delta_u = 1$. The remaining users, equivalent to $N_k - \delta$, remain unconnected.

Lemma 7.1. *The average number of subchannels per user connected to a k-th tier BS is*

$$\omega_k = \frac{p_k \delta}{\mathbb{E}[C_k]}. \tag{7.14}$$

Proof. The total average number of subchannels in all active k-th tier BSs is $\Omega_k = p_k \lambda_k \delta$. The total number of users connected by k-th tier BSs is $\mathrm{ACR}_k \times \lambda_u = \lambda_k \mathbb{E}[C_k]$. ∎

Corollary 7.1. *The average user rate in the HetNet is expressed as*

$$\mathcal{R}_u = \mathcal{B}\pi \sum_{k=1}^{K} \frac{p_k \lambda_k}{\mathbb{E}[C_k]} \Xi \left[e^{-\frac{\sigma^2(2^t-1)}{P_k L} z^{\alpha/2}} e^{-\pi \sum_{j=1}^{K} \lambda_j \widehat{P}_j^{2/\alpha} \mathcal{D}_j(t)\, z} \right] \quad [\text{bits/second}]. \tag{7.15}$$

In an interference-limited environment, the average user rate therefore simplifies to

$$\overline{\mathcal{R}}_u = \mathcal{B} \sum_{k=1}^{K} \frac{p_k}{\mathbb{E}[C_k]} \int_{t>0} \frac{\lambda_k P_k^{2/\alpha}}{\sum_{j=1}^{K} \lambda_j P_j^{2/\alpha} \mathcal{D}_j(t)} \mathrm{d}t. \tag{7.16}$$

Proof. The average rate of a typical user connected to a k-th tier BS is expressed as $\mathcal{R}_{u,k} = \omega_k \times \mathcal{R}_{ch,k}$. Using the law of total probability shown in (3.7), average user rate becomes $\mathcal{R}_u = \sum_{k=1}^{K} \mathcal{R}_{u,k} \times \mathcal{A}_k$. ∎

Corollary 7.2. *The average sum rate of the HetNet is then expressed as*

$$\mathcal{T} = \mathcal{B}\pi \sum_{k=1}^{K} p_k \lambda_k^2 \, \Xi \left[e^{-\frac{\sigma^2(2^t-1)}{P_k L} z^{\alpha/2}} e^{-\pi \sum_{j=1}^{K} \lambda_j \widehat{P}_j^{2/\alpha} \mathcal{D}_j(t)\, z} \right] \quad [\text{bits/sec}]. \tag{7.17}$$

In the interference-limited environment, the average sum rate therefore simplifies to

$$\overline{\mathcal{T}} = \mathcal{B} \sum_{k=1}^{K} \int_{t>0} \frac{p_k \lambda_k^2 P_k^{2/\alpha}}{\sum_{j=1}^{K} \lambda_j P_j^{2/\alpha} \mathcal{D}_j(t)} \mathrm{d}t. \tag{7.18}$$

Proof. The average sum rate of the k-th tier only is $\mathcal{T}_k = \mathcal{R}_{u,k} \times \mathrm{ACR}_k \times \lambda_u = \mathcal{R}_{u,k} \times \lambda_k \, \mathbb{E}[C_k]$. Then, the overall average sum rate is obtained as $\mathcal{T} = \sum_{k=1}^{K} \mathcal{T}_k$. The average sum rate can also be obtained directly from $\mathcal{T} = \mathrm{ACR} \times \lambda_u \times \mathcal{R}_u$. ∎

7.3.2 Coverage Probability

Theorem 7.3. *Coverage probability of a typical user in the HetNet is expressed as*

$$\mathcal{P}_c = \sum_{k=1}^{K} \pi \lambda_k \int_{z>0} e^{-\frac{T\sigma^2}{P_k L} z^{\alpha/2}} e^{-\pi \sum_{j=1}^{K} \lambda_j \widehat{P}_j^{2/\alpha} \mathcal{C}_j \, z} dz \tag{7.19}$$

where $\mathcal{C}_j = \widehat{\beta}_j^{2/\alpha} + p_j \mathcal{Z}(T, \alpha, \widehat{\beta}_j)$ *and* p_j *is the probability of a* j*-th tier BS being active.*

In the interference-limited environment, the coverage probability simplifies to

$$\mathcal{P}_c = \sum_{k=1}^{K} \frac{\lambda_k P_k^{2/\alpha}}{\sum_{j=1}^{K} \lambda_j P_j^{2/\alpha} \mathcal{C}_j}. \tag{7.20}$$

Proof. Apart from incorporating interference thinning, the proof is otherwise similar to that of Theorem 3.3. The Laplace transform of the thinned interference in this case is expressed as

$$\mathcal{L}_{I_j}(s_r) = \exp\left(-\pi p_j \lambda_j \widehat{P}_j^{2/\alpha} \mathcal{Z}(T, \alpha, \widehat{\beta}_j) \, r^2\right). \tag{7.21}$$

∎

7.4 Cell Size Distributions in Biased HetNets

A typical HetNet is characterized by a wide variation of cell sizes and shapes since BSs in different tiers are independently located and transmit at different power levels. Whereas the PPP-based homogeneous network resembles a PV tessellation, the HetNet layout is described by the *weighted* PV tessellation. The distribution of cell sizes in a HetNet are more difficult to formulate. The authors in [115] formulate cell size distributions for a conventional two-tier unbiased HetNet of macro and micro BSs using the gamma function. In this thesis, these approximated cell size distributions are extended to a conventional two-tier biased HetNet of macro BSs and small BSs to facilitate an investigation of the effect of conventional sleep mode on its energy performance.

Consider a PPP-based conventional two-tier HetNet in which the macrocell and small cell tiers are described by the tuples $(\Phi_b, \lambda_b, P_b, \beta_b)$ and $(\Phi_s, \lambda_s, P_s, \beta_s)$ respectively, where Φ_k, λ_k, P_k and

β_k are the respective PPP, BS density, transmit power and bias value of the k-th tier. Assume that the origin $(0,0)$ lies at the cell boundary between the nearest macro BS located at (x_b, y_b) and the nearest small BS located at (x_s, y_s). For a user located at the origin, long-term average biased received power (ABRP) from both the macro BS and small BS is similar i.e.

$$P_r = P_b\beta_b L\|d_b\|^{-\alpha} = P_s\beta_s L\|d_s\|^{-\alpha} \tag{7.22}$$

where $d_b = \sqrt{x_b^2 + y_b^2}$ and $d_s = \sqrt{x_s^2 + y_s^2}$. Hence $d_s = d_b\left(\frac{P_s\beta_s}{P_b\beta_b}\right)^{1/\alpha}$. Thus a macro BS located at (x_b, y_b) is equivalent to a small BS located at $\left(x_b\left(\frac{P_s\beta_s}{P_b\beta_b}\right)^{1/\alpha}, y_b\left(\frac{P_s\beta_s}{P_b\beta_b}\right)^{1/\alpha}\right)$.

Based on maximum ABRP connectivity, the PPP Φ_b of macro BSs can be expressed in terms of its *equivalent* PPP of micro BSs, denoted as $\overline{\Phi}_s$, as $\overline{\Phi}_s = \Phi_b\left(\frac{P_s\beta_s}{P_b\beta_b}\right)^{1/\alpha}$. This equivalent PPP is also homogeneous with the density $\overline{\lambda}_s = \lambda_b\left(\frac{P_b\beta_b}{P_s\beta_s}\right)^{2/\alpha}$ [115], [136]. Since $\overline{\Phi}_s$ is independent of Φ_s, their superposition $\overline{\Phi}_s + \Phi_s$ is still a homogeneous PPP according to Slivnyak's theorem [107]. To a typical user, the HetNet appears homogeneous with an equivalent small BS density of Λ_s, expressed as

$$\Lambda_s = \lambda_b\left(\frac{P_b\beta_b}{P_s\beta_s}\right)^{2/\alpha} + \lambda_s = \frac{\mathcal{S}}{\beta^{2/\alpha}}\lambda_b + \lambda_s \tag{7.23}$$

where $\mathcal{S} = \left(\frac{P_b}{P_s}\right)^{2/\alpha}$ expresses the power ratio of the respective transmit powers of macro BSs to micro BSs while $\beta = \frac{\beta_s}{\beta_b}$ expresses the ratio of the respective bias values of small BSs to macro BSs. If the macrocell and small cell tiers cover the fractions $\mathcal{A}_b$ and $\mathcal{A}_s$ of the network area respectively, A_b and A_s are expressed as

$$A_b = \frac{\mathcal{S}\lambda_b}{\mathcal{S}\lambda_b + \lambda_s\beta^{2/\alpha}} \quad \text{and} \quad A_s = \frac{\beta^{2/\alpha}\lambda_s}{\mathcal{S}\lambda_b + \lambda_s\beta^{2/\alpha}}. \tag{7.24}$$

Hence, the respective average sizes of a typical macrocell and small cell are expressed as

$$\overline{A}_b = \frac{\mathcal{S}}{\mathcal{S}\lambda_b + \lambda_s\beta^{2/\alpha}} \quad \text{and} \quad \overline{A}_s = \frac{\beta^{2/\alpha}}{\mathcal{S}\lambda_b + \lambda_s\beta^{2/\alpha}}. \tag{7.25}$$

For fixed BS densities λ_b and λ_s, (7.24)-(7.25) show that offloading traffic from macro BSs to small BSs can be enhanced by either increasing the bias ratio β or reducing the power ratio $\mathcal{S}$ (i.e. either increasing the transmit power of small BSs or reducing that of macro BSs).

Lemma 7.2. *The cell size distribution of the macrocell tier is expressed as*

$$f_b(x) = \frac{\mathcal{M}_b^{\mathcal{K}_b}}{\Gamma(\mathcal{K}_b)}x^{\mathcal{K}_b - 1}e^{-\mathcal{M}_b x} \tag{7.26}$$

where $\mathcal{M}_b = \frac{\mathcal{K}_b}{\overline{A}_b} = \frac{\mathcal{S}\lambda_b + \beta^{2/\alpha}\lambda_s}{\mathcal{S}}\mathcal{K}_b$, $\mathcal{K}_b = \mathcal{K}\cdot\frac{\lambda_b + k_{b1}\lambda_s}{\lambda_b + k_{b2}\lambda_s}$, $\mathcal{K} = 3.575$ *and* k_{b1} *and* k_{b2} *are constants.*

Similarly for the small cell tier, the cell size distribution is expressed as

$$f_s(x) = \frac{\mathcal{M}_s^{\mathcal{K}_s}}{\Gamma(\mathcal{K}_s)}x^{\mathcal{K}_s - 1}e^{-\mathcal{M}_s x} \tag{7.27}$$

where $\mathcal{M}_s = \frac{\mathcal{K}_s}{\overline{A}_s} = \frac{\mathcal{S}\lambda_b + \beta^{2/\alpha}\lambda_s}{\beta^{2/\alpha}}\mathcal{K}_s$, $\mathcal{K}_s = \mathcal{K} \cdot \frac{\lambda_s + k_{s1}\lambda_b}{\lambda_s + k_{s2}\lambda_b}$ *and* k_{s1} *and* k_{s2} *are constants.*

Proof. The proof is generally similar to that in [115, section IV(B)] which considers a conventional two-tier unbiased HetNet. This proof extends it to a two-tier biased HetNet.

Consider the gamma distribution function $f(x) = x^{k-1}\frac{e^{-x/\theta}}{\theta^k\Gamma(k)}$ where k and θ are the shape and scale parameters respectively. Then, $\mathbb{E}[x] = k\theta$ and $\mathrm{Var}[x] = k\theta^2$. Denote the parameters (k, θ) as $(\mathcal{K}_b, \theta_b)$ for the macrocell size distribution and $(\mathcal{K}_s, \theta_s)$ for the small cell size distribution. Hence from (7.25), $\mathcal{K}_b\theta_b = \overline{A}_b$ and $\mathcal{K}_s\theta_s = \overline{A}_s$. Parameters $(\mathcal{K}_b, \theta_b)$ and $(\mathcal{K}_s, \theta_s)$ depend on the BS density, transmit power and bias value of each tier. It is shown in [115, section IV(B)] that the macrocell and small cell size distributions are respectively expressed as

$$f_b(x) = x^{\mathcal{K}_b - 1}\frac{e^{-(\mathcal{K}_b/\overline{A}_b)x}}{(\overline{A}_b/\mathcal{K}_b)^{\mathcal{K}_b}\Gamma(\mathcal{K}_b)} \quad \text{and} \quad f_s(x) = x^{\mathcal{K}_s - 1}\frac{e^{-(\mathcal{K}_s/\overline{A}_s)x}}{(\overline{A}_s/\mathcal{K}_s)^{\mathcal{K}_s}\Gamma(\mathcal{K}_s)}. \tag{7.28}$$

For the biased HetNet, $\overline{A}_b$ and $\overline{A}_s$ are shown in (7.25). Substituting them into (7.28) and simplifying gives the results of $f_b(x)$ and $f_s(x)$. According to [115], [136], the parameters $\mathcal{K}_b$ and $\mathcal{K}_s$ are expressed in terms of $\mathcal{K} = 3.575$ as shown in (7.26) and (7.27) respectively. The parameters (k_{b_1}, k_{b_2}) and (k_{s_1}, k_{s_2}), which vary with the $\{\mathcal{S}, \beta\}$ set, are obtained numerically using fitting charts. Note the general resemblance between the HetNet distributions $f_b(x)$ and $f_s(x)$ and the cell size distribution $f_A(x)$ of the homogeneous network shown in (4.2). ∎

Since users are distributed according to the homogeneous PPP Φ_u of intensity λ_u, the number of users in a typical cell follows the Poisson distribution shown in (4.3). Therefore, the probability that a typical macrocell contains $N_b = n$ users is expressed as

$$\begin{aligned}\mathbb{P}(N_b = n) &= \int_0^\infty \mathbb{P}[N_b = n | X = x] f_b(x)\mathrm{d}x \\ &\overset{(a)}{=} \frac{\lambda_u^n \mathcal{M}_b^{\mathcal{K}_b}}{\Gamma(\mathcal{K}_b)n!}\int_0^\infty x^{n+\mathcal{K}_b-1}e^{(\lambda_u+\mathcal{M}_b)x}\mathrm{d}x \\ &= \frac{\lambda_u^n \mathcal{M}_b^{\mathcal{K}_b}\Gamma(n+\mathcal{K}_b)}{\Gamma(\mathcal{K}_b)n!(\lambda_u+\mathcal{M}_b)^{n+\mathcal{K}_b}}\end{aligned} \tag{7.29}$$

where (a) is solved using the identity [126, (3.381.4)]. Therefore, the probability p_b that a typical macro BS remains active is expressed as

$$p_b = 1 - \mathbb{P}(N_b = 0) = 1 - \left(\frac{\mathcal{M}_b}{\lambda_u + \mathcal{M}_b}\right)^{\mathcal{K}_b}. \tag{7.30}$$

Similarly, the probability that a typical small cell contains $N_s = n$ users is expressed as

$$\begin{aligned}\mathbb{P}(N_s = n) &= \int_0^\infty \mathbb{P}[N_s = n | X = x] f_s(x)\mathrm{d}x \\ &= \frac{\lambda_u^n \mathcal{M}_s^{\mathcal{K}_s}\Gamma(n+\mathcal{K}_s)}{\Gamma(\mathcal{K}_s)n!(\lambda_u+\mathcal{M}_s)^{n+\mathcal{K}_s}}.\end{aligned} \tag{7.31}$$

Therefore, the probability that a typical small BS remains active is expressed as

$$p_s = 1 - \mathbb{P}(N_s = 0) = 1 - \left(\frac{\mathcal{M}_s}{\lambda_u + \mathcal{M}_s}\right)^{\mathcal{K}_s}. \tag{7.32}$$

The probabilities p_b and p_s thin the aggregate interference suffered by a typical user in the HetNet. This has an effect on the average rate and coverage probability performances of the HetNet as shown in (7.9) and (7.19) respectively. These probabilities are used to analyze the effect of the prevailing user density on the average rate and coverage probability performance of the biased two-tier HetNet.

7.5 Numerical Results

Consider a conventional two-tier HetNet and assume that the macrocell tier BS density is known. This is a realistic assumption if an operator wants to densify an existing homogeneous macrocell network with small BSs for targeted capacity and coverage enhancement. The default parameters used to simulate the HetNet are shown in Table 7.1, unless otherwise stated. In addition, the distribution constants (k_{b_1}, k_{b_2}) and (k_{s_1}, k_{s_2}) of the macrocell and small cell tiers respectively vary with the bias ratio in the HetNet. Table 7.2 shows the distribution constants for bias ratios used in this thesis.

Table 7.1: Parameters used to obtain results

Parameters	**Value**
Network area and bandwidth	$A = 5\,\text{km} \times 5\,\text{km}$, $\mathcal{B} = 20\,\text{MHz}$
User density	$\lambda_u = 1 \times 10^{-3}\text{m}^{-2}$, $\sigma^2 = -110\,\text{dBm}$
Pathloss parameters	$\alpha = 4$, $L = -55\,\text{dBm}$
Macro BS density and bias	$\lambda_b = 5 \times 10^{-6}\,\text{m}^{-2}$, $\beta_b = 0\,\text{dB}$
Macro BS power parameters	$\mathcal{N}_b = 6$, $P_{0,b} = 130\,\text{W}$, $P_b = 20\,\text{W}$, $P_{sl,b} = 75\,\text{W}$
Small BS power parameters	$\mathcal{N}_s = 2$, $P_{0,b} = 56\,\text{W}$, $P_s = 2\,\text{W}$, $P_{sl,s} = 39\,\text{W}$
Additive noise parameters	$F = 10$, $T_a = 300\,\text{K}$

As small BSs are added to the unbiased HetNet, they offload users from the macrocell tier which lowers the average number of users per macrocell and increases the overall ACR of the HetNet as shown in Fig. 7.1. When $\lambda_s \gg \lambda_b$, most users will be served by the small cell tier and the macrocell tier is characterized by very low ACR_b. In the biased HetNet where $\beta_s > \beta_b$, cell association is biased to favor small BSs. Therefore, the offloading of users from the macrocell tier is enhanced which improves the overall ACR of the HetNet. This verifies the suitability of the biasing technique to achieve favorable load balancing in HetNets.

However, improper biasing can have an adverse effect on load balancing and other performance measures of the HetNet. For example, if cell association is excessively biased in favor of small BSs, many macro BSs may remain idle yet the small BSs are not able to connect all users which may negatively impact the ACR performance. Fig. 7.2 shows that there is an optimal

Table 7.2: Distribution constants in a biased two-tier HetNet

Bias ratio [dB]	k_{b_1}	k_{b_2}	k_{s_1}	k_{s_2}
0	0.411	0.167	2.533	5.195
2	0.432	0.139	2.931	5.105
4	0.437	0.135	2.941	5.091
6	0.462	0.121	3.104	4.976
8	0.482	0.101	3.334	4.665
10	0.521	0.092	3.506	4.552
12	0.523	0.089	3.594	4.109
14	0.526	0.087	3.701	3.713
16	0.529	0.082	3.839	3.356
18	0.532	0.079	3.947	3.184
20	0.536	0.075	4.181	2.947
22	0.551	0.069	4.202	2.923
24	0.582	0.062	4.233	2.891
26	0.632	0.053	4.381	2.801
28	0.651	0.041	4.462	2.779
30	0.686	0.043	4.583	2.755

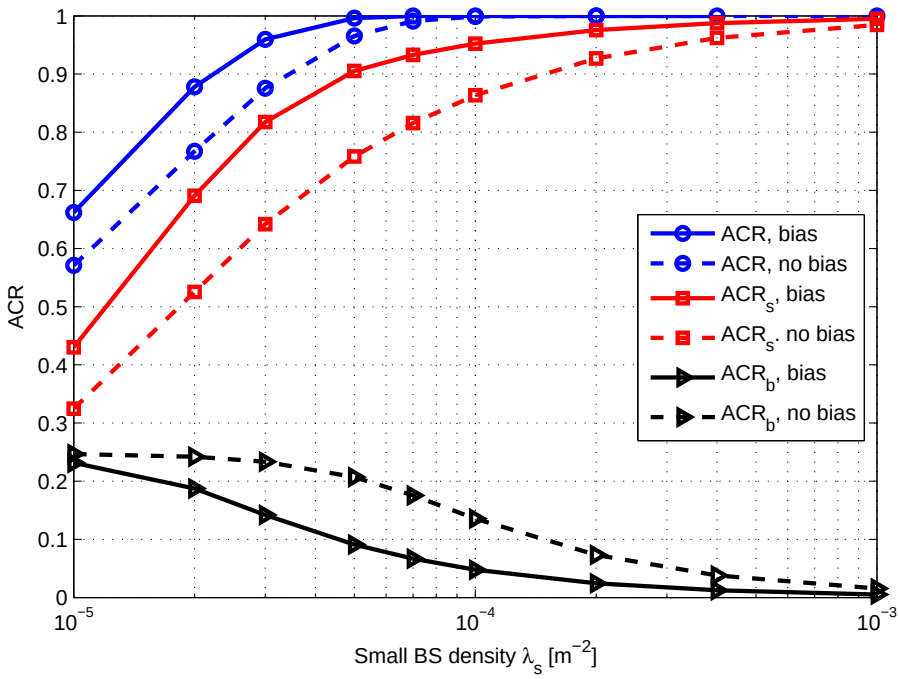


Figure 7.1: Variation of ACR with small BS density ($\lambda_u = 2 \times 10^{-3}\,\mathrm{m}^{-2}$, $\beta = 10$).

bias ratio at which the ACR performance is maximized. This optimal bias ratio and the ACR value achieved are both influenced by the prevailing user density. In general, if the small BS density is sufficient to connect all users, then any adverse effects of excessive biasing cannot be investigated using the ACR measure. Other measures such as average sum rate can also be used to investigate the effect of excessive biasing.

Intuitively, the effect of noise is to reduce the average subchannel rate in both biased and unbiased HetNets. As the small BS density increases, the network gradually becomes interference-limited and the effect of noise is negligible. When $\lambda_u \gg \{\lambda_b, \lambda_s\}$ (in the low λ_s range), then the probability of a macro BS and small BS being active are respectively approximated as $p_b \approx 1$ and $p_s \approx 1$. In this user density regime, the average subchannel rate in the unbiased

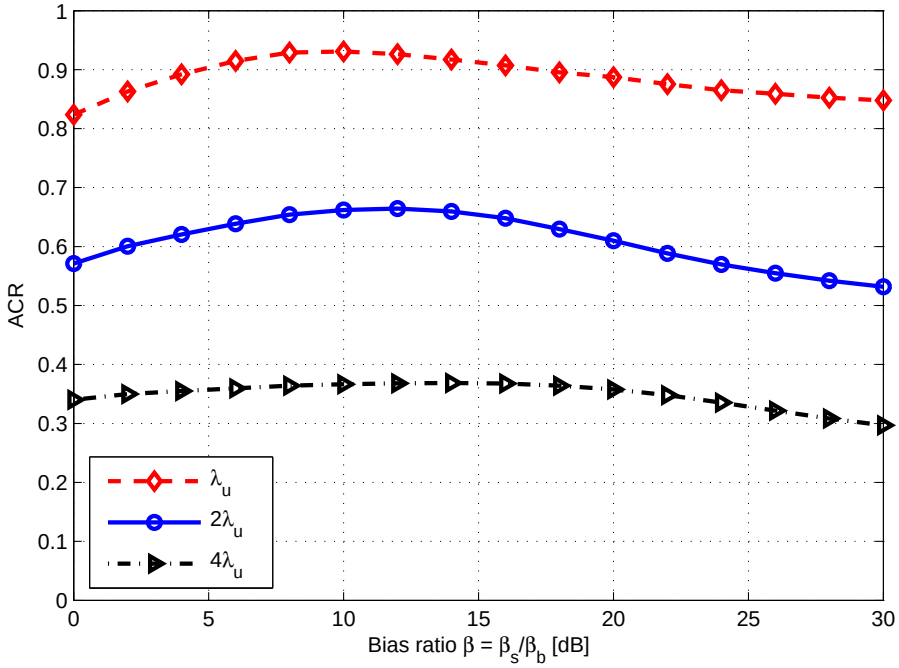


Figure 7.2: Variation of ACR with the bias ratio for $\lambda_s = 10^{-5}\,\mathrm{m}^{-2}$.

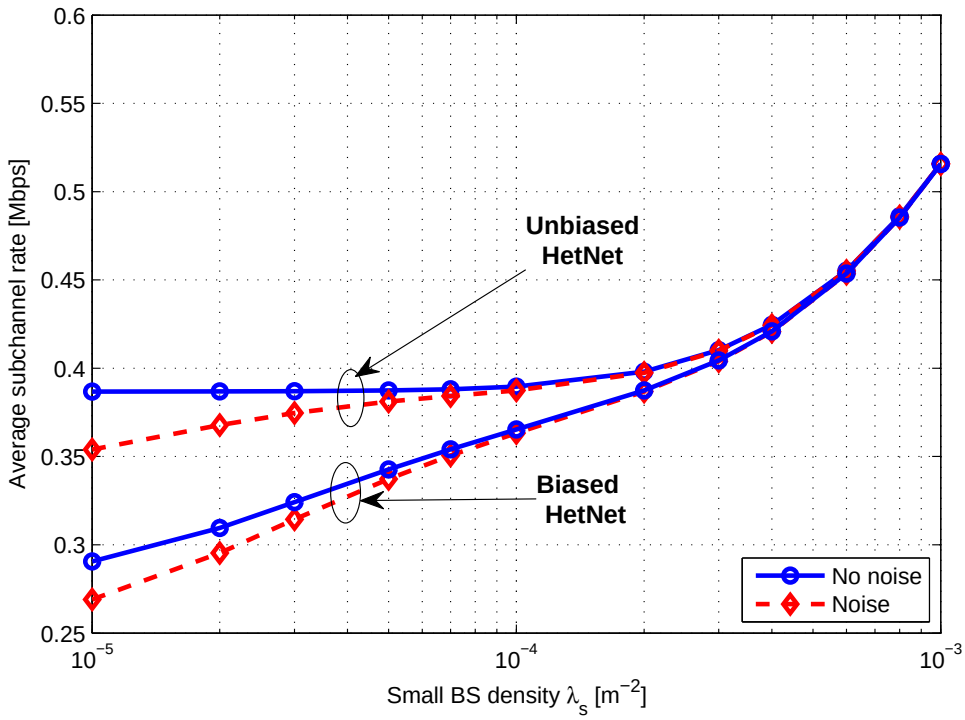


Figure 7.3: Average subchannel rate versus small BS density ($\lambda_u = 10^{-3}\,\mathrm{m}^{-2}$, $\beta = 20\,\mathrm{dB}$).

interference-limited HetNet is invariant with the BS densities as shown in Fig. 7.3. As discussed in [94], this is because any gain in the received signal from additional small BSs is counterbalanced by the additional interference. However, as the small BS density is increased further, the probabilities of idle macro BSs and small BSs eventually become significant which has the effect of thinning out the aggregate interference suffered by a typical user in the HetNet. Since the received signal strength increases with the small BS density, reduced interference increases the SINR and hence the average subchannel rate begins to increase as shown in Fig. 7.3.

In the biased HetNet, the average subchannel rate is lower because some users are forced to connect to small BSs even when macro BSs provide higher received signal strength. As

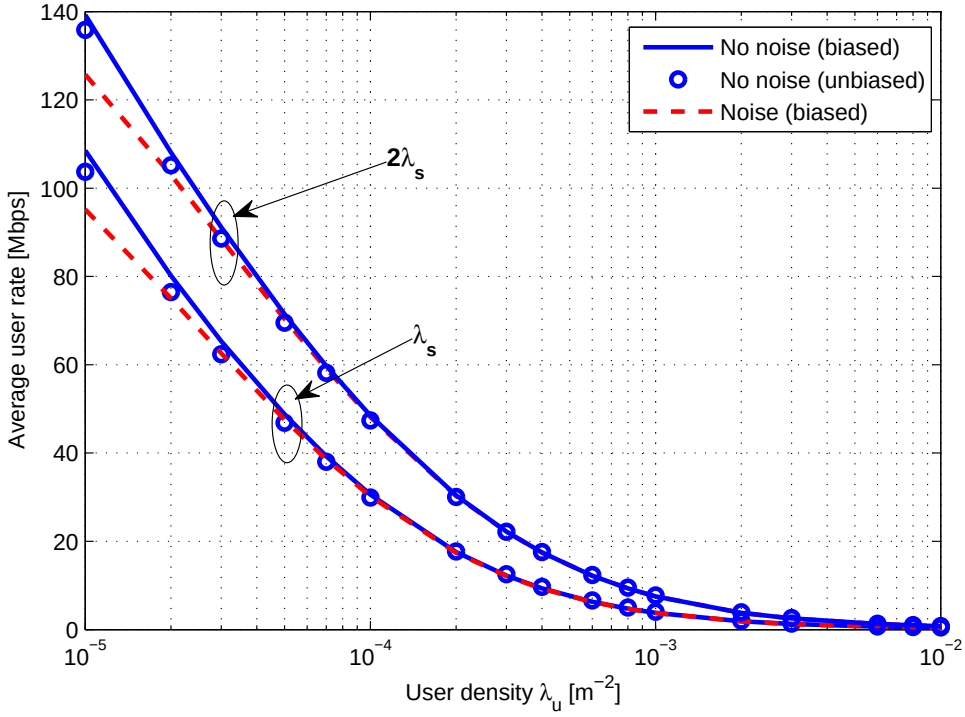


Figure 7.4: Average user rate versus user density ($\lambda_s = 10^{-4}\,\mathrm{m}^{-2}$, $\beta = 20\,\mathrm{dB}$).

a result, the average subchannel rate is no longer invariant but increases with the small BS density because as λ_s increases, distances between users and BSs reduce and the received signal strength provided by small BSs generally increases. As the small BS density increases further, the macrocell tier is gradually overwhelmed by the small cell tier which covers most of the network area. Therefore, the two-tier HetNet essentially resembles a 'homogeneous' network of small BSs. In this regime, biasing has no effect on the average subchannel rate and the performances of biased and unbiased HetNets merge as shown in Fig. 7.3.

When the user density is low, the density of idle BSs in both tiers is significant which greatly thins the aggregate interference and makes the effect of noise more pronounced as shown in Fig. 7.4. In addition, the average number of users per active BS is low which results in a very high average user rate. However, as the user density λ_u increases further, more BSs become active and the average number of users per BS increases. In other words, more users share the bandwidth which reduces the average rate per user. Since more BSs become active, the aggregate interference increases which gradually limits the effect of additive noise and eventually makes the HetNet interference-limited. At very high user density, many small BSs remain active and eventually overwhelm the remaining macro BSs making the HetNet essentially 'homogeneous'. As the effect of biasing becomes negligible, the average user rates of biased and unbiased HetNets eventually become identical.

Interestingly, although biasing reduces the average subchannel rate, it improves the average user rate by offloading some users to small BSs to enhance the overall average bandwidth per user. In addition, the average user rate increases with the small BS density over the whole λ_u-range because deploying more small BSs guarantees more bandwidth per connected user. This means that to maintain a given average user rate, the small BS density should be adaptable to changes in the user density. For example, the required small BS density can easily be obtained from (7.15) using numerical methods.

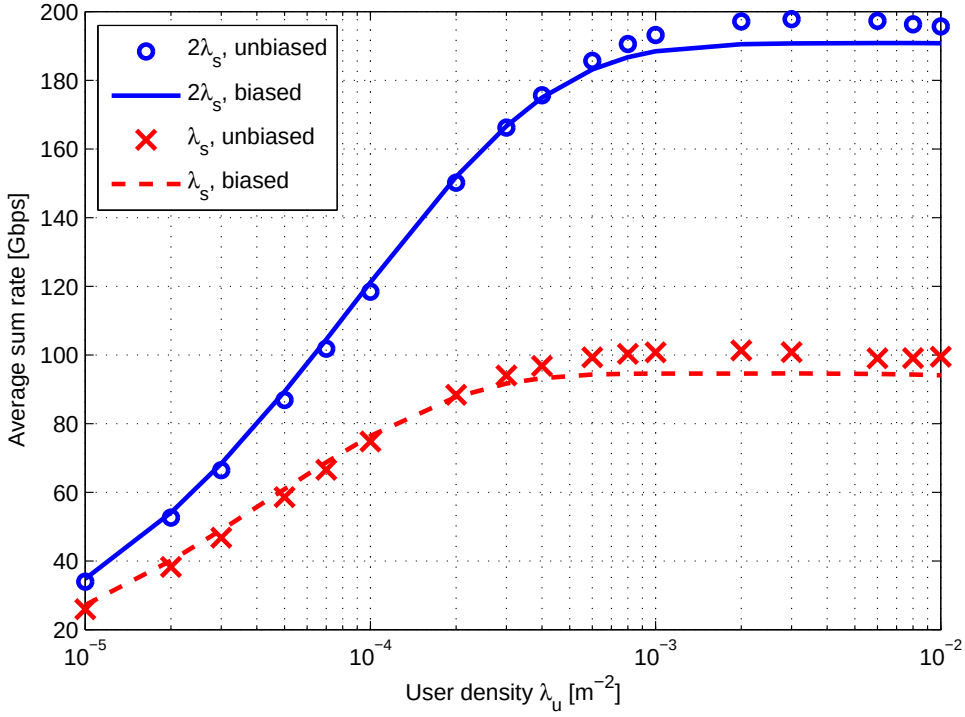


Figure 7.5: Variation of the average sum rate with user density ($\lambda_s = 10^{-4}\,\mathrm{m}^{-2}$, $\beta = 20\,\mathrm{dB}$).

The average sum rate of the HetNet also increases with the user density as shown in Fig. 7.5. This is because more BSs become active which reduces the average number of users per active BS and increases the average bandwidth per user. Since increasing the small BS density connects more users (higher ACR) and each user accesses more bandwidth, the overall average sum rate increases. However when $\lambda_u \gg \{\lambda_b, \lambda_s\}$, all BSs are active and there is no more gain in the average sum rate since any additional users simply share the same bandwidth with existing users. This upper bound on the sum rate is higher in the unbiased HetNet because biasing reduces the average subchannel rate which translates into reduced average sum rate.

In both biased and unbiased HetNets with conventional sleep mode, the EE of the HetNet increases with the small BS density up to a point beyond which it reduces as shown in Fig. 7.6. Although the average sum rate increases with the small BS density, the density of idle BSs in both tiers also increases. Eventually, the aggregate power consumption of BSs in sleep mode begins to adversely affect the EE of the HetNet. Hence, there is an optimal small BS density, denoted as $\lambda_s^\star$, at which the EE is maximized (denote the maximum value of EE as $\mathrm{EE}_{\max}$). If the user density increases, the HetNet achieves a higher $\mathrm{EE}_{\max}$ value and at a higher $\lambda_s^\star$. This is desirable because, besides the higher average sum rate and $\mathrm{EE}_{\max}$ achieved, a higher $\lambda_s^\star$ also guarantees a better ACR performance.

Furthermore, biasing enhances the EE in the low-small BS density range because it ensures that more macro BSs are put into sleep mode by offloading traffic to small BSs. Thus, although biasing reduces the average sum rate, it has a significant advantage in improving the EE of the HetNet. At higher λ_s, the network resembles a 'homogeneous' network of small BSs and biasing has no effect on the average sum rate and EE performance. For comparison purposes, if idle BSs are not put into sleep mode, the EE of both biased and unbiased HetNets reduces as λ_s increases as shown in Fig. 7.6. This is because even though the additional small BSs leave more macro BSs idle, the idle mode power consumption is still high which worsens the EE.

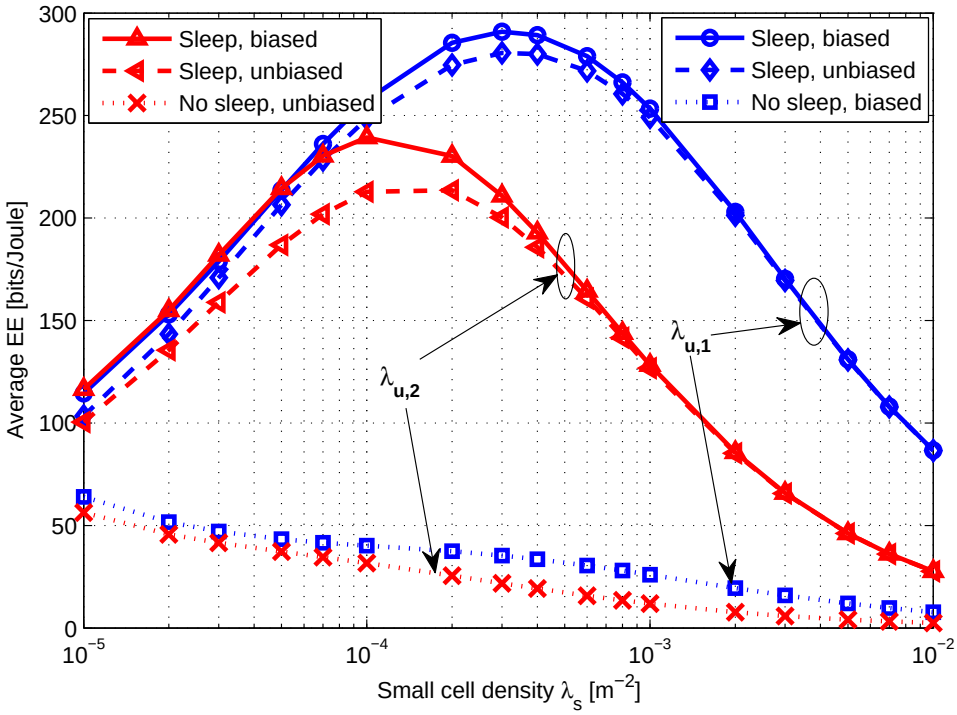


Figure 7.6: Variation of the average EE of the HetNet with the small BS density, where $\lambda_{u,1} = 5 \times 10^{-4}\,\mathrm{m}^{-2}$, $\lambda_{u,2} = 10^{-4}\,\mathrm{m}^{-2}$ and $\beta = 20\,\mathrm{dB}$.

This result emphasizes the importance of implementing sleep mode mechanisms to enhance the EE performance of HetNets.

7.6 Summary

This chapter analyzed user connectivity in HetNets and the effect of the prevailing user density on the realizable HetNet performance. When the combined BS densities in all tiers are not sufficient, the HetNet may get congested and some users will be blocked. Given the user density, the multi-user connectivity model helps to quantify the required BS densities to avoid congestion in the HetNet. Biasing cell association to the small cell tier reduces the average subchannel rate, and consequently the average sum rate, but it enhances the overall ACR as users are offloaded from congesting macro BSs to lightly-loaded small BSs. This offloading mechanism also has the advantage of enhancing the EE performance of the HetNet since more high-power macro BSs are put to sleep. Hence, the average rate performance is generally traded off for improved ACR and EE performance.

Chapter 8

Conclusions and Future Works

8.1 Conclusions

This thesis discussed various aspects that are involved in the design and deployment of energy efficient homogeneous networks and HetNets. Various performance measures such as coverage probability, average rate and energy consumption were investigated using mathematical analysis and Monte Carlo simulations. The analysis used PPPs to model the deployment of BSs and locations of users. This approach models irregular cell size and shapes that fairly similar to the layout of practical deployments. Since cells have highly varied sizes and shapes, the number of users per BS also highly varies and is generally influenced by the size of each cell and the prevailing user density. The multi-user connectivity model enables each BS to connect multiple users within its coverage area which facilitates a realistic study of sleep modes mechanisms by moving affected users to other neighboring BSs. The homogeneous network is optimized to determine the optimal combination of BS density and transmit power that minimizes its APC subject to both coverage and rate constraints. Analysis shows that optimal transmit power only depends on BS power consumption parameters and can therefore be predetermined.

The analysis in this thesis shows that sleep mode mechanisms are realistic solutions to adapt energy consumption to changes in user density. This is important because it ensures that energy is saved when traffic demand is low and significant resources are idle. With conventional sleep mode, all idle BSs are put to sleep to save energy. To reduce BS density further, the basic approach is random sleep mode where a given fraction of BSs are chosen randomly and put to sleep irrespective of their user occupancy. This has the undesired effect of minimizing average sum rate which consequently minimizes network EE. The proposed centralized strategic sleep mode prioritizes BSs with the least number of users to minimize the number of affected users and maximize their chance of getting reconnected in neighboring BSs. This scheme maximizes network EE performance. However this scheme is complex and difficult to manage due to its centralized nature. Distributed strategic sleep mode is proposed where the strategic algorithm is distributed in smaller clusters of BSs. Simulations show that its sub-optimal performance closely matches that of its optimal centralized counterpart.

Generally HetNets are more complex to analyze and optimize because of their multi-tier architecture. The most important challenge involves determining the best combination of BS density and transmit power per tier that achieves a given performance target. In the context of green HetNets, optimization involves determining its deployment configuration per tier that minimizes APC subject to appropriate coverage and rate constraints. Since HetNet performance is highly influenced by the user association scheme, three schemes are studied and compared with each other namely maximum ABRP scheme, minimum BTD scheme and maximum i-SINR scheme. When the HetNet is unbiased, the optimal deployment configuration is expressed in closed form for maximum ABRP and maximum i-SINR schemes. Generally however, deployment configuration is determined using numerical methods. Results show that maximum ABRP scheme has the best APC performance compared to the other schemes.

Biasing in HetNets is an important design technique that enables small BSs to offload traffic from potentially congesting macro BSs. Although these users may connect to less favorable small BSs, the overall average sum rate can potentially be enhanced if small BSs provide the users with more bandwidth. There are potential energy performance benefits as well; for instance, analysis showed that if a HetNet using maximum ABRP scheme is biased appropriately, it further reduces the APC at optimal deployment configuration compared to the unbiased HetNet. However with minimum BTD scheme, biasing in the appropriate range increases the APC compared to the unbiased HetNet. If idle BSs are identified and put to sleep, there is an optimal deployment configuration at which EE is maximized. Moreover, biasing enhances EE of the HetNet for certain combinations of macro BSs and small BSs in a two-tier HetNet.

8.2 Ideas for Future Work

This thesis has extensively studied the coverage probability, average rate and energy performance of homogeneous networks and general HetNets. However, there are several aspects and approaches to cellular network analysis that are not covered by this thesis and can therefore be considered for future works.

1. The work in this thesis is based on several assumptions that simplify analysis and generate more tractable results. The first idea for future work is about relaxing some of these assumptions to understand how they affect the important results on deployment configuration and energy performance of both homogeneous networks and HetNets. The first assumption to be relaxed is considering long-term shadowing in the analysis as explained in subsection 3.3.2 [105], [113], [114], [118].

   Our analysis of sleep mode assumes that each active BS has a full buffer i.e. all its subchannels are fully loaded. This assumption is commonly used in many works because it offers analytical simplicity and tractability and is generally realistic in highly dense environments. However it is rather pessimistic and provides a lower bound on network performance. Some existing works consider more practical traffic profiles as explained in subsection 3.3.3. It would interesting to understand how results improve when a practical traffic profile is considered.

2. Load balancing in HetNets is usually achieved via biasing users to connect to small BSs rather than congesting macro BSs. However optimizing the bias value associated with each tier is a complex problem. Moreover most works assume that the same bias value is associated to all BSs in a given tier. Such fixed biasing may not always achieve the intended load balancing objectives due to the rapid variability of traffic in space and time domains. Moreover the objective of cell association should be to determine an appropriate tradeoff between sum rate and energy performance of the HetNet. Even if bias values are varied periodically to respond to spatiotemporal changes in user density, the overall impact is still difficult to predict.

   Instead of relying on biasing which is difficult to optimize, a new cell association scheme is required where users do not only consider received signal strength (or even SINR) but also existing bandwidth resources at the candidate BSs. By taking into account the *perceived* rate that the user would get given both received signal and available time-frequency bandwidth resources, the sum rate is maximized which consequently enhances EE performance. This also avoids the difficult challenge of determining and optimizing a bias value for each BS to respond to changes in traffic within its local environment.

3. The other proposed research area considers the energy performance of future 5G systems which are meant to increase area SE 1000-fold compared to current 4G systems. To achieve this significant improvement, 5G systems will rely on several state-of-the-art technologies such as massive MIMO and cooperative networking techniques and will most likely use millimeter wave spectrum. Due to the high operating frequency and the required high reuse factor, 5G cells will generally be very small. Hence 5G networks will be characterized by an ultra-dense deployments of small BSs. All the technologies required to enhance capacity and QoS in 5G systems will consume significantly more energy than in current networks. As WiFi networks become ever more ubiquitous, they will aid 5G systems but offloading significant data traffic. Therefore it is necessary to design robust energy saving approaches and techniques such as sleep mode mechanisms and renewable energy sources that can reduce the energy costs of these ultra-dense networks without trading off their capacity provision. In addition, it is important to investigate the effect of WiFi networks on the energy consumption of 5G systems.

References

[1] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2015-2020," Cisco, Tech. Rep., February 2016. [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html

[2] J. Zander and P. Mahonen, "Riding the data tsunami in the cloud: Myths and challenges in future wireless access," *IEEE Communications Magazine*, vol. 51, no. 3, pp. 145–151, March 2013.

[3] S. Mclaughlin, P. M. Grant, J. Thompson, H. Haas, D. Laurenson, C. Khirallah, Y. Hou, and R. Wang, "Techniques for improving cellular radio base station energy efficiency," *IEEE Wireless Communications*, vol. 18, no. 5, pp. 10–17, October 2011.

[4] CISCO, "Digging for the new mobile gold: The next generation of mobile monetization," CISCO, Tech. Rep., May 2014. [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/mobile-internet/white-paper-c11-731757.pdf

[5] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Communications*, vol. 18, no. 3, pp. 10–21, June 2011.

[6] J. Andrews, "Seven ways that hetnets are a cellular paradigm shift," *IEEE Communications Magazine*, vol. 51, no. 3, pp. 136–144, March 2013.

[7] I. Hwang, B. Song, and S. Soliman, "A holistic view on hyper-dense heterogeneous and small cell networks," *IEEE Communications Magazine*, vol. 51, no. 6, pp. 20–27, June 2013.

[8] C. Li, J. Zhang, and K. B. Letaief, "Throughput and energy efficiency analysis of small cell networks with multi-antenna base stations," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2505–2517, May 2014.

[9] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive mimo: Benefits and challenges," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, October 2014.

[10] H. ElSawy and E. Hossain, "Two-tier HetNets with cognitive femtocells: Downlink performance modeling and analysis in a multichannel environment," *IEEE Transactions on Mobile Computing*, vol. 13, no. 3, pp. 649–663, March 2014.

[11] H. ElSawy, E. Hossain, and M. Haenggi, "Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 15, no. 3, pp. 996–1019, Third Quarter 2013.

[12] H. ElSawy, E. Hossain, and D. I. Kim, "Hetnets with cognitive small cells: user offloading and distributed channel access techniques," *IEEE Communications Magazine*, vol. 51, no. 6, pp. 28–36, June 2013.

[13] H. Tabassum, U. Siddique, E. Hossain, and M. J. Hossain, "Downlink performance of cellular systems with base station sleeping, user association, and scheduling," *IEEE Transactions on Wireless Communications*, vol. 13, no. 10, pp. 5752–5767, Oct 2014.

[14] D. Liu, L. Wang, Y. Chen, M. Elkashlan, K.-K. Wong, R. Schober, and L. Hanzo, "User association in 5G networks: A survey and an outlook," *Communications Surveys Tutorials, IEEE*, no. 99, pp. 1–26, January 2016.

[15] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013.

[16] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[17] O. Blume, D. Zeller, and U. Barth, "Approaches to energy efficient wireless access networks," in *4th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, March 2010, pp. 1–5.

[18] SMART2020, "Enabling the low carbon economy in the information age," Smart 2020, Tech. Rep., June 2008.

[19] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M. Imran, D. Sabella, M. Gonzalez, O. Blume, and A. Fehske, "How much energy is needed to run a wireless network?" *IEEE Wireless Communications*, vol. 18, no. 5, pp. 40–49, October 2011.

[20] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, "Cloud RAN for mobile networks - a technology overview," *IEEE Communications Surveys Tutorials*, vol. 17, no. 1, pp. 405–426, March 2015.

[21] J. Wu, Y. Zhang, M. Zukerman, and E. K. N. Yung, "Energy-efficient base-stations sleep-mode techniques in green cellular networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 17, no. 2, pp. 803–826, May 2015.

[22] J. Hoydis, M. Kobayashi, and M. Debbah, "Green small-cell networks," *IEEE Vehicular Technology Magazine*, vol. 6, no. 1, pp. 37–43, March 2011.

[23] R. Razavi and H. Claussen, "Urban small cell deployments: Impact on the network energy consumption," in *IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, April 2012, pp. 47–52.

[24] M. Shakir, K. Qaraqe, H. Tabassum, M.-S. Alouini, E. Serpedin, and M. Imran, "Green heterogeneous small-cell networks: Toward reducing the CO2 emissions of mobile communications industry using uplink power adaptation," *IEEE Communications Magazine*, vol. 51, no. 6, pp. 52–61, June 2013.

[25] E. Hossain and M. Hasan, "5G cellular: key enabling technologies and research challenges," *IEEE Instrumentation Measurement Magazine*, vol. 18, no. 3, pp. 11–21, June 2015.

[26] H. S. Dhillon, Y. Li, P. Nuggehalli, Z. Pi, and J. G. Andrews, "Fundamentals of heterogeneous cellular networks with energy harvesting," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2782–2797, May 2014.

[27] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with hybrid energy harvesting base station," *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, pp. 3412–3427, July 2013.

[28] O. G. Aliu, A. Imran, M. A. Imran, and B. Evans, "A survey of self organisation in future cellular networks," *IEEE Communications Surveys Tutorials*, vol. 15, no. 1, pp. 336–361, First 2013.

[29] H. Hu, J. Zhang, X. Zheng, Y. Yang, and P. Wu, "Self-configuration and self-optimization for LTE networks," *IEEE Communications Magazine*, vol. 48, no. 2, pp. 94–100, February 2010.

[30] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Communications Surveys Tutorials*, vol. 16, no. 4, pp. 1801–1819, November 2014.

[31] 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," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 46–54, June 2011.

[32] A. Fehske, G. Fettweis, J. Malmodin, and G. Biczok, "The global footprint of mobile communications: The ecological and economic perspective," *IEEE Communications Magazine*, vol. 49, no. 8, pp. 55–62, August 2011.

[33] 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," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 66–72, November 2010.

[34] H. Klessig, A. J. Fehske, and G. P. Fettweis, "Energy efficiency gains in interference-limited heterogeneous cellular mobile radio networks with random micro site deployment," in *34th IEEE Sarnoff Symposium*, May 2011, pp. 1–6.

[35] V. Chandrasekhar, J. Andrews, and A. Gatherer, "Femtocell networks: a survey," *IEEE Communications Magazine*, vol. 46, no. 9, pp. 59–67, September 2008.

[36] H. Claussen, L. T. Ho, and L. G. Samuel, "An overview of the femtocell concept," *Bell Labs Technical Journal*, vol. 13, no. 1, pp. 221–245, Spring 2008.

[37] J. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. Reed, "Femtocells: Past, present, and future," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 497–508, April 2012.

[38] D. Chambers, *Femtocell Primer*, 2nd ed. ThinkFemtocell, 2008.

[39] S. R. Saunders, C. Stuart, A. Giustina, R. R. Bhat, V. S. Rao, and R. Siegberg, *Femtocells: Opportunities and Challenges for Business and Technology*, 2nd ed. Wiley, June 2009.

[40] Informa Telecoms and Media, "Femtocell market status," Informa Telecoms and Media, Tech. Rep., 2011.

[41] E. Lang, S. Redana, and B. Raaf, "Business impact of relay deployment for coverage extension in 3GPP LTE-Advanced," in *IEEE International Conference on Communications Workshops*, June 2009, pp. 1–5.

[42] G. Gunter and F. Alagoz, "Green wireless communications via cognitive dimension: an overview," *IEEE Network*, vol. 25, no. 2, pp. 50–56, March 2011.

[43] G. Auer, I. Godor, L. Hevizi, M. A. Imran, J. Malmodin, P. Fazekas, G. Biczok, H. Holtkamp, D. Zeller, O. Blume, and R. Tafazolli, "Enablers for energy efficient wireless networks," in *IEEE 72nd Vehicular Technology Conference Fall (VTC 2010-Fall)*, September 2010, pp. 1–5.

[44] E. Commission, "Digital agenda: Global tech sector measures its carbon footprint," European Commission, Tech. Rep., March 2013. [Online]. Available: http://europa.eu/rapid/press-release_IP-13-231_en.htm

[45] EARTH, "Energy Aware Radio and neTwork tecHnologies," EARTH, Tech. Rep., March 2013. [Online]. Available: https://www.ict-earth.eu/

[46] J. Manner, M. Luoma, J. Ott, and J. Hamalainen, "Mobile networks unplugged," in *Proc. Ist International Conference on Energy Efficient Computing and Networking*, April 2010, pp. 71–74.

[47] L. Suarez, L. Nuaymi, and J.-M. Bonnin, "An overview and classification of research approaches in green wireless networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 142, pp. 1–18, April 2012.

[48] Alcatel-Lucent, "9900 wireless network guardian (white paper)," Alcatel-Lucent, Tech. Rep., August 2012. [Online]. Available: http://www3.alcatel-lucent.com/wps/DocumentStreamerServlet?LMSG_CABINET=Docs_and_Resource_Ctr&LMSG_CONTENT_FILE=White_Papers%2F9900_Wireless_Network_Guardian_EN_Tech_WhitePaper.pdf&REFERRER=lu.pub.rc.pg.whitepapers%20%7C%20White

[49] M. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Optimal energy savings in cellular access networks," in *IEEE International Conference on Communications Workshops*, June 2009, pp. 1–5.

[50] Z. Niu, "TANGO: Traffic-aware network planning and green operation," *IEEE Wireless Communications*, vol. 18, no. 5, pp. 25–29, October 2011.

[51] C. Peng, S.-B. Lee, S. Lu, H. Luo, and H. Li, "Traffic-driven power saving in operational 3g cellular networks," in *Proceedings of the 17th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '11, 2011, pp. 121–132.

[52] Y. Chen, S. Zhang, S. Xu, and G. Y. Li, "Fundamental trade-offs on green wireless networks," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 30–37, June 2011.

[53] G. He, S. Zhang, Y. Chen, and S. Xu, "Fundamental tradeoffs and evaluation methodology for future green wireless networks," in *1st IEEE International Conference on Communications in China Workshops (ICCC)*, August 2012, pp. 74–78.

[54] H. Claussen, I. Ashraf, and L. T. W. Ho, "Dynamic idle mode procedures for femtocells," *Bell Labs Technical Journal*, vol. 15, no. 2, pp. 95–116, September 2010.

[55] A. Conte, "Power consumption of base stations," in *Proc. TREND Plenary Meeting*, February 2012, pp. 1–19.

[56] GreenTouch, "Greentouch technical solutions for energy efficient mobile networks," GreenTouch, Tech. Rep. 1, August 2015. [Online]. Available: https://s3-us-west-2.amazonaws.com/belllabs-microsite-greentouch/uploads/documents/White%20Paper%20GreenTouch%20Mobile%20Technologies%20August%202015%20-%20v1.pdf

[57] H. Claussen, L. T. W. Ho, and F. Pivit, "Effects of joint macrocell and residential picocell deployment on the network energy efficiency," in *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications*, September 2008, pp. 1–6.

[58] Z. Hasan, H. Boostanimehr, and V. K. Bhargava, "Green cellular networks: A survey, some research issues and challenges," *IEEE Communications Surveys Tutorials*, vol. 13, no. 4, pp. 524–540, November 2011.

[59] Huawei, "20,000 hybrid wind/solar-powered base stations," Huawei, Tech. Rep., 2012. [Online]. Available: http://pr.huawei.com/en/media-kit/advertisement/hw-u_202293-wind-solar-hybrid-nature.htm#.Vz2SEkxrjIU

[60] I. Ashraf, F. Boccardi, and L. Ho, "Sleep mode techniques for small cell deployments," *IEEE Communications Magazine*, vol. 49, no. 8, pp. 72–79, August 2011.

[61] W. Guo and T. O'Farrell, "Green cellular network: Deployment solutions, sensitivity and tradeoffs," in *Wireless Advanced (WiAd)*, June 2011, pp. 42–47.

[62] W. Li, W. Zheng, Y. Xie, and X. Wen, "Clustering based power saving algorithm for self-organized sleep mode in femtocell networks," in *15th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, September 2012, pp. 379–383.

[63] I. Ashraf, L. Ho, and H. Claussen, "Improving energy efficiency of femtocell base stations via user activity detection," in *IEEE Wireless Communications and Networking Conference (WCNC)*, April 2010, pp. 1–5.

[64] G. Y. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient wireless communications: Tutorial, survey, and open issues," *IEEE Wireless Communications*, vol. 18, no. 6, pp. 28–35, December 2011.

[65] Y. Zou, J. Zhu, and R. Zhang, "Exploiting network cooperation in green wireless communication," *IEEE Transactions on Communications*, vol. 61, no. 3, pp. 999–1010, March 2013.

[66] C. Xiong, G. Y. Li, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient resource allocation in OFDMA networks," in *IEEE Global Telecommunications Conference (GLOBECOM)*, December 2011, pp. 1–5.

[67] D. Feng, C. Jiang, G. Lim, L. J. Cimini, G. Feng, and G. Y. Li, "A survey of energy-efficient wireless communications," *IEEE Communications Surveys Tutorials*, vol. 15, no. 1, pp. 167–178, February 2013.

[68] M. Li, P. Li, X. Huang, Y. Fang, and S. Glisic, "Energy consumption optimization for multihop cognitive cellular networks," *IEEE Transactions on Mobile Computing*, vol. 14, no. 2, pp. 358–372, February 2015.

[69] L. Hevizi and I. Godor, "Power savings in mobile networks by dynamic base station sectorization," in *IEEE 22nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, September 2011, pp. 2415–2417.

[70] X. Wang, P. Krishnamurthy, and D. Tipper, "Cell sleeping for energy efficiency in cellular networks: Is it viable?" in *IEEE Wireless Communications and Networking Conference (WCNC)*, April 2012, pp. 2509–2514.

[71] Y. S. Soh, T. Quek, M. Kountouris, and H. Shin, "Energy efficient heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 5, pp. 840–850, May 2013.

[72] J. Peng, P. Hong, and K. Xue, "Stochastic analysis of optimal base station energy saving in cellular networks with sleep mode," *IEEE Communications Letters*, vol. 18, no. 4, pp. 612–615, April 2014.

[73] J. Wu, S. Zhou, and Z. Niu, "Traffic-aware base station sleeping control and power matching for energy-delay tradeoffs in green cellular networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 8, pp. 4196–4209, August 2013.

[74] D. Tsilimantos, J.-M. Gorce, and E. Altman, "Stochastic analysis of energy savings with sleep mode in ofdma wireless networks," in *Proceedings of IEEE International Conference on Computer Communications*, April 2013, pp. 1097–1105.

[75] L. Xiang, F. Pantisano, R. Verdone, X. Ge, and M. Chen, "Adaptive traffic load-balancing for green cellular networks," in *22nd IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, September 2011, pp. 41–45.

[76] E. Mugume and D. K. C. So, "Sleep mode mechanisms in dense small cell networks," in *IEEE International Conference on Communications (ICC)*, June 2015, pp. 192–197.

[77] Z. Niu, Y. Wu, J. Gong, and Z. Yang, "Cell zooming for cost-efficient green cellular networks," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 74–79, November 2010.

[78] G. Micallef, P. Mogensen, and H. O. Scheck, "Cell size breathing and possibilities to introduce cell sleep mode," in *European Wireless Conference (EW)*, April 2010, pp. 111–115.

[79] K. Son, H. Kim, Y. Yi, and B. Krishnamachari, "Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1525–1536, September 2011.

[80] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Multiple daily base station switch-offs in cellular networks," in *4th International Conference on Communications and Electronics (ICCE)*, August 2012, pp. 245–250.

[81] D. Cao, S. Zhou, C. Zhang, and Z. Niu, "Energy saving performance comparison of coordinated multi-point transmission and wireless relaying," in *IEEE Global Telecommunications Conference (GLOBECOM 2010)*, December 2010, pp. 1–5.

[82] K. Dufkova, M. Bjelica, B. Moon, L. Kencl, and J. Y. L. Boudc, "Energy savings for cellular network with evaluation of impact on data traffic performance," in *European Wireless Conference (EW)*, April 2010, pp. 916–923.

[83] M. M. Butt, B. Schubert, M. Kurras, K. Brner, T. Haustein, and L. Thiele, "On the energy-bandwidth trade-off in green wireless networks: System level results," in *Communications in China Workshops (ICCC), 2012 1st IEEE International Conference on*, August 2012, pp. 91–95.

[84] V. Rodoplu and T. H. Meng, "Bits-per-joule capacity of energy-limited wireless networks," *IEEE Transactions on Wireless Communications*, vol. 6, no. 3, pp. 857–865, March 2007.

[85] A. J. Fehske, P. Marsch, and G. P. Fettweis, "Bit per joule efficiency of cooperating base stations in cellular networks," in *IEEE Globecom Workshops*, December 2010, pp. 1406–1411.

[86] B. Badic, T. O'Farrrell, P. Loskot, and J. He, "Energy efficient radio access architectures for green radio: Large versus small cell size deployment," in *70th IEEE Vehicular Technology Conference Fall (VTC 2009-Fall)*, Sept 2009, pp. 1–5.

[87] E. Mugume, W. Prawatmuang, and D. K. C. So, "Cooperative spectrum sensing for green cognitive femtocell network," in *IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, Sept 2013, pp. 2368–2372.

[88] F. Cao and Z. Fan, "The tradeoff between energy efficiency and system performance of femtocell deployment," in *7th International Symposium on Wireless Communication Systems (ISWCS)*, Sept 2010, pp. 315–319.

[89] Y. Hou and D. I. Laurenson, "Energy efficiency of high QoS heterogeneous wireless communication network," in *72nd IEEE Vehicular Technology Conference Fall (VTC 2010-Fall)*, Sept 2010, pp. 1–5.

[90] ETSI, "Mobility enhancements in heterogeneous networks," European Telecommunications Standards Institute (ETSI), TS 102.706, August 2009. [Online]. Available: http://www.etsi.org/deliver/etsi_ts/102700_102799/102706/01.01.01_60/ts_102706v010101p.pdf

[91] A. Khandekar, N. Bhushan, J. Tingfang, and V. Vanghi, "LTE-Advanced: Heterogeneous networks," in *European Wireless Conference (EW)*, April 2010, pp. 978–982.

[92] J. G. Andrews, S. Singh, Q. Ye, X. Lin, and H. S. Dhillon, "An overview of load balancing in hetnets: Old myths and open problems," *IEEE Wireless Communications*, vol. 21, no. 2, pp. 18–25, April 2014.

[93] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "User association for load balancing in heterogeneous cellular networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 6, pp. 2706–2716, June 2013.

[94] H.-S. Jo, Y. J. Sang, P. Xia, and J. Andrews, "Heterogeneous cellular networks with flexible cell association: A comprehensive downlink SINR analysis," *IEEE Transactions on Wireless Communications*, vol. 11, no. 10, pp. 3484–3495, October 2012.

[95] J. B. Rao and A. O. Fapojuwo, "Analysis of spectrum efficiency and energy efficiency of heterogeneous wireless networks with intra-/inter-RAT offloading," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 7, pp. 3120–3139, July 2015.

[96] E. Mugume and D. K. C. So, "Capacity and energy efficiency analysis of dense HetNets with biasing," in *26th IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, September 2015.

[97] E. Mugume, D. K. C. So, and E. Alsusa, "Energy efficient deployment of dense heterogeneous cellular networks," in *IEEE Global Communications Conference (GLOBECOM)*, December 2015, pp. 1–6.

[98] 3GPP, "Mobility enhancements in heterogeneous networks," 3rd Generation Partnership Project (3GPP), TR 36.839, September 2012. [Online]. Available: http://www.3gpp.org/dynareport/36839.htm

[99] S. Sadr and R. S. Adve, "Handoff rate and coverage analysis in multi-tier heterogeneous networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 5, pp. 2626–2638, May 2015.

[100] K. Okino, T. Nakayama, C. Yamazaki, H. Sato, and Y. Kusano, "Pico cell range expansion with interference mitigation toward LTE-Advanced heterogeneous networks," in *IEEE International Conference on Communications Workshops*, June 2011, pp. 1–5.

[101] T. Nakamura, S. Nagata, A. Benjebbour, Y. Kishiyama, T. Hai, S. Xiaodong, Y. Ning, and L. Nan, "Trends in small cell enhancements in LTE Advanced," *IEEE Communications Magazine*, vol. 51, no. 2, pp. 98–105, February 2013.

[102] Q. Han, B. Yang, G. Miao, C. Chen, X. Wang, and X. Guan, "Backhaul-aware user association and resource allocation for energy-constrained HetNets," *IEEE Transactions on Vehicular Technology*, vol. PP, no. 99, pp. 1–13, March 2016.

[103] A. de Domenico, V. Savin, and D. Ktenas, "A backhaul-aware cell selection algorithm for heterogeneous cellular networks," in *IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, Sept 2013, pp. 1688–1693.

[104] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Prentice Hall PTR, 2001.

[105] A. Shojaeifard, K. A. Hamdi, E. Alsusa, D. K. C. So, and J. Tang, "A unified model for the design and analysis of spatially-correlated load-aware hetnets," *IEEE Transactions on Communications*, vol. 62, no. 11, pp. 1–16, November 2014.

[106] J. G. Andrews, F. Baccelli, and R. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122–3134, November 2011.

[107] D. Stoyan, W. Kendall, and J. Mecke, *Stochastic Geometry and Its Applications*, 2nd ed. John Wiley and Sons, 1996.

[108] H. Dhillon, R. Ganti, F. Baccelli, and J. Andrews, "Modeling and analysis of k-tier downlink heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 550–560, April 2012.

[109] M. Haenggi, J. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1029–1046, September 2009.

[110] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes*, 2nd ed. Verlag, New York: Springer, 2003.

[111] F. Baccelli and B. Blaszczyszyn, *Stochastic Geometry and Wireless Networks.* NOW Publishers, 2010.

[112] M. Haenggi and R. K. Ganti, *Interference in Large Wireless Networks.* NOW Publishers, 2009.

[113] H. S. Dhillon and J. G. Andrews, "Downlink rate distribution in heterogeneous cellular networks under generalized cell selection," *IEEE Wireless Communications Letters*, vol. 3, no. 1, pp. 42–45, February 2014.

[114] H. P. Keeler, B. Baszczyszyn, and M. K. Karray, "Sinr-based k-coverage probability in cellular networks with arbitrary shadowing," in *IEEE International Symposium on Information Theory*, July 2013, pp. 1167–1171.

[115] D. Cao, S. Zhou, and Z. Niu, "Optimal combination of base station densities for energy-efficient two-tier heterogeneous cellular networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 4350–4362, September 2013.

[116] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Coverage and ergodic rate in K-tier downlink heterogeneous cellular networks," in *49th Annual Allerton Conference on Communication, Control, and Computing*, September 2011, pp. 1627–1632.

[117] J. Peng, P. Hong, and K. Xue, "Energy-aware cellular deployment strategy under coverage performance constraints," *IEEE Transactions on Wireless Communications*, vol. 14, no. 1, pp. 69–80, January 2015.

[118] P. Madhusudhanan, J. G. Restrepo, Y. Liu, T. X. Brown, and K. R. Baker, "Downlink performance analysis for a generalized shotgun cellular system," *IEEE Transactions on Wireless Communications*, vol. 13, no. 12, pp. 6684–6696, December 2014.

[119] K. Stamatiou and M. Haenggi, "Traffic management in random cellular networks," in *Information Theory and Applications Workshop (ITA)*, Feb 2014, pp. 1–5.

[120] R. W. Heath and M. Kountouris, "Modeling heterogeneous network interference," in *Information Theory and Applications Workshop (ITA)*, February 2012, pp. 17–22.

[121] K. Hosseini, W. Yu, and R. Adve, "A stochastic analysis of network MIMO systems," *IEEE Transactions on Signal Processing*, vol. PP, no. 99, pp. x–x, x 2016.

[122] S. Biswas, J. Xue, F. Khan, and T. Ratnarajah, "On the capacity of correlated massive MIMO systems using stochastic geometry," in *IEEE International Symposium on Information Theory (ISIT)*, June 2015, pp. 2603–2607.

[123] A. T. Inc., "Noise Figure Measurement Accuracy – The Y-Factor Method, Application Note 57-2," Agilent Technologies Inc., Tech. Rep., February 2014. [Online]. Available: http://cp.literature.agilent.com/litweb/pdf/5952-3706E.pdf

[124] H. S. Dhillon, R. K. Ganti, and J. G. Andrews, "A tractable framework for coverage and outage in heterogeneous cellular networks," in *Information Theory and Applications Workshop (ITA)*, February 2011, pp. 1–6.

[125] E. Pineda and D. Crespo, "Temporal evolution of the domain structure in a poisson-voronoi nucleation and growth transformation," *American Physical Society: Physical Review E*, 2008.

[126] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. ELSEVIER, 2007.

[127] R. L. Burden and J. D. Faires, *Numerical Analysis: 2.1 - The Bisection Method*, 9th ed. Brooks/Cole, Cengage Learning, 2011.

[128] J. G. Andrews, A. Ghosh, and R. Muhamed, *Fundamentals of WiMAX: Understanding Broadband Wireless Networking*. Prentice Hall, 2007.

[129] S. Sarkar, R. Ganti, and M. Haenggi, "Optimal base station density for power efficiency in cellular networks," in *IEEE International Conference on Communications (ICC)*, June 2014, pp. 4054–4059.

[130] J. Weitzen, M. Li, E. Anderland, and V. Eyuboglu, "Large-scale deployment of residential small cells," *Proceedings of the IEEE*, vol. 101, no. 11, pp. 2367–2380, Nov 2013.

[131] X. Wang, P. Krishnamurthy, and D. Tipper, "Cell sleeping for energy efficiency in cellular networks: Is it viable?" in *IEEE Wireless Communications and Networking Conference (WCNC)*, April 2012, pp. 2509–2514.

[132] E. Mugume and D. K. C. So, "Spectral and energy efficiency analysis of dense small cell networks," in *IEEE Vehicular Technology Conference (VTC Spring)*, May 2015, pp. 1–5.

[133] Y. S. Soh, T. Q. S. Quek, and M. Kountouris, "Dynamic sleep mode strategies in energy efficient cellular networks," in *IEEE International Conference on Communications (ICC)*, June 2013, pp. 3131–3136.

[134] S. Lee and K. Huang, "Coverage and economy of cellular networks with many base stations," *IEEE Communications Letters*, vol. 16, no. 7, pp. 1038–1040, July 2012.

[135] J. Yu, Y. Liu, and C. Yin, "Energy-efficient base station deployment in heterogeneous cellular network with QoS constraint," in *IEEE 24th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC Workshops)*, Sept 2013, pp. 62–30.

[136] D. Cao, S. Zhou, and Z. Niu, "Optimal base station density for energy-efficient heterogeneous cellular networks," in *IEEE International Conference on Communications (ICC)*, June 2012, pp. 4379–4383.

[137] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, 2nd ed. Cambridge University Press, 2005.

[138] 3GPP, "Evolved universal terrestrial radio access (E-UTRA) and evolved universal terrestrial radio access network (E-UTRAN)," 3rd Generation Partnership Project (3GPP), TS 36.300, March 2012. [Online]. Available: http://www.3gpp.org/dynareport/36300.htm

[139] Anritsu, "LTE Resources Guide." [Online]. Available: http://www.cs.columbia.edu/6181/hw/anritsu.pdf

[140] C. Mehlfuhrer, M. Wrulich, J. Ikuno, D. Bosanska, and M. Rupp, "Simulating the long term evolution physical layer," in *17th European Signal Processing Conference*, Aug 2009, pp. 1471–1478.

Appendix A

Wireless Communication

Wireless communication is one of the most vibrant sectors in the communications industry today. In comparison to wired communication, wireless communication enables operators to provide reliable and ubiquitous connectivity to subscribers without significant infrastructure costs. Mobile cellular communication in particular has seen tremendous growth over the last decade and this trend is expected to continue in the foreseeable future. However, the wireless channel continues to present a stiff challenge to the reliability of mobile communication.

A.1 The Wireless Channel

The wireless channel between a transmitter and receiver can vary from a line-of-sight (LOS) to a non-line-of-sight (NLOS) channel obstructed by barriers such as buildings, terrain and foliage. Modeling such propagation channels is a very complex task since they are very unpredictable. In most cases, analytical models are rather inaccurate and channel modeling is instead achieved empirically. However, empirical models require measurements in a specific location and may not be consistent in a different environment or after a certain period of time. Support for mobility of users in some mobile wireless systems further complicates the challenge since there is a need to support the handover of ongoing sessions between different BSs. In addition, the speed of motion determines how fast the wireless channel changes [104], [137].

Homogeneous cellular networks are mobile networks that consist of only BSs of the same type. A typical homogeneous cellular network is a deployment of high power macro BSs that are positioned in planned locations to provide coverage and capacity over a defined geographical area. The area covered by a single BS is called a cell. For simulation purposes, cells are often represented as a lattice of hexagonal or square regions with the BS located at the centre of each region. However in real networks, BSs are not located on a grid but their locations are influenced by factors such as user density, capacity demand, propagation environment and acquisition of site leases. Furthermore, BSs may be configured to transmit at different power levels which further contributes to the irregular shapes and sizes of cells in real networks. Moreover, overlap between neighboring cells is essential to support the handover of mobile users in ongoing sessions [104], [137].

The main advantage of this cellular architecture is to facilitate the spatial reuse of the scarce frequency resources throughout the network. However, this introduces the potential for signal interference whose strength is influenced by the frequency reuse factor and network optimization effort. This interference may be intra-cell (between users in the same cell) or inter-cell (between users in different cells which use the same frequency band). In practice, interference is managed by a combination of proper planning of BS locations, signal processing techniques in the receiver, and continuous optimization of the network through frequency planning, changing transmit powers, antenna tilts and azimuths, etc [104], [137].

When a transmit antenna broadcasts an electromagnetic wave towards one or more receive antennas, the wave is likely to suffer any of the three basic mechanisms namely reflection, diffraction and scattering. Reflection occurs when a signal impinges on an object or surface whose dimensions are much larger than its wavelength such as walls and buildings, cars, surface of the earth, etc. At the point of impact between the incident signal and reflecting surface, the signal may totally be reflected but some of it may be absorbed or transmitted into the reflecting object (this is considered lost). Therefore, the amount of signal loss due to reflection depends on the nature of the reflecting surfaces and how often reflection occurs before the signal reaches the receiver [104], [137].

Scattering is a phenomenon that occurs when a signal wave travels in a medium that consists of irregular objects whose size is small compared to the signal wavelength. For example, a rough wall or foliage may cause significant scattering of the signal such that multiple signal components with reduced signal strength spread out and travel in numerous directions [104]. On the other hand, diffraction may occur when the transmission path is obstructed by a surface or object with sharp edges. The signal waves bend over the edges and propagate into the space behind the object even though a LOS path does not exist. However, signal strength deteriorates significantly as the diffracted signal propagates deeper into the shadowed region [104].

In cellular networks, most transmitter-receiver relationships are NLOS, especially in urban areas where high-rise buildings cause significant signal reflection and diffraction. In such an environment, the question of where to locate the BS is very important but difficult to solve. In general, signal strength deteriorates with the transmission distance and other losses add onto this distance-dependent loss. Propagation mechanisms tend to force the signal to follow multiple paths of varying lengths and channel responses. These multipath components reach the receiver at different times with different amplitudes, phase angles and angles of arrival. The instantaneous sum of the received multipath components highly varies over short time periods since they add constructively and destructively. This phenomenon called multipath fading or small-scale fading causes deep fades over small distances (fraction of a wavelength) and/or time durations (on the order of milliseconds) [104], [137].

A.2 Modeling Radio Channels

In order to understand the impact of the wireless channel on a transmitted signal, attempts are made to design analytical or empirical models that can describe this impact [104]. Analytical

models are developed purely based on mathematical principles that describe the propagation of the wave. For empirical models, actual field measurements are carried out and curves or mathematical equations are formulated to recreate the measured results. As a result, empirical models are more accurate than analytical models because they take into account all factors, whether known and unknown. However, the application of an empirical model at frequencies or environments other than those used to formulate it may be invalid unless new supporting measurements are carried out in the new environment and/or at the new frequency [104].

A.3 Large Scale Modeling

Large scale models try to predict the mean signal strength at a given distance from the transmitter. In general, they express the received power as a function of the transmit power, associated transmit and receive antenna gains and the pathloss (all values in dB); thus

$$P_r(d) = P_t + G_t + G_r - PL(d) \tag{A.1}$$

where $P_r(d)$ is the received power at a propagation distance d in metres, P_t is the transmit power, G_t is the transmit antenna gain, G_r is the receive antenna gain and $PL(d)$ is the pathloss over the distance d. Since the other parameters are known, large scale modeling concentrates on finding the pathloss component [104]-[137]. Empirical and analytical models both predict that the mean signal strength reduces logarithmically with transmission distance. For an arbitrary propagation distance d, the mean large scale pathloss can be expressed using a pathloss exponent n as [104]:

$$\overline{PL}(d)[\text{dB}] = \overline{PL}(d_0) + 10n\log_{10}\left(\frac{d}{d_0}\right) \tag{A.2}$$

where d_0 is a close-in reference distance that is determined from measurements near the transmitter and n indicates the rate of increase of pathloss with distance. The pathloss exponent varies depending on the propagation environment; for example $n = 2$ in free space but changes in other environments. Pathloss over the reference distance can be determined from field measurements at distance d_0 [104].

A.4 Log-normal Shadowing

The model in (A.2) predicts that the average pathloss at a distance d from the transmitter is the same in all directions. In reality however, the received power at any two points located the same distance from a transmitter may be vastly different from the average value predicted by (A.2). Measurements have shown that the pathloss $PL(d)$ at a distance d from the transmitter is random and log-normally (normal in dB) distributed about the average value predicted in

(A.2). Therefore, the actual pathloss can be expressed as [104]

$$PL(d)[\text{dB}] = \overline{PL}(d) + X_\sigma = \overline{PL}(d_0) + 10n\log_{10}\left(\frac{d}{d_0}\right) + X_\sigma \tag{A.3}$$

where X_σ is a zero-mean Gaussian-distributed random variable (in dB) with a standard deviation of σ (in dB). The log-normal shadowing model uses the random variable X_σ to account for the random shadowing effects on any particular propagation path.

A.5 Indoor Propagation

Indoor and outdoor propagation differ from each other in several ways. Indoor propagation distances are likely to be much shorter and the clutter is greater and more unpredictable over such short distances [104]. Indoor environments vary from large buildings with open interiors such as auditoriums, to those with more complex structures having numerous interior partitions and floors. Indoor propagation is mainly influenced by the type and layout of the building, construction materials, purpose for which the building is used, etc. Although indoor propagation is also influenced by reflection, diffraction and scattering mechanisms, these mechanisms vary more aggressively. For example, scenarios which may cause large variations in the received power include whether the exterior and interior windows and doors are open or shut, height of the building, location of outdoor or indoor antenna, etc [104].

A.6 Small Scale Modeling

Small scale models characterize the rapid fluctuation (or fading) of the received signal strength over very short distances and/or time durations. Fading is caused by multipath propagation where various components of the same signal take different paths and arrive at the receiver at different times and with different amplitudes and phases [104]. Combining these components at the receiver results into a signal with a highly fluctuating amplitude due to the fact that the components add up constructively and destructively depending on their phase angles. Multipath propagation is most prevalent in an urban environment due to high-rise buildings which reduce any chance of LOS propagation while providing many surfaces for signal reflection [104].

In mobile networks, the relative motion between the mobile terminal and the BS causes an apparent shift in the frequency of each multipath component. This frequency shift, called Doppler shift, is a function of the speed and direction of motion with respect to the direction of arrival of each multipath component. Even if the mobile terminal is stationary, other objects in the channel may be in motion which still induces Doppler shifts. In addition, the components that reach the receiver at different times cause time dispersion (echos) of the signal since they suffer different delays [104].

The envelope of the sum of two quadrature Gaussian noise signals forms a Rayleigh distribution. In small scale fading, the Rayleigh distribution is commonly used to describe the statistical

time varying nature of the received signal envelope of a flat fading channel. In case there is a dominant nonfading component (such as a LOS component) in addition to the random multipath components, the received signal envelope obeys a Rician distribution. At the output of the envelope detector, this nonfading component can be seen as a DC component in the random multipath. As the nonfading component becomes weaker, the Rician faded signal envelope gradually forms a Rayleigh distribution [104].

The multipath power delay profile represents the relative received power as a function of the delay suffered by the multipath signal components with respect to a reference time delay. Small scale fading can be categorized depending on how the channel affects the transmitted signal. This is influenced by signal parameters (e.g. signal bandwidth) and channel parameters (e.g. delay spread). Multipath time delay dispersion of the channel leads to either flat or frequency selective fading [104]. In flat fading, the channel has a constant gain and linear phase response over a bandwidth that is greater than the signal bandwidth. This channel bandwidth, also called the coherence bandwidth, is a statistical measure of the bandwidth over which the channel is essentially *flat*. Therefore, the spectral characteristics of the signal are preserved at the receiver. Flat fading channels are also referred to as narrowband channels because the signal bandwidth is narrower than the coherence bandwidth.

In frequency selective fading, the channel has a constant gain and linear phase response over a bandwidth that is smaller than the signal bandwidth. In other words, coherence bandwidth is smaller than the signal bandwidth. Frequency selective fading causes time dispersion of the information symbols within the channel which distorts the received signal and causes inter-symbol interference (ISI). In the frequency domain, the channel impacts different frequency components of the same signal with different gains. It is difficult to solve frequency selective fading and often requires a complex receiver. Frequency selective channels are also called wide-band channels because the coherence bandwidth is smaller than the signal bandwidth [104]. In orthogonal frequency division multiplexing (OFDM), the wideband channel is divided into many narrowband subcarriers such that each subcarrier suffers flat fading. Flat fading is easier to deal with at the receiver using simple error correction and equalization techniques as opposed to frequency selective fading.

Appendix B

Long Term Evolution – An Overview

LTE technology was developed by 3GPP and was first released in 2009 as LTE Rel-8. It was designed to provide much higher capacity compared to 3G and earlier systems [138]. This is because LTE can use larger frequency bands and higher-order MIMO spatial processing schemes. In theory, LTE can provide 100 Mbps and 50 Mbps on the DL and UL respectively over a 20 MHz channel bandwidth [139]. LTE uses OFDMA on the DL and single carrier frequency division multiple access (SC-FDMA) on the UL.

B.1 Bandwidth Characteristics

In LTE, OFDM splits the bandwidth into many orthogonal subcarriers of 15 kHz each [139]. Each subcarrier is modulated using QPSK, 16-QAM or 64-QAM depending on the SINR received on it. To maintain orthogonality, the symbol duration on each subcarrier is $66.7\,\mu s$ (1/15 kHz). Therefore, each subcarrier can support a symbol rate of 15 ksps (kilo-symbols per second). Assuming 64-QAM modulation, this gives a maximum data rate of 90 kbps per subcarrier. OFDMA then assigns each user a number of subcarriers based on their data requirements and switches off any unassigned subcarriers to save energy and reduce interference [139]. Subcarriers may be grouped in packs of twelve to form one resource block. Assuming a bandwidth of 20 MHz which contains 100 resource blocks (see Table B.1), the maximum throughput expected from this bandwidth is $90 \times 12 \times 100\,\text{kbps} = 108\,\text{Mbps}$.

SC-FDMA is a pre-coded version of OFDMA which has a lower peak-to-average-power ratio (PAPR) compared to OFDM [139]. As opposed to OFDMA where each subcarrier carries unique data, SC-FDMA spreads the data across multiple subcarriers. It is preferred on the UL because the lower PAPR allows a simpler PA design, reduces UE battery drain and improves cell edge performance. However, it requires a complex receiver design which is not a problem on the UL but is impractical on the DL [139].

Table B.1: LTE bandwidths and corresponding number of RBs

Channel BW (MHz)	1.4	3	5	10	15	20
Number of RBs	6	15	25	50	75	100

B.2 Adaptive Modulation and Coding

LTE implements adaptive modulation and coding (AMC) to take advantage of channel conditions and maximize throughput. AMC, also called link adaptation, is a technique where the modulation and coding schemes are varied based on the radio link conditions. Via measurement reports, the UE informs the network of the channel conditions on the DL. As the radio link improves (high SINR), a higher modulation scheme such as 64-QAM is used to maximize the data rate. Similarly, a higher code rate is possible since the probability of bit errors is low at high SINR. If the link is very poor (low SINR), a lower modulation scheme such as QPSK is used because it can withstand higher levels of interference and noise. This case also requires more redundant bits to be transmitted with the data, which lowers the code rate. Below a certain SINR threshold, the user is in outage.

Table B.2: LTE CQI Table

CQI Index	**Modulation**	**Code rate × 1024**	**Efficiency**
0	out of range		
1	QPSK	78	0.1523
2	QPSK	120	0.2344
3	QPSK	193	0.3770
4	QPSK	308	0.6016
5	QPSK	449	0.8770
6	QPSK	602	1.1758
7	16-QAM	378	1.4766
8	16-QAM	490	1.9141
9	16-QAM	616	2.4063
10	64-QAM	466	2.7305
11	64-QAM	567	3.3223
12	64-QAM	666	3.9023
13	64-QAM	772	4.5234
14	64-QAM	873	5.1152
15	64-QAM	948	5.5547

The quality of each radio link is expressed in terms of a Channel Quality Indicator (CQI). Depending on the CQI, a specific modulation scheme and effective code rate are chosen for data transmission. In LTE, AMC must ensure a 10% block error rate (BLER) or less [140]. Based on this figure of merit, it is possible to map the CQI values to the SINR range required to achieve this BLER. LTE defines 15 CQI values. Each CQI is approximately 2 dB from its neighboring CQIs [140]. Table B.2 shows all the 15 CQIs and their corresponding modulation scheme, code rate and efficiency.