ENERGY-EFFICIENT DESIGN OF HETEROGENEOUS CELLULAR NETWORKS USING STOCHASTIC GEOMETRY

JIA CHENLONG

NATIONAL UNIVERSITY OF SINGAPORE

2015

ENERGY-EFFICIENT DESIGN OF HETEROGENEOUS CELLULAR NETWORKS USING STOCHASTIC GEOMETRY

JIA CHENLONG

(*B. Eng. Shandong University*)

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING
NATIONAL UNIVERSITY OF SINGAPORE

2015

Declaration

I hereby declare that this thesis is my original work and it has been written by me in its entirety.

I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Jia Chenlong

17 Dec. 2015

Acknowledgements

First of all, I want to express my sincere gratitude and appreciation to my supervisor Professor Teng Joon Lim for his great support and invaluable guidance throughout my Ph.D course at the National University of Singapore. I have benefitted tremendously from his great professional vision and logical way of thinking. Without his continual advice, direction, encouragement and patience, this thesis would certainly not be possible.

My thanks also go to my colleagues Can Chen, Utku Tefek, Shixin Luo, Shuowen Zhang, Tianyu Song, Xun Zhou, Zheng Guo, Tong Wu, Liang Liu, Yu Wang, Gaofeng Wu, Fan Cheng, Jie Xu, Yong Zeng, Suzhi Bi, Yinghao Guo, Yang Hu, Huanhuan Zheng, Hong Xing, Qian Wang, Mingwei Wu, and many others, for their friendship.

At last, but at most, I wish to express my infinite love and thankfulness to my parents, Xiaolin Li and Yuehua Jia, and my beautiful girlfriend, Zheng Xin, for their caring, support, devotion and unselfish love.

Table of Contents

Summary

Deploying small cells within the existing macro cellular networks brings tremendous improvement in network capacity and also ensures large scale network coverage. However, the increased number of base stations triggers escalation of network energy consumption. The energy-efficient design of heterogeneous cellular networks (HCNs) consisting of different types of base stations therefore has drawn significant attention recently. Due to its accuracy and tractability, stochastic geometry analysis has been widely used as a powerful tool to study HCNs. This thesis aims to provide various stochastic geometry based frameworks to investigate the energy-efficient design of HCNs by reducing power consumption in the downlink.

Firstly, the thesis starts with addressing joint resource partitioning and cell load adaptation design that reduces HCN power consumption. We consider a two-tier HCN with multiple macro and small-cell base stations that can be put into sleep mode to reduce energy cost. With resource partitioning, an entire tier of base stations is muted on a fraction of the transmission resources. Furthermore, cell load adaptation strategy is used to determine the set of users served on the partitioned resources. To jointly analyse resource partitioning and cell load adaptation, a tractable framework is proposed. We use stochastic geometry analysis to characterize network performance by modelling the two tiers of base stations as independent Poisson point processes (PPPs). Based on the tractable throughput characterization, we further solve a non-convex problem to get the optimal resource partitioning and cell load adaptation rule that minimizes network power consumption with throughput constraints. The solution provides valuable guidelines on energy-efficient HCN design.

Then, we investigate the power saving problem in HCNs with adaptive macro base station transmit power. We consider a basic setup of a single macrocell with multiple deployed femtocells. Circular exclusion zones are assumed around each femtocell base station to mitigate inter-tier interference to the macro users. In this case, the macro base station can adjust the transmit power according to its cell load, which depends on the sizes of exclusion zones. To build the relationship between cell load, or equivalently the exclusion zone size, and macro base station transmit power, we adopt appropriate approximations of the stochastic geometry based characterization of cell edge user outage probability constraints. By applying bisection search algorithm, we determine the optimal exclusion zone radius that minimizes the average transmit power of the macro base station.

Lastly, we study how macro base station deployment helps in power reduction in two-tier HCNs with multi-antenna base stations. Although increasing the base station density helps improve quality of service, the increased number of active base stations raises network power consumption and hence reduces energy efficiency. Thus, we propose a strategy to reduce the number of active base stations by deploying macro base stations and at the same time muting some of the small-cell base stations. The base station muting takes inter-tier dependence into consideration where only the small cells located within the macro base station coverage areas are turned off. This inter-tier dependence is realistic but analytically complicated. Moreover, it remains a challenging task to characterize HCN performance by taking multi-antenna beamforming at base stations into consideration. To address these problems, we use stochastic geometry to give a framework that provides tractable spectral efficiency approximations. Solving power minimization problem based on the proposed framework offers guidelines on determining the density and coverage area of macro base stations for the proposed base station deployment scheme.

List of Figures

List of Abbreviations

ABS	Almost Blank Subframe
BF	Beamforming
BS	Base Station
CDF	Cumulative Distribution Function
CF	Characteristic Function
CSCG	Circularly Symmetric Complex Gaussian
CSI	Channel State Information
DoF	Degree of Freedom
eICIC	Enhanced Inter-Cell Interference Coordination
fBS	Femto Base Station
FFR	Fractional Frequency Reuse
HCN	Heterogeneous Cellular Network
HCP	Hardcore Process
HPPP	Homogeneous Poisson Point Process
ICIC	Inter-Cell Interference Coordination
ICT	Information and Communication Technology
i.i.d.	Independent and Identically Distributed
LHS	Left Hand Side
IN	Interference Nulling
LT	Laplace Transform
mBS	Micro Base Station
MBS	Macro Base Station
MGF	Moment Generating Function

PCP	Poisson Cluster Process
PDF	Probability Density Function
PHP	Poisson Hole Process
PL	Path Loss
PPP	Poisson Point Process
PSD	Power Spectral Density
QoS	Quality of Service
RHS	Right hand side
Rx	Receiver
SINR	Signal-to-Interference-Plus-Noise-Ratio
SIR	Signal-to-Interference-Ratio
SNR	Signal-to-Noise-Ratio
TDMA	Time-Division-Multiple-Access
Tx	Transmitter
UE	User Equipment
ZF	Zero Forcing

List of Symbols

Throughout this thesis, scalars are denoted by lower-case letters, vectors are denoted by bold-face lower-case letters, and matrices are denoted by bold-face upper-case letters. Additionally, we use calligraphic fonts to denote sets and blackboard bold fonts to denote probabilities. Also, we define the following symbols:

$\mathbb{C}$	the Complex Number Space
$\mathbb{R}$	the Real Number Space
$\mathbb{R}^x$	the $x \times 1$ Real Vector Space
$A \backslash B$	the Set $\{x \vert x \in A \text{ and } x \notin B\}$
$\vert A \vert$	the Cardinality of Set A
	the Area of Domain A in $\mathbb{R}^2$
$\Pr(A)$	the Probability of Event A
$\Pr(A \vert B)$	the Probability of Event A Conditioned on Event B
$\mathbb{E}_X[\cdot]$	the Operation of Averaging over Random Variable X
$\bar{X}$	the Expected Value of Random Variable X
$\mathcal{CN}(\boldsymbol{x}, \boldsymbol{\Sigma})$	the Distribution of a CSCG Random Vector with
	Mean Vector $\boldsymbol{x}$ and Covariance Matrix $\boldsymbol{\Sigma}$
$\Vert \boldsymbol{x} \Vert$	the Euclidean Norm of a Complex Vector $\boldsymbol{x}$
$\lfloor x \rfloor$	the Floor Function, the Maximum Integer Smaller than or
	Equals to x.
$\boldsymbol{I}_N$	the $N \times N$ Identity Matrix
$\boldsymbol{M}^H$	the Conjugate Transpose of $\boldsymbol{M}$
$\boldsymbol{M}^T$	the Transpose of $\boldsymbol{M}$

Chapter 1

Introduction

The data traffic usage in cellular networks has been experiencing an explosion in recent years. To meet the skyrocketing demands for larger network capacity, a simple yet powerful solution is to increase the density of base stations and hence shrink the cell size in cellular systems [3]. By deploying low-power base stations within the existing macro cellular networks, coverage and capacity can be enhanced in the areas covered by the overlaid small-cells and the overall network coverage can be guaranteed by the macro base stations (MBSs). Therefore, such a heterogeneous cellular network (HCN) architecture has drawn significant research attention and been recognized as a key technology for future 5G wireless networks. More detailed surveys of the recent development and applications of HCNs can be found in [4–9].

In practice, the small-cells are typically deployed over a large area in an ad-hoc manner, and therefore each sub-region within the network will have a different deployment pattern. Due to the random nature of base station deployment, it is too complicated to design a transmission strategy for users in the network by taking the exact interactions among all base station configurations into consideration. Additionally, the average network performance is of interest if we want to have a simple characterization of the entire network. Without a tractable analytical framework, the study of HCNs can only be done through massive Monte Carlo simulations, which are computationally expensive. This is where stochastic geometry comes to the rescue. As a statistical tool, stochastic geometry is powerful in modelling networks with random topologies [10]. Using stochastic geometry for the modelling, analysis and design of HCNs gives accurate and analytically tractable

results [11, 12]. A systematic introduction to stochastic geometry can be found in [13–15].

Stochastic geometry offers a powerful tool in the design and analysis of HCNs, and therefore of future generations of wireless networks. However, it is still necessary to choose a network performance metric to focus on when optimizing HCNs. The increase in base station density proposed for HCNs can easily result in a higher energy usage than that of single-tier systems in use today. Some studies, such as the SMART 2020 report [16], have shown that if energy-efficient technologies were not adopted there would be a dramatic escalation of energy consumption due to rapidly increasing data traffic demands in communication networks. The growing energy cost of information and communication technology (ICT) not only exacerbates global environmental degradation, but also presents a major obstacle to the growth of the telecommunication industry. As a result, green design has been spurred and is emphasized in current and future 5G network designs [17,18]. According to [19], base stations account for a large fraction of energy consumption in cellular networks. For this reason, there has been great momentum to investigate power saving strategies on the base station side. However, there exist intricate trade-offs between base station power consumption and quality of service (QoS) [20]. Consequently, it is practically meaningful but theoretically challenging to investigate the base station power saving problem in HCNs. This thesis is devoted to proposing stochastic geometry based frameworks for analytically investigating the base station power minimization problem in HCNs with QoS constraints.

1.1 Stochastic Geometry Modelling of Heterogeneous Cellular Networks

In HCNs, as depicted in Fig. 1.1, low power small-cell base stations (e.g. micro base station, pico base station, femto base station) are overlaid onto macro networks. Unlike the MBSs deployed by the operators in a well-planned manner

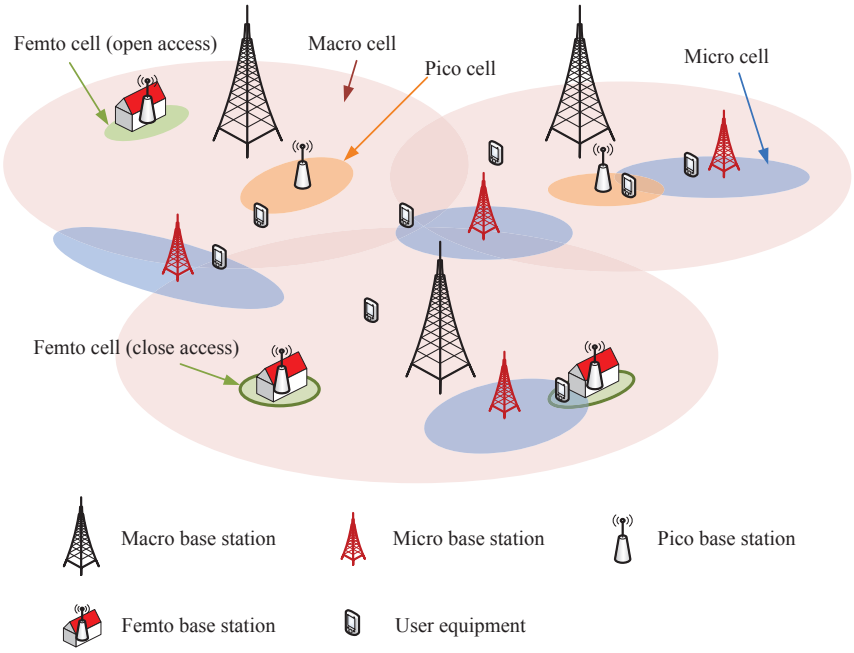


Figure 1.1: The heterogeneous cellular network.

for large area coverage, small-cell base stations (small-cell BSs) are typically placed in a flexible and targeted manner to satisfy local data traffic requirements. Consequently, the use of small-cells brings about additional topological randomness and complicated interactions between different types of nodes. In fact, even for the well-planned macro cellular networks, the real base station placement, as can be seen in Fig. 1.2, is between a deterministic regular grid deployment and an individually independent random deployment. The spatial randomness in HCNs typically makes accurate network simulations computationally challenging, because these will require complete and exact location information for all nodes, as well as three-dimensional topological maps with accurate channel models. Due to these complicated interactions between different nodes in HCNs, studies based on Monte Carlo simulations consume huge amount of time and computation resources. Even if simulations are performed, the results have problems regarding repeatability, which means they cannot be generalized to provide guidance for the large scale network optimal design. As a result, it is useful to conduct analytical

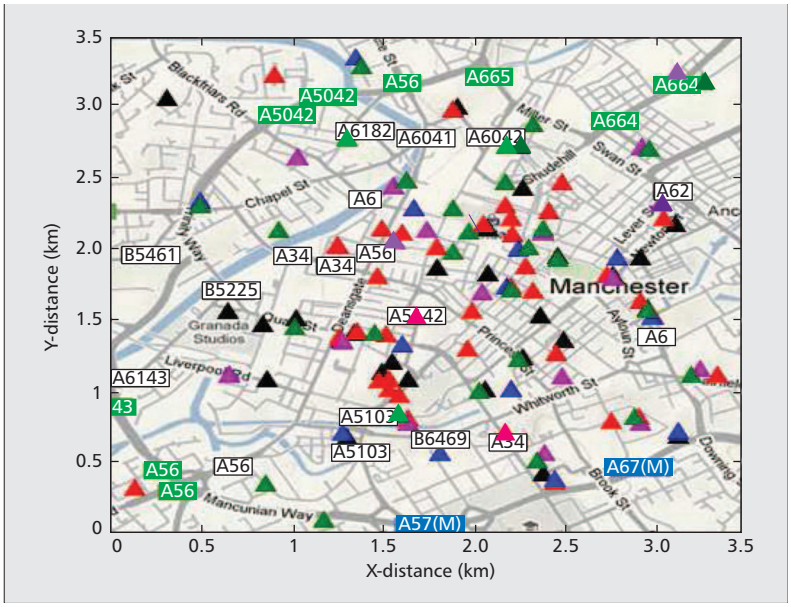


Figure 1.2: A snapshot of base station locations from a part of Manchester, United Kingdom [1]. The base station information can be obtained from [2].

investigations on HCNs. For network-wide performance evaluation and optimal design, a tractable system-level modelling technique is required. Unfortunately, most of the existing analytical network models are oversimplified and highly inaccurate [21–24]. Therefore, a modelling tool for HCNs that achieves a good balance between accuracy and tractability is called for.

As a statistical mathematical tool, stochastic geometry captures the spatial randomness in network node locations. Based on the random network deployment model, stochastic geometry analysis investigates the quantities of interest (e.g. SINR, interference, outage probability) seen from a generic node by averaging over all random network topologies [11]. Combined with its advantages described below, stochastic geometry appears to be a handy tool for HCN characterization.

1.1.1 Poisson Point Processes

To capture the spatial randomness, point processes are used to abstract the positions of network entities in stochastic geometry analysis. Choosing a convenient

point process will lead to an analytically tractable network model and simple precise expressions for the quantities of interest. The Poisson point process (PPP) is the simplest one that meets the above needs, and is defined as follows.

Definition 1.1.1. *(Poisson point process (PPP) [13]): A random set of points $X \subset \mathbb{R}^2$ is said to be a PPP of intensity $\lambda > 0$ on the plane if it satisfies the conditions:*

1. *For mutually disjoint domains of $\mathbb{R}^2$ $D_1, \ldots, D_k$, the random variables $X(D_1), \ldots, X(D_k)$ are mutually independent, where $X(D)$ denotes the number of points of X inside domain D.*

2. *For any bounded domain $D \subset \mathbb{R}^2$ we have that for every $k \geq 0$*

$$\Pr\left(X(D) = k\right) = e^{-\lambda |D|} \frac{\left(\lambda |D|\right)^k}{k!}, \tag{1.1}$$

where $|D|$ denotes the area of domain D.

Specifically, if the node intensity λ is a constant independent of spatial location, the PPP is a homogeneous Poisson point process (HPPP); otherwise, it is called an inhomogeneous PPP. Starting from a PPP, various kinds of more complicated point processes such as Poisson cluster process (PCP) and Matérn hardcore process (Matérn HCP) can be obtained [25].

Due to its independence property, PPP provides a good balance for the trade-off between simplicity and accuracy in network modelling and hence is the most popular point process used in stochastic geometry analysis. In the pioneering work [26], PPP modelling was applied to a single-tier cellular network, and showed good accuracy and tractability for coverage and rate analysis. From the baseline model in [26], stochastic geometry analyses for HCNs were performed, with the multiple tiers of base stations modelled as independent PPPs in [27–29]. A PPP based two-tier HCN model is depicted in Fig. 1.3.

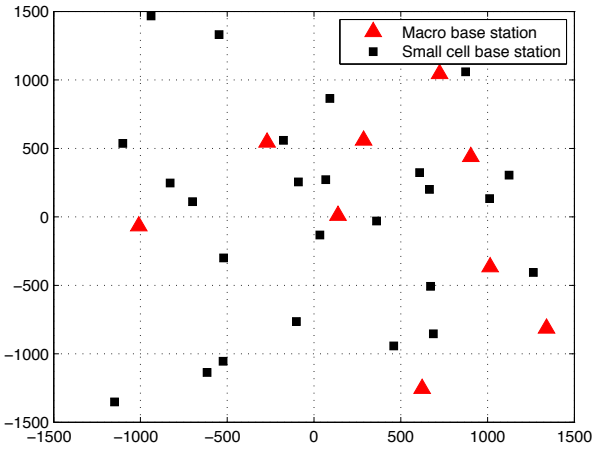


Figure 1.3: The two-tier HCN modelled by independent PPPs in a 3000m×3000m region. The macro base station intensity is 1 point/km^2, and the small-cell base station intensity is 3 points/km^2

1.1.2 Interference Related Performance Analysis

For simplicity, a network may be designed with limited coordination among base stations, in which case interference in large scale cellular networks is non-negligible or even dominates the noise term. Interference determines many key network performance metrics such as outage probability and throughput. Based on the point process abstraction of network node locations, the aggregated interference I_{agg} at a generic node is a spatio-temporal stochastic process, which can be characterized using stochastic geometry analysis.

A key tool that enables interference characterization is the Laplace transform (LT) of the probability density function (PDF) of I_{agg}. The LT of I_{agg} is defined as

$$\mathcal{L}_{I_{\text{agg}}}(s) = \mathbb{E}\left[e^{-sI_{\text{agg}}}\right], \tag{1.2}$$

which is conceptually equivalent to the moment generating function (MGF) or the characteristic function (CF) of I_{agg}. With a well-chosen point process network approximation (e.g. PPP), exact LT, MGF, or CF expressions can

be found. Consequently, the moments of I_{agg} can be calculated as $\mathbb{E}\left[I_{\text{agg}}^n\right] = (-1)^n \frac{\mathrm{d}^n \mathcal{L}_{I_{\text{agg}}}(s)}{\mathrm{d}s^n}\Big|_{s=0}$. In the literature, interference related network performance metrics, especially outage probability and throughput, can be evaluated by several techniques that utilize the LT, MGF or CF. These techniques are summarized as follows.

1. **Utilizing Rayleigh fading assumption**

   By assuming Rayleigh fading on the desired link, the distribution of the generic user's signal-to-interference-plus-noise-ratio (SINR) can be determined exactly. Specifically, being conditioned on path loss and assuming flat Rayleigh fading between the transmitter and the receiver, the received desired signal power, denoted as S_{des}, is an exponential random variable with mean $\frac{1}{\mu}$ taking transmit power and path loss into account. With the additive noise power denoted by N_{add}, the cumulative distribution function (CDF) of the receiver SINR is [26]

   $$\Pr\left(\text{SINR} \leq \theta\right) = \Pr\left(\frac{S_{\text{des}}}{I_{\text{agg}} + N_{\text{add}}} \leq \theta\right) = 1 - e^{-\theta\mu N_{\text{add}}} \mathcal{L}_{I_{\text{agg}}}(\theta\mu). \tag{1.3}$$

   For certain point processes, e.g. the HPPP, the LT of I_{agg} is known, and hence the above SINR CDF can be found. With the exact SINR distribution determined, outage probability and rate related network metrics (based on the Shannon formula) can be obtained. Because of its tractability, this performance evaluating technique has been widely used in the literatures [27,28,30]. However, there exists a trade-off between tractability and accuracy when we apply this technique to analyzing networks with general channel fadings.

2. **Approximating interference by known distributions**

   Although the moments of I_{agg} can be obtained by calculating the derivatives of its LT (1.2), there is still no known expression for the PDF of I_{agg} when channel fading is not Rayleigh. To address this issue, in this

technique, the distribution of I_{agg} is approximated by a known distribution, where the parameters of the approximate PDF can be determined using the LT of I_{agg}. In [29], the interference distribution in PPP based HCN model was approximated by a gamma distribution. In [31], log-normal and shifted log-normal distributions were used to approximate interference in HCNs. One limitation of this technique is that uncertainty exists in selecting the approximate distribution. Simulation tests are required to select the distribution with sufficient approximation accuracy. Additionally, the approximated interference distribution typically has a complicated PDF expression that prevents its application in network optimization analysis.

3. **Considering dominant interferers**

   If the signal power drops quickly with increasing distance, i.e. the path loss exponent is large, I_{agg} can be tightly lower bounded by only considering the dominant interferers. This technique gives lower bound approximations for the outage probability [32]. And then, with the knowledge of LT, MGF, or CF of I_{agg}, outage probability upper bound approximations can be obtained by applying the Markov inequality, Chebyshev inequality, or the Chernoff bound [33]. However, the approximation approach is only applicable to the high path loss exponent case and its accuracy depends on how the dominant interferers are determined.

1.2 Related Works and Challenges

This thesis focuses on using stochastic geometry to investigate power saving problems in downlink HCNs. We therefore first provide an overview of stochastic geometry formulations for interference control analysis in the literature followed by a discussion of the potential challenges. The related works and open problems in power saving strategies and the fundamental tradeoffs in HCN energy reduction analysis are then introduced.

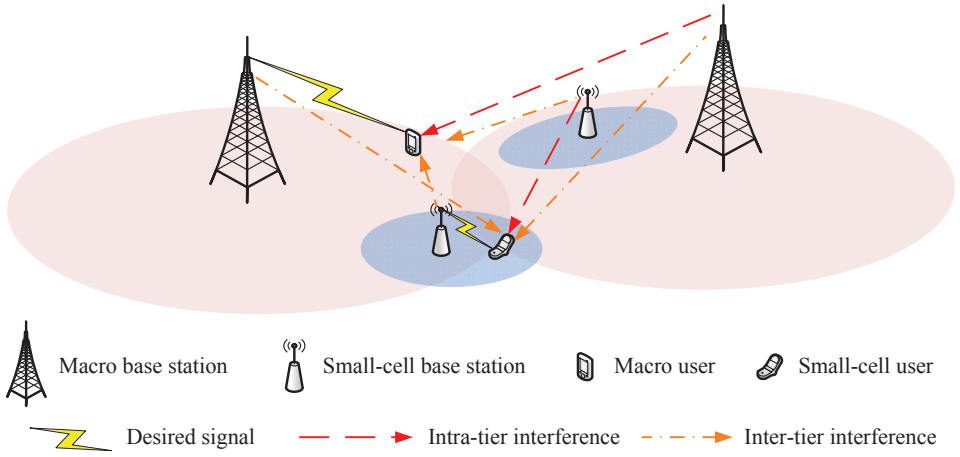


Figure 1.4: Interference in a two-tier HCN with co-channel assignment of base stations.

1.2.1 Interference Control

Related works

As depicted in Fig. 1.4, there are intra-tier interference (or co-tier interference) and inter-tier interference (or cross-tier interference) in the HCN downlink with co-channel base station assignment. Various kinds of stochastic geometry based frameworks have been proposed to investigate the intra- and inter-tier interference control problems. Basically, the interference control schemes in prior published works were realized through resource allocation and/or inter-cell interference coordination (ICIC), which are respectively discussed as follows.

By allocating orthogonal transmission resources to users, interference in cellular networks can be eliminated. For homogeneous macro cellular networks, a stochastic geometry framework for analyzing fractional frequency reuse (FFR) was proposed in [34], where the neighbouring macrocells used orthogonal channels to serve their corresponding edge users and hence reduce inter-cell interference. Based on the proposed FFR scheme, the outage probability was derived in closed-form expression in [34]. For HCNs, various kinds of resource allocation schemes have been investigated using stochastic geometry analysis. In [35], a partial spectrum reuse

scheme was proposed, where each small-cell BS randomly chose a fraction of the spectrum of the MBSs' to transmit on. The random spectrum reuse scheme lowered the probability of multiple small-cells accessing the same spectrum and hence reduced the inter-tier interference level for MBSs. Additionally, stochastic geometry analysis was conducted in [35] to determine the optimal spectrum reuse fraction that minimized network power consumption. In [36], two spectrum allocation schemes in HCNs were investigated where the different tiers of base stations transmitted either on the same spectrum or on orthogonal channels. For the two spectrum allocation schemes, stochastic geometry frameworks were proposed to find the corresponding user association probability and outage probability. The optimal user association and spectrum allocation schemes were then determined by solving a rate coverage based fair utility optimization problem.

Besides resource allocation, interference in cellular networks can be eliminated by ICIC, which includes base station clustering and coordinated transmission. One simple ICIC approach is to let several base stations form a transmission cluster and allow only one base station in that cluster to transmit at a time. An application of this ICIC scheme in homogeneous cellular networks (or single tier cellular networks) was described in [37], which is equivalent to using the ICIC for intra-tier interference management in HCNs. In [37], the K nearest base stations to the generic user formed a cluster within which only the nearest base station transmitted to the user. In the high spectral efficiency regime, it was shown that this ICIC scheme resulted in a better outage performance than the selective combining based intra-cell diversity transmission with larger bandwidth [37]. Interference avoidance in a two-tier HCN using ICIC was investigated in [38], where the outage probability expressions for the intra- and inter-tier interference mitigation methods were determined, respectively. In [38], with circular shaped exclusion zones deployed around each transmitting MBS, the interfering MBSs and micro base stations within the exclusion zones were correspondingly muted to remove the intra- and inter-tier interference. Instead of muting the co-clustered interfering base stations, ICIC can also be achieved

through multi-antenna beamforming. In [39], base stations were clustered according to a lattice grid. Inter-cell interference within a cluster was mitigated through beamforming. A stochastic geometry framework was proposed to show how outage probability scaled with the number of cooperating base stations in each cluster. In [40], the spectral efficiency expression was derived using stochastic geometry analysis on a network with interference reduced through dynamic K base stations clustering and coordinated beamforming.

By jointly applying resource allocation and ICIC, enhanced inter-cell interference coordination (eICIC) is proposed to boost the performance of HCNs [41]. Using eICIC, base stations adjust their coverage areas so that users close to inter-tier boundaries are reallocated to different tiers of base stations. Additionally, only specific tier of base stations are allowed to use some fraction of the transmission resources to serve those reallocated users. In [42], a joint resource partitioning and offloading scheme in HCNs was proposed where partitioned resources were allocated to the users that were offloaded from the macrocells to the small-cells to eliminate inter-tier interference. Using stochastic geometry analysis, tractable SINR and rate coverage probabilities for the proposed scheme were derived. Numerical results in [42] verified that the joint use of offloading and resource partitioning improved cell edge coverage performance in HCNs.

Challenges

Most of the existing stochastic geometry analyses aim to derive expressions for the interference related QoS metrics such as outage probability or capacity expressions. The derived expressions are typically too complicated to be applied in optimization problems. Thus, non-trivial manipulations are needed to approximate the performance metrics to make them amenable to optimal network design. Furthermore, in stochastic geometry analysis, interference control schemes are mainly used to improve coverage and capacity performance. It still remains unclear how to jointly use interference control and cell load adaptation strategies to improve

HCN energy efficiency. Moreover, most of the existing optimal eICIC design strategies [43–47] are based on semi-analytical approaches. It still remains an open problem to design eICIC scheme in HCNs using analytically tractable methods like stochastic geometry.

In the following chapters, we will tackle the above mentioned problems by tractably designing cell load adaptation coupled with various interference control schemes to reduce the power consumption in downlink HCNs with QoS constraints.

1.2.2 Power Saving in HCNs

Related works

As shown in [48], a base station's power consumption depends on not only the transmit power, but also the transmission-independent power consumption due to signal processing, battery backup, site cooling, etc. According to the base station power consumption model [48], power minimization in the downlink of cellular networks can be achieved by reducing base station transmit power and/or muting base stations with QoS constraints on SINR [49], spectral efficiency [50], throughput [51], delay [52], etc.

A centralized transmit power minimization and user admission control problem in a wireless network with multiple interfering single input single output (SISO) links was investigated in [49]. In this work, based on linear programming relaxation of the NP-hard problem, a link removal approach was proposed to minimize the total transmit power with the maximum number of users supported at the specified SINR targets. In [53], resource allocation scheme in an HCN was designed in a centralized manner to minimize the base station transmit power with given data rate requirements. Convex relaxation techniques were used to solve the mixed integer nonlinear resource allocation problem. In [54], transmit power minimization in a femtocell network was realized through distributed resource allocation with user throughput constraints. An efficient algorithm was derived to approximately solve the integer programming based resource allocation problem.

In [55], outage probability constrained transmit power minimization problems were solved for two-tier HCNs with perfect and imperfect channel state information (CSI), respectively. Applying convex relaxation strategies on the outage probability constraints, suboptimal base station beamforming vectors that solved the power minimization problems were determined in [55].

In addition to reducing base station transmit power, base station sleeping is another option for lowering network power consumption. As an energy-efficient load adaptation strategy, cell zooming [56] allows lightly loaded base stations to be put into sleep mode, and have their users offloaded to neighbouring cells. For networks with deterministic deployment, power minimization through base station sleeping and/or cell zooming is a combinatorial problem and therefore can only be solved through heuristic algorithms. With traffic load and node location information, the sleeping mechanism of femtocells in a two tier HCN was studied in [57] using Markov decision processes (MDPs). In [58], a distributed base station on/off algorithm was proposed, where each base station took its impact on the neighbouring cells into account. For HCNs with random network topology, stochastic geometry frameworks were proposed in [59] and [60] to correspondingly investigate the minimum densities of active base stations that met SINR and rate coverage probability requirements. In [59] and [60], independent PPPs were used to model HCNs, and all base stations in the investigated HCNs were assumed to transmit on the same channel.

Challenges

The derived algorithms in most of the above mentioned literatures are based on deterministic network deployment and require perfect or partial information of node locations and channel coefficients [48, 53–55, 57, 58]. However, assuming that every base station in an HCN know other nodes' locations and global CSI is not realistic because of the tremendous cost to get the information. Even if the node positions and CSI are perfectly known, determining the resource allocation and base station selection schemes that minimize power consumption in HCNs with various QoS

constraints is typically a challenging task where only sub-optimal solutions can be found. Indeed, some of the works [59, 60] proposed stochastic geometry frameworks for HCN analysis, which were based on point process network abstractions and did not require global CSI. However, in the above mentioned stochastic geometry analyses, the base station power saving was achieved by homogeneously reducing the active node density without jointly taking interference control and cell load adaptation into consideration. Moreover, geometric dependence was ignored in [59, 60] where base stations were allowed to be located arbitrarily close to each other.

In the subsequent chapters of this thesis, with interference control and cell load adaptation considered, we will use stochastic geometry analysis to investigate power saving strategies in HCNs, which involve base station muting, transmit power adaptation, and the geometrically dependent base station deployment.

1.3 Contributions and Organization of the Thesis

1.3.1 Major Contributions of the Thesis

The main contributions of this thesis are given in the following three sub-sections.

Three New HCN Frameworks for Joint Load and Interference Control Analysis

The first contribution of this thesis is to propose three stochastic geometry frameworks that jointly analyse load adaptation and interference control in HCNs.

In Chapter 2, we consider applying biased user association rule and time domain resource partitioning scheme in a two-tier HCN with multiple macro and small-cell base stations. Specifically, users located close to inter-tier cell boundaries are served on time slots in which only one tier of base stations is allowed to transmit. Using

stochastic geometry analysis, we investigate the effects of changing user association bias factor and resource partitioning fraction on controlling cell load and inter-tier interference. It is shown that under given throughput requirements, the joint load adaptation and resource partitioning achieves better coverage performance and lower network power consumption over solutions which rely only on load control.

In Chapter 3, a model with single macrocell and multiple femtocells is used. Circular shaped macrocell exclusion zones are deployed centered on femto base stations (fBSs) within which macro users are served on inter-tier interference free channels. The relationship between MBS transmit power and the size of exclusion zones is investigated through stochastic geometry outage probability characterization. It is shown that MBS transmit power is reduced by using circular shaped exclusion zones to avoid interference.

We propose a framework to investigate HCN with multi-antenna base stations in Chapter 4. Each small-cell BS uses a beamforming strategy to transmit to its designated user and mitigate its interference to the neighbouring cells. MBSs with non-overlapping circular shaped coverage areas are deployed within the small-cell network. Each MBS adopts zero forcing beamforming to simultaneously serve the small-cell users within its coverage, and the corresponding small-cell BSs are therefore muted to save energy. We study the proposed inhomogeneous base station muting scheme by characterizing QoS performance through stochastic geometry analysis. Because the density and coverage radius of the deployed MBSs determine cell load and interference level, we investigate the effects of the density and coverage radius of macrocells on the characterized QoS.

New QoS Metric Approximations

Another contribution of this thesis is that we derive closed-form approximations for QoS metrics that are not only accurate in capturing their dependency on certain key parameters, but also useful for optimal network design.

The exact characterization of a typical user's average throughput involves

a weighted sum of the cell load PDF, which gives no closed-form expression. To address this problem, in Chapter 2 we approximate the user throughput as the coverage spectral efficiency divided by the average cell load. Based on the approximate QoS metric, the relationship between association bias factor and resource partitioning fraction can be found.

In Chapter 3, the cell edge user outage probability is considered as the QoS metric. The outage probability constraints provide implicit expressions for the MBS transmit power. With a low SINR target, we derive explicit approximate transmit power expressions. The approximate expressions provide insights into the impact of exclusion zone radius and correspondingly allocated bandwidth on the MBS transmit power.

In Chapter 4, due to the inhomogeneous small-cell muting scheme and the multi-antenna beamforming at base stations, exact characterization of user spectral efficiency is a challenging task. To solve this problem, we first approximate multiple tiers of base stations as independent PPPs. Next, the spectral efficiency is approximated by calculating the expected values of the desired and interference signals. With its accuracy verified through simulation, the approximate spectral efficiency expression reveals the effects of MBS density and coverage radius.

Optimal Power Saving HCN Design

Last but not least, based on the proposed HCN frameworks and QoS approximations, we provide optimal network parameter design methods that minimize network power consumption.

To reduce power consumption by muting base stations on the partitioned resources, in Chapter 2 we determine the optimal user association bias and the resource partitioning fraction by finding the feasible sets of these parameters for average throughput constraints. It is shown that there may exist more than one association bias that achieve the minimum base station power consumption. Among all feasible bias values, the one that maximizes network coverage probability is then

determined.

In Chapter 3, with outage probability threshold for cell edge users, the optimal exclusion zone radius that minimizes MBS transmit power in the single macrocell model is obtained. We consider two cases where users in exclusion zones are either served by the MBS or offloaded to the nearest femtocells. We show that for both cases, the optimal exclusion zone radius should be set to the maximum achievable values that satisfy the outage probability constraints.

In Chapter 4, due to the MBS deployment and small-cell muting scheme, the network power consumption is related with the density and coverage radius of the deployed MBSs. By equating the approximate spectral efficiency to a given certain target value, the mapping from MBS density to the coverage radius can be determined. Therefore, an algorithm is proposed to determine the optimal MBS coverage radius and hence the MBS density that minimize the network power consumption.

1.3.2 Thesis Organization

The remainder of this thesis is organized as follows. In Chapter 2, we investigate joint design of cell load adaptation and resource partitioning in two-tier HCNs with sleep mode base stations. Chapter 3 studies optimal exclusion zone design that minimizes MBS transmit power within a macrocell overlaid with multiple fBSs. In Chapter 4, multi-antenna MBSs are deployed to replace parts of the coverage area of small-cells, where the optimal density and coverage radius of the MBSs are determined to minimize network power consumption. Finally, Chapter 5 concludes the thesis and discusses potential areas of future work.

Chapter 2

Joint Resource Partitioning and User Association with Sleep-Mode Base Stations

2.1 Introduction

As shown in [61], allocating well-combined orthogonal and identical transmission time/frequency resources to different tiers of base stations in HCNs provides high network throughput. In addition, it has been verified that intelligent cell load adaptation within and between base station tiers results in performance gains in terms of reduced base station power consumption [62], improved user coverage [63], throughput [64] and various utilities [65]. Driven by the desire for cellular networks that deliver vastly improved quality of experience and quality of service to users at a much higher energy efficiency (in Joules per bit) [19,66], the joint use of resource allocation and cell load adaptation becomes an appealing approach for HCN design.

Due to their limited coverage areas and high deployment density, small-cells in HCNs are typically lightly loaded [67, 68]. As a result, each small-cell BS only needs to deliver data on a fraction of the total available transmission resources to meet its designated users' needs. The remaining transmission resources are called almost blank subframes (ABSs) because they are not used by the small-cells for data transmission. By using ABSs together with cell load adaptation techniques, enhanced inter-cell interference coordination (eICIC) frameworks [69] have been proposed to improve network performance. Specifically, using appropriate user association schemes, a subset of users in the congested macrocells that are located

close to inter-tier cell boundaries can be determined and served on ABSs. Through resource partitioning, the selected users allocated ABSs are either served by the MBSs or offloaded to the small-cells, where inter-tier interference can be eliminated. The performance gains of HCNs adopting eICIC were verified in [70–73] through numerical tests. Tractable stochastic geometry frameworks were then proposed in [42, 74] for resource partitioning coupled with load adaptation in HCNs using eICIC. Based on single macrocell multiple small-cells [74] and multiple macrocells multiple small-cells [42] setups, the user outage probabilities and average spectral efficiencies were characterized. However, the optimal design of cell load control and resource allocation for eICIC were not investigated in [42, 74]. The optimal cell load adaptation strategy was studied in [75] based on a deterministic HCN setup. In [75], a cell cooperative scheduling scheme for eICIC was proposed, where the small-cells expanded their coverage areas to help offload macro users. The design of resource allocation for eICIC was proposed in [76]. In [76], users suffering from severe inter-tier interference were served on ABSs at reduced interference. Using stochastic geometry characterization of dominant interferers, the set of victim users was determined and the number of required ABSs was obtained by solving a throughput maximization problem. Although cell load adaptation and resource allocation design were correspondingly discussed in [75] and [76], they were not investigated in a combined manner. By proposing algorithms that jointly allocated ABSs and adapted cell load through user association, throughput and fairness optimization problems were solved in [46, 64]. However, only semi-analytical approaches were derived. As a result, an analytically tractable design of the optimal strategy for joint resource partitioning and load adaptation still remains unknown. Additionally, to the best of our knowledge, none of the above mentioned works conduct energy efficiency or power saving analysis in HCNs with eICIC.

To address the problems mentioned above, in this chapter, we consider a resource partitioning and user association scheme in two-tier HCNs, where the locations of all base stations and users are modeled as independent PPPs. With

resource partitioning, we propose a user association scheme that adjusts cell load on different resources to improve network performance. Users that experience high inter-tier interference are served on resources on which only one tier of base stations are active. The inactive base stations are put into sleep mode to reduce energy consumption. Based on this framework, we formulate energy minimization and coverage probability maximization problems with throughput constraints. Solving the problems leads to the jointly optimal resource partitioning and user association strategy. The discoveries provide guidelines on energy efficient eICIC design in HCNs.

The remainder of this chapter is structured as follows. The system model is described in Section 2.2. The resource allocation and user association scheme is proposed and base station power consumption models are characterised. Based on the proposed model, we give expressions for association probability and coverage probability in Section 2.3 and then derive the average throughput constraints for each user set. In Section 2.4, we formulate the optimization problems to minimize the BS power consumption and maximize coverage probability over the entire network and provide a method to find the optimum solutions. Section 2.5 gives numerical results that verify our analytical discoveries. Finally, Section 2.6 concludes the chapter.

2.2 System Model

In this chapter, we consider the downlink of a two-tier network, where each tier consists of base stations of the same type. Without loss of generality, let MBSs constitute tier 1 and the small-cell BSs be tier 2 base stations. Base stations in the k-th tier are assumed to form a HPPP Φ_k with intensity λ_k. The spatial distribution of user equipments (UEs) is another HPPP Φ_{u} with constant intensity λ_{u}. Moreover, Φ_1, Φ_2 and Φ_{u} are independent.

The total spectrum resource has a bandwidth W. All tier k base stations are assumed to have constant transmit power spectral density (PSD) S_k over W.

We assume the base station density is large enough that the interference power dominates the additive noise. In the rest of this chapter we will therefore ignore the additive noise. A typical UE i at distance $d_{i,j}$ away from a base station j in the k-th tier has the received signal power $\mathrm{P}_k^{\mathrm{rx}}(d_{i,j}) = H_{i,j} g_k(d_{i,j}) \mathrm{S}_k \Delta_i$, where $H_{i,j}$ is an exponential random variable with unit mean modelling the power attenuation due to Rayleigh fading, $g_k(d_{i,j})$ is the path loss (PL) from the k-th tier base station j to UE i, and Δ_i is the bandwidth allocated to UE i. As in [61, 63, 77], we characterize the PL model as

$$g_k(d_{i,j}) = K d_{i,j}^{-\alpha_k},$$

where K is a constant propagation loss and α_k is the k-th tier PL exponent factor. By assuming identical PL environments in both tiers, i.e. $\alpha_1 = \alpha_2 = \alpha$, we can benefit from more compact and useful expressions of user association and coverage probabilities, which will be given in the next section. Furthermore, the value of PL exponent for urban area cellular radio environment typically lies within a small range between 2.7 and 3.5 [78, Table 4.2]. Therefore, the approximation of $\alpha_1 = \alpha_2 = \alpha$ does not cause much loss in accuracy and it has already been adopted in many previous papers [27, 35, 77]. For simplicity, in the rest of the chapter, we use the PL model notation $g(d_{i,j}) = K d_{i,j}^{-\alpha}$ without the subscript k.

2.2.1 Resource Partitioning and User Association

Our model applies a user association and resource partitioning scheme similar to the one proposed in [42].

Resource partitioning

For clarity, the term "resources" used in this chapter refers to a set of time/frequency 3GPP resource elements. Throughout this chapter, a resource element is said to be "shared" by the two tiers of base stations if both the MBSs

and small-cell BSs are active on it to transmit to the target users. Similarly, we denote the resource elements that are deliberately allocated to one of the two base station tiers as "unshared" resources in the rest of the chapter.

Let η be the fraction of resources shared by the MBSs and small-cell BSs. The remaining $1-\eta$ fraction of the resources are unshared resources, which may be allocated to MBSs or small-cell BSs, unlike in [42]. η is called the resource partitioning factor.

User association

We assume that user association discussed in this chapter is based on the pilot/reference signal power. The average pilot signal power received by user i from the k-th tier base station j is $\mathbb{E}_{H_{i,j}}\left[\mathrm{P}_k^{\mathrm{rx}}(d_{i,j})\right] = g(d_{i,j})\mathrm{S}_k\Delta_i$. For the convenience of load adaptation, the average received pilot signal power is multiplied by a bias factor, which is called the biased received power [63]. Furthermore, without loss of generality, the bias factor for macro (tier 1) base stations is 1 and a bias factor B_2 is used for small-cell (tier 2) base stations.

In the literature, for example [67], one intention of deploying small-cells in the existing macro cellular networks is to let the macrocells ensure basic coverage and use small-cells to provide high data transmission. To reflect this, in this chapter, two kinds of rate requirements for users are determined based on the biased received power from base stations. If the biased received power at a user from the macro tier is higher than that from the small-cell tier, we denote the user's target rate as C_1. Otherwise, the target rate is C_2. We assume that users with larger biased received power from small-cells than from macrocells have a higher rate requirement, which means $C_1 \leq C_2$. Specifically, the target rate C_j for a typical user can be determined using the biased received power association:

$$C_j = \begin{cases} C_1, & \text{if } \mathrm{S}_1 g(d_1) \geq \mathrm{B}_2\mathrm{S}_2 g(d_2) \\ C_2, & \text{if } \mathrm{B}_2\mathrm{S}_2 g(d_2) > \mathrm{S}_1 g(d_1) \end{cases} \tag{2.1}$$

where d_k denotes the distance between the typical user and its nearest base station in tier k, and B_2 is the association bias factor for tier 2 (Here, we set the bias factor for tier 1 to be 0 dB). B_2 determines how many UEs can be served with rate requirements C_1 and C_2. As B_2 increases, more UEs will be served by small-cell BSs with target rate C_2. Note that a user measures the received pilot signal power from the small-cell BSs and MBSs over the same bandwidth, and therefore the comparison of signal powers is equivalent to the above comparison of PSDs.

We let $\mathcal{U}_2$ denote the set of users with target rate C_2. The set of users with target rate C_1 is splitted into two disjoint sets $\mathcal{U}_1$ and $\mathcal{U}_\mathrm{D}$, where $\mathcal{U}_1$ consists of users closer to the macro-cell centres and $\mathcal{U}_\mathrm{D}$ represents the set of users located closer to the cell boundaries between the two tiers. A user in $\mathcal{U}_\mathrm{D}$ suffers from severe inter-tier interference if both tiers of base stations transmit on its allocated resources. To determine whether a user belongs to $\mathcal{U}_1$ or $\mathcal{U}_\mathrm{D}$, another association bias factor B_1 is introduced, where $\mathrm{B}_1 \geq \mathrm{B}_2$. The mapping of user i with rate requirement C_j to sets $\mathcal{U}_1$, $\mathcal{U}_2$ and $\mathcal{U}_\mathrm{D}$ is summarized as follows:

$$i \in \begin{cases} \mathcal{U}_1, & \text{if } C_j = C_1 \text{ and } g(d_1)\mathrm{S}_1 \geq \mathrm{B}_1 g(d_2)\mathrm{S}_2 \\ \mathcal{U}_\mathrm{D}, & \text{if } C_j = C_1 \text{ and } \mathrm{B}_1 g(d_2)\mathrm{S}_2 > g(d_1)\mathrm{S}_1 \geq \mathrm{B}_2 g(d_2)\mathrm{S}_2 \\ \mathcal{U}_2, & \text{if } C_j = C_2 \text{ and } \mathrm{B}_2 g(d_2)\mathrm{S}_2 > g(d_1)\mathrm{S}_1, \end{cases} \tag{2.2}$$

By changing the value of B_1, we can control the number of users in $\mathcal{U}_1$ and $\mathcal{U}_\mathrm{D}$. Set $\mathcal{U}_\mathrm{D}$ will be empty only if $\mathrm{B}_1 = \mathrm{B}_2$.

To resolve the inter-tier interference problem, we combine the user association rule in (2.2) with resource partitioning, where $\mathcal{U}_1$ and $\mathcal{U}_2$ users are served by MBSs and small-cell BSs on the shared resources, respectively. The unshared resources are allocated to users in $\mathcal{U}_\mathrm{D}$.[1] Since either small-cell BSs or MBSs can be muted on the unshared resources, we will consider the two cases separately. For simplicity, in the

[1]By increasing B_1/decreasing B_2, the inter-tier interference level at users in $\mathcal{U}_1$/$\mathcal{U}_2$ can be reduced. Due to the difference in coverage areas, the number of $\mathcal{U}_1$ users in each macrocell is typically larger than the number of $\mathcal{U}_2$ users in each small-cell. Therefore, reducing inter-tier interference for $\mathcal{U}_1$ users is more favourable. Thus, it is typically desirable to let $\mathrm{B}_1 \geq 1$ (0 dB), but B_2 is not restricted to be smaller than 1.

subsequent analysis we use *Case 1* to denote the scenario where UEs in $\mathcal{U}_{\mathrm{D}}$ associate with MBSs and use $\mathcal{U}_{\mathrm{D1}}$ to represent the set $\mathcal{U}_{\mathrm{D}}$. If small-cell BSs are assigned to serve users in $\mathcal{U}_{\mathrm{D}}$ on the unshared resources, it is similar to the offloading scheme in [42]. This scenario is denoted as *Case 2* and $\mathcal{U}_{\mathrm{D2}}$ is used to represent $\mathcal{U}_{\mathrm{D}}$. For clarity, the considered system is illustrated in Fig. 2.1.

Some additional remarks are listed as follows:

1. In this chapter, we consider the scenario where only users with target rate C_1 are offloaded to the unshared resources. We would also like to point out that the method proposed in this chapter can be easily generalised to analyse the case where both users with required rate C_1 and C_2 can be allocated to the unshared resources. The details are not provided here in order not to detract from the main contributions of this chapter.

2. The traditional fully shared resource allocation scheme is a special case of our proposed model with $\mathrm{B}_1 = \mathrm{B}_2$, which means $\mathcal{U}_{\mathrm{D}} = \emptyset$ in (2.2). Thus, all resources are shared among users in $\mathcal{U}_1$ and users in $\mathcal{U}_2$.

3. By setting $\mathrm{B}_1 = \infty$, $\mathcal{U}_1 = \emptyset$. If MBSs are selected to serve users in $\mathcal{U}_{\mathrm{D}}$ ($\mathcal{U}_{\mathrm{D}} = \mathcal{U}_{\mathrm{D1}}$), we have the MBSs and small-cell BSs occupy orthogonal resource elements. This is in fact the traditional unshared resource allocation scheme. On the other hand, if small-cell BSs are used to transmit to $\mathcal{U}_{\mathrm{D}}$ ($\mathcal{U}_{\mathrm{D}} = \mathcal{U}_{\mathrm{D2}}$) users and $\mathrm{B}_1 = \infty$, MBSs can be completely muted on all resources and all users are served by only small-cell BSs. In that case, the two tier heterogeneous network becomes a homogeneous network consisting of only small-cell BSs.

4. For simplicity, we treat B_2 as a given constant. For any association bias B_2, we present the method of finding B_1 to optimize network performance in terms of power consumption and UE coverage. Joint optimization of B_1 and B_2 will be discussed in the future.

5. We assume that all base stations are active – if in fact a k-th tier base station is

on with probability p_k, then, without loss of generality, we replace the Poisson density λ_k with $p_k\lambda_k$.

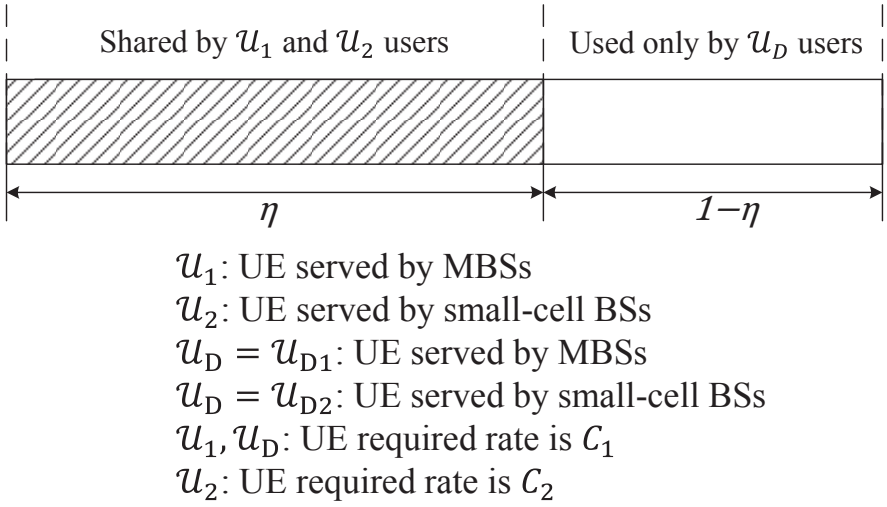


Figure 2.1: Resource allocation scheme.

2.2.2 Base Station Power Consumption Model

A simple power consumption model based on measurements done on real hardware was given in [79]. Using that model, the power consumption of MBSs and small-cell BSs during downlink transmission are given respectively by

$$\mathrm{P}_{\mathrm{MBS}} = a_{\mathrm{M}}\mathrm{P}_{\mathrm{M}} + b_{\mathrm{M}}, \quad \mathrm{P}_{\mathrm{scBS}} = a_{\mathrm{sc}}\mathrm{P}_{\mathrm{sc}} + b_{\mathrm{sc}}. \tag{2.3}$$

In the above models, P_{M} and P_{sc} are the transmit powers of MBSs and small-cell BSs, respectively. The coefficients a_{M} and a_{sc} account for the power consumption that scales with the transmit power. The terms b_{M} and b_{sc} represent the transmission-independent power consumption due to signal processing, battery backup, site cooling, etc. The model in (2.3) reflects the fact that the average power consumption of a base station comprises both transmit power and non-transmit power. In [79], the authors gave the numerical values for the parameters $(a_{\mathrm{M}}, b_{\mathrm{M}})$ and $(a_{\mathrm{sc}}, b_{\mathrm{sc}})$ as $a_{\mathrm{M}} = 22.6$, $b_{\mathrm{M}} = 412.4$W and $a_{\mathrm{sc}} = 5.5$, $b_{\mathrm{sc}} = 32.0$W.

The average base station power consumption per unit area is therefore given by

$$\mathrm{P}_{\mathrm{net}} = \lambda_1 \mathrm{P}_{\mathrm{MBS}} + \lambda_2 \mathrm{P}_{\mathrm{scBS}}. \tag{2.4}$$

Whether time or frequency resource partitioning is used impacts the formulation of energy saving problem, as explained next.

Frequency domain resource partitioning

For frequency domain resource partitioning, muting base stations on unshared bands will reduce transmit power but the non-transmit power consumption remains unchanged. If $\mathcal{U}_{\mathrm{D}} = \mathcal{U}_{\mathrm{D1}}$ (i.e. *Case 1*), the MBSs will transmit over the entire band of W Hz while the small-cell BSs will transmit over a band of ηW Hz, thus $\mathrm{P}_{\mathrm{M}} = \mathrm{S}_1 W$ and $\mathrm{P}_{\mathrm{sc}} = \mathrm{S}_2 W\eta$. We use $\mathrm{P}_{\mathrm{net}}^{\mathrm{f1}}$ to denote the power consumption $\mathrm{P}_{\mathrm{net}}$ in this case. On the other hand, when $\mathcal{U}_{\mathrm{D}} = \mathcal{U}_{\mathrm{D2}}$ (*Case 2*), we have $\mathrm{P}_{\mathrm{M}} = \mathrm{S}_1 W\eta$, $\mathrm{P}_{\mathrm{sc}} = \mathrm{S}_2 W$ and $\mathrm{P}_{\mathrm{net}} = \mathrm{P}_{\mathrm{net}}^{\mathrm{f2}}$. The average base station power consumption per unit area is therefore, from (2.4), given respectively by

$$\mathrm{P}_{\mathrm{net}}^{\mathrm{f1}} = \lambda_1 \left[a_{\mathrm{M}} \mathrm{S}_1 W + b_{\mathrm{M}}\right] + \lambda_2 \left[a_{\mathrm{sc}} \mathrm{S}_2 W\eta + b_{\mathrm{sc}}\right], \qquad \text{if } \mathcal{U}_{\mathrm{D}} = \mathcal{U}_{\mathrm{D1}} \tag{2.5}$$

$$\mathrm{P}_{\mathrm{net}}^{\mathrm{f2}} = \lambda_1 \left[a_{\mathrm{M}} \mathrm{S}_1 W\eta + b_{\mathrm{M}}\right] + \lambda_2 \left[a_{\mathrm{sc}} \mathrm{S}_2 W + b_{\mathrm{sc}}\right], \qquad \text{if } \mathcal{U}_{\mathrm{D}} = \mathcal{U}_{\mathrm{D2}} \tag{2.6}$$

Time domain resource partitioning

For time domain resource partitioning, base stations can be put into sleep mode on the unshared time slots to reduce both transmit and non-transmit power consumption, which in fact is a generalization of eICIC in LTE discussed in [80]. The transmit power for tier 1 and 2 base stations are respectively $\mathrm{P}_{\mathrm{M}} = \mathrm{S}_1 W$ and $\mathrm{P}_{\mathrm{sc}} = \mathrm{S}_2 W$. For clarity, we use $\mathrm{P}_{\mathrm{net}} = \mathrm{P}_{\mathrm{net}}^{\mathrm{t1}}$ for *Case 1* and $\mathrm{P}_{\mathrm{net}} = \mathrm{P}_{\mathrm{net}}^{\mathrm{t2}}$ for *Case 2*.

Therefore,

$$\mathrm{P}_{\mathrm{net}}^{\mathrm{t1}} = \lambda_1 \left[a_{\mathrm{M}} \mathrm{S}_1 W + b_{\mathrm{M}}\right] + \eta\lambda_2 \left[a_{\mathrm{sc}} \mathrm{S}_2 W + b_{\mathrm{sc}}\right], \qquad \text{if } \mathcal{U}_{\mathrm{D}} = \mathcal{U}_{\mathrm{D1}} \tag{2.7}$$

$$\mathrm{P}_{\mathrm{net}}^{\mathrm{t2}} = \eta\lambda_1 \left[a_{\mathrm{M}} \mathrm{S}_1 W + b_{\mathrm{M}}\right] + \lambda_2 \left[a_{\mathrm{sc}} \mathrm{S}_2 W + b_{\mathrm{sc}}\right], \qquad \text{if } \mathcal{U}_{\mathrm{D}} = \mathcal{U}_{\mathrm{D2}} \tag{2.8}$$

Although the $\mathrm{P}_{\mathrm{net}}$ expressions for resource partitioning in frequency and time domains are different, minimizing $\mathrm{P}_{\mathrm{net}}$ in both cases can be done similarly. To be more concrete and concise, we will only elaborate on time domain resource partitioning in the rest of this chapter. Using $\mathrm{P}_{\mathrm{net}}^{\mathrm{t1}}$ and $\mathrm{P}_{\mathrm{net}}^{\mathrm{t2}}$ expressions (2.7) and (2.8), we will give the optimal scheme that minimizes network-wide average power consumption under certain rate constraints.

2.3 Throughput Characterization

In this section, user throughput constraints are characterized and then used in the next section to find the optimal resource allocation scheme. To derive throughput constraints, user association probability and coverage probability in sets $\mathcal{U}_1, \mathcal{U}_2, \mathcal{U}_{\mathrm{D1}}$ and $\mathcal{U}_{\mathrm{D2}}$ are required.

Let $\mathbb{A}_j$ ($j \in \{1, 2, \mathrm{D1}, \mathrm{D2}\}$) denote the user association probability for set $\mathcal{U}_j$. The expressions of $\mathbb{A}_j$ are determined in the following lemma.

Lemma 2.3.1. *The probability that a randomly selected UE belongs to set* $\mathcal{U}_j$, $j \in \{1, 2, \mathrm{D1}, \mathrm{D2}\}$ *can be obtained as*

$$\begin{aligned}
\mathbb{A}_1 &= \frac{\lambda_1}{\lambda_1 + \left(\frac{\mathrm{B}_1\mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}} \lambda_2}, \\
\mathbb{A}_2 &= \frac{\lambda_2}{\lambda_2 + \left(\frac{\mathrm{S}_1}{\mathrm{B}_2\mathrm{S}_2}\right)^{\frac{2}{\alpha}} \lambda_1}, \\
\mathbb{A}_{\mathrm{D1}} = \mathbb{A}_{\mathrm{D2}} &= \frac{\lambda_1}{\lambda_1 + \left(\frac{\mathrm{B}_2\mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}} \lambda_2} - \frac{\lambda_1}{\lambda_1 + \left(\frac{\mathrm{B}_1\mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}} \lambda_2}.
\end{aligned} \tag{2.9}$$

(Note that it does not matter whether the unshared resources are used by the macro or small-cell BSs.)

Proof. The derivation of $\mathbb{A}_j$ expressions in (2.9) follows that of Lemma 1 of [28]. □

The coverage probability of a typical user i in set $\mathcal{U}_j$ is defined as $\mathbb{P}_j = \Pr\left(\text{SIR} \geq \Upsilon | \, i \in \mathcal{U}_j\right)$, where Υ is the target signal-to-interference-ratio (SIR). Recall that we ignored additive noise in the system model. This assumption will simplify the analysis and closed-form expressions can then be obtained.

Lemma 2.3.2. *For a typical user $i \in \mathcal{U}_j$, the coverage probabilities $\mathbb{P}_j$ are given in* (2.10) – (2.13) *for $j \in \{1, 2, \text{D1}, \text{D2}\}$.*

$$\mathbb{P}_1 = \frac{\lambda_1 + \left(\frac{\text{B}_1\text{S}_2}{\text{S}_1}\right)^{\frac{2}{\alpha}} \lambda_2}{\lambda_1 \left[1 + \rho\left(\Upsilon, \alpha\right)\right] + \lambda_2 \left(\frac{\text{B}_1\text{S}_2}{\text{S}_1}\right)^{\frac{2}{\alpha}} \left[1 + \rho\left(\frac{\Upsilon}{\text{B}_1}, \alpha\right)\right]}, \tag{2.10}$$

$$\mathbb{P}_2 = \frac{\lambda_2 + \left(\frac{\text{S}_1}{\text{B}_2\text{S}_2}\right)^{\frac{2}{\alpha}} \lambda_1}{\lambda_2 \left[1 + \rho\left(\Upsilon, \alpha\right)\right] + \lambda_1 \left(\frac{\text{S}_1}{\text{B}_2\text{S}_2}\right)^{\frac{2}{\alpha}} \left[1 + \rho\left(\Upsilon \text{B}_2, \alpha\right)\right]}, \tag{2.11}$$

$$\mathbb{P}_{\text{D1}} = \frac{\left(\lambda_1 + \left(\frac{\text{B}_1\text{S}_2}{\text{S}_1}\right)^{\frac{2}{\alpha}} \lambda_2\right)\left(\lambda_1 + \left(\frac{\text{B}_2\text{S}_2}{\text{S}_1}\right)^{\frac{2}{\alpha}} \lambda_2\right)}{\left(\lambda_1 \left[1 + \rho\left(\Upsilon, \alpha\right)\right] + \lambda_2 \left(\frac{\text{B}_1\text{S}_2}{\text{S}_1}\right)^{\frac{2}{\alpha}}\right)\left(\lambda_1 \left[1 + \rho\left(\Upsilon, \alpha\right)\right] + \lambda_2 \left(\frac{\text{B}_2\text{S}_2}{\text{S}_1}\right)^{\frac{2}{\alpha}}\right)}, \tag{2.12}$$

$$\mathbb{P}_{\text{D2}} = \frac{\left(\lambda_1 + \left(\frac{\text{B}_1\text{S}_2}{\text{S}_1}\right)^{\frac{2}{\alpha}} \lambda_2\right)\left(\lambda_1 + \left(\frac{\text{B}_2\text{S}_2}{\text{S}_1}\right)^{\frac{2}{\alpha}} \lambda_2\right)}{\left(\lambda_1 + \left[1 + \rho\left(\Upsilon, \alpha\right)\right] \lambda_2 \left(\frac{\text{B}_1\text{S}_2}{\text{S}_1}\right)^{\frac{2}{\alpha}}\right)\left(\lambda_1 + \left[1 + \rho\left(\Upsilon, \alpha\right)\right] \lambda_2 \left(\frac{\text{B}_2\text{S}_2}{\text{S}_1}\right)^{\frac{2}{\alpha}}\right)}, \tag{2.13}$$

where $\rho\left(\Upsilon, \alpha\right) = \Upsilon^{\frac{2}{\alpha}} \int_{\Upsilon^{-\frac{2}{\alpha}}}^{\infty} \frac{1}{1+u^{\frac{\alpha}{2}}} \text{d}u$.

Proof. The results follow from Lemma 2 of [42] and closed-form expressions were obtained by letting $\alpha_1 = \alpha_2 = \alpha$ and setting the noise power to zero and then completing the integrals in that lemma. □

The overall coverage probability is used as the metric to characterize the network coverage performance. Based on Lemmas 2.3.1 and 2.3.2, the overall

coverage probability can be calculated as

$$\mathbb{P}_{\mathrm{c}} = \begin{cases} \sum\limits_{j\in\{1,2,\mathrm{D1}\}} \mathbb{P}_j \mathbb{A}_j & \text{for } \textit{Case 1} \\ \sum\limits_{j\in\{1,2,\mathrm{D2}\}} \mathbb{P}_j \mathbb{A}_j & \text{for } \textit{Case 2} \end{cases} \tag{2.14}$$

We define the coverage spectrum efficiency for users in set $\mathcal{U}_j$ as

$$r_j = \log_2(1+\Upsilon)\,\mathbb{P}_j. \tag{2.15}$$

r_j is the average achieved throughput if users in outage are assumed to not transmit at all, while users that are not in outage transmit at $\log_2(1+\Upsilon)$ bps/Hz.

The cell load for a base station is defined as follows. On the shared resources, we use N_1 to denote the number of $\mathcal{U}_1$ users within a tagged macrocell, and let N_2 be the number of $\mathcal{U}_2$ users in a small-cell. In *Case 1*, MBSs are allowed to transmit on the unshared resources. The load of an MBS on the unshared resources, consisting of $\mathcal{U}_{\mathrm{D1}}$ users, is denoted as N_{D1}. In *Case 2*, the number of $\mathcal{U}_{\mathrm{D2}}$ users served by a small-cell BS on the unshared resources is N_{D2}. The mean cell load averaged over the entire network is denoted by $\bar{\mathrm{N}}_j$ ($j \in \{1,2,\mathrm{D1},\mathrm{D2}\}$). According to [42] and [63],

$$\bar{\mathrm{N}}_j = \begin{cases} 0 & \text{if } \mathbb{A}_j = 0 \\ 1 + \frac{1.28\lambda_{\mathrm{u}}}{\lambda_{m(j)}}\mathbb{A}_j & \text{otherwise,} \end{cases} \tag{2.16}$$

where we used the mapping: $m(j) = \begin{cases} 1 & \text{if } j \in \{1,\mathrm{D1}\} \\ 2 & \text{if } j \in \{2,\mathrm{D2}\} \end{cases}$.

In this chapter, we assume that all $\mathcal{U}_j$ users associated with a particular base station are allocated equal resources, which can be achieved by round-robin scheduling in the time-sharing resource allocation scheme. Then, we characterize the average user rate in a similar manner to what was done in [28, 81, 82]. For a typical user in set $\mathcal{U}_j$ ($j \in \{1,2,\mathrm{D1},\mathrm{D2}\}$) and given SIR target Υ, the user's average

throughput is defined as

$$\bar{T}_j = \frac{W_j r_j}{\bar{\mathrm{N}}_j} = \frac{W_j \log_2(1+\Upsilon)\,\mathbb{P}_j}{\bar{\mathrm{N}}_j}, \tag{2.17}$$

where $W_1 = W_2 = \eta W$ and $W_{\mathrm{D1}} = W_{\mathrm{D2}} = (1-\eta)W$. Using this average user throughput expression, we do not need to integrate the Laplace functional of a PPP, which leads to the tractability of the optimization problem in this chapter. As discussed in Section 2.2.1, we assumed that the target rate is C_1 for $\mathcal{U}_1 \bigcup \mathcal{U}_{\mathrm{D1}}$ in *Case 1* and $\mathcal{U}_1 \bigcup \mathcal{U}_{\mathrm{D2}}$ in *Case 2*, and C_2 for $\mathcal{U}_2$. A user's average throughput in set $\mathcal{U}_j$ should be no smaller than the corresponding target rate. Therefore, the following throughput constraints can be applied for set $\mathcal{U}_j$.

$$\begin{cases} \eta W \mathbb{P}_j \log_2(1+\Upsilon) \geq C_j \bar{\mathrm{N}}_j & \text{if } j \in \{1,2\} \\ (1-\eta) W \mathbb{P}_j \log_2(1+\Upsilon) \geq C_1 \bar{\mathrm{N}}_j, & \text{if } j \in \{\mathrm{D1},\mathrm{D2}\} \end{cases} \tag{2.18}$$

where the expressions of $\mathbb{P}_j$ and $\bar{\mathrm{N}}_j$ are given in Lemma 2.3.2 and (2.16). Additionally, we can observe from (2.9) and (2.10)-(2.13) that $\mathbb{A}_j$ and $\mathbb{P}_j$ all depend on B_1. Also, (2.16) shows that $\bar{\mathrm{N}}_j$ is related with $\mathbb{A}_j$. Thus, both the left and right hand sides of (2.18) depend on B_1.

We acknowledge that there are other metrics to measure throughput. For example, the fifth and the median percentile throughput were used in [42]. However, these performance metrics result in tremendous mathematical complexity. Hence, the resource partitioning and user association strategy that achieves the fifth or the median percentile throughput requirements can only be found numerically. In that case, it is hard to model and analyse the problem. On the contrary, based on the user average rate characterization (2.17), which has been adopted in [28, 81] and [82], the average throughput constraints (2.18) used in this chapter have the merit of mathematical tractability.

2.4 Power Minimization and Coverage Improvement

In this section, we derive resource allocation and user association schemes that are used to minimize power consumption and maximize coverage probability. It is possible to shut down (put to sleep) MBSs or small-cell BSs on the unshared resources. In this section, we will first formulate and solve the optimization problems for the two cases separately. Then we give a discussion of how to determine which of the two cases should be used.

2.4.1 *Case 1*: Unshared Resources Allocated to MBSs

By optimizing over the resource sharing fraction η and user association threshold B_1, we can reduce the network power consumption $\mathrm{P}_{\mathrm{net}}^{\mathrm{t1}}$ and improve user coverage $\mathbb{P}_{\mathrm{c}}$. The values of η and B_1 should satisfy the minimum throughput constraints expressed in (2.18). Hence, a feasible (η, B_1) set $\mathcal{F}_1$ can be described by

$$\begin{aligned}\mathcal{F}_1 = \{(\eta, \mathrm{B}_1) \ \text{ s.t. } & \eta W \mathbb{P}_1 \log_2(1+\Upsilon) \geq C_1 \bar{\mathrm{N}}_1, \\ & \eta W \mathbb{P}_2 \log_2(1+\Upsilon) \geq C_2 \bar{\mathrm{N}}_2, \\ & (1-\eta) W \mathbb{P}_{\mathrm{D1}} \log_2(1+\Upsilon) \geq C_1 \bar{\mathrm{N}}_{\mathrm{D1}}, \\ & \mathrm{B}_1 > \mathrm{B}_2\}. \end{aligned} \tag{2.19}$$

We can rewrite $\mathcal{F}_1$ as $\{(\eta, \mathrm{B}_1) | \max\{\eta_1(\mathrm{B}_1), \eta_2\} \leq \eta \leq \eta_{\mathrm{D1}}(\mathrm{B}_1), \mathrm{B}_1 > \mathrm{B}_2\}$, where $\eta_1(\mathrm{B}_1) = \frac{C_1 \bar{\mathrm{N}}_1}{W \mathbb{P}_1 \log_2(1+\Upsilon)}$ is the minimum resource percentage required to serve $\mathcal{U}_1$ users in each cell. Similarly, $\eta_2 = \frac{C_2 \bar{\mathrm{N}}_2}{W \mathbb{P}_2 \log_2(1+\Upsilon)}$ is the minimum fraction of resources that should be allocated to $\mathcal{U}_2$ users to achieve the target rate. Finally, $\eta_{\mathrm{D1}}(\mathrm{B}_1) = 1 - \frac{C_1 \bar{\mathrm{N}}_{\mathrm{D1}}}{W \mathbb{P}_{\mathrm{D1}} \log_2(1+\Upsilon)}$ is derived from $\mathcal{U}_{\mathrm{D1}}$ user rate requirements.

The feasible region $\mathcal{F}_1$ is non-empty if and only if the upper bound $\eta_{\mathrm{D1}}(\mathrm{B}_1)$ is bigger than the lower bounds $\eta_1(\mathrm{B}_1)$ and η_2. Since $\eta_1(\mathrm{B}_1)$ and $\eta_{\mathrm{D1}}(\mathrm{B}_1)$ are functions of B_1, a feasible set for B_1 can be derived. The feasible values of B_1 must satisfy

both

$$\eta_{\mathrm{D1}}(\mathrm{B}_1) \geq \eta_1(\mathrm{B}_1) \tag{2.20}$$

$$\text{and } \eta_{\mathrm{D1}}(\mathrm{B}_1) \geq \eta_2. \tag{2.21}$$

Let $\mathcal{M}_1$ denote the set of B_1 values that satisfies $\eta_{\mathrm{D1}}(\mathrm{B}_1) \geq \eta_1(\mathrm{B}_1)$. By finding the first order derivative of $\eta_1(\mathrm{B}_1)$, we can get $\frac{\mathrm{d}\eta_1(\mathrm{B}_1)}{\mathrm{dB}_1} \leq 0$ and hence, $\eta_1(\mathrm{B}_1)$ is a non-increasing function of B_1. However, $\eta_{\mathrm{D1}}(\mathrm{B}_1)$ is not necessarily monotonic, since the sign of $\frac{\mathrm{d}\eta_{\mathrm{D1}}(\mathrm{B}_1)}{\mathrm{dB}_1}$ depends on the value of other parameters (node density, transmit power, SIR threshold, etc.). Thus, a method is proposed as follows to determine the set $\mathcal{M}_1$.

Corollary 2.4.1. *The range of* B_1 *values that satisfies* $\eta_{\mathrm{D1}}(\mathrm{B}_1) \geq \eta_1(\mathrm{B}_1)$ *is the set* $\mathcal{M}_1 = \{\mathrm{B}_1 | \, \mathrm{B}_1 \geq b_1\}$*, where* b_1 *is given by*

1. $b_1 = \infty$*, if* $\lim\limits_{x\to\infty} \eta_{\mathrm{D1}}(x) < \lim\limits_{x\to\infty} \eta_1(x)$.
2. $b_1 = \mathrm{B}_2$*, if* $\lim\limits_{x\to\mathrm{B}_2} \eta_{\mathrm{D1}}(x) \geq \lim\limits_{x\to\mathrm{B}_2} \eta_1(x)$.
3. b_1 *is the single root of* $\eta_{\mathrm{D1}}(x) - \eta_1(x) = 0$*, if* $\lim\limits_{x\to\mathrm{B}_2} \eta_{\mathrm{D1}}(x) < \lim\limits_{x\to\mathrm{B}_2} \eta_1(x)$ *and* $\lim\limits_{x\to\infty} \eta_{\mathrm{D1}}(x) \geq \lim\limits_{x\to\infty} \eta_1(x)$.

Proof. Please see Appendix A. □

Constraint $\eta_{\mathrm{D1}}(\mathrm{B}_1) \geq \eta_2$ defines another range of B_1, which is denoted by $\mathcal{M}_2$. To find $\mathcal{M}_2$, we first let $x = \mathrm{B}_1^{\frac{2}{\alpha}}$. Then, $\eta_{\mathrm{D1}}(\mathrm{B}_1) \geq \eta_2$ can be rewritten in terms of x using Lemma 2.3.1, Lemma 2.3.2 and (2.16). As the derivation is tedious, we only give the final result as follows:

$$ux^2 + vx + w \leq 0, \tag{2.22}$$

where the coefficients u, v and w are

$$u = q^2\lambda_2^2 + \frac{1.28q^2\lambda_2^2\lambda_{\mathrm{u}}}{\lambda_1 + \lambda_2 q \mathrm{B}_2^{\frac{2}{\alpha}}} - \frac{q^2\lambda_2^2(1-\eta_2)\left(\lambda_1 + \lambda_2 q \mathrm{B}_2^{\frac{2}{\alpha}}\right)}{\mathrm{Z}_0\left(\lambda_1\left(1+\rho\left(\Upsilon,\alpha\right)\right) + \lambda_2 q \mathrm{B}_2^{\frac{2}{\alpha}}\right)}$$

$$v = q\lambda_1\lambda_2\left[\left(2+\rho\left(\Upsilon,\alpha\right)\right)\left(1 + \frac{1.28\lambda_{\mathrm{u}}}{\lambda_1 + \lambda_2 p \mathrm{B}_2^{\frac{2}{\alpha}}}\right) - \frac{1.28\lambda_{\mathrm{u}}}{\lambda_1}\right] - \frac{2q\lambda_1\lambda_2(1-\eta_2)\left(\lambda_1 + \lambda_2 q \mathrm{B}_2^{\frac{2}{\alpha}}\right)}{\mathrm{Z}_0\left(\lambda_1\left(1+\rho\left(\Upsilon,\alpha\right)\right) + \lambda_2 q \mathrm{B}_2^{\frac{2}{\alpha}}\right)}$$

$$w = \left(1+\rho\left(\Upsilon,\alpha\right)\right)\left[\left(1 + \frac{1.28\lambda_{\mathrm{u}}}{\lambda_1 + \lambda_2 q \mathrm{B}_2^{\frac{2}{\alpha}}}\right)\lambda_1^2 - 1.28\lambda_{\mathrm{u}}\lambda_1\right] - \frac{\lambda_1^2(1-\eta_2)\left(\lambda_1 + \lambda_2 q \mathrm{B}_2^{\frac{2}{\alpha}}\right)}{\mathrm{Z}_0\left(\lambda_1\left(1+\rho\left(\Upsilon,\alpha\right)\right) + \lambda_2 q \mathrm{B}_2^{\frac{2}{\alpha}}\right)},$$

where $\mathrm{Z}_0 = \frac{C_1}{W\log_2(1+\Upsilon)}$ and $q = \left(\frac{\mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}}$.

Let the solution set of x for (2.22) be $\mathcal{X}$. The following results can be used to determine $\mathcal{X}$.

Corollary 2.4.2. *The set $\mathcal{X}$ is determined according to the sign of u as follows.*

1. $u > 0$: *If* $v^2 - 4uw < 0$, $\mathcal{X} = \emptyset$. *Otherwise,* $\mathcal{X} = \left\{x : \frac{-v-\sqrt{v^2-4uw}}{2u} \le x \le \frac{-v+\sqrt{v^2-4uw}}{2u}\right\}$.
2. $u < 0$: *If* $v^2 - 4uw < 0$, $\mathcal{X} = \mathbb{R}$. *Otherwise,* $\mathcal{X} = \left\{x : x \le \frac{-v-\sqrt{v^2-4uw}}{2u}, x \ge \frac{-v+\sqrt{v^2-4uw}}{2u}\right\}$.
3. $u = 0$: *If* $v > 0$, $\mathcal{X} = \left\{x : x \le -\frac{w}{v}\right\}$. *If* $v < 0$, $\mathcal{X} = \left\{x : x \ge -\frac{w}{v}\right\}$.

Since the expressions of u, v and w are known, $\mathcal{X}$ for specific parameter settings can be derived using Corollary 2.4.2. Then using the relation $x = \mathrm{B}_1^{\frac{2}{\alpha}}$, we obtain the feasible set $\mathcal{M}_2$ for constraint (2.21) from $\mathcal{X}$. The feasible set of B_1 satisfying both (2.20) and (2.21) is finally $\mathcal{M}_1 \bigcap \mathcal{M}_2$.

With the feasible sets of η and B_1 determined, the energy reduction problem can be stated as

$$(\mathrm{P1a}): \quad \underset{(\eta,\mathrm{B}_1)\in\mathcal{F}_1}{\text{minimize}} \quad \mathrm{P}_{\mathrm{net}}^{\mathrm{t1}}$$

Since $\mathrm{P}_{\mathrm{net}}^{\mathrm{t1}} = \lambda_1 \left[a_{\mathrm{M}} \mathrm{S}_1 W + b_{\mathrm{M}}\right] + \eta \lambda_2 \left[a_{\mathrm{sc}} \mathrm{S}_2 W + b_{\mathrm{sc}}\right]$ increases with η, problem (P1a) is solved at the minimum η in $\mathcal{F}_1$, i.e. $\eta^* = \max\left\{\eta_1(\mathrm{B}_1^*), \eta_2\right\}$, where $(\eta^*, \mathbf{B}_1^*)$ is the globally optimal solution of (P1a). Using the expressions of $\mathbb{A}_j$ and $\mathbb{P}_j$ ($j \in \{1, 2, \mathrm{D1}\}$), the following results are derived:

Theorem 2.4.1. *When UEs in $\mathcal{U}_{\mathrm{D}}$ are served by the MBSs, η^* and B_1^* are given by the following expressions. For simplicity, we denote $\max\limits_{x \in \mathcal{M}_1 \bigcap \mathcal{M}_2}(x)$ by b_2.*

1. *If $\mathcal{M}_1 \bigcap \mathcal{M}_2 \neq \emptyset$ and $b_2 < \infty$, $\eta^* = \eta_2 = \frac{C_2 \bar{\mathrm{N}}_2}{W \mathbb{P}_2 \log_2(1+\Upsilon)}$ and B_1^* can take any value in the set $\mathcal{M}_1 \bigcap \mathcal{M}_2$ that satisfies $\eta_1(\mathrm{B}_1^*) \leq \eta_2$. In other words, $\mathrm{P}_{\mathrm{net}}^{\mathrm{t1}}$ is a constant for all values of B_1 in this range. The optimal power consumption per unit area is $\mathrm{P}_{\mathrm{net}}^{\mathrm{t1}*} = \lambda_1 \left[a_{\mathrm{M}} \mathrm{S}_1 W + b_{\mathrm{M}}\right] + \eta_2 \lambda_2 \left[a_{\mathrm{sc}} \mathrm{S}_2 W + b_{\mathrm{sc}}\right]$.*

2. *If $\mathcal{M}_1 \bigcap \mathcal{M}_2 \neq \emptyset$ and $b_2 = \infty$, MBSs and small-cell BSs occupy orthogonal resources, i.e. a fully unshared scheme is used. The optimal $\mathrm{B}_1^* = \infty$ and $\eta^* = \frac{C_2 \bar{N}_2}{W \mathbb{P} \log_2(1+\Upsilon)}$ with $\mathbb{P} = \frac{\lambda_1 + \left(\frac{\mathrm{B}_2 \mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}} \lambda_2}{\lambda_1 + [1+\rho(\Upsilon,\alpha)]\lambda_2 \left(\frac{\mathrm{B}_2 \mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}}}$. The minimum power consumption per unit area in this case is $\mathrm{P}_{\mathrm{net}}^{\mathrm{t1}*} = (1-\eta^*)\lambda_1 \left[a_{\mathrm{M}} \mathrm{S}_1 W + b_{\mathrm{M}}\right] + \eta^* \lambda_2 \left[a_{\mathrm{sc}} \mathrm{S}_2 W + b_{\mathrm{sc}}\right]$.*

3. *If $\mathcal{M}_1 \bigcap \mathcal{M}_2 = \emptyset$, resource partitioning is infeasible. In this case, if throughput constraints in sets $\mathcal{U}_1$ and $\mathcal{U}_2$ can be satisfied with $\eta = 1$ and $\mathrm{B}_1 = \mathrm{B}_2$, then the fully shared scheme is feasible and $\mathrm{P}_{\mathrm{net}}^{\mathrm{t1}*} = \lambda_1 \left[a_{\mathrm{M}} \mathrm{S}_1 W + b_{\mathrm{M}}\right] + \lambda_2 \left[a_{\mathrm{sc}} \mathrm{S}_2 W + b_{\mathrm{sc}}\right]$. Otherwise, the network cannot support the given rate requirement.*

Proof. Firstly, we consider the case $\mathcal{M}_1 \bigcap \mathcal{M}_2 \neq \emptyset$. According to the discussion above Corollary 2.4.1, $\eta_1(\mathrm{B}_1)$ decreases as B_1 increases. Thus, η^* is the larger of $\eta_1(b_2)$ and η_2, where $b_2 = \max\limits_{\mathrm{B}_1 \in \mathcal{M}_1 \bigcap \mathcal{M}_2}(\mathrm{B}_1)$. From Corollary 2.4.1, $\mathcal{M}_1 = \{\mathrm{B}_1 | \, \mathrm{B}_1 \geq b_1\}$. Hence, the value of b_2 is determined by constraint (2.21). We can conclude that $\eta_{\mathrm{D1}}(b_2) = \eta_2$, when $b_2 < \infty$. In addition, quantity b_2 also satisfies constraint (2.20) where $\eta_{\mathrm{D1}}(b_2) \geq \eta_1(b_2)$. Thus, it can be shown that $\eta_1(b_2) \leq \eta_2$ for all $b_2 < \infty$ and the minimum η therefore equals η_2. Secondly, if $b_2 = \infty$ is feasible, all the UEs with required rate C_1 can be served by MBSs on the unshared

resources (i.e. $\mathrm{B}_1 = \infty$). The two tiers of base stations will not cause inter-tier interference to each other. For small-cell BSs, the coverage probability $\mathbb{P}$ is derived in [81] and hence the minimum required resource fraction η^* equals $\frac{C_2\bar{\mathrm{N}}_2}{W\mathbb{P}\log_2(1+\Upsilon)}$. In that case, a fully unshared allocation scheme is used to achieve a minimum power consumption $\mathrm{P}_{\mathrm{net}}^{\mathrm{t1}*} = (1-\eta^*)\lambda_1\left[a_{\mathrm{ma}}\mathrm{S}_1 W + b_{\mathrm{ma}}\right] + \eta^*\lambda_2\left[a_{\mathrm{sm}}\mathrm{S}_2 W + b_{\mathrm{sm}}\right]$. Thirdly, it is possible that no feasible region exists for B_1, i.e. $\mathcal{M}_1 \bigcap \mathcal{M}_2 = \emptyset$. In that scenario, resource partitioning cannot be applied on the network with the given parameter settings. Instead, MBSs and small-cell BSs should adopt the fully shared scheme to allocate the resources. □

As shown in Theorem 2.4.1, when resource partitioning is feasible (results 1 and 2 in Theorem 2.4.1), the optimal η that minimizes network power consumption is determined by the small-cell user's throughput constraint. Furthermore, if all the macro users can be served on the unshared resources, which is the second result in Theorem 2.4.1, we can let MBSs and small-cell BSs operate on orthogonal resources to completely eliminate inter-tier interference.

According to our user rate definition, the average user throughput $\bar{T}_j$ is approximated by $\frac{\mathbb{P}_j}{\bar{\mathrm{N}}_j} W_j \log_2(1+\Upsilon)$, where $\mathbb{P}_j$ is the average probability of set $\mathcal{U}_j$ users' SIR exceeding the target value Υ, i.e. the event that these users are served. Additionally, as can be seen from Theorem 2.4.1, when the optimal $\eta^* = \eta_2$, B_1^* takes any value in the set $\mathcal{M}_1 \bigcap \mathcal{M}_2$ that satisfies $\eta_1(\mathrm{B}_1^*) \leq \eta_2$. Thus, there exist more than one B_1^* values that solve Problem (P1a). By further taking coverage maximization into account we can now determine a unique optimal value of B_1. The coverage maximization problem is

$$(\mathrm{P1b}): \quad \underset{(\eta,\mathrm{B}_1)\in\mathcal{F}_1}{\text{maximize}} \quad \mathbb{P}_{\mathrm{c}}.$$

From (2.14), the overall coverage probability $\mathbb{P}_{\mathrm{c}}$ only depends on B_1 and is a non-decreasing function of B_1. As a result, to solve (P1b), we simply select the maximum value of B_1 from its feasible range. Based on the discussion in the proof

of Theorem 2.4.1, the optimum η for (P1a) can also be achieved when B_1 takes the biggest value in its feasible set $\mathcal{M}_1 \bigcap \mathcal{M}_2$. Thus, we can simultaneously minimize the network power consumption and maximize UE coverage probability for *Case 1* (using MBSs to serve UEs in $\mathcal{U}_{\mathrm{D}}$). When resource partitioning is feasible, the optimal B_1^* is $\mathrm{B}_1^* = \max_{x \in \mathcal{M}_1 \bigcap \mathcal{M}_2}(x)$ and η^* is obtained using Theorem 2.4.1.

2.4.2 *Case 2*: Unshared Resources Allocated to Small-Cell BSs

Similar to the analysis for *Case 1*, a feasible (η, B_1) set $\mathcal{F}_2$ can be determined using the minimum throughput constraints for UEs in $\mathcal{U}_1$, $\mathcal{U}_2$ and $\mathcal{U}_{\mathrm{D}2}$. Since small-cell BSs are activated on the unshared resources, we have the throughput constraint $(1-\eta)W\mathbb{P}_{\mathrm{D}1}\log_2(1+\Upsilon) \geq C_1\bar{\mathrm{N}}_{\mathrm{D}1}$ in (2.19) replaced by $(1-\eta)W\mathbb{P}_{\mathrm{D}2}\log_2(1+\Upsilon) \geq C_1\bar{\mathrm{N}}_{\mathrm{D}2}$. The throughput constraints for UEs in $\mathcal{U}_1$ and $\mathcal{U}_2$ remain the same as those in (2.19). Thus, the feasible set $\mathcal{F}_2$ is derived as

$$\begin{aligned}\mathcal{F}_2 = \{(\eta, \mathrm{B}_1) \text{ s.t. } & \eta W\mathbb{P}_1\log_2(1+\Upsilon) \geq C_1\bar{\mathrm{N}}_1, \\ & \eta W\mathbb{P}_2\log_2(1+\Upsilon) \geq C_2\bar{\mathrm{N}}_2, \\ & (1-\eta)W\mathbb{P}_{\mathrm{D}2}\log_2(1+\Upsilon) \geq C_1\bar{\mathrm{N}}_{\mathrm{D}2}, \\ & \mathrm{B}_1 > \mathrm{B}_2 \}, \end{aligned} \tag{2.23}$$

The set $\mathcal{F}_2$ is re-expressed as $\{(\eta, \mathrm{B}_1) | \max\{\eta_1(\mathrm{B}_1), \eta_2\} \leq \eta \leq \eta_{\mathrm{D}2}(\mathrm{B}_1), \mathrm{B}_1 > \mathrm{B}_2\}$ with $\eta_1(\mathrm{B}_1) = \frac{C_1\bar{\mathrm{N}}_1}{W\mathbb{P}_1\log_2(1+\Upsilon)}$, $\eta_2 = \frac{C_2\bar{\mathrm{N}}_2}{W\mathbb{P}_2\log_2(1+\Upsilon)}$ and $\eta_{\mathrm{D}2}(\mathrm{B}_1) = 1 - \frac{C_1\bar{\mathrm{N}}_{\mathrm{D}2}}{W\mathbb{P}_{\mathrm{D}2}\log_2(1+\Upsilon)}$. The feasible set of B_1 is determined by

$$\eta_{\mathrm{D}2}(\mathrm{B}_1) \geq \eta_1(\mathrm{B}_1), \ \eta_{\mathrm{D}2}(\mathrm{B}_1) \geq \eta_2. \tag{2.24}$$

The procedure for finding the feasible set is similar to that in *Case 1*. From constraints in (2.24), two sets $\mathcal{M}_1'$ and $\mathcal{M}_2'$ can be respectively found. The η upper bound $\eta_{\mathrm{D}2}$ is a monotonically decreasing function of B_1, unlike in *Case 1* where the

monotonicity of η_{D1} depends on specific parameter settings. Hence, finding $\mathcal{M}_1'$ and $\mathcal{M}_2'$ is much easier for *Case 2*. Specifically, $\mathcal{M}_1' = \{b_1' \leq \mathrm{B}_1\}$ and $\mathcal{M}_2' = \{\mathrm{B}_1 \leq b_2'\}$, where b_1' can be obtained by applying Corollary 2.4.1 with minor modifications and b_2' is given as

1. $b_2' = 0$, if $\lim\limits_{x \to \mathrm{B}_2} \eta_{\mathrm{D2}}(x) < \eta_2$.
2. $b_2' = \infty$, if $\lim\limits_{x \to \infty} \eta_{\mathrm{D2}}(x) > \eta_2$.
3. Otherwise, the value of b_2' is obtained by solving $\eta_{\mathrm{D2}}(x) - \eta_2 = 0$.

The feasible range of B_1 is then $\mathcal{M}_1' \cap \mathcal{M}_2' = \{b_1' \leq \mathrm{B}_1 \leq b_2'\}$. Similarly, the energy reduction and coverage improvement problems for *Case 2* can be stated as

$$(\mathrm{P2a}): \quad \underset{(\eta, \mathrm{B}_1) \in \mathcal{F}_2}{\text{minimize}} \quad \mathrm{P}_{\mathrm{net}}^{\mathrm{t2}}, \qquad (\mathrm{P2b}): \quad \underset{(\eta, \mathrm{B}_1) \in \mathcal{F}_2}{\text{maximize}} \quad \mathbb{P}_{\mathrm{c}}.$$

The optimal (η^*, B_1^*) for (P2a) and (P2b) are denoted as $(\eta_a^*, \mathrm{B}_{1a}^*)$ and $(\eta_b^*, \mathrm{B}_{1b}^*)$, respectively.

As discussed for (P1a), solving (P2a) is equivalent to minimizing the term $\max\limits_{\mathrm{B}_1 \in \mathcal{M}_1' \bigcap \mathcal{M}_2'} (\eta_1(\mathrm{B}_1), \eta_2)$. The optimal solution for (P2a) is given in the following theorem. Due to its similarity to Theorem 2.4.1, we omit the proof of it.

Theorem 2.4.2. *When UEs in $\mathcal{U}_{\mathrm{D}}$ are served by the small-cell BSs, the optimal η_a^* and B_{1a}^* for (P2a) are found from the intersection of feasible sets $\mathcal{M}_1'$ and $\mathcal{M}_2'$.*

1. *If $\mathcal{M}_1' \bigcap \mathcal{M}_2' \neq \emptyset$ and $b_2' < \infty$, the optimal η_a^* for (P2a) equals η_2, and B_{1a}^* can take any value in the set $\mathcal{M}_1' \bigcap \mathcal{M}_2'$ that satisfies $\eta_1(\mathrm{B}_{1a}^*) \leq \eta_2$. The optimal power consumption per unit area is calculated as $\mathrm{P}_{\mathrm{net}}^{\mathrm{t2}*} = \eta_2 \lambda_1 \left[a_{\mathrm{M}} \mathrm{S}_1 W + b_{\mathrm{M}}\right] + \lambda_2 \left[a_{\mathrm{sc}} \mathrm{S}_2 W + b_{\mathrm{sc}}\right]$.*
2. *If $\mathcal{M}_1' \bigcap \mathcal{M}_2' \neq \emptyset$ and $b_2' = \infty$, all the UEs can be served by small-cell BSs, i.e. turning down MBSs on all resources, and $\mathrm{P}_{\mathrm{net}}^{\mathrm{t2}*} = \lambda_2 \left[a_{\mathrm{sc}} \mathrm{S}_2 W + b_{\mathrm{sc}}\right]$.*
3. *If $\mathcal{M}_1' \bigcap \mathcal{M}_2' = \emptyset$, resource partitioning is infeasible. In that case, if the fully shared scheme is feasible, $\eta = 1$ and $\mathrm{P}_{\mathrm{net}}^{\mathrm{t2}*} = \lambda_1 \left[a_{\mathrm{M}} \mathrm{S}_1 W + b_{\mathrm{M}}\right] +$*

$\lambda_2 [a_{\text{sc}} \text{S}_2 W + b_{\text{sc}}]$. *Otherwise, the network cannot support the given rate requirement.*

Similar as Theorem 2.4.1, Theorem 2.4.2 shows that the optimal η that minimizes network power consumption for *Case 2* is also determined by the small-cell user throughput constraint. However, result 2 in Theorem 2.4.2 is different from that in Theorem 2.4.1. With macro users offloaded to the nearest small-cells in this case, if $\text{B}_1 = \infty$ is feasible, we can use only small-cells to serve all users and mute all the MBSs to minimize the network power.

Optimizing coverage probability is not as straightforward in *Case 2* as in *Case 1*. Two results are given below to solve (P2b). Firstly, when all the users are served by small-cell BSs, the macro tier can be completely muted on all the resources. Then the overall coverage probability is denoted as $\mathbb{P}_\infty$. Referring to the derivation of coverage probability for homogeneous networks in [26], we can obtain $\mathbb{P}_\infty = \frac{1}{1+\rho(\Upsilon,\alpha)}$. The second result is summarized in the following lemma.

Lemma 2.4.1. *With resource partitioning adopted, the overall coverage probability is calculated using* (2.14) *as* $\mathbb{P}_\text{c} = \sum_{j\in\{1,2,\text{D2}\}} \mathbb{P}_j \mathbb{A}_j$. *If there exists a* B_c *that satisfies*

$$\frac{1}{\lambda_1 + (1+\rho(\Upsilon,\alpha))\,\lambda_2 \left(\frac{\text{S}_2\text{B}_\text{c}}{\text{S}_1}\right)^{\frac{2}{\alpha}}} = \frac{(\text{B}_\text{c}/(\Upsilon+\text{B}_\text{c}))^{\frac{1}{2}}}{\lambda_1(1+\rho(\Upsilon,\alpha)) + \lambda_2 \left(\frac{\text{S}_2\text{B}_\text{c}}{\text{S}_1}\right)^{\frac{2}{\alpha}} \left(1+\rho(\frac{\Upsilon}{\text{B}_\text{c}},\alpha)\right)}, \tag{2.25}$$

the coverage probability $\mathbb{P}_\text{c}$ *is a unimodal function with its maximum value achieved at* $\text{B}_1 = \text{B}_\text{c}$. *Otherwise,* $\mathbb{P}_\text{c}$ *is monotonic in* B_1.

Proof. Please refer to Appendix B. □

Based on Lemma 2.4.1, the optimal $\text{B}_1 = b_0$ that maximizes $\mathbb{P}_\text{c} = \sum_{j\in\{1,2,\text{D2}\}} \mathbb{P}_j \mathbb{A}_j$ can be found within the feasible range $b_1' \le \text{B}_1 \le b_2'$. The value of b_0 is determined as

1. If $\text{B}_\text{c} \le b_1'$, $b_0 = b_1'$.

2. If $b_1' < \mathrm{B_c} \leq b_2'$, $b_0 = \mathrm{B_c}$.

3. If $\mathrm{B_c} > b_2'$, $b_0 = b_2'$.

4. If $\mathrm{B_c}$ does not exist, $b_0 = \underset{\{b_1', b_2'\}}{\arg\max} \mathbb{P}_\mathrm{c}$.

Note that when $b_2' = \infty$, the network becomes homogeneous and there is a sudden change from $\mathbb{P}_\mathrm{c}$ to $\mathbb{P}_\infty$. Therefore, when $b_2' = \infty$ the optimal B_{1b}^* for (P2b) is selected from b_0 and ∞ that maximizes $\mathbb{P}_\mathrm{c}$. Finding the optimal ${\mathrm{B_{1b}}}^*$ and η_b^* for (P2b) is summarized in the following theorem.

Theorem 2.4.3. *When UEs in $\mathcal{U}_\mathrm{D}$ are served by the small-cell BSs and $\mathcal{M}_1' \bigcap \mathcal{M}_2' \neq \emptyset$, the optimal η_b^* and B_{1b}^* for (P2b) are*

1. $\mathrm{B}_{1b}^* = b_0$ *and* $\eta_b^* = \max\{\eta_1(b_0), \eta_2\}$, *when* $b_2' < \infty$.

2. $\mathrm{B}_{1b}^* = b_0$ *and* $\eta_b^* = \max\{\eta_1(b_0), \eta_2\}$, *when* $b_2' = \infty$ *and* $\mathbb{P}_\mathrm{c}|_{\mathrm{B}_1 = b_0} \geq \mathbb{P}_\infty$.

3. $\mathrm{B}_{1b}^* = \infty$ *and the network is homogeneous consisting of only small-cell BSs, when* $b_2' = \infty$ *and* $\mathbb{P}_\mathrm{c}|_{\mathrm{B}_1 = b_0} < \mathbb{P}_\infty$.

With the user association bias B_1 set to its optimal value that maximizes overall coverage probability, the corresponding resource partitioning fraction factor η that achieves the minimum network power consumption is determined in Theorem 2.4.3. The optimal values $(\eta_b^*, \mathrm{B}_{1\mathrm{b}}^*)$ determined by Theorem 2.4.3 may not equal to the optimal values $(\eta_a^*, \mathrm{B}_{1\mathrm{a}}^*)$ obtained using Theorem 2.4.1. Therefore, solving the power minimization problem (P2a) and the coverage maximization problem(P2b) for *Case 2* may not always give the same solutions.

2.4.3 Additional Comments

In the above two subsections, we showed how to find the optimal user association for load adaptation and resource allocation schemes in *Case 1* and *Case 2*. According to the discussion for *Case 1*, the minimum power consumption per unit area and maximum overall network coverage probability can be achieved simultaneously,

where $\mathrm{B}_1^* = \max_{x \in \mathcal{M}_1 \bigcap \mathcal{M}_2}(x)$ and η^* is obtained using Theorem 2.4.1. In *Case 2*, however, the optimum power consumption and coverage probability may be achieved at different (η, B_1) pairs. In other words, problems (P2a) and (P2b) do not always have a common solution. The optimal $(\eta_a^*, \mathrm{B}_{1a}^*)$ for (P2a) and $(\eta_b^*, \mathrm{B}_{1b}^*)$ for (P1b) are given in Theorems 2.4.2 and 2.4.3, respectively.

Whether to assign the unshared resources to the MBSs (*Case 1*) or to the small-cell BSs (*Case 2*) thus depends on the relative importance of power consumption and network coverage. To find the optimal (η, B_1) that minimizes network power consumption, we can first solve (P1a) and (P2a) separately and then choose the smaller optimal power. Similarly, for maximizing network coverage, we can find the optimal coverage probabilities derived from (P1b) and (P2b) and then choose the better one.

We can extend the analysis to a more general case, where the unshared resources allocated to user set $\mathcal{U}_\mathrm{D}$ are further divided into two parts. The MBSs could have exclusive access to one part of the unshared resources, while the small-cell BSs have exclusive access to the other. MBSs and small-cell BSs can jointly access the shared resources as before. The problems discussed in the above sections are special cases of this general model. According to Theorems 2.4.1 and 2.4.2, it can be concluded that when resource partitioning is feasible the optimal resource sharing factor η equals η_2. Note that the value of η_2 is the percentage of resources allocated to UEs in $\mathcal{U}_2$ to achieve their target rate. As the expression $\eta_2 = \frac{C_2 \bar{\mathrm{N}}_2}{W \mathbb{P}_2 \log_2(1+\Upsilon)}$ below (2.19) shows, η_2 only depends on B_2. Similarly, for the general case, the optimal η for feasible resource partitioning only depends on B_2. Therefore, with fixed B_2, we only need to consider the problem of resource allocation and user association within the user set $\mathcal{U}_\mathrm{D}$. Since there is no inter-tier interference in $\mathcal{U}_\mathrm{D}$, this problem can be solved using the method proposed in [81] with minor changes.

2.5 Numerical Results

In this section, numerical results are provided to verify the analytical results derived in the above section. The bandwidth and base station transmit power spectral density are set as follows: $W = 10$ MHz, $\mathrm{S}_1 = 2 \times 10^{-3}$ mW/Hz and $\mathrm{S}_2 = 1 \times 10^{-4}$ mW/Hz. According to this parameter setting, the transmit power of MBS is upper bounded by 20 W and the transmit power of small-cell BS is no larger than 100 mW, which are valid assumptions according to [79]. In simulation, the assumed PL model has parameters $\alpha = 3.8$ and $K = -30$ dB. These parameters have been shown to model the practical PL environment well [83].

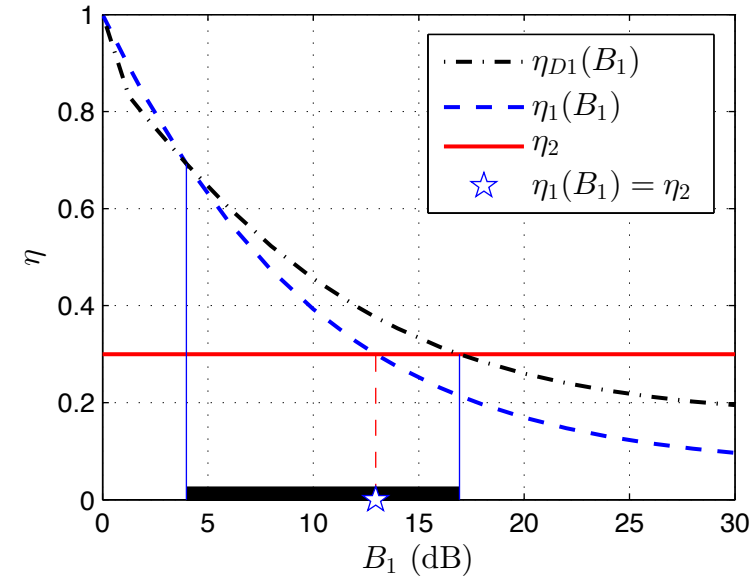
According to our resource partitioning and user allocation policy, MBSs and small-cell BSs use the shared η fraction of resources to serve users in $\mathcal{U}_1$ and $\mathcal{U}_2$, respectively. When *Case 1* is considered, small-cell BSs are switched to sleep mode on the unshared resources and only MBSs transmit to the $\mathcal{U}_{\mathrm{D1}}$ users. On the other hand, in *Case 2*, users who have been allocated the unshared resources are served by small-cell BSs. In this scenario, MBSs are muted on the unshared resources. The effects of association bias B_1 on the size of user set $\mathcal{U}_j$ ($j \in \{1, 2, \mathrm{D1}, \mathrm{D2}\}$) can be found in Section 2.2.1.

In Fig. 2.2, user density λ_{u} and base station density λ_1, λ_2 are set as $\lambda_{\mathrm{u}} = 15\lambda_1 = 3\lambda_2 = \frac{15}{500^2\pi}$ units/m^2. These settings mean that MBSs have on average 500 m coverage radius, and an average of 4 small-cells and 20 UEs are within the coverage of each MBS. The value of B_2 is fixed at 1 (0 dB). The SIR threshold $\Upsilon = 0.5$ and small-cells need at least 30% of the resources to support the required rate for UEs in $\mathcal{U}_2$ (i.e., $\eta_2 = 0.3$, where $\eta_2 = \frac{C_2 \tilde{\mathrm{N}}_2}{W \mathbb{P}_2 \log_2(1+\Upsilon)}$ has been defined below (2.19)). In Fig. 2.2 (a), we show how the boundaries of η ($\eta_1(\mathrm{B}_1)$, $\eta_{\mathrm{D1}}(\mathrm{B}_1)$ and η_2) change with B_1 in *Case 1*. The upper bound $\eta_{\mathrm{D1}}(\mathrm{B}_1)$ and the two lower bounds $\eta_1(\mathrm{B}_1)$ and η_2 are derived from the feasible (η, B_1) set $\mathcal{F}_1$ in (2.19). According to (2.20) and (2.21), the feasible range of B_1 is the one that makes $\mathcal{F}_1$ non-empty. By comparing the η bounds in Fig. 2.2 (a) we can get the feasible B_1 range, which is $[3.97, 16.93]$ dB, and we highlight this interval in bold on the x-axis for clarity.

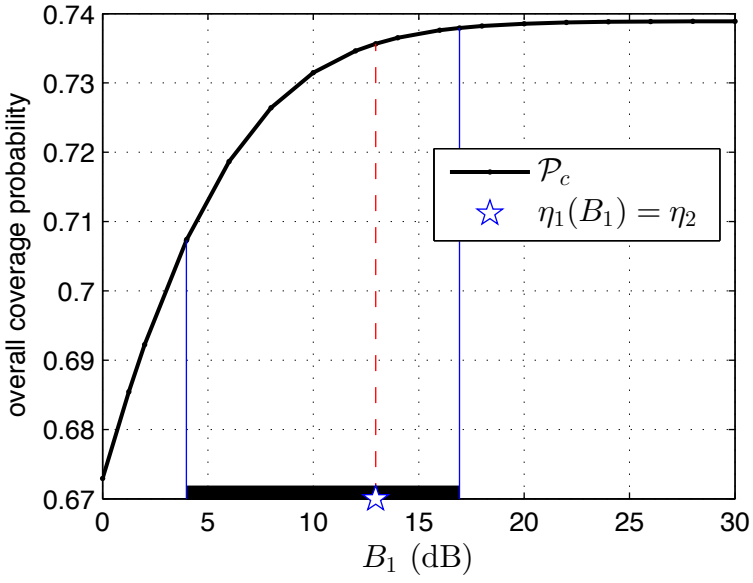
With the feasible B_1 set determined, we are able to apply Theorem 2.4.1 to find the minimum power consumption for *Case 1*. From Theorem 2.4.1, the minimum η is selected to minimize $\mathrm{P}_{\mathrm{net}}^{\mathrm{t1}}$. According to Fig. 2.2 (a), $\eta_1(\mathrm{B}_1) > \eta_2$ for $\mathrm{B}_1 < 12.96$ dB (indicated by pentagram) and $\eta_1(\mathrm{B}_1) \leq \eta_2$ otherwise. Since η_2 is independent of B_1, the minimum power consumption is achieved at $\eta = \eta_2$ with $\mathrm{B}_1 \in [12.96, 16.93]$ dB. From the discussion under Theorem 2.4.1, we know that the overall coverage probability $\mathbb{P}_{\mathrm{c}}$ for *Case 1* is a non-decreasing function of B_1, as shown in Fig. 2.2 (b). It can be observed from Fig. 2.2 (b) that the optimum overall coverage probability is achieved at the maximum value of B_1 within its feasible range determined in Fig. 2.2 (a).

Under the same parameter settings, a similar analysis is adopted for *Case 2* to determine the feasible range of B_1 and the maximum achievable coverage probability as shown in Fig. 2.3 (a) and Fig. 2.3 (b). From Fig. 2.3 (a), we can know that the feasible interval for B_1 is $[1.38, \infty)$ dB. By setting B_1 to infinity, we can have all MBSs muted and all UEs served by small-cells so that the network is homogeneous and consumes the least amount of energy. However, as shown in Fig. 2.3 (b), the overall coverage probability $\mathbb{P}_{\mathrm{c}}$ for *Case 2* is not an increasing function of B_1. When $\mathrm{B}_1 = \infty$, the coverage probability is denoted as $\mathbb{P}_{\infty}$, which is smaller than the maximum $\mathbb{P}_{\mathrm{c}}$ achieved at B_{c} (5.86 dB). Therefore, a tradeoff between reducing energy consumption and improving coverage probability exists in *Case 2*.

Fig. 2.4 and Fig. 2.5 are obtained by changing Υ from 0.5 to 10. Similarly, the feasible range of B_1 is $[7.73, 10.31] \bigcup [17, \infty]$ dB for *Case 1* and $[2.57, 16.16]$ dB for *Case 2*. From Fig. 2.4 (a), we note that B_1 can be set to ∞ for *Case 1*, which means fully unshared resource allocation is adopted and the minimum amount of power is consumed. With $\mathrm{B}_1 = \infty$, Fig. 2.4 (b) shows that the overall coverage probability achieves the maximum value at $\mathrm{B}_1 = \infty$. Specifically, in Fig. 2.4 (b) we denote the value of $\mathbb{P}_c$ at $\mathrm{B}_1 = \infty$ as $\mathbb{P}_{\infty}$. Using the analysis in [81], we have $\mathbb{P}_{\infty} = \frac{\lambda_1}{\lambda_1[1+\rho(\Upsilon,\alpha)]+\lambda_2\left(\frac{\mathrm{B}_2\mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}}} + \frac{\left(\frac{\mathrm{B}_2\mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}}\lambda_2}{\lambda_1+[1+\rho(\Upsilon,\alpha)]\lambda_2\left(\frac{\mathrm{B}_2\mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}}}$. In Fig. 2.5 (a), the feasible B_1 is upper bounded by 16.6 dB for *Case 2*. In that scenario, we cannot mute

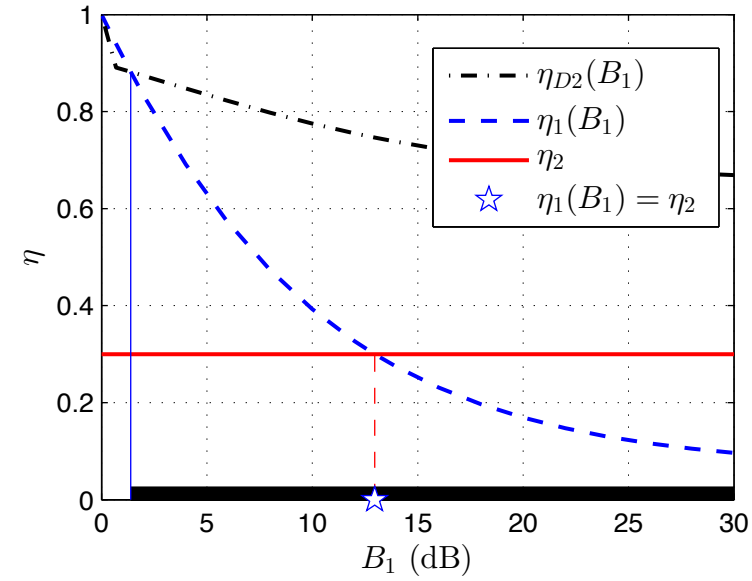


(a)

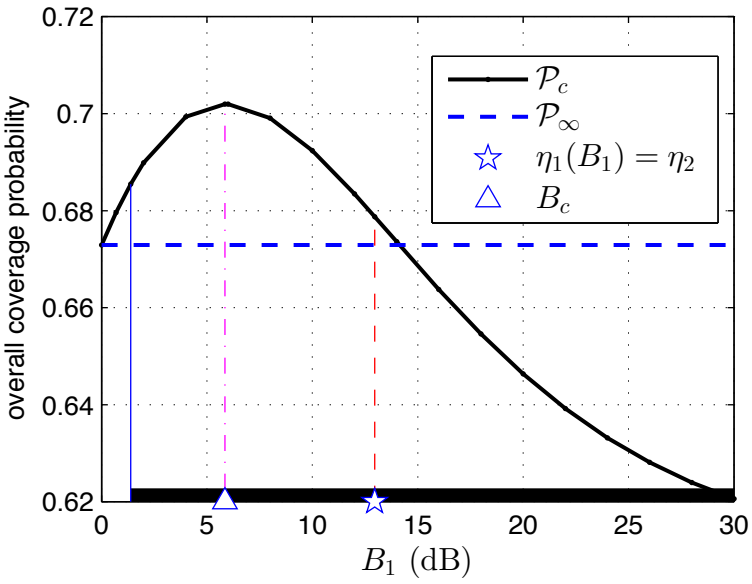


(b)

Figure 2.2: Effects of B_1 on η and overall coverage probability $\mathbb{P}_c$ with SIR threshold $\Upsilon = 0.5$ for *Case 1*. The thick horizontal lines in the figures indicate the feasible ranges of B_1.



(a)



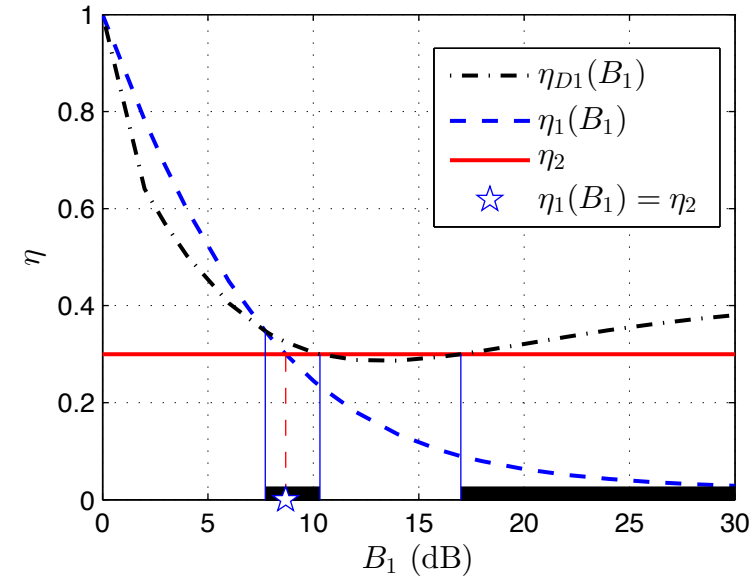
(b)

Figure 2.3: Effects of B_1 on η and overall coverage probability $\mathbb{P}_c$ with SIR threshold $\Upsilon = 0.5$ for *Case 2*. The thick horizontal lines in the figures indicate the feasible ranges of B_1.

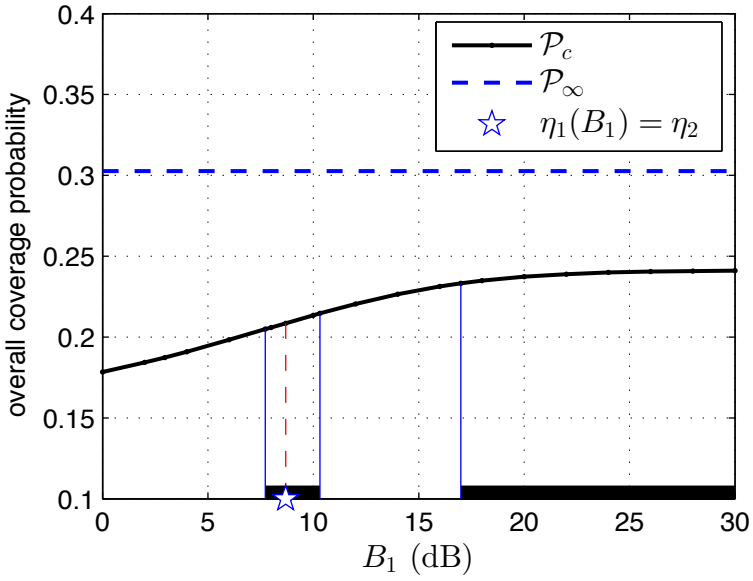
all the MBSs. According to Theorem 2.4.2, we can draw the conclusion that the minimum power consumption for *Case 2* is achieved when $\mathrm{B}_1 \in \mathcal{F}_2$ and $\eta_1(\mathrm{B}_1) \leq \eta_2$, which is reflected in Fig. 2.5 (a) as the highlighted range $\mathrm{B}_1 \in [8.69, 16.16]$ dB. Furthermore, the value of B_c (8.9 dB) that maximizes $\mathbb{P}_\mathrm{c}$ in Fig. 2.5 (b) also lies within this interval. Thus, for the parameter settings in Fig. 2.4 and Fig. 2.5, the minimum power consumption and maximum coverage probability will be achieved simultaneously for both *Case 1* and *Case 2*. From the discussions of Fig. 2.2 – Fig. 2.5, we can conclude that the feasible ranges of B_1 depend on the values of other parameters (in this case the parameter is Υ). It also supports our claim in Section 2.4.3 that simultaneously achieving the minimum power consumption and maximum coverage probability is always valid for *Case 1* but cannot be guaranteed for *Case 2*.

According to Lemma 2.3.1 and Lemma 2.3.2, the population of user set $\mathcal{U}_j$ has a great effect on our decision of selecting the optimal resource allocation and user association schemes. Since B_2 and base station density ratio $\frac{\lambda_2}{\lambda_1}$ are the two key factors, we show how the optimal scheme varies with B_2 and $\frac{\lambda_2}{\lambda_1}$ in Fig. 2.6 and Fig. 2.7, respectively. In those figures, we set $\Upsilon = 0.5$, $\lambda_1 = \frac{1}{500\pi^2}$ units/m^2 and the required rates are $C_1 = 300$ kbps and $C_2 = 900$ kbps. The value of B_2 varies from -2 dB to 4 dB and $\frac{\lambda_2}{\lambda_1} = 5$ in Fig. 2.6. On the other hand, B_2 is fixed at 1 and $\frac{\lambda_2}{\lambda_1} \in [3.5, 10]$ in Fig. 2.7. Recalling the four optimization problems (P1a), (P1b), (P2a) and (P2b) in Section 2.4, different optimal (η, B_1) pairs are found accordingly. Each optimal (η, B_1) pair individually results in its power consumption and coverage probability. From Theorems 2.4.1 and 2.4.2, there may exist more than one value of B_1 that minimizes network power for (P1a) and (P2a). Thus, for each of these two problems we select the B_1 from the optimal values, which achieves the highest coverage probability.

The power consumptions and coverage probabilities of problem (P1a), (P1b), (P2a) and (P2b) are compared with respect to B_2 in Fig. 2.6. According to Section 2.2.1, B_2 controls the population in user classes with required rates C_1 and C_2. As

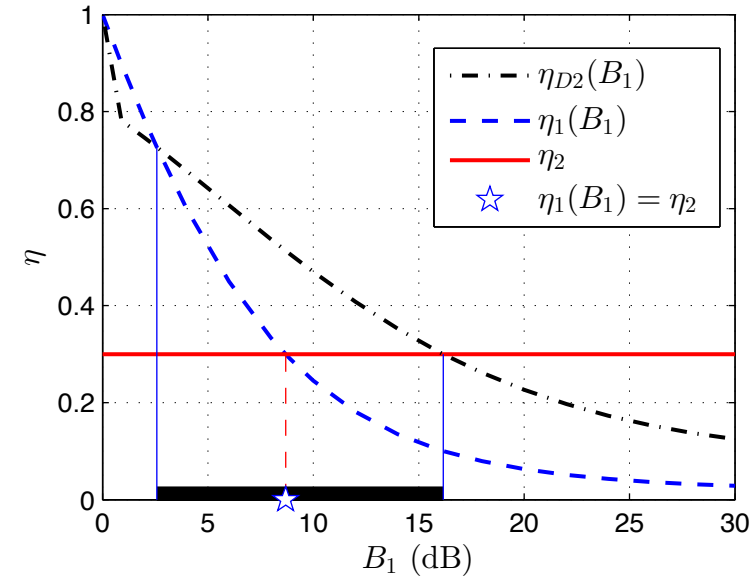


(a)

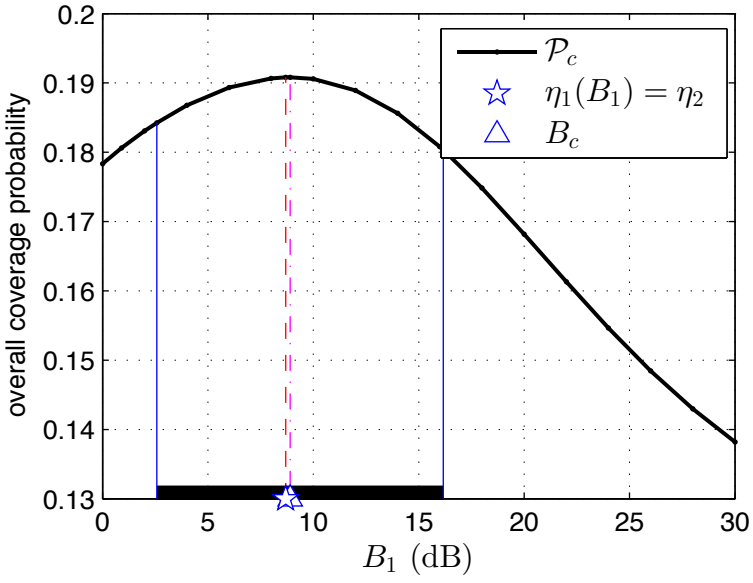


(b)

Figure 2.4: Effects of B_1 on η and overall coverage probability $\mathbb{P}_c$ with SIR threshold $\Upsilon = 10$ for *Case 1*. The thick horizontal lines in the figures indicate the feasible ranges of B_1.



(a)

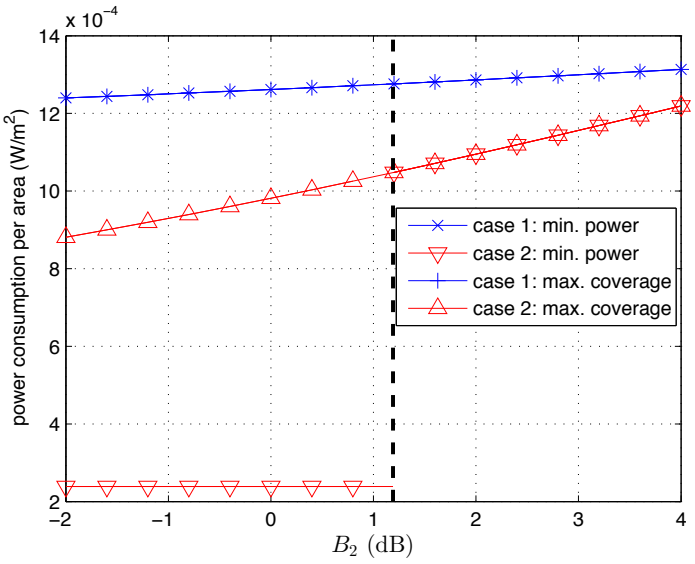


(b)

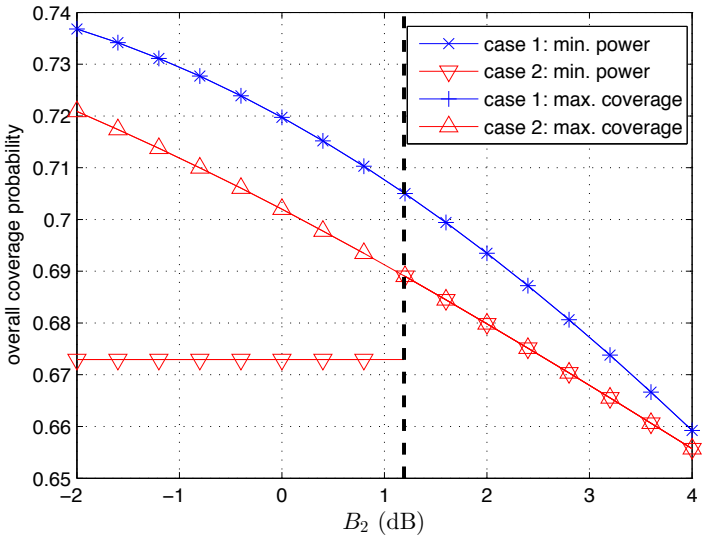
Figure 2.5: Effects of B_1 on η and overall coverage probability $\mathbb{P}_c$ with SIR threshold $\Upsilon = 10$ for *Case 2*. The thick horizontal lines in the figures indicate the feasible ranges of B_1.

B_2 increases, more users will be served at rate C_2 by small-cell BSs. Hence, the size of set $\mathcal{U}_2$ grows, which is reflected by a larger value of $\mathbb{A}_2$. On the other hand, due to the limited transmit power of small-cell BSs and path loss effect, the average distance of UEs in $\mathcal{U}_2$ increases with B_2, resulting in a decrease in $\mathbb{P}_2$. In order to achieve the throughput requirements, the value of η is raised to allocate more resources to UEs in $\mathcal{U}_2$. Since the network power consumption grows with η, the curves in Fig. 2.6 (a) all increase with B_2. From the previous discussion we know that power minimization and coverage maximization can be achieved at the same (η, B_1) pair in *Case 1*. Therefore, the power consumption in Fig. 2.6 (a) and coverage probability in Fig. 2.6 (b) are the same for the crossed- and x-labeled lines. In these two curves, the fraction η of the resources shared with small-cell BSs increases with B_2. For *Case 2*, however, the power minimization and coverage maximization are not guaranteed to be achieved simultaneously. In Fig. 2.6 (a), when $B_2 < 1.19$ dB the optimal energy saving scheme is obtained by setting $B_1 = \infty$, which means using only small-cell BSs to serve all the UEs. However, the coverage optimal strategy is to allow MBSs stay awake on some shared fraction of resources (the coverage performance can be seen in Fig. 2.6 (b)). For B_2 exceeds 1.19 dB, to serve UEs at required rate C_2 consumes too many resources that makes muting all MBSs impossible. Thus, MBSs are awakened on the shared resources and a sudden change is observed at $B_2 = 1.19$ dB for the *Case 2* minimum power curve and maximum coverage probability curve in Fig. 2.6 (a) and Fig. 2.6 (b) respectively.

In Fig. 2.7, we show how power consumption and coverage probability change with the number of small-cell BSs. The analyses for Fig. 2.7 (a) and 2.7 (b) are similar to those for Fig. 2.6 (a) and 2.6 (b). Due to the limited coverage area of small-cell BSs, when the ratio of $\frac{\lambda_2}{\lambda_1}$ is low, only a small group of UEs associate with small-cell BSs and therefore resource sharing is required. However, when the number of small-cells is sufficiently high, we can either adopt fully unshared resource allocation in *Case 1* ($\frac{\lambda_2}{\lambda_1} > 7.99$) or mute all MBSs in *Case 2* ($\frac{\lambda_2}{\lambda_1} > 4.75$). As observed from Fig. 2.6 and Fig. 2.7, choosing *Case 1* will generally consume higher energy



(a)



(b)

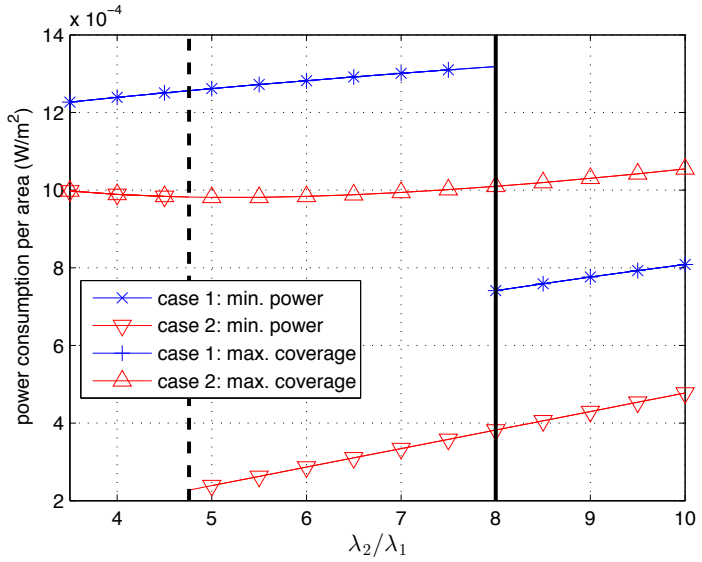
Figure 2.6: Effects of B_2 on (a) network power consumption and (b) coverage performance.

than choosing *Case 2* due to the lower energy consumption of small cells. However, in terms of coverage maximization, *Case 1* may outperform *Case 2* at most of the points in Fig. 2.6 and 2.7. This is because the signal received by small-cell edge users is low in strength and thus reduces the coverage probability.

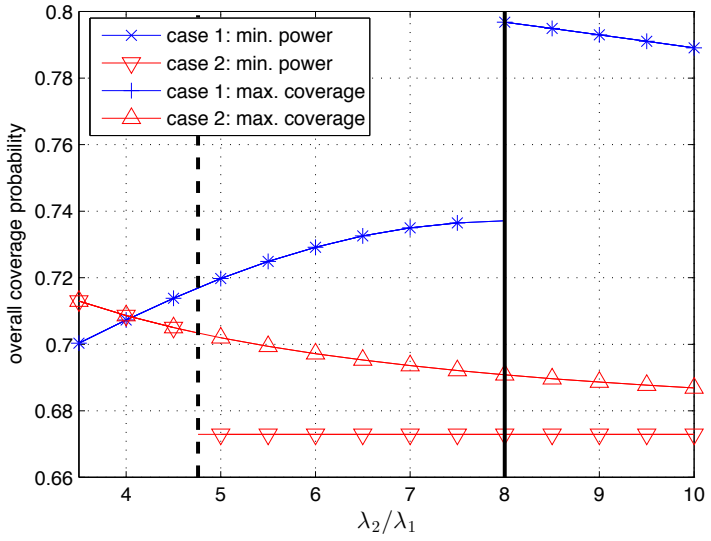
Although a stochastic geometry model that assumes purely random network deployment is adopted in this chapter, the derived results still give practical guidelines for real network designs. In practical network designs, MBSs could gather the load information of the small-cells within their coverage areas and then determine the optimal association bias $B_1{}^*$ and the resource partitioning factor η^* using the theorems proposed in Section 2.4 for both *Case 1* and *Case 2*. If *Case 1* gives a better performance in terms of power consumption or user coverage than *Case 2* does, a $1-\eta^*$ fraction of the resources is used by only MBSs, and the small-cell BSs are notified to go to sleep mode on these resources. Otherwise, the MBSs are muted on the unshared resources. The association biases are then transmitted to the users through the control channel so that each user can determine which tier to associate with and which set of resource elements will be used.

2.6 Conclusion

In this chapter, we provided a theoretical framework to study resource partitioning and load adaptation strategies in two-tier HCNs. Using stochastic geometry, closed-form expressions of coverage probability and average user throughput were obtained for each user set. The minimization of average base station power consumption and maximization of overall coverage probability with given throughput constraints were formulated separately. Optimal solutions were found by finding feasible sets of association bias factor B_1 and resource partition fraction η. Numerical results verified that the proposed optimal resource allocation strategy helps in reducing network power consumption and improving user coverage.



(a)



(b)

Figure 2.7: Effects of base station densities on (a) network power consumption and (b) coverage performance.

Chapter 3

Power Saving Design of Femtocell Exclusion Zones

3.1 Introduction

In Chapter 2, we investigated resource partitioning coupled with cell load adaptation to reduce HCN power consumption by muting base stations on the partitioned resources. In this chapter, we elaborate on a femtocell network and study a more challenging HCN power minimization problem based on designing exclusion zones which are circular shaped areas deployed around the femtocells. During solving this problem, resource partitioning and user association are jointly investigated and MBS transmit power varied with cell load is considered.

The main purpose of using exclusion zones is to eliminate simultaneous transmissions by base stations that cause severe interference to each other. Based on stochastic geometry modelling, several kinds of exclusion zones were proposed for HCN uplinks [84] and downlinks [29, 31, 77, 85]. The user outage probability and spectral efficiency were characterized with closed-form expressions in these works. However, the works in [29, 31, 84, 85] did not investigate the optimal design of exclusion zones to improve network energy efficiency. In [77], an energy-efficient design method was proposed for exclusion zones in a two-tier HCN, where any two small-cells located within a separation distance were not allowed to transmit simultaneously. Based on stochastic geometry analysis, the optimal separation distance, i.e. the radius of the exclusion zone, that minimized network power consumption was determined in [77]. Nevertheless, inter-tier interference

mitigation was not discussed and the base station transmit power was assumed to be independent of the cell load in [77]. It is shown that adaptively changing the transmit power of base stations according to cell load and user QoS may bring about additional performance gain [86–89], and therefore must be an important aspect of the energy-efficient HCN design. In [86], an uplink power control scheme was proposed to mitigate aggregated interference in femtocell networks. The design of HCN downlink power control schemes were investigated in [87,88] to correspondingly reduce network power consumption and improve throughput. In [89], a downlink transmit power control scheme was proposed in femtocell networks with exclusion zones to help users achieve outage probability requirements. However, transmit power adaptation algorithms in [86–89] were heuristic and analytically intractable. Therefore, a tractable framework that considers energy-efficient exclusion zone design together with adaptive transmit power control for HCNs is desired.

To address the above mentioned problems, we propose an exclusion zone design strategy in a two-tier system consisting of single MBS and multiple femtocells. According to [90,91], there are two kinds of access control mechanisms for femtocells, namely *closed access* and *open access*. For femtocells operating in closed access, only femto users can access the fBSs and macro users can only be served by the MBSs. For open access femtocells, a fBS can communicate with the macro users as long as the users are within the femtocell coverage. Therefore, for femtocells working under closed and open access modes, we correspondingly let macro users in the exclusion zones be served by the MBS and fBSs on inter-tier interference-free sub-bands. In this chapter, we propose a circular shaped exclusion zone model, which results in a user association rule different from the one based on the biased downlink average received power used in Chapter 2 and many other works [63,81,92]. Using the proposed exclusion zone model, stochastic geometry analysis is adopted to characterize cell loads and coverage probabilities in closed-form expressions. Additionally, we allow the MBS to adapt the transmit power according to its cell load and optimally design the exclusion zone size and allocated bandwidth, which

minimize the MBS power consumption.

The remainder of this chapter is structured as follows. The system model is described in Section 3.2, where the exclusion zone deployment and corresponding user association schemes are discussed. Based on the proposed exclusion zone model, we characterize cell load as well as success probability in Section 3.3. In Section 3.4, we formulate and solve the optimization problem that minimizes the MBS's transmit power by designing the size of exclusion zones and the partitioned bandwidth. Numerical results are given and discussed in Section 3.5. Finally, Section 3.6 concludes this chapter.

3.2 System Model

3.2.1 Base Station Deployment and Channel Model

In this chapter, we consider a system with one circular macrocell of radius $\mathrm{R_M}$ and multiple femtocells within it, which has also been used in [85, 93, 94]. This is a reasonable approximation if we assume orthogonal channel allocations among neighbouring macrocells, which limits macro-macro interference levels and hence allows each macrocell to be treated in isolation. We assume the MBS is at the origin and label it as the transmitter (Tx) node 0. Femtocells are assumed to be installed and powered on randomly by end-users, and therefore fBSs operating on the MBS's channels are modelled as an HPPP Φ_f with intensity λ_f. The femtocells are labeled as $\{1, 2, 3, \dots\}$. Each femtocell coverage area is a circle with radius $\mathrm{R_f}$ centered on the fBS. Without loss of generality, fBS 1 is assumed to be the nearest fBS to the typical user i and the distance between fBS 1 and user i is denoted as $d_{i,1}$. We define

$$\begin{cases} \text{if } d_{i,1} < \mathrm{R_f}, \text{ user } i \text{ is a femto user;} \\ \text{if } d_{i,1} \geq \mathrm{R_f}, \text{ user } i \text{ is a macro user.} \end{cases} \tag{3.1}$$

Within the coverage of fBSs, femto users are uniformly deployed with intensity $\lambda_{\mathrm{u}}^{\mathrm{f}}$. The macro users with intensity $\lambda_{\mathrm{u}}^{\mathrm{M}}$ are randomly located outside the coverage of femtocells following an uniform distribution. Note that the deployment of femtocells aims to improve QoS for indoor users [95]. Therefore, for simplicity, we assume all femto users are located indoors and all macro users are located outdoors in this chapter.

Let $\mathrm{P_M}$ and $\mathrm{P_f}$ be the transmit power of MBS and fBS, respectively. In the following analysis, we assume that the MBS adjusts its transmit power $\mathrm{P_M}$ according to the cell load while the fBSs transmit at constant power $\mathrm{P_f}$. The justifications of this assumption are as follows. Firstly, from the base station power consumption model in [79], the transmit power accounts for a large amount of the total power consumption of an MBS. Moreover, it is shown in [79] that the major power consumption of each fBS is transmission-independent. Therefore, letting each fBS adapt its transmit power $\mathrm{P_f}$ results in limited power reduction for the fBS. Secondly, the MBS typically serves more users than each fBS does, and the variance of user number in the macrocell is larger than that in a femtocell. The former implies that $\mathrm{P_M} \gg \mathrm{P_f}$, while the latter means that the optimal $\mathrm{P_f}$ for different fBSs will not change much from one deployment to another. Due to these two facts, we adapt $\mathrm{P_M}$ according to the cell load, but not $\mathrm{P_f}$, to reduce the MBS power consumption.

Similar to the channel model discussion in Chapter 2, flat Rayleigh fading and distance based PL are considered to account for signal attenuation. The effects of shadow fading can be taken into account by using the method proposed in [96]. However, for simplicity, the discussion in this chapter does not consider the random shadow fading. The power attenuation due to Rayleigh fading between receiver (Rx) i and Tx j is denoted as an exponential random variable $H_{i,j}$ with unit mean. According to [97], the PL models can be obtained as follows with wall penetration loss considered for outdoor-indoor and indoor-indoor wireless propagation.

MBS to macro users. Let the distance between macro user i and the MBS be $d_{i,0}$. The PL model is denoted as $g_{\mathrm{M}}(d_{i,0}) = K d_{i,0}^{-\alpha_{\mathrm{o}}}$, where K is the fixed

propagation loss and α_{o} is the outdoor PL exponent.

MBS to femto users. For a typical femto user j, the PL model should take wall penetration loss into account. The PL model is given as $g_{\mathrm{f}}^{\mathrm{M}}(d_{j,0}) = \omega K d_{j,0}^{-\alpha_{\mathrm{oi}}}$, where ω is the wall penetration loss arising from outdoor to indoor propagation; α_{oi} is the outdoor to indoor propagation PL exponent. Usually, outdoor propagation distance is larger than the one during indoor transmission. Therefore, for simplicity, we use the outdoor part to approximate the outdoor-indoor propagation and let $\alpha_{\mathrm{oi}} = \alpha_{\mathrm{o}}$.

fBS to macro users. We denote the distance between macro user i and fBS k as $d_{i,k}$. Correspondingly, the PL is $g_{\mathrm{M}}^{\mathrm{f}}(d_{i,k}) = \omega K d_{i,k}^{-\alpha_{\mathrm{o}}}$.

fBS to designated femto users. Without loss of generality, the distance between femto user j and its designated fBS 1 is denoted as $d_{j,1}$. The PL model is $g_{\mathrm{f}}(d_{j,1}) = K d_{j,1}^{-\alpha_{\mathrm{in}}}$. In this model, indoor PL exponent is α_{in}.

fBS to neighbouring femto users. The PL between femto user j and an interfering fBS k is $g_{\mathrm{f}}^{\mathrm{f}}(d_{j,k}) = \omega^2 K d_{j,k}^{-\alpha_{\mathrm{o}}}$. The wall penetration loss ω^2 is due to assuming that two femto users in different femtocells are separated by two walls on average [85, 94]. This is admittedly a rather crude approximation.

3.2.2 Exclusion Zone Setup

For closed access femtocells, the fBSs only serve the covered indoor users. For open access femtocells, on the other hand, the fBSs are able to offload and then provide service to the macro users that are within range. Therefore, in the subsequent analysis, we will respectively elaborate on the two femtocell access modes to investigate the exclusion zone design problem.

Resource Partitioning

Let the total bandwidth be W. We divide the total spectrum into two disjoint parts each with bandwidth ηW and $(1-\eta)W$, where η is defined as the resource partitioning factor with its value within the range $[0, 1]$. For simplicity, we name

sub-bands occupying the bandwidth ηW as *shared sub-bands*. Correspondingly, the remaining sub-bands with bandwidth $(1-\eta)W$ are called *unshared sub-bands*. Both the macro and femto base stations are allowed to simultaneously transmit on the shared sub-bands. On the other hand, the unshared sub-bands are allocated to only one tier of base stations to serve macro users or femto users so that inter-tier interference is non-existent on these sub-bands. Due to the fact that femtocells are typically lightly loaded, we only consider the case where the unshared sub-bands are used to serve macro users in the remainder of this chapter. In order not to detract from the main contributions of this chapter, the generalized analyses for the cases where both macro and femto users are allowed to access the unshared sub-bands are not discussed.

Exclusion Zone and User Association Rule

For each femtocell, we propose an exclusion zone around the fBS's coverage as a circular ring with inner radius $\mathrm{R_f}$ and outer radius $\mathrm{R_o}$ centered on the fBS. Based on the circular exclusion zone setup, a user association rule will then be determined, which is different from the previously proposed user association scheme in Chapter 2 where the biased signal powers received from the MBS and the nearest fBS are compared. The user association rule is described as follows.

Recall that in (3.1), if the distance $d_{i,1}$ between typical user i and its nearest fBS (fBS 1) is smaller than $\mathrm{R_f}$, user i is a femto user; otherwise, user i is a macro user. For simplicity, we use $\mathcal{U}_\mathrm{f}$ to denote the set of all femto users. The set of macro users will be further divided based on the distances between macro users and fBSs. Due to the power-law PL model, a user experiences less interference from an interfering base station if it is further away from that base station. Therefore, for macro users located within the femtocell exclusion zones, they may suffer from severe inter-tier interference if the MBS and fBSs simultaneously transmit on the same channel. For macro users located out of the exclusion zones, they see only a limited level of inter-tier interference. As a result, we use $\mathcal{U}_\mathrm{D}$ and $\mathcal{U}_\mathrm{M}$ to denote

the set of macro users within and out of the exclusion zones, respectively. For open access femtocells, macro users in $\mathcal{U}_{\mathrm{D}}$ will be offloaded to and served by the corresponding nearest fBSs. We use $\mathcal{U}_{\mathrm{D1}}$ to denote $\mathcal{U}_{\mathrm{D}}$. For closed access femtocells, macro users in $\mathcal{U}_{\mathrm{D}}$ are served by the MBS. We use $\mathcal{U}_{\mathrm{D2}}$ to represent $\mathcal{U}_{\mathrm{D}}$. Specifically, based on $d_{i,1}$, the user association rule can be summarised as:

$$\begin{cases} \text{If } d_{i,1} \leq \mathrm{R_f}, \text{ then } i \in \mathcal{U}_{\mathrm{f}}; \\ \text{If } \mathrm{R_o} \geq d_{i,1} > \mathrm{R_f}, \text{ then } i \in \mathcal{U}_{\mathrm{D}} \ (\mathcal{U}_{\mathrm{D1}} \text{ or } \mathcal{U}_{\mathrm{D2}}); \\ \text{If } d_{i,1} > \mathrm{R_o}, \text{ then } i \in \mathcal{U}_{\mathrm{M}}. \end{cases} \tag{3.2}$$

Since macro users in $\mathcal{U}_{\mathrm{D}}$ are more sensitive to the inter-tier interference, according to the resource partitioning discussion in previous section, we let $\mathcal{U}_{\mathrm{D}}$ users access to the unshared sub-bands. The shared sub-bands are allocated to the MBS and fBSs to let them communicate to $\mathcal{U}_{\mathrm{M}}$ and $\mathcal{U}_{\mathrm{f}}$ users, respectively. The exclusion zone model is illustrated in Fig. 3.1.

Note that the association rule in (3.2) only requires the information of the distances between users and their nearest fBSs. In practical implementation, to determine whether a macro user is served on the shared or unshared sub-bands, we can first allow the fBSs to send pilot signals with power $\mathrm{P_f}$ to the nearby macro users. Then, if the average received power at a macro user from the nearest femtocell is below $\mathrm{P_f} g_{\mathrm{M}}^{\mathrm{f}}(\mathrm{R_o})$, the macro user will be served by the MBS on the shared sub-bands. Otherwise, the macro user will be served on the unshared sub-bands.

3.3 Cell Load and Success Probability Characterization

This section derives the average cell load and user success probability expressions, which are used in subsequent analysis to determine the optimal exclusion zone parameters.

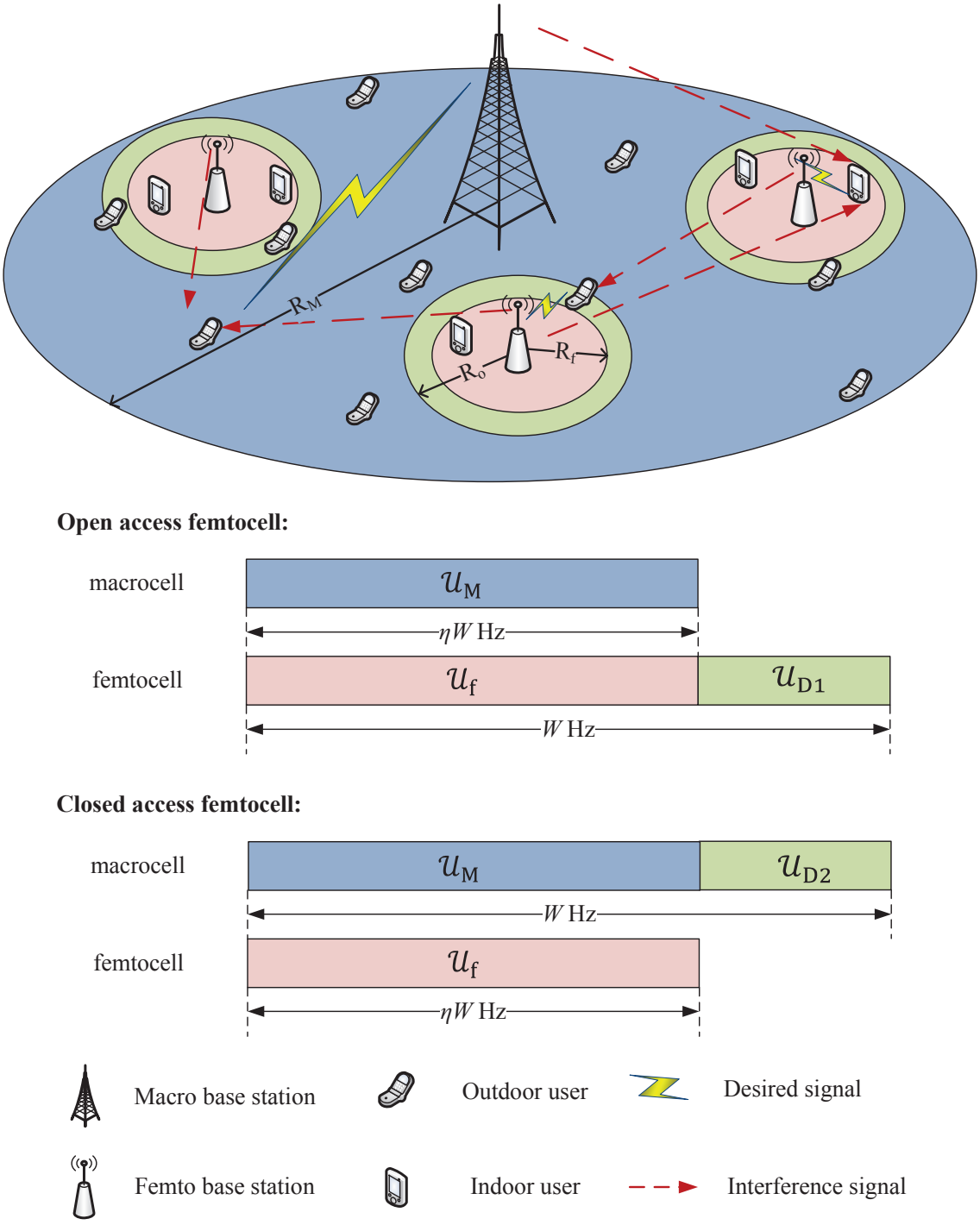


Figure 3.1: Resource partitioning and exclusion zone model.

3.3.1 User Association Probability and Cell Load

By assuming fBSs as PPP Φ_{f}, we can approximate the probability of users belonging to sets $\mathcal{U}_{\mathrm{M}}$, $\mathcal{U}_{\mathrm{f}}$ and $\mathcal{U}_{\mathrm{D}}$ (i.e. $\mathcal{U}_{\mathrm{D1}}$ or $\mathcal{U}_{\mathrm{D2}}$) based on the proposed cell association rule (3.2). The user association probabilities and expected cell loads for networks with open and closed access femtocells are given in the following lemma.

Lemma 3.3.1. *The probability that a typical user i belongs to $\mathcal{U}_k$ ($k \in \{\mathrm{M}, \mathrm{f}, \mathrm{D1}, \mathrm{D2}\}$), also known as its $\mathcal{U}_k$-association probability, is defined as $\mathbb{A}_{\mathrm{k}} = \Pr(i \in \mathcal{U}_k)$. It is approximated as*

$$\mathbb{A}_{\mathrm{M}} = e^{-\pi\lambda_{\mathrm{f}}\mathrm{R}_{\mathrm{o}}^2}, \quad \mathbb{A}_{\mathrm{f}} = 1 - e^{-\pi\lambda_{\mathrm{f}}\mathrm{R}_{\mathrm{f}}^2}, \quad \mathbb{A}_{\mathrm{D1}} = \mathbb{A}_{\mathrm{D2}} = e^{-\pi\lambda_{\mathrm{f}}\mathrm{R}_{\mathrm{f}}^2} - e^{-\pi\lambda_{\mathrm{f}}\mathrm{R}_{\mathrm{o}}^2}. \tag{3.3}$$

For a tagged base station, the number of users in set $\mathcal{U}_k$ ($k \in \{\mathrm{M}, \mathrm{f}, \mathrm{D1}, \mathrm{D2}\}$) within the base station's coverage is a random variable N_k. The expected values of cell load N_{M} and N_{f} are given as

$$\bar{\mathrm{N}}_{\mathrm{M}} = \lambda_{\mathrm{u}}^{\mathrm{M}}\pi\mathrm{R}_{\mathrm{M}}^2 e^{-\pi\lambda_{\mathrm{f}}\mathrm{R}_{\mathrm{o}}^2}, \quad \bar{\mathrm{N}}_{\mathrm{f}} = \frac{\lambda_{\mathrm{u}}^{\mathrm{f}}}{\lambda_{\mathrm{f}}}\left(1 - e^{-\pi\lambda_{\mathrm{f}}\mathrm{R}_{\mathrm{f}}^2}\right). \tag{3.4}$$

If all femtocells are under open access mode, N_{D1} is the number of $\mathcal{U}_{\mathrm{D1}}$ users served by each fBS. The expected value of N_{D1} is

$$\bar{\mathrm{N}}_{\mathrm{D1}} = \frac{\lambda_{\mathrm{u}}^{\mathrm{M}}}{\lambda_{\mathrm{f}}}\left(e^{-\pi\lambda_{\mathrm{f}}\mathrm{R}_{\mathrm{f}}^2} - e^{-\pi\lambda_{\mathrm{f}}\mathrm{R}_{\mathrm{o}}^2}\right). \tag{3.5}$$

If all femtocells are under closed access mode, N_{D2} is the number of $\mathcal{U}_{\mathrm{D2}}$ users served by the MBS. The expected value of N_{D2} is

$$\bar{\mathrm{N}}_{\mathrm{D2}} = \lambda_{\mathrm{u}}^{\mathrm{M}}\pi\mathrm{R}_{\mathrm{M}}^2\left(e^{-\pi\lambda_{\mathrm{f}}\mathrm{R}_{\mathrm{f}}^2} - e^{-\pi\lambda_{\mathrm{f}}\mathrm{R}_{\mathrm{o}}^2}\right). \tag{3.6}$$

Proof. For the association probability $\mathbb{A}_{\mathrm{M}}$, the event "user $i \in \mathcal{U}_{\mathrm{M}}$" is equivalent to "the distance between user i and its nearest fBS is bigger than R_{o}". The latter event is in turn equivalent to there being no fBSs within a distance of R_{o} of user

i, i.e. in an area of $\pi \mathrm{R}_{\mathrm{o}}^2$. Therefore, given that the fBSs form a 2-D PPP, we have $\Pr(i \in \mathcal{U}_{\mathrm{M}}) = e^{-\pi \lambda_{\mathrm{f}} \mathrm{R}_{\mathrm{o}}^2}$. The probabilities $\mathbb{A}_{\mathrm{f}}$, $\mathbb{A}_{\mathrm{D1}}$ and $\mathbb{A}_{\mathrm{D2}}$ can be derived similarly.

Given a macro user, the probability that the user associates with the MBS on the shared sub-bands is $\frac{\mathbb{A}_{\mathrm{M}}}{\mathbb{A}_{\mathrm{M}}+\mathbb{A}_{\mathrm{D1}}}$ (or equivalently, $\frac{\mathbb{A}_{\mathrm{M}}}{\mathbb{A}_{\mathrm{M}}+\mathbb{A}_{\mathrm{D2}}}$), and the number of macro users within the macrocell is a Poisson random variable with expected value $\pi \mathrm{R}_{\mathrm{M}}^2 \lambda_{\mathrm{u}}^{\mathrm{M}} (\mathbb{A}_{\mathrm{M}} + \mathbb{A}_{\mathrm{D1}})$. Thus, the average number of users in $\mathcal{U}_{\mathrm{M}}$ is $\bar{\mathrm{N}}_{\mathrm{M}} = \pi \mathrm{R}_{\mathrm{M}}^2 \lambda_{\mathrm{u}}^{\mathrm{M}} \mathbb{A}_{\mathrm{M}}$. To calculate the cell load of each fBS on the shared sub-bands, we first determine the expected value of the size of set $\mathcal{U}_{\mathrm{f}}$, which equals to $\lambda_{\mathrm{u}}^{\mathrm{f}} \pi \mathrm{R}_{\mathrm{M}}^2 \mathbb{A}_{\mathrm{f}}$. Next, the expected number of fBSs within the macrocell is $\pi \mathrm{R}_{\mathrm{M}}^2 \lambda_{\mathrm{f}}$. Noting that each fBS has almost the same load, we therefore have the expected femtocell load on the shared sub-bands approximated as $\bar{\mathrm{N}}_{\mathrm{f}} = \frac{\pi \mathrm{R}_{\mathrm{M}}^2 \lambda_{\mathrm{u}}^{\mathrm{M}} \mathbb{A}_{\mathrm{f}}}{\pi \mathrm{R}_{\mathrm{M}}^2 \lambda_{\mathrm{f}}}$. The derivations of $\bar{\mathrm{N}}_{\mathrm{D1}}$ in (3.5) and $\bar{\mathrm{N}}_{\mathrm{D2}}$ in (3.6) follow the derivations of $\bar{\mathrm{N}}_{\mathrm{f}}$ and $\bar{\mathrm{N}}_{\mathrm{M}}$, respectively. □

3.3.2 SINR and Success Probability

Given the prescribed target rates C_{M} for macro users and C_{f} for femto users, an outage event occurs for a certain user when the used channel cannot support the user's target rate. In the following analysis, we use the same C_{M} for all macro users, which means they all have the same rate requirement. Similarly, all femto users have the same target rate C_{f}. Without loss of generality, we assume $C_{\mathrm{M}} \leq C_{\mathrm{f}}$, because femtocells are typically devoted to providing high QoS for the subscribed users.

Open Access Femtocell

For the open access femtocell downlink analysis, users within the same cell occupy orthogonal sub-bands to eliminate intra-cell interference. For simplicity, we use γ_i^{M}, γ_i^{f}, and γ_i^{D1} to denote the SINRs for the typical user i in set $\mathcal{U}_{\mathrm{M}}$, $\mathcal{U}_{\mathrm{f}}$, and $\mathcal{U}_{\mathrm{D1}}$, respectively. Assuming that user i's nearest fBS is labeled as 1, the expressions

of γ_i^{M}, γ_i^{f} and γ_i^{D1} are given as follows.

$$\gamma_i^{\mathrm{M}} = \frac{\mathrm{P_M} H_{i,0} g_{\mathrm{M}}(d_{i,0})}{\sum\limits_{j\in\Phi_{\mathrm{f}}} \mathrm{P_f} H_{i,j} g_{\mathrm{M}}^{\mathrm{f}}(d_{i,j}) + \sigma^2}, \tag{3.7}$$

$$\gamma_i^{\mathrm{f}} = \frac{\mathrm{P_f} H_{i,1} g_{\mathrm{f}}(d_{i,1})}{\mathrm{P_M} H_{i,0} g_{\mathrm{f}}^{\mathrm{M}}(d_{i,0}) + \sum\limits_{j\in\Phi_{\mathrm{f}}\setminus\{1\}} \mathrm{P_f} H_{i,j} g_{\mathrm{f}}^{\mathrm{f}}(d_{i,j}) + \sigma^2} \tag{3.8}$$

$$\gamma_i^{\mathrm{D1}} = \frac{\mathrm{P_f} H_{i,1} g_{\mathrm{M}}^{\mathrm{f}}(d_{i,1})}{\sum\limits_{j\in\Phi_{\mathrm{f}}\setminus\{1\}} \mathrm{P_f} H_{i,j} g_{\mathrm{M}}^{\mathrm{f}}(d_{i,j}) + \sigma^2}, \tag{3.9}$$

where σ^2 is the variance of additive noise.

An outage event occurs if a user cannot achieve the predetermined target rate. Conditioned on the given cell load of user i's serving base station and the distances between the typical user i and its designated base station, the success probabilities for user i in $\mathcal{U}_{\mathrm{M}}$, $\mathcal{U}_{\mathrm{f}}$ and $\mathcal{U}_{\mathrm{D1}}$ are defined as

$$\mathbb{P}_{\mathrm{M}}(\mathrm{N_M}, d_{i,0}) \triangleq \Pr\left(\frac{\eta W}{\mathrm{N_M}} \log\left(1+\gamma_i^{\mathrm{M}}\right) \geq C_{\mathrm{M}} \middle| \, \mathrm{N_M}, d_{i,0} \right) \tag{3.10}$$

$$\mathbb{P}_{\mathrm{f}}(\mathrm{N_f}, d_{i,1}) \triangleq \Pr\left(\frac{\eta W}{\mathrm{N_f}} \log\left(1+\gamma_i^{\mathrm{f}}\right) \geq C_{\mathrm{f}} \middle| \, \mathrm{N_f}, d_{i,1} \right) \tag{3.11}$$

$$\mathbb{P}_{\mathrm{D1}}(\mathrm{N_{D1}}, d_{i,1}) \triangleq \Pr\left(\frac{(1-\eta) W}{\mathrm{N_{D1}}} \log\left(1+\gamma_i^{\mathrm{D1}}\right) \geq C_{\mathrm{M}} \middle| \, \mathrm{N_{D1}}, d_{i,1} \right). \tag{3.12}$$

Using the SINR expressions in (3.7)–(3.9), the above conditional success probabilities are expressed in the following lemma.

Lemma 3.3.2. *For network with open access fBSs, the expressions of conditional*

success probabilities $\mathbb{P}_{\mathrm{M}}(\mathrm{N_M}, d_{i,0})$, $\mathbb{P}_{\mathrm{f}}(\mathrm{N_f}, d_{i,1})$ *and* $\mathbb{P}_{\mathrm{D1}}(\mathrm{N_{D1}}, d_{i,1})$ *are*

$$\mathbb{P}_{\mathrm{M}}(\mathrm{N_M}, d_{i,0}) = \exp\left(-\pi\lambda_{\mathrm{f}}\mathrm{R_o^2}\frac{1}{\mathrm{Z_1}}\int_{\mathrm{Z_1}}^{\infty}\frac{1}{1+u^{\frac{\alpha_{\mathrm{o}}}{2}}}\mathrm{d}u - \frac{\Upsilon_{\mathrm{M}}\sigma^2}{\mathrm{P_M}Kd_{i,0}^{-\alpha_{\mathrm{o}}}}\right) \tag{3.13}$$

$$\mathbb{P}_{\mathrm{f}}(\mathrm{N_f}, d_{i,1}) = \mathbb{E}_{d_{i,0}}\left[\left(1+\frac{\Upsilon_{\mathrm{f}}\omega\mathrm{P_M}d_{i,0}^{-\alpha_{\mathrm{o}}}}{\mathrm{P_f}d_{i,1}^{-\alpha_{\mathrm{in}}}}\right)^{-1}\times\exp\left(-\frac{\Upsilon_{\mathrm{f}}\sigma^2}{\mathrm{P_f}Kd_{i,1}^{-\alpha_{\mathrm{in}}}}\right)\right.$$
$$\left.\times\exp\left(-\pi\lambda_{\mathrm{f}}\left(\omega^2\Upsilon_{\mathrm{f}}\right)^{\frac{2}{\alpha_{\mathrm{o}}}}d_{i,1}^{\frac{2\alpha_{\mathrm{in}}}{\alpha_{\mathrm{o}}}}\int_{\left(\omega^2\Upsilon_{\mathrm{f}}\right)^{\frac{1}{\alpha_{\mathrm{o}}}}d_{i,1}^{\frac{\alpha_{\mathrm{in}}}{\alpha_{\mathrm{o}}}-1}}^{\infty}\frac{1}{1+u^{\frac{\alpha_{\mathrm{o}}}{2}}}\mathrm{d}u\right)\right] \tag{3.14}$$

$$\mathbb{P}_{\mathrm{D1}}(\mathrm{N_{D1}}, d_{i,1}) = \exp\left(-\pi\lambda_{\mathrm{f}}\Upsilon_{\mathrm{D1}}^{\frac{2}{\alpha_{\mathrm{o}}}}d_{i,1}^2\int_{\Upsilon_{\mathrm{D1}}^{-\frac{2}{\alpha_{\mathrm{o}}}}}^{\infty}\frac{1}{1+u^{\frac{\alpha_{\mathrm{o}}}{2}}}\mathrm{d}u - \frac{\Upsilon_{\mathrm{D1}}\sigma^2}{\omega\mathrm{P_f}Kd_{i,1}^{-\alpha_{\mathrm{o}}}}\right) \tag{3.15}$$

where $\Upsilon_{\mathrm{M}} = 2^{\frac{C_{\mathrm{M}}\mathrm{N_M}}{\eta W}}-1$, $\Upsilon_{\mathrm{f}} = 2^{\frac{C_{\mathrm{f}}\mathrm{N_f}}{\eta W}}-1$ *and* $\Upsilon_{\mathrm{D1}} = 2^{\frac{C_{\mathrm{M}}\mathrm{N_{D1}}}{(1-\eta)W}}-1$. $\mathrm{Z_1} = \left[\frac{\mathrm{R_o}\mathrm{P_M}^{\frac{1}{\alpha_{\mathrm{o}}}}}{d_{i,0}(\Upsilon_{\mathrm{M}}\omega\mathrm{P_f})^{\frac{1}{\alpha_{\mathrm{o}}}}}\right]^2$.

When the Υ_{M}, Υ_{f} *and* Υ_{D1} *are small, which are typically true for users near the cell edge, and the values of* $\mathbb{P}_{\mathrm{M}}(\mathrm{N_M}, d_{i,0})$, $\mathbb{P}_{\mathrm{f}}(\mathrm{N_f}, d_{i,1})$ *and* $\mathbb{P}_{\mathrm{D1}}(\mathrm{N_{D1}}, d_{i,1})$ *are sufficiently large,* (3.13)–(3.15) *can be further approximated as*

$$\mathbb{P}_{\mathrm{M}}(\mathrm{N_M}, d_{i,0}) \approx \widetilde{\mathbb{P}}_{\mathrm{M}}(\mathrm{N_M}, d_{i,0}) = \exp\left(-\frac{2\pi\lambda_{\mathrm{f}}\mathrm{R_o^{2-\alpha_o}}\omega\mathrm{P_f}\Upsilon_{\mathrm{M}}}{(\alpha_{\mathrm{o}}-2)d_{i,0}^{-\alpha_{\mathrm{o}}}\mathrm{P_M}} - \frac{\Upsilon_{\mathrm{M}}\sigma^2}{\mathrm{P_M}Kd_{i,0}^{-\alpha_{\mathrm{o}}}}\right) \tag{3.16}$$

$$\mathbb{P}_{\mathrm{f}}(\mathrm{N_f}, d_{i,1}) \approx \widetilde{\mathbb{P}}_{\mathrm{f}}(\mathrm{N_f}, d_{i,1}) = \frac{2}{\alpha_{\mathrm{o}}}\left(\frac{\Upsilon_{\mathrm{f}}\omega\mathrm{P_M}\mathrm{R_M^{-\alpha_o}}}{\mathrm{P_f}d_{i,1}^{-\alpha_{\mathrm{in}}}}\right)^{\frac{2}{\alpha_{\mathrm{o}}}}\Gamma\left(-\frac{2}{\alpha_{\mathrm{o}}}, \frac{\Upsilon_{\mathrm{f}}\omega\mathrm{P_M}\mathrm{R_M^{-\alpha_o}}}{\mathrm{P_f}d_{i,1}^{-\alpha_{\mathrm{in}}}}\right)$$
$$\times\exp\left(-\frac{2}{\alpha_{\mathrm{o}}-2}\pi\lambda_{\mathrm{f}}\omega^2\Upsilon_{\mathrm{f}}d_{i,1}^{\alpha_{\mathrm{in}}-\alpha_{\mathrm{o}}+2} - \frac{\Upsilon_{\mathrm{f}}\sigma^2}{\mathrm{P_f}Kd_{i,1}^{-\alpha_{\mathrm{in}}}}\right) \tag{3.17}$$

$$\mathbb{P}_{\mathrm{D1}}(\mathrm{N_{D1}}, d_{i,1}) \approx \widetilde{\mathbb{P}}_{\mathrm{D1}}(\mathrm{N_{D1}}, d_{i,1}) = \exp\left(-\frac{2}{\alpha_{\mathrm{o}}-2}\pi\lambda_{\mathrm{f}}\Upsilon_{\mathrm{D1}}d_{i,1}^2 - \frac{\Upsilon_{\mathrm{D1}}\sigma^2}{\omega\mathrm{P_f}Kd_{i,1}^{-\alpha_{\mathrm{o}}}}\right), \tag{3.18}$$

where $\Gamma(\cdot,\cdot)$ *is the incomplete gamma function [98].*

Proof. Please refer to Appendix C. □

Closed Access Femtocell

For closed access femtocells, the users in set $\mathcal{U}_{\mathrm{D2}}$ are served by the MBS on the unshared sub-bands. Because of the different interference levels on the shared and unshared sub-bands, the users in set $\mathcal{U}_{\mathrm{M}}$ and $\mathcal{U}_{\mathrm{D2}}$ are correspondingly served by the MBS with transmit power P_{M1} and P_{M2}. Similar as the analysis in Section 3.3.2, the SINR expressions of γ_i^{M}, γ_i^{f} and γ_i^{D2} are given as

$$\gamma_i^{\mathrm{M}} = \frac{\mathrm{P}_{\mathrm{M1}} H_{i,0} g_{\mathrm{M}}(d_{i,0})}{\sum\limits_{j\in\Phi_{\mathrm{f}}} \mathrm{P}_{\mathrm{f}} H_{i,j} g_{\mathrm{M}}^{\mathrm{f}}(d_{i,j}) + \sigma^2}, \tag{3.19}$$

$$\gamma_i^{\mathrm{f}} = \frac{\mathrm{P}_{\mathrm{f}} H_{i,1} g_{\mathrm{f}}(d_{i,1})}{\mathrm{P}_{\mathrm{M1}} H_{i,0} g_{\mathrm{f}}^{\mathrm{M}}(d_{i,0}) + \sum\limits_{j\in\Phi_{\mathrm{f}}\setminus\{1\}} \mathrm{P}_{\mathrm{f}} H_{i,j} g_{\mathrm{f}}^{\mathrm{f}}(d_{i,j}) + \sigma^2} \tag{3.20}$$

$$\gamma_i^{\mathrm{D2}} = \frac{\mathrm{P}_{\mathrm{M2}} H_{i,0} g_{\mathrm{M}}(d_{i,0})}{\sigma^2}. \tag{3.21}$$

Conditioned on the given cell load of user i's serving base station and the distances between the typical user i and its designated base station, the success probabilities for user i in $\mathcal{U}_{\mathrm{M}}$, $\mathcal{U}_{\mathrm{f}}$ and $\mathcal{U}_{\mathrm{D2}}$ are

$$\mathbb{P}_{\mathrm{M}}(\mathrm{N}_{\mathrm{M}}, d_{i,0}) \triangleq \Pr\left(\frac{\eta W}{\mathrm{N}_{\mathrm{M}}} \log\left(1+\gamma_i^{\mathrm{M}}\right) \geq C_{\mathrm{M}} \middle| \, \mathrm{N}_{\mathrm{M}}, d_{i,0} \right) \tag{3.22}$$

$$\mathbb{P}_{\mathrm{f}}(\mathrm{N}_{\mathrm{f}}, d_{i,1}) \triangleq \Pr\left(\frac{\eta W}{\mathrm{N}_{\mathrm{f}}} \log\left(1+\gamma_i^{\mathrm{f}}\right) \geq C_{\mathrm{f}} \middle| \, \mathrm{N}_{\mathrm{f}}, d_{i,1} \right) \tag{3.23}$$

$$\mathbb{P}_{\mathrm{D2}}(\mathrm{N}_{\mathrm{D2}}, d_{i,0}) \triangleq \Pr\left(\frac{(1-\eta) W}{\mathrm{N}_{\mathrm{D2}}} \log\left(1+\gamma_i^{\mathrm{D2}}\right) \geq C_{\mathrm{M}} \middle| \, \mathrm{N}_{\mathrm{D2}}, d_{i,0} \right). \tag{3.24}$$

The expressions of $\mathbb{P}_{\mathrm{M}}(\mathrm{N}_{\mathrm{M}}, d_{i,0})$, $\mathbb{P}_{\mathrm{f}}(\mathrm{N}_{\mathrm{f}}, d_{i,1})$, and $\mathbb{P}_{\mathrm{D2}}(\mathrm{N}_{\mathrm{D2}}, d_{i,0})$ in (3.22)–(3.24) are derived in the following lemma.

Lemma 3.3.3. *For network with closed access fBSs, the expressions of conditional*

success probabilities $\mathbb{P}_{\mathrm{M}}(\mathrm{N}_{\mathrm{M}}, d_{i,0})$, $\mathbb{P}_{\mathrm{f}}(\mathrm{N}_{\mathrm{f}}, d_{i,1})$ *and* $\mathbb{P}_{\mathrm{D2}}(\mathrm{N}_{\mathrm{D2}}, d_{i,0})$ *are*

$$\mathbb{P}_{\mathrm{M}}(\mathrm{N}_{\mathrm{M}}, d_{i,0}) = \exp\left(-\pi\lambda_{\mathrm{f}}\mathrm{R}_{\mathrm{o}}^2\frac{1}{\mathrm{Z}_1}\int_{\mathrm{Z}_1}^{\infty}\frac{1}{1+u^{\frac{\alpha_{\mathrm{o}}}{2}}}\mathrm{d}u - \frac{\Upsilon_{\mathrm{M}}\sigma^2}{\mathrm{P}_{\mathrm{M1}}Kd_{i,0}^{-\alpha_{\mathrm{o}}}}\right) \tag{3.25}$$

$$\begin{aligned}\mathbb{P}_{\mathrm{f}}(\mathrm{N}_{\mathrm{f}}, d_{i,1}) = \mathbb{E}_{d_{i,0}}\Bigg[&\left(1+\frac{\Upsilon_{\mathrm{f}}\omega\mathrm{P}_{\mathrm{M1}}d_{i,0}^{-\alpha_{\mathrm{o}}}}{\mathrm{P}_{\mathrm{f}}d_{i,1}^{-\alpha_{\mathrm{in}}}}\right)^{-1}\times\exp\left(-\frac{\Upsilon_{\mathrm{f}}\sigma^2}{\mathrm{P}_{\mathrm{f}}Kd_{i,1}^{-\alpha_{\mathrm{in}}}}\right)\\ &\times\exp\left(-\pi\lambda_{\mathrm{f}}\left(\omega^2\Upsilon_{\mathrm{f}}\right)^{\frac{2}{\alpha_{\mathrm{o}}}}d_{i,1}^{\frac{2\alpha_{\mathrm{in}}}{\alpha_{\mathrm{o}}}}\int_{\left(\omega^2\Upsilon_{\mathrm{f}}\right)^{\frac{1}{\alpha_{\mathrm{o}}}}d_{i,1}^{\frac{\alpha_{\mathrm{in}}}{\alpha_{\mathrm{o}}}-1}}^{\infty}\frac{1}{1+u^{\frac{\alpha_{\mathrm{o}}}{2}}}\mathrm{d}u\right)\Bigg]\end{aligned} \tag{3.26}$$

$$\mathbb{P}_{\mathrm{D2}}(\mathrm{N}_{\mathrm{D2}}, d_{i,0}) = \exp\left(-\frac{\Upsilon_{\mathrm{D2}}\sigma^2}{\mathrm{P}_{\mathrm{M2}}Kd_{i,0}^{-\alpha_{\mathrm{o}}}}\right) \tag{3.27}$$

where $\Upsilon_{\mathrm{M}} = 2^{\frac{C_{\mathrm{M}}\mathrm{N}_{\mathrm{M}}}{\eta W}}-1$, $\Upsilon_{\mathrm{f}} = 2^{\frac{C_{\mathrm{f}}\mathrm{N}_{\mathrm{f}}}{\eta W}}-1$ *and* $\Upsilon_{\mathrm{D2}} = 2^{\frac{C_{\mathrm{M}}\mathrm{N}_{\mathrm{D2}}}{(1-\eta)W}}-1$. $\mathrm{Z}_1 = \left[\frac{\mathrm{R}_{\mathrm{o}}\mathrm{P}_{\mathrm{M1}}^{\frac{1}{\alpha_{\mathrm{o}}}}}{d_{i,0}(\Upsilon_{\mathrm{M}}\omega\mathrm{P}_{\mathrm{f}})^{\frac{1}{\alpha_{\mathrm{o}}}}}\right]^2$.

When the Υ_{M} *and* Υ_{f} *are small, and the values of* $\mathbb{P}_{\mathrm{M}}(\mathrm{N}_{\mathrm{M}}, d_{i,0})$ *and* $\mathbb{P}_{\mathrm{f}}(\mathrm{N}_{\mathrm{f}}, d_{i,1})$ *are sufficiently large,* (3.25) *and* (3.26) *can be further approximated as*

$$\mathbb{P}_{\mathrm{M}}(\mathrm{N}_{\mathrm{M}}, d_{i,0}) \approx \widetilde{\mathbb{P}}_{\mathrm{M}}(\mathrm{N}_{\mathrm{M}}, d_{i,0}) = \exp\left(-\frac{2\pi\lambda_{\mathrm{f}}\mathrm{R}_{\mathrm{o}}^{2-\alpha_{\mathrm{o}}}\omega\mathrm{P}_{\mathrm{f}}\Upsilon_{\mathrm{M}}}{(\alpha_{\mathrm{o}}-2)d_{i,0}^{-\alpha_{\mathrm{o}}}\mathrm{P}_{\mathrm{M1}}} - \frac{\Upsilon_{\mathrm{M}}\sigma^2}{\mathrm{P}_{\mathrm{M1}}Kd_{i,0}^{-\alpha_{\mathrm{o}}}}\right) \tag{3.28}$$

$$\begin{aligned}\mathbb{P}_{\mathrm{f}}(\mathrm{N}_{\mathrm{f}}, d_{i,1}) \approx \widetilde{\mathbb{P}}_{\mathrm{f}}(\mathrm{N}_{\mathrm{f}}, d_{i,1}) = &\frac{2}{\alpha}\left(\frac{\Upsilon_{\mathrm{f}}\omega\mathrm{P}_{\mathrm{M1}}\mathrm{R}_{\mathrm{M}}^{-\alpha_{\mathrm{o}}}}{\mathrm{P}_{\mathrm{f}}d_{i,1}^{-\alpha_{\mathrm{in}}}}\right)^{\frac{2}{\alpha_{\mathrm{o}}}}\Gamma\left(-\frac{2}{\alpha_{\mathrm{o}}}, \frac{\Upsilon_{\mathrm{f}}\omega\mathrm{P}_{\mathrm{M1}}\mathrm{R}_{\mathrm{M}}^{-\alpha_{\mathrm{o}}}}{\mathrm{P}_{\mathrm{f}}d_{i,1}^{-\alpha_{\mathrm{in}}}}\right)\\ &\times\exp\left(-\frac{2}{\alpha_{\mathrm{o}}-2}\pi\lambda_{\mathrm{f}}\omega^2\Upsilon_{\mathrm{f}}d_{i,1}^{\alpha_{\mathrm{in}}-\alpha_{\mathrm{o}}+2} - \frac{\Upsilon_{\mathrm{f}}\sigma^2}{\mathrm{P}_{\mathrm{f}}Kd_{i,1}^{-\alpha_{\mathrm{in}}}}\right)\end{aligned} \tag{3.29}$$

where $\Gamma(\cdot,\cdot)$ *is the incomplete gamma function [98].*

Due to its similarity to Lemma 3.3.2, the proof of Lemma 3.3.3 is omitted.

3.4 Exclusion Zone Design: MBS Transmit Power Minimization

As discussed in Section 3.2, the MBS transmit power accounts for a large portion of the HCN downlink power consumption, and considering MBS transmit

power adaptation according to the cell load could benefit network energy efficiency improvement. In this section, we will investigate the optimal exclusion zone design problem that controls cell load and hence minimizes the MBS transmit power for any given fBS transmit power $\mathrm{P_f}$. Specifically, we will elaborate on jointly determining the size of the exclusion zone and the allocated bandwidth. The analyses for open and closed access femtocells will be conducted separately.

3.4.1 Open Access Femtocells

When femtocells operate in open access, the MBS only transmits on the shared sub-bands. As a result, the MBS transmit power is $\eta \mathrm{P_M}$. Note that the MBS has the ability to adjust its transmit power according to its random cell load $\mathrm{N_M}$, which means the value of $\eta \mathrm{P_M}$ varies with $\mathrm{N_M}$. Thus, the averaged MBS transmit power by taking all possible cell load instances into account is of interest, which can be expressed as $\mathbb{E}_{\mathrm{N_M}}[\eta \mathrm{P_M}]$. To minimize $\mathbb{E}_{\mathrm{N_M}}[\eta \mathrm{P_M}]$, we resort to finding the optimal exclusion zone radius $\mathrm{R_o}$ and the resource partitioning factor η while letting users meet certain QoS requirements. User success probability (or equivalently the outage probability) is used as the QoS metric when minimizing the MBS's transmit power. Given the target rates C_{M} and C_{f}, the success probabilities of all users should be no smaller than a predetermined value θ. Using the closed-form approximations in Lemma 3.3.2, the success probability constraints can be characterized. Specifically, the above described MBS transmit power minimization problem is formulated as follows.

$$\underset{\mathrm{R_o},\eta}{\text{minimize}} \quad \mathbb{E}_{\mathrm{N_M}}[\eta \mathrm{P_M}] \tag{3.30a}$$

$$\text{subject to} \quad \min_{\forall i \in \mathcal{U}_{\mathrm{M}}} \left\{ \widetilde{\mathbb{P}}_{\mathrm{M}}(\mathrm{N_M}, d_{i,0}) \right\} \geq \theta, \quad 0 \leq \mathrm{N_M} \leq \mathrm{N_M^{max}} \tag{3.30b}$$

$$\min_{\forall i \in \mathcal{U}_{\mathrm{f}}} \left\{ \widetilde{\mathbb{P}}_{\mathrm{f}}(\mathrm{N_f}, d_{i,1}) \right\} \geq \theta, \quad 0 \leq \mathrm{N_f} \leq \mathrm{N_f^{max}} \tag{3.30c}$$

$$\min_{\forall i \in \mathcal{U}_{\mathrm{D1}}} \left\{ \widetilde{\mathbb{P}}_{\mathrm{D1}}(\mathrm{N_{D1}}, d_{i,1}) \right\} \geq \theta, \quad 0 \leq \mathrm{N_{D1}} \leq \mathrm{N_{D1}^{max}} \tag{3.30d}$$

$$0 \leq \eta \leq 1, \quad \mathrm{R_o} \geq \mathrm{R_f} \tag{3.30e}$$

where $\mathrm{N}_k^{\max}$ is the maximum number of users in set $\mathcal{U}_k$ that can be simultaneously served by a base station. We assume that the value of $\mathrm{N}_k^{\max}$ is much larger than the corresponding average cell load $\bar{\mathrm{N}}_k$ so that there is only negligible probability that all users within the cell cannot be simultaneously served.

Constraints (3.30b)–(3.30d) show that all served users have their success probabilities exceed the threshold θ. For this worst case analysis, we introduce the concept of *cell edge user* for user set $\mathcal{U}_k$, which is defined as a suppositional user node located where success probability would be the lowest for users within $\mathcal{U}_k$. Then, the constraints in (3.30b)–(3.30d) can be re-expressed by cell edge user success probabilities. Using (3.16)–(3.18), the cell edge user success probabilities can be determined as:

$$\min_{\forall i \in \mathcal{U}_{\mathrm{M}}} \left\{ \widetilde{\mathbb{P}}_{\mathrm{M}}(\mathrm{N}_{\mathrm{M}}, d_{i,0}) \right\} \geq \theta \Leftrightarrow \widetilde{\mathbb{P}}_{\mathrm{M}}(\mathrm{N}_{\mathrm{M}}, \mathrm{R}_{\mathrm{M}}) \geq \theta \tag{3.31}$$

$$\min_{\forall i \in \mathcal{U}_{\mathrm{f}}} \left\{ \widetilde{\mathbb{P}}_{\mathrm{f}}(\mathrm{N}_{\mathrm{f}}, d_{i,1}) \right\} \geq \theta \Leftrightarrow \widetilde{\mathbb{P}}_{\mathrm{f}}(\mathrm{N}_{\mathrm{f}}, \mathrm{R}_{\mathrm{f}}) \geq \theta \tag{3.32}$$

$$\min_{\forall i \in \mathcal{U}_{\mathrm{D1}}} \left\{ \widetilde{\mathbb{P}}_{\mathrm{D1}}(\mathrm{N}_{\mathrm{D1}}, d_{i,1}) \right\} \geq \theta \Leftrightarrow \widetilde{\mathbb{P}}_{\mathrm{D1}}(\mathrm{N}_{\mathrm{D1}}, \mathrm{R}_{\mathrm{o}}) \geq \theta \tag{3.33}$$

From (3.16), we can show that $\widetilde{\mathbb{P}}_{\mathrm{M}}(\mathrm{N}_{\mathrm{M}}, \mathrm{R}_{\mathrm{M}})$ is an increasing function of P_{M}. Therefore, for given N_{M}, R_{o} and η, a lower bound of P_{M}, denoted as $L(\mathrm{N}_{\mathrm{M}}, \mathrm{R}_{\mathrm{o}}, \eta)$, is derived from the right hand side (RHS) of (3.31) using the $\widetilde{\mathbb{P}}_{\mathrm{M}}(\mathrm{N}_{\mathrm{M}}, \mathrm{R}_{\mathrm{M}})$ expression in (3.16):

$$\mathrm{P}_{\mathrm{M}} \geq L(\mathrm{N}_{\mathrm{M}}, \mathrm{R}_{\mathrm{o}}, \eta) = \left(2^{\frac{C_{\mathrm{M}}\mathrm{N}_{\mathrm{M}}}{\eta W}} - 1\right) \left(\frac{2\pi\lambda_{\mathrm{f}}\mathrm{R}_{\mathrm{o}}^{2-\alpha_{\mathrm{o}}}\omega\mathrm{P}_{\mathrm{f}}}{(\alpha_{\mathrm{o}}-2)\mathrm{R}_{\mathrm{M}}^{-\alpha_{\mathrm{o}}} \ln\frac{1}{\theta}} + \frac{\sigma^2}{K\mathrm{R}_{\mathrm{M}}^{-\alpha_{\mathrm{o}}} \ln\frac{1}{\theta}} \right). \tag{3.34}$$

From (3.17), we can obtain that $\widetilde{\mathbb{P}}_{\mathrm{f}}(\mathrm{N}_{\mathrm{f}}, \mathrm{R}_{\mathrm{f}})$ is a decreasing function of P_{M}. As a result, the RHS of (3.32) gives an upper bound of P_{M}, which is denoted as $U(\mathrm{N}_{\mathrm{f}}, \eta)$. However, the exact expression of $U(\mathrm{N}_{\mathrm{f}}, \eta)$ is hard to get. Additionally, constraint (3.30d) is independent of P_{M}.

When Problem (3.30) is feasible, we have $L(\mathrm{N}_{\mathrm{M}}, \mathrm{R}_{\mathrm{o}}, \eta) \leq U(\mathrm{N}_{\mathrm{f}}, \eta)$ for all $\mathrm{N}_{\mathrm{M}} \in [0, \mathrm{N}_{\mathrm{M}}^{\max}]$ and $\mathrm{N}_{\mathrm{f}} \in [0, \mathrm{N}_{\mathrm{f}}^{\max}]$. The minimum MBS transmit power $\eta\mathrm{P}_{\mathrm{M}}$ for given

$\mathrm{R_o}$ and η therefore equals to its lower bound $\eta L(\mathrm{N_M}, \mathrm{R_o}, \eta)$. Thus, we can replace $\mathbb{E}_{\mathrm{N_M}}[\eta \mathrm{P_M}]$ in (3.30a) by a function $f(\mathrm{R_o}, \eta)$ defined as

$$f(\mathrm{R_o}, \eta) = \mathbb{E}_{\mathrm{N_M}}\left[\eta L(\mathrm{N_M}, \mathrm{R_o}, \eta)\right]. \tag{3.35}$$

The expression of $f(\mathrm{R_o}, \eta)$ can be approximately characterized using the following lemma.

Lemma 3.4.1. *The function $f(\mathrm{R_o}, \eta)$ in (3.35) is approximated as $\hat{f}(\mathrm{R_o}, \eta)$, which is given as*

$$\begin{aligned}\hat{f}(\mathrm{R_o}, \eta) =& \eta \left(\exp\left(\pi \lambda_{\mathrm{u}}^{\mathrm{M}} \mathrm{R_M^2} e^{-\pi \lambda_{\mathrm{f}} \mathrm{R_o^2}} \left(2^{\frac{C_{\mathrm{M}}}{\eta W}} - 1\right)\right) - 1\right) \\ &\times \left(\frac{2\pi \lambda_{\mathrm{f}} \mathrm{R_o^{2-\alpha_o}} \omega \mathrm{P_f}}{(\alpha_{\mathrm{o}} - 2) R_{\mathrm{M}}^{-\alpha_{\mathrm{o}}} \ln \frac{1}{\theta}} + \frac{\sigma^2}{K \mathrm{R_M^{-\alpha_o}} \ln \frac{1}{\theta}}\right)\end{aligned} \tag{3.36}$$

Proof. Please see Appendix D. □

With the objective function (3.30a) replaced by $f(\mathrm{R_o}, \eta)$ using its approximated expression $\hat{f}(\mathrm{R_o}, \eta)$, we will then reformulate constraints (3.30b)–(3.30d) to determine the feasible region of $(\mathrm{R_o}, \eta)$. Recall that $\mathrm{P_M} = L(\mathrm{N_M}, \mathrm{R_o}, \eta)$ and $L(\mathrm{N_M}, \mathrm{R_o}, \eta) \leq U(\mathrm{N_f}, \eta)$ are feasible, where $L(\mathrm{N_M}, \mathrm{R_o}, \eta)$ and $U(\mathrm{N_M}, \mathrm{R_o}, \eta)$ are derived from constraints (3.30b) and (3.30c), respectively. Therefore, constraint (3.30b) is always achieved with equality when $\mathrm{P_M} = L(\mathrm{N_M}, \mathrm{R_o}, \eta)$. Furthermore, it can be verified using (3.16) and (3.17) that $L(\mathrm{N_M}, \mathrm{R_o}, \eta)$ increases with $\mathrm{N_M}$ and $U(\mathrm{N_f}, \eta)$ decreases with $\mathrm{N_f}$. As a result, we can set $L(\mathrm{N_M^{max}}, \mathrm{R_o}, \eta) \leq U(\mathrm{N_f^{max}}, \eta)$ to make constraint (3.30c) feasible, which is equivalent to $\widetilde{\mathbb{P}}_{\mathrm{f}}(\mathrm{N_f^{max}}, \mathrm{R_f})\Big|_{\mathrm{P_M} = L(\mathrm{N_M^{max}}, \mathrm{R_o}, \eta)} \geq \theta$. For constraint (3.30d), the cell edge probability $\widetilde{\mathbb{P}}_{\mathrm{D1}}(\mathrm{N_{D1}}, \mathrm{R_o})$ decreases with increasing $\mathrm{N_{D1}}$. Thus, constraint (3.30d) can be

rewritten as $\widetilde{\mathbb{P}}_{\mathrm{D1}}(\mathrm{N}_{\mathrm{D1}}^{\max}, \mathrm{R_o}) \geq \theta$. We therefore reformulate Problem (3.30) as

$$\underset{\mathrm{R_o},\eta}{\text{minimize}} \quad \hat{f}(\mathrm{R_o}, \eta) \tag{3.37a}$$

$$\text{subject to} \quad \widetilde{\mathbb{P}}_{\mathrm{f}}(\mathrm{N}_{\mathrm{f}}^{\max}, \mathrm{R_f})\Big|_{\mathrm{P_M}=L(\mathrm{N}_{\mathrm{M}}^{\max},\mathrm{R_o},\eta)} \geq \theta \tag{3.37b}$$

$$\widetilde{\mathbb{P}}_{\mathrm{D1}}(\mathrm{N}_{\mathrm{D1}}^{\max}, \mathrm{R_o}) \geq \theta \tag{3.37c}$$

$$0 \leq \eta \leq 1, \quad \mathrm{R_o} \geq \mathrm{R_f} \tag{3.37d}$$

With the reformulated Problem (3.37), we are then able to find the optimal values of η^* and $\mathrm{R_o^*}$. For given $\mathrm{R_o}$, the optimal η is determined using the following corollary.

Corollary 3.4.1. *The value of $\hat{f}(\mathrm{R_o}, \eta)$ in (3.36) decreases with increasing η.*

Proof. Taking derivative of $\hat{f}(\mathrm{R_o}, \eta)$ with respect to η, we have

$$\begin{aligned}\frac{\partial \hat{f}(\mathrm{R_o}, \eta)}{\partial \eta} = &\Big[\Big(\exp\Big(\pi\lambda_{\mathrm{u}}^{\mathrm{M}}\mathrm{R_M^2}e^{-\pi\lambda_{\mathrm{f}}\mathrm{R_o^2}}\Big(2^{\frac{C_{\mathrm{M}}}{\eta W}}-1\Big)\Big)-1\Big)\\ &-\pi\lambda_{\mathrm{u}}^{\mathrm{M}}\mathrm{R_M^2}e^{-\pi\lambda_{\mathrm{f}}\mathrm{R_o^2}}\frac{C_{\mathrm{M}}\ln 2}{\eta W}2^{\frac{C_{\mathrm{M}}}{\eta W}}\exp\Big(\pi\lambda_{\mathrm{u}}^{\mathrm{M}}\mathrm{R_M^2}e^{-\pi\lambda_{\mathrm{f}}\mathrm{R_o^2}}\Big(2^{\frac{C_{\mathrm{M}}}{\eta W}}-1\Big)\Big)\Big]\\ &\times\left(\frac{2\pi\lambda_{\mathrm{f}}\mathrm{R_o^{2-\alpha_o}}\omega\mathrm{P_f}}{(\alpha_{\mathrm{o}}-2)\mathrm{R_M^{-\alpha_o}}\ln\frac{1}{\theta}}+\frac{\sigma^2}{K\mathrm{R_M^{-\alpha_o}}\ln\frac{1}{\theta}}\right)\end{aligned}$$

Additionally, it can be verified that $\frac{\partial^2 \hat{f}(\mathrm{R_o},\eta)}{\partial \eta^2} \geq 0$. Note that $\frac{\partial \hat{f}(\mathrm{R_o},\eta)}{\partial \eta}\Big|_{\eta=\infty} = 0$, we therefore have $\frac{\partial \hat{f}(\mathrm{R_o},\eta)}{\partial \eta} \leq 0$ for all η, which means that $\hat{f}(\mathrm{R_o}, \eta)$ is a decreasing function of η. □

By calculating the first order derivative with respect to η, it can be verified that the left hand side (LHS) of (3.37b) is an increasing function of η and the LHS of (3.37c) is a decreasing function of η. Therefore, an implicit lower bound of η, denoted as $\Delta_1(\mathrm{R_o})$, is obtained from (3.37b), and a η upper bound $\Delta_2(\mathrm{R_o})$ is determined using (3.37c). Using the expressions of $\widetilde{\mathbb{P}}_{\mathrm{D1}}(\mathrm{N}_{\mathrm{D1}}^{\max}, \mathrm{R_o})$ in (3.18), we have

the expression of $\Delta_2(\mathrm{R_o})$ as

$$\Delta_2(\mathrm{R_o}) = 1 - \frac{\mathrm{N_{D1}^{max}} C_{\mathrm{M}}}{W \log_2\left(1 + \frac{\ln\frac{1}{\theta}}{\frac{2}{\alpha_o - 2}\pi\lambda_{\mathrm{f}}\mathrm{R_o^2} + \frac{\sigma^2}{\omega \mathrm{P_f} K}\mathrm{R_o^{\alpha_o}}}\right)} \tag{3.38}$$

According to Corollary 3.4.1, we can replace η in Problem (3.37) by $\Delta_2(\mathrm{R_o})$ and then solve for the optimal $\mathrm{R_o^*}$ with the constraint $\Delta_1(\mathrm{R_o}) \leq \Delta_2(\mathrm{R_o})$. Problem (3.37) then becomes a single variable optimization problem.

$$\underset{\mathrm{R_o}}{\text{minimize}} \quad \hat{f}(\mathrm{R_o}, \Delta_2(\mathrm{R_o})) \tag{3.39a}$$

$$\text{subject to} \quad \Delta_1(\mathrm{R_o}) \leq \Delta_2(\mathrm{R_o}) \tag{3.39b}$$

$$0 \leq \Delta_2(\mathrm{R_o}) \leq 1, \quad \mathrm{R_o} \geq \mathrm{R_f} \tag{3.39c}$$

Unfortunately, determining $\mathrm{R_o^*}$ still remains a challenging task that involves calculating complicated derivatives of $\hat{f}(\mathrm{R_o}, \Delta_2(\mathrm{R_o}))$, $\Delta_1(\mathrm{R_o})$ and $\Delta_2(\mathrm{R_o})$. To propose a simplified method to determine $\mathrm{R_o^*}$, we can exploit the fact that cell edge users' spectrum efficiencies are much smaller than 1 in practical network settings [83]. When a low average spectral efficiency target is assumed for the cell edge users in set $\mathcal{U}_{\mathrm{M}}$, i.e. $\frac{C_{\mathrm{M}}\bar{\mathrm{N}}_{\mathrm{M}}}{\eta W}$, that satisfies $\frac{C_{\mathrm{M}}\bar{\mathrm{N}}_{\mathrm{M}}}{\eta W} = \frac{C_{\mathrm{M}}\bar{\mathrm{N}}_{\mathrm{M}}}{\Delta_2(\mathrm{R_o})W} \ll 1$, we have the following approximation.

Corollary 3.4.2. *When* $\frac{C_{\mathrm{M}}\bar{\mathrm{N}}_{\mathrm{M}}}{\Delta_2(\mathrm{R_o})W} \ll 1$, *the function* $\hat{f}(\mathrm{R_o}, \Delta_2(\mathrm{R_o}))$ *in* (3.39a) *can be further approximated as*

$$\widetilde{f}(\mathrm{R_o}) = \pi\lambda_{\mathrm{u}}^{\mathrm{M}}\mathrm{R_M^2} e^{-\pi\lambda_{\mathrm{f}}\mathrm{R_o^2}} \frac{C_{\mathrm{M}}}{W} \ln 2 \left(\frac{2\pi\lambda_{\mathrm{f}}\mathrm{R_o^{2-\alpha_o}}\omega\mathrm{P_f}}{(\alpha_o - 2)\mathrm{R_M^{-\alpha_o}} \ln\frac{1}{\theta}} + \frac{\sigma^2}{K\mathrm{R_M^{-\alpha_o}} \ln\frac{1}{\theta}}\right), \tag{3.40}$$

which is a decreasing function of $\mathrm{R_o}$.

Proof. From the assumption $\frac{C_{\mathrm{M}}\bar{\mathrm{N}}_{\mathrm{M}}}{\Delta_2(\mathrm{R_o})W} \ll 1$, we can then get that $\frac{C_{\mathrm{M}}}{\Delta_2(\mathrm{R_o})W} \ll 1$. By

applying the approximation $2^x - 1 \overset{x\to 0}{\approx} x \ln 2$, we have

$$\begin{aligned}\hat{f}(\mathrm{R_o}, \Delta_2(\mathrm{R_o})) &\overset{(a)}{\approx} \eta \left(2^{\left(\pi\lambda_{\mathrm{u}}^{\mathrm{M}}\mathrm{R}_{\mathrm{M}}^2 e^{-\pi\lambda_{\mathrm{f}}\mathrm{R}_{\mathrm{o}}^2}\frac{C_{\mathrm{M}}}{\eta W}\right)} - 1\right)\left(\frac{2\pi\lambda_{\mathrm{f}}\mathrm{R}_{\mathrm{o}}^{2-\alpha_{\mathrm{o}}}\omega\mathrm{P_f}}{(\alpha_{\mathrm{o}}-2)\mathrm{R}_{\mathrm{M}}^{-\alpha_{\mathrm{o}}}\ln\frac{1}{\theta}} + \frac{\sigma^2}{K\mathrm{R}_{\mathrm{M}}^{-\alpha_{\mathrm{o}}}\ln\frac{1}{\theta}}\right)\\ &\overset{(b)}{\approx} \pi\lambda_{\mathrm{u}}^{\mathrm{M}}\mathrm{R}_{\mathrm{M}}^2 e^{-\pi\lambda_{\mathrm{f}}\mathrm{R}_{\mathrm{o}}^2}\frac{C_{\mathrm{M}}}{W}\ln 2\left(\frac{2\pi\lambda_{\mathrm{f}}\mathrm{R}_{\mathrm{o}}^{2-\alpha_{\mathrm{o}}}\omega\mathrm{P_f}}{(\alpha_{\mathrm{o}}-2)\mathrm{R}_{\mathrm{M}}^{-\alpha_{\mathrm{o}}}\ln\frac{1}{\theta}} + \frac{\sigma^2}{K\mathrm{R}_{\mathrm{M}}^{-\alpha_{\mathrm{o}}}\ln\frac{1}{\theta}}\right),\end{aligned}$$

where (a) uses $2^{\frac{C_{\mathrm{M}}}{\Delta_2(\mathrm{R_o})W}} - 1 \approx \frac{C_{\mathrm{M}}}{\Delta_2(\mathrm{R_o})W}\ln 2$, and (b) applies $2^{\frac{C_{\mathrm{M}}\bar{\mathrm{N}}_{\mathrm{M}}}{\Delta_2(\mathrm{R_o})W}} - 1 \approx \frac{C_{\mathrm{M}}\bar{\mathrm{N}}_{\mathrm{M}}}{\Delta_2(\mathrm{R_o})W}\ln 2$. □

It can be observed from Corollary 3.4.2 that $\hat{f}(\mathrm{R_o}, \Delta_2(\mathrm{R_o}))$ decreases with increasing $\mathrm{R_o}$ when user spectral efficiency target is small. Thus, solving Problem (3.39) is equivalent to finding the maximum $\mathrm{R_o}$ that satisfies (3.39b) and (3.39c). Because $\Delta_2(\mathrm{R_o})$ is a decreasing function of $\mathrm{R_o}$, we can then obtain $\mathrm{R_o^*}$ by solving the equation $\Delta_1(\mathrm{R_o^*}) = \Delta_2(\mathrm{R_o^*})$, which can be done through a bisection search. The optimal η^* then equals to $\Delta_2(\mathrm{R_o^*})$.

3.4.2 Closed Access Femtocells

With closed access femtocells, users in $\mathcal{U}_{\mathrm{M}}$ are served by the MBS on the shared sub-bands with transmit power $\eta\mathrm{P_{M1}}$. Additionally, users in $\mathcal{U}_{\mathrm{D2}}$ are served by the MBS on the unshared sub-bands with transmit power $(1-\eta)\mathrm{P_{M2}}$. The values of $\eta\mathrm{P_{M1}}$ and $(1-\eta)\mathrm{P_{M2}}$ change according to cell loads $\mathrm{N_M}$ and $\mathrm{N_{D2}}$, respectively. As a result, the total MBS transmit power averaged over cell loads in $\mathcal{U}_{\mathrm{M}}$ and $\mathcal{U}_{\mathrm{D2}}$ is represented as $\mathbb{E}_{\mathrm{N_M},\mathrm{N_{D2}}}\left[\eta\mathrm{P_{M1}} + (1-\eta)\mathrm{P_{M2}}\right]$. In addition, the user success probability constraints are characterized using the approximations derived in Lemma 3.3.3. Similar as the open access femtocell analysis, the MBS transmit power minimization problem for

closed access femtocells is formulated as

$$\underset{\mathrm{R_o},\eta}{\text{minimize}} \quad \mathbb{E}_{\mathrm{N_M},\mathrm{N_{D2}}}\left[\eta \mathrm{P_{M1}} + (1-\eta)\mathrm{P_{M2}}\right] \tag{3.41a}$$

$$\text{subject to} \quad \min_{\forall i \in \mathcal{U}_\mathrm{M}} \left\{\widetilde{\mathbb{P}}_\mathrm{M}(\mathrm{N_M}, d_{i,0})\right\} \geq \theta, \quad 0 \leq \mathrm{N_M} \leq \mathrm{N_M^{max}} \tag{3.41b}$$

$$\min_{\forall i \in \mathcal{U}_\mathrm{f}} \left\{\widetilde{\mathbb{P}}_\mathrm{f}(\mathrm{N_f}, d_{i,1})\right\} \geq \theta, \quad 0 \leq \mathrm{N_f} \leq \mathrm{N_f^{max}} \tag{3.41c}$$

$$\min_{\forall i \in \mathcal{U}_\mathrm{D2}} \left\{\mathbb{P}_\mathrm{D2}(\mathrm{N_{D2}}, d_{i,0})\right\} \geq \theta, \quad 0 \leq \mathrm{N_{D2}} \leq \mathrm{N_{D2}^{max}} \tag{3.41d}$$

$$0 \leq \eta \leq 1, \quad \mathrm{R_o} \geq \mathrm{R_f} \tag{3.41e}$$

where $\widetilde{\mathbb{P}}_\mathrm{M}(\mathrm{N_M}, d_{i,0})$ is given in (3.28), $\widetilde{\mathbb{P}}_\mathrm{f}(\mathrm{N_f}, d_{i,1})$ is defined in (3.29), and $\mathbb{P}_\mathrm{D2}(\mathrm{N_{D2}}, d_{i,1})$ is expressed using (3.27).

The worst case success probability constraints in (3.41b)–(3.41d) can be reformulated by the cell edge user success probabilities.

$$\min_{\forall i \in \mathcal{U}_\mathrm{M}} \left\{\widetilde{\mathbb{P}}_\mathrm{M}(\mathrm{N_M}, d_{i,0})\right\} \geq \theta \Leftrightarrow \widetilde{\mathbb{P}}_\mathrm{M}(\mathrm{N_M}, \mathrm{R_M}) \geq \theta \tag{3.42}$$

$$\min_{\forall i \in \mathcal{U}_\mathrm{f}} \left\{\widetilde{\mathbb{P}}_\mathrm{f}(\mathrm{N_f}, d_{i,1})\right\} \geq \theta \Leftrightarrow \widetilde{\mathbb{P}}_\mathrm{f}(\mathrm{N_f}, \mathrm{R_f}) \geq \theta \tag{3.43}$$

$$\min_{\forall i \in \mathcal{U}_\mathrm{D2}} \left\{\mathbb{P}_\mathrm{D2}(\mathrm{N_{D2}}, d_{i,0})\right\} \geq \theta \Leftrightarrow \mathbb{P}_\mathrm{D2}(\mathrm{N_{D2}}, \mathrm{R_M}) \geq \theta. \tag{3.44}$$

Similar to the analysis for open access femtocells, we can derive a $\mathrm{P_{M1}}$ lower bound $L_1(\mathrm{N_M}, \mathrm{R_o}, \eta)$ from the RHS of (3.42), a $\mathrm{P_{M1}}$ upper bound $U_1(\mathrm{N_f}, \mathrm{R_o}, \eta)$ from the RHS of (3.43), and a $\mathrm{P_{M2}}$ lower bound $L_2(\mathrm{N_{D2}}, \eta)$ from the RHS of (3.44). The expressions of $L_1(\mathrm{N_M}, \mathrm{R_o}, \eta)$ and $L_2(\mathrm{N_{D2}}, \eta)$ are given as

$$L_1(\mathrm{N_M}, \mathrm{R_o}, \eta) = \left(2^{\frac{C_\mathrm{M}\mathrm{N_M}}{\eta W}} - 1\right)\left(\frac{2\pi\lambda_\mathrm{f}\mathrm{R_o}^{2-\alpha_\mathrm{o}}\omega\mathrm{P_f}}{(\alpha_\mathrm{o}-2)\mathrm{R_M}^{-\alpha_\mathrm{o}}\ln\frac{1}{\theta}} + \frac{\sigma^2}{K\mathrm{R_M}^{-\alpha_\mathrm{o}}\ln\frac{1}{\theta}}\right), \tag{3.45}$$

$$L_2(\mathrm{N_{D2}}, \eta) = \left(2^{\frac{C_\mathrm{M}\mathrm{N_{D2}}}{(1-\eta)W}} - 1\right)\frac{\sigma^2}{K\mathrm{R_M}^{-\alpha_\mathrm{o}}\ln\frac{1}{\theta}}. \tag{3.46}$$

The $\mathrm{P_{M1}}$ upper bound $U_1(\mathrm{N_f}, \mathrm{R_o}, \eta)$ has implicit expression.

When Problem (3.41) is feasible, we have $L_1(\mathrm{N_M}, \mathrm{R_o}, \eta) \leq U_1(\mathrm{N_f}, \mathrm{R_o}, \eta)$ for $\mathrm{N_M} \in [0, \mathrm{N_M^{max}}]$ and $\mathrm{N_f} \in [0, \mathrm{N_f^{max}}]$. The MBS transmit power is lower bounded by

$\eta L_1(\mathrm{N_M}, \mathrm{R_o}, \eta)+(1-\eta)L_2(\mathrm{N_{D2}}, \eta)$. We can then replace $\mathbb{E}_{\mathrm{N_M},\mathrm{N_{D2}}}\left[\eta \mathrm{P_{M1}} + (1-\eta)\mathrm{P_{M2}}\right]$ in (3.41a) by

$$g(\mathrm{R_o}, \eta) = \mathbb{E}_{\mathrm{N_M},\mathrm{N_{D2}}}\left[\eta L_1(\mathrm{N_M}, \mathrm{R_o}, \eta) + (1-\eta)L_2(\mathrm{N_{D2}}, \eta)\right], \tag{3.47}$$

with $g(\mathrm{R_o}, \eta)$ approximately expressed using the following lemma.

Lemma 3.4.2. *The function $g(\mathrm{R_o}, \eta)$ in (3.47) is approximated as $\hat{g}(\mathrm{R_o}, \eta)$, which is given as*

$$\hat{g}(\mathrm{R_o}, \eta) = \frac{(1-\eta)\sigma^2}{K\mathrm{R_M}^{-\alpha_o}\ln\frac{1}{\theta}}\left[\exp\left(\pi\lambda_u^{\mathrm{M}}\mathrm{R_M^2}\left(e^{-\pi\lambda_f\mathrm{R_f^2}} - e^{-\pi\lambda_f\mathrm{R_o^2}}\right)\left(2^{\frac{C_M}{(1-\eta)W}}-1\right)\right)-1\right] + \eta\left[\exp\left(\pi\lambda_u^{\mathrm{M}}\mathrm{R_M^2}e^{-\pi\lambda_f\mathrm{R_o^2}}\left(2^{\frac{C_M}{\eta W}}-1\right)\right)-1\right]\left(\frac{2\pi\lambda_f\mathrm{R_o}^{2-\alpha_o}\omega\mathrm{P_f}}{(\alpha_o-2)\mathrm{R_M}^{-\alpha_o}\ln\frac{1}{\theta}} + \frac{\sigma^2}{K\mathrm{R_M}^{-\alpha_o}\ln\frac{1}{\theta}}\right) \tag{3.48}$$

We omit the proof of Lemma 3.4.2 due to its similarity to Lemma 3.4.1.

Using the similar method for deriving Problem (3.37) in open access case, we can reformulate Problem (3.41) as

$$\underset{\mathrm{R_o},\eta}{\text{minimize}} \quad \hat{g}(\mathrm{R_o}, \eta) \tag{3.49a}$$

$$\text{subject to} \quad \widetilde{\mathbb{P}}_{\mathrm{f}}(\mathrm{N_f^{max}}, \mathrm{R_f})\Big|_{\mathrm{P_{M1}}=L_1(\mathrm{N_M^{max}},\mathrm{R_o},\eta)} \geq \theta \tag{3.49b}$$

$$0 \leq \eta \leq 1, \quad \mathrm{R_o} \geq \mathrm{R_f} \tag{3.49c}$$

When the cell edge users in $\mathcal{U}_{\mathrm{M}}$ and $\mathcal{U}_{\mathrm{D2}}$ have small average spectrum efficiency targets that satisfy $\frac{C_M\bar{\mathrm{N}}_{\mathrm{M}}}{\eta W} \ll 1$ and $\frac{C_M\bar{\mathrm{N}}_{\mathrm{D2}}}{(1-\eta)W} \ll 1$, we can further simplify $\hat{g}(\mathrm{R_o}, \eta)$ and then solve for the optimal η^* and $\mathrm{R_o^*}$.

Corollary 3.4.3. *When $\frac{C_M\bar{\mathrm{N}}_{\mathrm{M}}}{\eta W} \ll 1$ and $\frac{C_M\bar{\mathrm{N}}_{\mathrm{D2}}}{(1-\eta)W} \ll 1$, the function $\hat{g}(\mathrm{R_o}, \eta)$ in (3.48)*

can be further approximated as follows

$$\begin{aligned}
&\hat{g}(\mathrm{R_o},\eta)\\
\overset{(a)}{\approx}\ &\check{g}(\mathrm{R_o},\eta) = \frac{(1-\eta)\sigma^2}{K\mathrm{R_M}^{-\alpha_o}\ln\frac{1}{\theta}}\left(2^{\pi\lambda_u^M\mathrm{R_M}^2\left(e^{-\pi\lambda_f\mathrm{R_f}^2}-e^{-\pi\lambda_f\mathrm{R_o}^2}\right)\frac{C_M}{(1-\eta)W}}-1\right)\\
&\quad+\eta\left(2^{\left(\pi\lambda_u^M\mathrm{R_M}^2e^{-\pi\lambda_f\mathrm{R_o}^2}\frac{C_M}{\eta W}\right)}-1\right)\left(\frac{2\pi\lambda_f\mathrm{R_o}^{2-\alpha_o}\omega\mathrm{P_f}}{(\alpha_o-2)\mathrm{R_M}^{-\alpha_o}\ln\frac{1}{\theta}}+\frac{\sigma^2}{K\mathrm{R_M}^{-\alpha_o}\ln\frac{1}{\theta}}\right) \qquad (3.50)\\
\overset{(b)}{\approx}\ &\widetilde{g}(\mathrm{R_o}) = \pi\lambda_u^M\mathrm{R_M}^2\left(e^{-\pi\lambda_f\mathrm{R_f}^2}-e^{-\pi\lambda_f\mathrm{R_o}^2}\right)\frac{C_M}{W}\ln 2\frac{\sigma^2}{K\mathrm{R_M}^{-\alpha_o}\ln\frac{1}{\theta}}\\
&\quad+\pi\lambda_u^M\mathrm{R_M}^2e^{-\pi\lambda_f\mathrm{R_o}^2}\frac{C_M}{W}\ln 2\left(\frac{2\pi\lambda_f\mathrm{R_o}^{2-\alpha_o}\omega\mathrm{P_f}}{(\alpha_o-2)\mathrm{R_M}^{-\alpha_o}\ln\frac{1}{\theta}}+\frac{\sigma^2}{K\mathrm{R_M}^{-\alpha_o}\ln\frac{1}{\theta}}\right). \qquad (3.51)
\end{aligned}$$

Proof. The step (a) uses $2^{\frac{C_M}{(1-\eta)W}}-1\approx\frac{C_M}{(1-\eta)W}\ln 2$ and $2^{\frac{C_M}{\eta W}}-1\approx\frac{C_M}{\eta W}\ln 2$, and step (b) applies $2^{\frac{C_M\bar{N}_{D2}}{(1-\eta)W}}-1\approx\frac{C_M\bar{N}_{D2}}{(1-\eta)W}\ln 2$ and $2^{\frac{C_M\bar{N}_M}{\eta W}}-1\approx\frac{C_M\bar{N}_M}{\eta W}\ln 2$. □

By calculating the second order derivative with respect to η, we can show that $\check{g}(\mathrm{R_o},\eta)$ in (3.50) is a convex function of η. Additionally, as discussed in the open access femtocell case, constraint (3.49b) gives an implicit lower bound of η. Thus, for given $\mathrm{R_o}$, we can replace the objective function $\hat{g}(\mathrm{R_o},\eta)$ in Problem (3.49) by $\check{g}(\mathrm{R_o},\eta)$ in (3.50) and found the optimal η denoted as $\Delta^*(\mathrm{R_o})$.

With all η in Problem (3.49) replaced by $\Delta^*(\mathrm{R_o})$, we then use the $\hat{g}(\mathrm{R_o},\eta)$ approximation $\widetilde{g}(\mathrm{R_o})$ in (3.51) to determine the optimal $\mathrm{R_o^*}$. It can be obtained from (3.51) that the optimal $\mathrm{R_o^*}$ equals to the maximum feasible value $\mathrm{R_o^{max}}$, which makes constraint (3.49b) achieve equality. Similar to the open access case, one dimensional search algorithms can be used to determine the maximum achievable $\mathrm{R_o^{max}}$. With $\mathrm{R_o^*}=\mathrm{R_o^{max}}$ found, η^* equals to $\Delta^*(\mathrm{R_o^*})$.

3.5 Numerical Results

In this section, simulation and numerical results are used to validate the proposed exclusion zone model. The parameters for PL models defined in Section

3.2.1 are set as: $K = 10^{-3}$, $\alpha_{\rm o} = 4$, $\alpha_{\rm in} = 3$ and $\omega = -5$ dB. The radii of the MBS and fBSs are $\mathrm{R_M} = 500$ m and $\mathrm{R_f} = 20$ m, respectively. The fBS intensity is 100 BS/km^2. Additionally, the intensities of macro and femto users used in this section are $\lambda_{\rm u}^{\rm M} = 200$ users/km^2 and $\lambda_{\rm u}^{\rm f} = 2$ users/400m^2. The macro users target rate $C_{\rm M}$ is 200 kbps, and the femto user target rate $C_{\rm f} = 500$ kbps.

In Fig. 3.2, the success probabilities for cell edge users in $\mathcal{U}_k$ ($k \in \{\mathrm{M}, \mathrm{f}, \mathrm{D1}, \mathrm{D2}\}$) are simulated and then compared with the analytical results derived in Lemma 3.3.2 and Lemma 3.3.3. We set the resource partitioning factor $\eta = 0.7$. Moreover, the total transmit power of the MBS is fixed at $\mathrm{P_M} = 43$ dBm and all fBSs transmit with $\mathrm{P_f} = 13$ dBm. The simulated results in Fig. 3.2 are obtained using Monte Carlo methods. The simulation experiments are built on the MATLAB platform. In the simulation, we randomly deploy the network for 1000 times. For each network realization, users and fBSs are randomly deployed following Poisson point processes with corresponding intensities. Within each network deployment, we randomly select cell edge users for all base stations and generate random channel fading coefficients 100 times for each cell edge user. If a cell edge user belonging to set $\mathcal{U}_k$ ($k \in \{\mathrm{M}, \mathrm{f}, \mathrm{D1}, \mathrm{D2}\}$) has its rate exceeding the rate target, we mark it as a successful transmission event for user set $\mathcal{U}_k$. By averaging the number of successful transmissions in set $\mathcal{U}_k$ over all network and channel realizations, we get the corresponding simulated success probabilities[1]. The analytical curves in Fig. 3.2 are obtained by calculating cell edge user success probabilities using (3.16), (3.17), (3.18) and (3.27). As shown in Fig. 3.2, the analytical probability expressions derived in Lemma 3.3.2 and Lemma 3.3.3 match well with the simulated results.

The optimal design of resource partitioning factor η for open access femtocells is investigated in Fig. 3.3. We let the success probability threshold $\theta = 0.8$ and the maximum number of served users in set $\mathcal{U}_k$ be $\mathrm{N}_k^{\max} = 2\bar{\mathrm{N}}_k$. With $\mathrm{R_o} = 50, 70, 90$ m, we correspondingly determine the feasible ranges of η using (3.39b). Within the

[1]For each point in Fig. 3.2, one complete simulation experiment is obtained by averaging over 1000 random network deployments $\times$ 100 channel realizations. For all the points on the curve indicating $\mathcal{U}_{\rm M}$ user success probability in Fig. 3.2, we repeat simulation experiments 100 times and obtained that the 95% confidence interval is smaller than 0.0025.

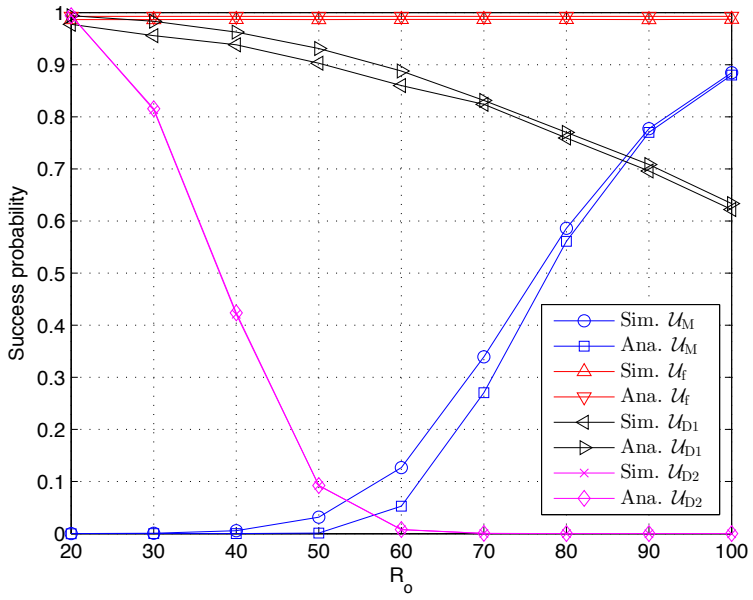


Figure 3.2: The cell edge user success probabilities for different $\mathrm{R_o}$ with $\mathrm{P_M} = 43$ dBm, $\mathrm{P_f} = 13$ dBm, and $\eta = 0.7$.

η feasible range for each $\mathrm{R_o}$, we plot the $\mathbb{E}_{\mathrm{N_M}}[\eta \mathrm{P_M}]$ lower bound $f(\mathrm{R_o}, \eta)$ in (3.35). The values of $f(\mathrm{R_o}, \eta)$ are numerically obtained by averaging over 1000 network realizations. Additionally, we also give the approximations of $f(\mathrm{R_o}, \eta)$ using the expression $\hat{f}(\mathrm{R_o}, \eta)$ in (3.36). From Fig. 3.3, we can observe that the gaps between $f(\mathrm{R_o}, \eta)$ and $\hat{f}(\mathrm{R_o}, \eta)$ are small for all three $\mathrm{R_o}$ settings. In addition, both the curves of $f(\mathrm{R_o}, \eta)$ and $\hat{f}(\mathrm{R_o}, \eta)$ decrease with increasing η, which validates Corollary 3.4.1 and shows that the optimal η that minimizes MBS transmit power is achieved at the maximum feasible point for the open access femtocell case.

To find the optimal exclusion zone radius $\mathrm{R_o}$ for open access femtocells, we plot the values of $f(\mathrm{R_o}, \eta^*)$ according to different $\mathrm{R_o}$ in Fig. 3.4. According to the discussion in Section 3.4.1 and the numerical results in Fig. 3.3, the optimal η^* for the open access femtocell case is achieved at its maximum achievable value $\Delta_2(\mathrm{R_o})$ (3.38). In Fig. 3.4, the numerically obtained $f(\mathrm{R_o}, \Delta_2(\mathrm{R_o}))$ curve is compared with the analytical results $\hat{f}(\mathrm{R_o}, \Delta_2(\mathrm{R_o}))$ and $\widetilde{f}(\mathrm{R_o})$ in Corollary 3.4.2. Fig. 3.4 shows that $\hat{f}(\mathrm{R_o}, \Delta_2(\mathrm{R_o}))$ provides accurate approximation of $f(\mathrm{R_o}, \Delta_2(\mathrm{R_o}))$, and the gap

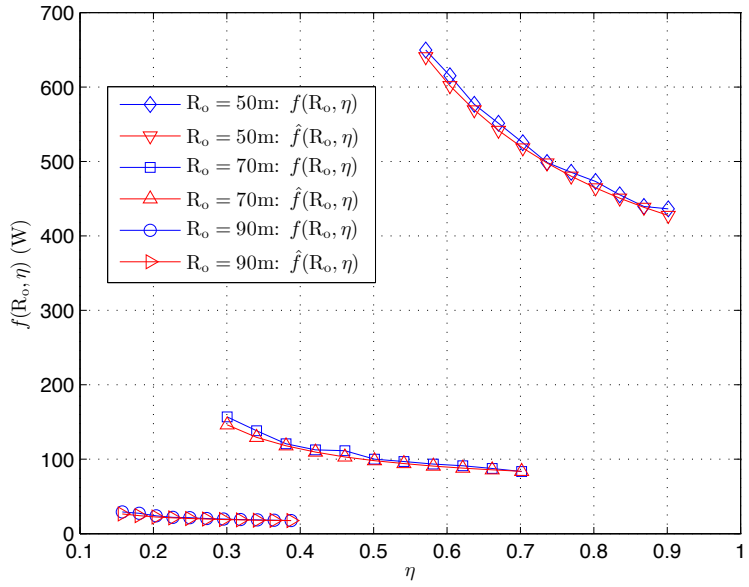


Figure 3.3: Effects of resource partitioning factor η on $f(\mathrm{R_o},\eta)$ and $\hat{f}(\mathrm{R_o},\eta)$.

between $f\left(\mathrm{R_o},\Delta_2(\mathrm{R_o})\right)$ and $\widetilde{f}\left(\mathrm{R_o}\right)$ decreases as $\mathrm{R_o}$ increases. Fig. 3.4 also validates our discussion that the optimal $\mathrm{R_o}$ for the open access femtocell case is obtained at its maximum achievable value.

Similar as the open access femtocell case, we investigate the optimal design of resource partitioning factor η and exclusion zone radius $\mathrm{R_o}$ for closed access femtocells in Fig. 3.5 and Fig. 3.6. The parameter settings are the same as those for the open access case. Fig. 3.5 shows that $\hat{g}(\mathrm{R_o},\eta)$ in (3.48) and $\check{g}(\mathrm{R_o},\eta)$ in (3.50) provide good approximations for determining the optimal η that minimizes $g(\mathrm{R_o},\eta)$ in (3.49a). Furthermore, it can be observed from Fig. 3.5 that $g(\mathrm{R_o},\eta)$, $\hat{g}(\mathrm{R_o},\eta)$, and $\check{g}(\mathrm{R_o},\eta)$ are convex functions of η with fixed $\mathrm{R_o}$.

In Fig. 3.6, we plot the curves of $g(\mathrm{R_o},\eta^*)$, $\hat{g}(\mathrm{R_o},\eta^*)$, $\check{g}(\mathrm{R_o},\eta^*)$ and $\widetilde{g}(\mathrm{R_o})$. With given $\mathrm{R_o}$, the optimal η^* that minimizes $g(\mathrm{R_o},\eta)$ is found through exhaustive search. The optimal values of η^* for $\hat{g}(\mathrm{R_o},\eta)$ and $\check{g}(\mathrm{R_o},\eta)$ are obtained by numerically calculating the corresponding first order derivatives. Although there exist gaps between $g(\mathrm{R_o},\eta^*)$ and its approximations using $\hat{g}(\mathrm{R_o},\eta^*)$, $\check{g}(\mathrm{R_o},\eta^*)$ and $\widetilde{g}(\mathrm{R_o})$, all the curves in Fig. 3.6 drop with increasing $\mathrm{R_o}$, which verified our discussion that

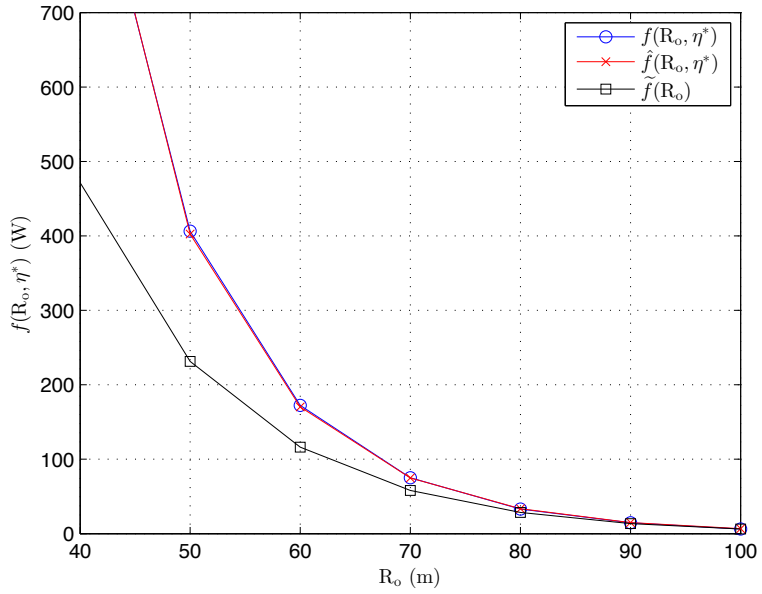


Figure 3.4: Effects of exclusion zone radius $\mathrm{R_o}$ on $f(\mathrm{R_o}, \eta^*)$, $\hat{f}(\mathrm{R_o}, \eta^*)$ and $\widetilde{f}(\mathrm{R_o})$.

the optimal $\mathrm{R_o}$ for closed access femtocells equals to its maximum feasible value.

Based on the discussion in Section 3.4, we can determine the optimal $\mathrm{R_o^*}$ using a bisection search algorithm and then obtain the optimal η^*. In Fig. 3.7, we plot the variations of the minimum average MBS transmit power with $\mathrm{R_o^*}$ and η^* according to fBS intensity for open and closed access femtocells, respectively. For comparison, we also give the average MBS transmit power for the no exclusion zone scenarios where the MBS and fBSs transmit on orthogonal sub-bands (denoted as "orthogonal" in the figure) or both transmit on the entire band (denoted as "fully shared" in the figure). The femtocell intensity λ_f varies from 50 fBSs/km^2 to 150 fBSs/km^2. The other parameters used to generate Fig. 3.7 are the same as before. From Fig. 3.7, the MBS transmit power for the fully shared case increases with the femtocell intensity. This is due to the growing inter-tier interference caused by the increasing number of fBSs transmitting on the MBS's band. On the other hand, with orthogonal resource allocation or exclusion zone deployment, the effects of inter-tier interference can be mitigated. Additionally, as shown in Fig. 3.7, using the optimally designed exclusion zones in open access femtocell networks give the minimum MBS transmit

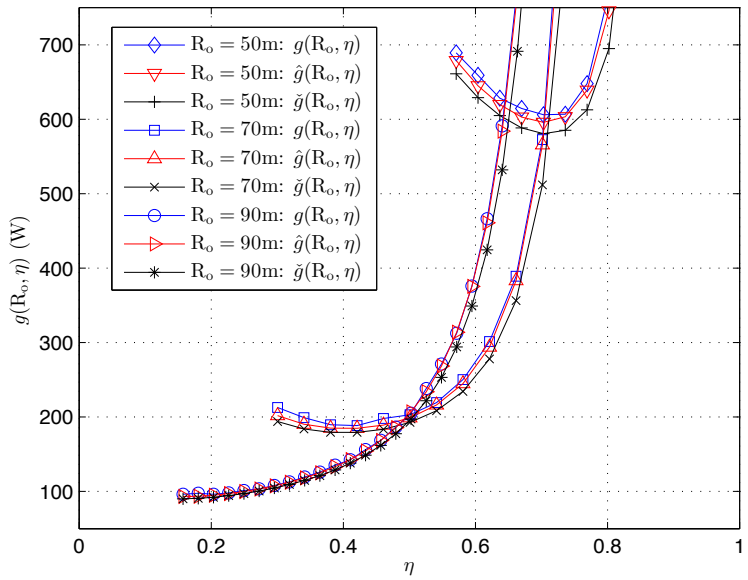


Figure 3.5: Effects of resource partitioning factor η on $g(\mathrm{R_o}, \eta)$, $\hat{g}(\mathrm{R_o}, \eta)$ and $\check{g}(\mathrm{R_o}, \eta)$.

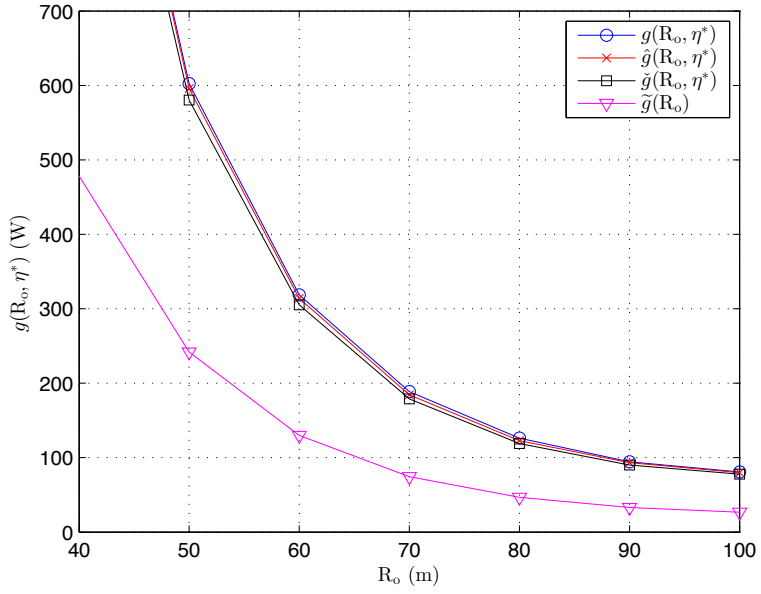


Figure 3.6: Effect of exclusion zone radius $\mathrm{R_o}$ on $g(\mathrm{R_o}, \eta^*)$, $\hat{g}(\mathrm{R_o}, \eta^*)$, $\check{g}(\mathrm{R_o}, \eta^*)$ and $\widetilde{g}(\mathrm{R_o})$.

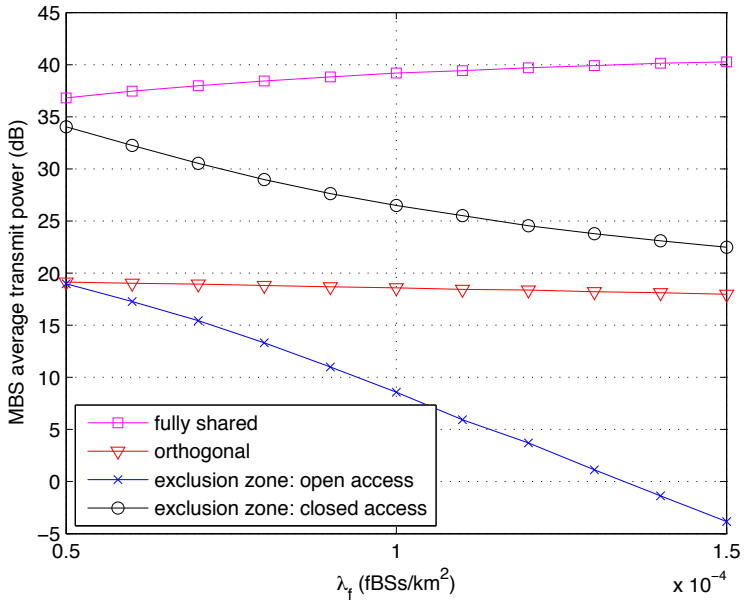


Figure 3.7: The effects of femtocell intensity λ_f on the minimum MBS average transmit power for different fBS access modes and resource allocation schemes.

power, which is due to the reduced macrocell load by offloading macro users to femtocells. For the closed access femtocell with exclusion zones, as the number of fBSs grows, more macro users are covered by the exclusion zones and served by the MBS on the unshared sub-bands. Therefore, the MBS transmit power for the closed access case converges to the one for orthogonal case as λ_f grows.

In Fig. 3.8, the effects of fBS transmit power P_f on the MBS transmit power are depicted. The femtocell intensity λ_f is fixed at 100 fBSs/km^2 and P_f changes from 5 dBm to 15 dBm. Similar to Fig. 3.7, the optimal designed exclusion zones for open and closed access femtocells are considered and compared with the fully shared and orthogonal resource allocation scenarios. From Fig. 3.8, for the fully shared case, the MBS transmit power grows as the fBS transmit power P_f increases. However, the other three curves in Fig. 3.8 show that the MBS transmit power is insensitive to the variation of P_f with efficient interference control using exclusion zones or orthogonal sub-band allocation for the MBS and fBSs.

From Fig. 3.7 and Fig. 3.8, it can be in turn concluded that with the same

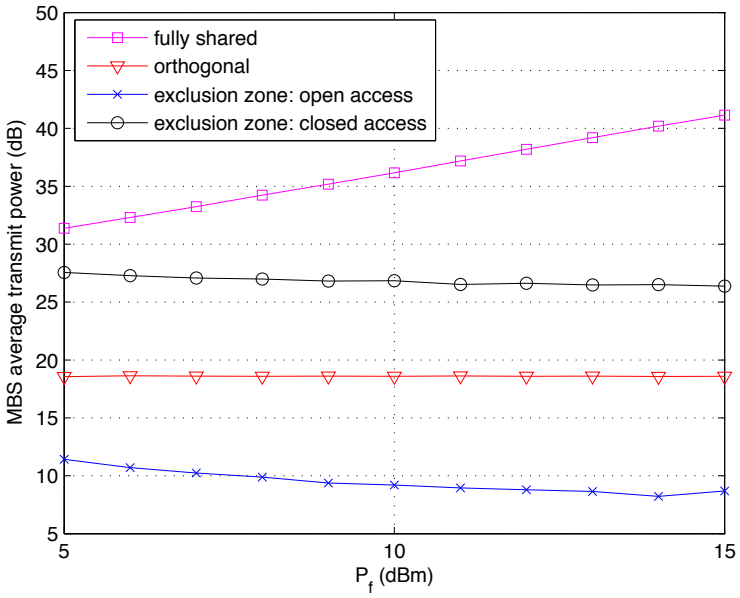


Figure 3.8: The effects of fBS transmit power P_f on the minimum MBS average transmit power for different fBS access modes and resource allocation schemes.

MBS transmit power the proposed open access femtocell exclusion zone model results in the highest cell edge user success probability. This conclusion aligns with existing discoveries in [42, 70] that the joint use of cell load adaptation and resource partitioning improves cell edge user coverage.

3.6 Conclusion

In this chapter, we investigated exclusion zone design coupled with resource partitioning in femtocell networks. Based on the proposed exclusion zone model, tractable expressions of average cell load and success probability were derived. Subsequently, we provided the methods to determine the exclusion zone sizes and allocated bandwidths that minimize MBS transmit power for open and closed access femtocells, respectively. Simulation and numerical results verified that the proposed exclusion zone design strategies reduce MBS power consumption while guaranteeing a certain cell edge user success probability.

Chapter 4

Energy-Efficient Multi-Antenna Two-Tier Base Station Deployment

4.1 Introduction

In previous chapters, we have studied two network power consumption minimization problems with QoS (throughput and outage probability, respectively) constraints, which are correspondingly solved through muting base stations on the partitioned resources and adjusting base station transmit power according to the cell load. Additionally, interference control in previous discussions is achieved by allocating orthogonal transmission resources to interfering transmitters. In this chapter, we investigate HCN energy efficiency improvement from a new perspective by smartly deploying multi-antenna base stations. Specifically, we propose an energy efficient MBS deployment strategy in a network consisting of small-cell BSs to reduce the number of active base stations and hence lower the network power consumption. Moreover, due to the efficient QoS gains by equipping base stations with multiple antennas [99], we will incorporate multi-antenna beamforming into the analysis to eliminate interference in the network.

The effects of base station deployment on heterogeneous network energy efficiency were investigated in [79], where the MBSs and small-cell BSs were regularly deployed on a grid. However, besides the unrealistic regular grid network abstraction, the discoveries in [79] were all based on simulation and were unlikely to be usable for optimizing the key parameters that maximized network performance gain. Using stochastic geometry analysis, tractable approaches were proposed

in [59, 60] to determine the node densities of different kinds of base stations that minimized network power consumption with outage probability constraints. In [59, 60], multiple tiers of base stations were modelled as independent PPPs that ignored geometric dependence where two base stations could be located arbitrarily close to each other. Furthermore, the energy efficient base station deployment strategies investigated in [59, 60, 79] did not take multi-antenna beamforming into consideration.

The effects of multi-antenna transmission and geometric dependence on network performance have been investigated through stochastic geometry analysis in the literature. For homogeneous networks with multi-antenna base stations, the energy efficiency analysis was conducted in [100]. By modelling base station locations as an HPPP, tractable outage probability expressions were derived in [100]. The benefits of coordinated beamforming in improving homogeneous network coverage and throughput were investigated in [40, 101]. Stochastic geometry frameworks for analysing HCNs consisting of multi-antenna base stations have also been proposed. In [94], the expressions of user outage probability were derived in a two-tier femtocell network with zero forcing beamforming at each base station. With a single macrocell overlaid with multiple femtocells, the effects of inter-tier dependence on user outage performance were investigated in [94], which showed that the MBS and fBSs should not be deployed too close to each other otherwise they might cause unacceptable inter-tier interference. In [102], with multiple multi-antenna MBSs and small-cell BSs, tractable outage probabilities considering inter-tier and intra-tier dependence were correspondingly derived. Nevertheless, the optimal design of the key parameters that improve network energy efficiency was not discussed in [94,102]. Furthermore, the use of coordinated beamforming in HCNs was not considered. As a result, it still remains an open problem to jointly study the effects of geometrically dependent base station deployment and coordinated beamforming on HCN energy efficiency.

To tackle the above mentioned problems, in this chapter, energy efficient

MBS deployment in a two-tier HCN with multi-antenna base stations is analysed. We propose a tractable stochastic geometry framework to model and analyse the network by jointly taking geometric dependence and coordinated beamforming into consideration. For geometric dependence, the deployed macrocells have non-overlapping coverage areas and small-cells within the coverage of macro base stations are muted to reduce power consumption. In addition, small-cell BSs adopt coordinated beamforming to reduce interference within the network. Based on the proposed framework, we derive closed-form expressions of user average spectral efficiencies. Downlink power minimization problem with user rate constraints is formulated and solved to determine the optimal cell size and density of the deployed MBSs. The solution provides guidelines for energy efficient base station deployment in multi-antenna HCNs.

The organization of this chapter is as follows. The system model is described in Section 4.2. Section 4.3 gives the method to approximate user spectral efficiency. Based on the derived spectral efficiency approximations, Section 4.4 explains how to determine the density and coverage radius for the MBSs to minimize network power consumption. Numerical results are given in Section 4.5. At last, Section 4.6 concludes the chapter.

4.2 System Model

We first consider a small-cell network consisting of only micro base stations (mBSs). The positions of mBSs follow an HPPP Φ_{m} of intensity λ_{m}. Each mBS is equipped with $\mathrm{N_m}$ transmit antennas and has total transmit power $\mathrm{P_m}$. We assume users are each equipped with a single receive antenna, and each mBS serves one user in each time slot using time division multiple access (TDMA)[1]. Furthermore, to ensure high received SNR at user nodes, we assume that every mBS serves all users within distance $\mathrm{R_m}$. As a result, the probability $p_{\bar{\mathrm{c}}}$ of a generic user not covered

[1]The derivations in this chapter can be extended to other orthogonal multiple acess methods.

by any mBSs can be calculated according to the HPPP assumption:

$$p_{\bar{c}} = \exp\left(-\pi\lambda_{\mathrm{m}}\mathrm{R}_{\mathrm{m}}^2\right). \tag{4.1}$$

To ensure that users are covered by mBSs with high probability, we can set $p_{\bar{c}} \leq \epsilon$ with ϵ being the maximum allowable probability of a user being located in a coverage hole. Therefore, with only mBS deployment, the required mBS intensity to achieve high coverage probability is

$$\lambda_{\mathrm{m}} \geq \frac{\ln\frac{1}{\epsilon}}{\pi\mathrm{R}_{\mathrm{m}}^2}. \tag{4.2}$$

4.2.1 Macro Base Station Deployment

To reduce the number of active base stations, we deploy MBSs within the small-cell network to replace some of the mBSs. Specifically, all the mBSs within the coverage areas of the MBSs are muted and have their users served by the corresponding nearest MBSs. To ensure that the maximum number of mBSs can be muted within each MBS's coverage, we propose that the MBSs are deployed in a "sparse" manner, where each deployed MBS has a circular coverage area with radius R_{M} that does not overlap with the coverage of other MBSs. Typically, we assume that $\mathrm{R}_{\mathrm{M}} \geq \mathrm{R}_{\mathrm{m}}$. The "sparse" deployment of MBSs can be modelled by a Matérn hardcore process (Matérn HCP), where the minimum distance between two MBSs is $2\mathrm{R}_{\mathrm{M}}$ to model the sparsity condition that their coverage areas do not overlap. The transmit power of an MBS is P_{M}, and the number of transmit antennas at each MBS is denoted as N_{M}. We assume that N_{M} equals to the average number of mBS antennas covered by each MBS, i.e. $\mathrm{N}_{\mathrm{M}} = \lfloor\lambda_{\mathrm{m}}\pi\mathrm{R}_{\mathrm{M}}^2\mathrm{N}_{\mathrm{m}}\rfloor$ with $\lfloor x\rfloor$ denoting the maximum integer smaller than or equal to x. This assumption ensures that the average number of active antennas within the network is invariant to the deployment of MBSs. The network structure is shown in Fig. 4.1.

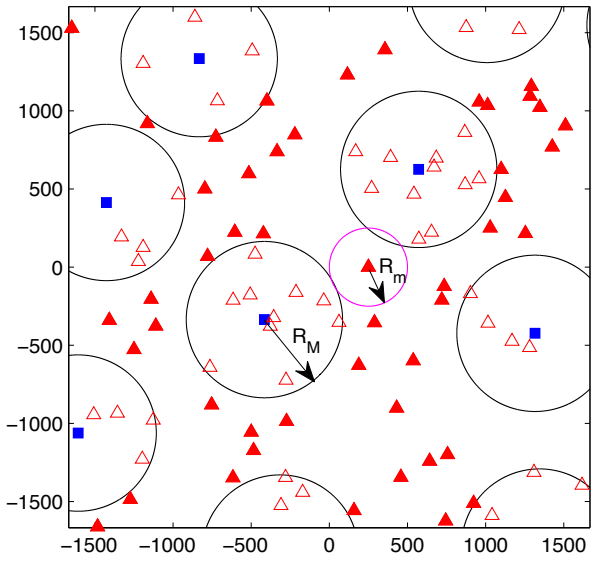


Figure 4.1: The two tier network deployment with geometric dependence. The solid squares are MBSs. The solid triangles are active mBSs. The hollow triangles are muted mBSs.

4.2.2 Transmission Beamforming

We assume that mBSs adopt coordinated beamforming to help mitigate inter-cell interference. On the other hand, each MBS replaces multiple mBSs within its coverage, and therefore each MBS serves multiple users at a time through beamforming that avoids intra-cell interference. In this work, we assume interference nulling beamforming (IN-BF) for mBSs and zero forcing beamforming (ZF-BF) for MBSs. The beamforming strategies for each mBS and MBS are discussed in details as follows.

mBS Beamforming: IN-BF

The channel vector between a user i at position x and an mBS j at position y is $\boldsymbol{f}_{i,j} = \boldsymbol{h}_{i,j}\sqrt{\beta_{i,j}}$, where $\boldsymbol{h}_{i,j} \in \mathbb{C}^{\mathrm{N_m}\times 1}$ is the channel fading coefficient, and the elements of $\boldsymbol{h}_{i,j}$ are i.i.d. $\mathcal{CN}(0,1)$. $\beta_{i,j} = K(\|x-y\|)^{-\alpha}$ is the large-scale PL with constant propagation loss K and exponent factor $\alpha > 2$ [78].

Through exploiting the available spatial degrees of freedom (DoF) at a transmitter, we can drive the sum of undesired signals received at a user node to zero. Note that each mBS serves only one user at a time. As a result, by applying IN-BF, we can make use of the spatial DoF at the mBSs that not only let each mBS serve its designated user, but also let the mBS null its interference to a subset of users in the neighbouring cells [101, 103]. In this chapter, IN-BF is realized as follows. Firstly, all the active users will broadcast a request signal to the nearby mBSs for interference mitigation. Each user's request will be detected by the mBSs in neighbouring cells located within a distance $\mathrm{R}_{\mathrm{det}}$ from the user. Typically, we have $\mathrm{R}_{\mathrm{det}} \geq \mathrm{R}_{\mathrm{m}}$. Next, each mBS will do beamforming using the channel state information (CSI) of the designated user combined with the CSI obtained from the request-detected users.

Assuming a typical user 0 and its designated mBS b, the channel vector between user 0 and b is $\boldsymbol{f}_{0,b} = \boldsymbol{h}_{0,b}\sqrt{\beta_{0,b}}$. The mBS also receives interference nulling requests from $\mathrm{N}_{\mathrm{req}}$ users in neighbouring cells, and the detected users are labeled as $\{1, 2, \ldots, \mathrm{N}_{\mathrm{req}}\}$. Note that the mBS can null its interference to at most $\mathrm{N}_{\mathrm{m}} - 1$ users. If $\mathrm{N}_{\mathrm{req}} > \mathrm{N}_{\mathrm{m}} - 1$, the mBS will randomly suppress interference to $\mathrm{N}_{\mathrm{m}} - 1$ detected users. Without loss of generality, when $\mathrm{N}_{\mathrm{req}} > \mathrm{N}_{\mathrm{m}} - 1$, we assume that the mBS will suppress interference to the first $\mathrm{N}_{\mathrm{m}} - 1$ detected users. Let the channel matrix between mBS b and the detected users be $\boldsymbol{F} = [\boldsymbol{f}_{1,b}, \boldsymbol{f}_{2,b}, \ldots, \boldsymbol{f}_{\min\{\mathrm{N}_{\mathrm{req}}, \mathrm{N}_{\mathrm{m}}-1\},b}]$. The beamforming vector for mBS b is [103]

$$\boldsymbol{w}_b = \frac{\left(\boldsymbol{I}_{\mathrm{N}_{\mathrm{m}}} - \boldsymbol{F}\left(\boldsymbol{F}^H\boldsymbol{F}\right)^{-1}\boldsymbol{F}^H\right)\boldsymbol{f}_{0,b}}{\left\|\left(\boldsymbol{I}_{\mathrm{N}_{\mathrm{m}}} - \boldsymbol{F}\left(\boldsymbol{F}^H\boldsymbol{F}\right)^{-1}\boldsymbol{F}^H\right)\boldsymbol{f}_{0,b}\right\|}$$

where $\boldsymbol{I}_{\mathrm{N}_{\mathrm{m}}}$ is the $\mathrm{N}_{\mathrm{m}} \times \mathrm{N}_{\mathrm{m}}$ identity matrix.

We would like to point out that the maximum ratio combining beamforming (MRC-BF) can be seen as a special case of IN-BF by setting $\mathrm{R}_{\mathrm{det}} = 0$.

MBS Beamforming: ZF-BF

If an MBS p covers M mBSs, the MBS serves in total M users of the covered mBSs at one time slot. Unlike the IN-BF where the spatial DoF afforded by multi-antenna transmission is used to mitigate inter-cell interference, we let each MBS adopt traditional single cell ZF-BF to serve all covered users on the same band at the same time. The M users served by MBS p are labeled as $\{0, 1, \ldots, M-1\}$. Note that the number of transmit antennas at each MBS is $\mathrm{N_M} = \lfloor \lambda_{\mathrm{m}} \pi \mathrm{R_M^2} \mathrm{N_m} \rfloor$. The number of users that MBS p can serve is $\min\{\mathrm{N_M}, \mathrm{M}\}$. Without loss of generality, if $M > \mathrm{N_M}$, the MBS only serves users with labels $\{0, 1, \ldots, \mathrm{N_M} - 1\}$. Let the channel matrix between MBS p and its served users be denoted as $\boldsymbol{G} = \left[\boldsymbol{g}_{0,p}, \boldsymbol{g}_{1,p}, \ldots, \boldsymbol{g}_{\min\{M-1,\mathrm{N_M}-1\},p}\right]^T$, where $\boldsymbol{g}_{i,p} = \boldsymbol{h}_{i,p}\sqrt{\beta_{i,p}} \in \mathbb{C}^{\mathrm{N_M}\times 1}$ is the channel vector between user i and MBS p. The MBS p's beamforming matrix is

$$\boldsymbol{V} = \boldsymbol{G}^H \left(\boldsymbol{G}\boldsymbol{G}^H\right)^{-1}. \tag{4.3}$$

The beamforming vector $\mathbf{v}_{i,p}$ for user $i \in \{0, 1, \ldots, \min\{M-1, \mathrm{N_M}-1\}\}$ is the normalized $(i+1)$th column of the matrix $\boldsymbol{V}$.

4.3 Spectral Efficiency Analysis

4.3.1 User Served by mBS

Let user i be served by mBS b. The received signal $y_{i,b}$ at user i is a sum of the intended signal transmitted by mBS b, the intra-tier interference from other active mBSs that do not null interference to user i, the inter-tier interference from MBSs,

and the additive noise.

$$y_{i,b} = \underbrace{\sqrt{\mathrm{P_m}}\boldsymbol{f}_{i,b}^H \boldsymbol{w}_b s_b}_{\text{intended signal}} + \underbrace{\sum_{j\in\Phi_{\mathrm{m}}^{(\mathrm{a})}\setminus\{b\}} \sqrt{\mathrm{P_m}}\boldsymbol{f}_{i,j}^H \boldsymbol{w}_j s_j}_{\text{intra-tier interference}} + \underbrace{\sum_{k\in\Phi_{\mathrm{M}}^{(\mathrm{a})}} \sum_{m\in\mathcal{M}_k} \sqrt{\frac{\mathrm{P_M}}{|\mathcal{M}_k|}}\boldsymbol{g}_{i,k}^H \boldsymbol{v}_{m,k} s_{m,k}}_{\text{inter-tier interference}} + \underbrace{n_0}_{\text{additive noise}} \tag{4.4}$$

where $\boldsymbol{w}_b$ and $\boldsymbol{v}_{m,k}$ are the IN-BF vector of mBS b and the ZF-BF vector of MBS k for its served user m, respectively. $\Phi_{\mathrm{m}}^{(\mathrm{a})}$ and $\Phi_{\mathrm{M}}^{(\mathrm{a})}$ denote the sets of active mBSs and MBSs, respectively. The signal transmitted by mBS b to its intended user is denoted by a complex scalar s_b with $\mathbb{E}\left[|s_b|^2\right] = 1$. The set of users served by the kth MBS is denoted as $\mathcal{M}_k$ with cardinality $|\mathcal{M}_k|$, and the signal for user m transmitted by MBS k is $s_{m,k}$ with unit average power, i.e. $\mathbb{E}\left[|s_{m,k}|^2\right] = 1$. The additive noise n_0 is a circularly symmetric complex Gaussian (CSCG) random variable with variance σ^2.

With the received signal determined in (4.4), the SINR of user i associated with mBS b is given by

$$\Upsilon_{i,b}^{\mathrm{mBS,IN}} = \frac{\mathrm{P_m}\beta_{i,b}\left\|\boldsymbol{f}_{i,b}^H \boldsymbol{w}_b\right\|^2}{\sum_{j\in\Phi_{\mathrm{m}}^{(\mathrm{a})}\setminus\{b\}} \mathrm{P_m}\left\|\boldsymbol{f}_{i,j}^H \boldsymbol{w}_j\right\|^2 + \sum_{k\in\Phi_{\mathrm{M}}^{(\mathrm{a})}} \sum_{m\in\mathcal{M}_k} \frac{\mathrm{P_M}}{|\mathcal{M}_k|}\left\|\boldsymbol{g}_{i,k}^H \boldsymbol{v}_{m,k}\right\|^2 + \sigma^2}. \tag{4.5}$$

For simplicity, we use the following more descriptive notations:

$$S_{i,b}^{\mathrm{mBS,IN}} = \mathrm{P_m}\beta_{i,b}\left\|\boldsymbol{f}_{i,b}^H \boldsymbol{w}_b\right\|^2, \tag{4.6}$$

$$I_{i,\mathrm{intra}}^{\mathrm{mBS,IN}} = \sum_{j\in\Phi_{\mathrm{m}}^{(\mathrm{a})}\setminus\{b\}} \mathrm{P_m}\left\|\boldsymbol{f}_{i,j}^H \boldsymbol{w}_j\right\|^2, \tag{4.7}$$

$$I_{i,\mathrm{inter}}^{\mathrm{MBS,ZF}} = \sum_{k\in\Phi_{\mathrm{M}}^{(\mathrm{a})}} \sum_{m\in\mathcal{M}_k} \frac{\mathrm{P_M}}{|\mathcal{M}_k|}\left\|\boldsymbol{g}_{i,k}^H \boldsymbol{v}_{m,k}\right\|^2, \tag{4.8}$$

where $S_{i,b}^{\mathrm{mBS,IN}}$, $I_{i,\mathrm{intra}}^{\mathrm{mBS,IN}}$ and $I_{i,\mathrm{inter}}^{\mathrm{MBS,ZF}}$ correspondingly represent the intended signal,

intra-tier interference and inter-tier interference powers received by user i.

Let the distance between user i and mBS b be $d_{i,b}$, and then the PL on the desired link is $\beta_{i,b} = Kd_{i,b}^{-\alpha}$. The ergodic spectral efficiency of the investigated user i conditioned on $d_{i,b}$ is therefore defined as

$$\bar{r}_{i,b}^{\mathrm{mBS,IN}}\left(d_{i,b}\right) = \mathbb{E}\left[\log_2\left(1+\Upsilon_{i,b}^{\mathrm{mBS,IN}}\right)\middle|\, \beta_{i,b} = Kd_{i,b}^{-\alpha}\right]. \tag{4.9}$$

Note that obtaining closed-form expression for $\bar{r}_{i,b}^{\mathrm{mBS,IN}}$ is a challenging task. As a result, we resort to analytically tractable approximations for user ergodic spectral efficiency. In this work, we propose a new performance metric named *virtual spectral efficiency*, which is defined as

$$\hat{r}_{i,b}^{\mathrm{mBS,IN}}\left(d_{i,b}\right) = \log_2\left(1+\hat{\Upsilon}_{i,b}^{\mathrm{mBS,IN}}\left(d_{i,b}\right)\right), \tag{4.10}$$

where $\hat{\Upsilon}_{i,b}^{\mathrm{mBS,IN}}\left(d_{i,b}\right)$ is the *virtual SINR* of user i with given $d_{i,b}$. The virtual SINR $\hat{\Upsilon}_{i,b}^{\mathrm{mBS,IN}}\left(d_{i,b}\right)$ is calculated as

$$\hat{\Upsilon}_{i,b}^{\mathrm{mBS,IN}}\left(d_{i,b}\right) = \frac{\mathbb{E}\left[S_{i,b}^{\mathrm{mBS,IN}}\middle|\, \beta_{i,b} = Kd_{i,b}^{-\alpha}\right]}{\mathbb{E}\left[I_{i,\mathrm{intra}}^{\mathrm{mBS,IN}} + I_{i,\mathrm{inter}}^{\mathrm{MBS,ZF}} + \sigma^2\middle|\, \beta_{i,b} = Kd_{i,b}^{-\alpha}\right]} \tag{4.11}$$

According to the base station deployment scheme proposed in Section 4.2, the positions of base stations no longer follow HPPPs. Additionally, the coordinated IN-BF at mBSs further complicates the user SINR analysis. Thus, to derive closed form expressions for the virtual SINR $\hat{\Upsilon}_{i,b}^{\mathrm{mBS,IN}}\left(d_{i,b}\right)$, we first make the following assumptions.

Assumption 1. *The locations of the MBSs follow a Matérn HCP obtained from a stationary parent PPP of intensity* λ_{p} *with the nodes retained only if they are at distance at least* $2\mathrm{R_M}$ *from all other points. This Matérn HCP can be then approximated by a PPP of intensity* $\widetilde{\lambda}_{\mathrm{M}} = p\lambda_{\mathrm{p}}$ *with* $p = \frac{1-\exp\left(-\lambda_{\mathrm{p}}\pi(2\mathrm{R_M})^2\right)}{\lambda_{\mathrm{p}}\pi(2\mathrm{R_M})^2}$ *and* $\widetilde{\lambda}_{\mathrm{M}} \leq \frac{1}{\pi(2\mathrm{R_M})^2}$.

This assumption follows the approximation in [104]. Note that a Matérn HCP is obtained from a parent PPP by enforcing a minimum distance between any two nodes ($2\mathrm{R_M}$ for our analysis). Denoting the parent PPP node intensity as λ_p, a dependent thinning strategy is proposed in [104] to approximate the intensity of the generated Matérn HCP with separation distance $2\mathrm{R_M}$ as $p\lambda_\mathrm{p}$. The thinning probability p is

$$p = \int_0^1 \exp\left(-\lambda_\mathrm{p}\pi\left(2\mathrm{R_M}\right)^2 t\right) \mathrm{d}t = \frac{1-\exp\left(-\lambda_\mathrm{p}\pi\left(2\mathrm{R_M}\right)^2\right)}{\lambda_\mathrm{p}\pi\left(2\mathrm{R_M}\right)^2}. \tag{4.12}$$

Therefore, in this work we approximate the deployed MBSs as a PPP with intensity $\widetilde{\lambda}_\mathrm{M} = p\lambda_\mathrm{p} = \frac{1-\exp\left(-\lambda_\mathrm{p}\pi(2\mathrm{R_M})^2\right)}{\pi(2\mathrm{R_M})^2}$. Note that for any $\lambda_\mathrm{p} > 0$, $\widetilde{\lambda}_\mathrm{M}$ is always smaller than $\frac{1}{\pi(2\mathrm{R_M})^2}$.

Assumption 2. *Each mBS is able to process all the received interference nulling requests, i.e. there exists negligible probability that the number of detected users at a mBS exceeds the mBS's number of antennas* $\mathrm{N_m}$.

Furthermore, the average number of detected users at a mBS is approximated in the following lemma.

Lemma 4.3.1. *Assume each user sends an interference nulling request and can be detected by the mBSs located within distance* $\mathrm{R_{det}}$. *The average number of user requests detected by a mBS is denoted as* $\bar{\mathrm{N}}_\mathrm{det}$, *where*

$$\bar{\mathrm{N}}_\mathrm{det} \approx \lambda_\mathrm{m}\pi\left(\mathrm{R}_\mathrm{det}^2 - \frac{1}{2}\mathrm{R}_\mathrm{m}^2\right). \tag{4.13}$$

Proof. With no MBS deployed, denote the number of mBSs that receive the request sent from a typical user i as N_i. N_i is a random variable with expected value $\bar{\mathrm{N}}_i$. According to the assumption in Section 4.2, each mBS serves only one user. Without MBS deployment, the total number of requests sent from all users equals to the total number of requests detected by all mBSs. Thus, we adopt the approximation $\bar{\mathrm{N}}_\mathrm{det} \approx \bar{\mathrm{N}}_i$.

To calculate $\bar{\mathrm{N}}_i$, we first denote the distance between user i and its associated mBS b as $d_{i,b}$. Since mBS b is the nearest mBS to user i and $d_{i,b} \leq \mathrm{R_m}$, we have the PDF of $d_{i,b}$ as

$$f_{d_{i,b}}(d_0) = \frac{2d_0}{\mathrm{R_m^2}} \mathbf{1}_{d_{i,b} \leq \mathrm{R_m}}, \tag{4.14}$$

where $\mathbf{1}_x$ is the indicator function with value 1 if condition x is true, and value 0 otherwise.

Conditioned on $d_{i,b} = d_0$, the expected value of N_i is

$$\mathbb{E}\left[\mathrm{N}_i \middle| d_{i,b} = d_0\right] = \lambda_\mathrm{m} \pi \left(\mathrm{R_{det}^2} - d_0^2\right). \tag{4.15}$$

The expected value $\bar{\mathrm{N}}_i$ is then calculated by averaging $\mathbb{E}\left[\mathrm{N}_i \middle| d_{i,b} = d_0\right]$ over $d_{i,b}$, which can be obtained as

$$\bar{\mathrm{N}}_i = \int_0^{\mathrm{R_m}} \mathbb{E}\left[\mathrm{N}_i \middle| d_{i,b} = d_0\right] f_{d_{i,b}}(d_0) \mathrm{d}d_0 = \lambda_\mathrm{m} \pi \left(\mathrm{R_{det}^2} - \frac{1}{2}\mathrm{R_m^2}\right) \tag{4.16}$$

□

Based on Assumption 1, Assumption 2, and Lemma 4.3.1, the approximated expression of $\hat{\Upsilon}_{i,b}^{\mathrm{mBS,IN}}(d_{i,b})$ defined in (4.11) is given in the following lemma.

Lemma 4.3.2. *The virtual SINR $\hat{\Upsilon}_{i,b}^{mBS,IN}(d_{i,b})$ is approximated as $\widetilde{\Upsilon}_{i,b}^{mBS,IN}(d_{i,b})$, which can be expressed as*

$$\widetilde{\Upsilon}_{i,b}^{mBS,IN}(d_{i,b}) = \frac{\widetilde{S}_{i,b}^{mBS,IN}(d_{i,b})}{\widetilde{I}_{i,intra}^{mBS,IN} + \widetilde{I}_{i,inter}^{MBS,ZF} + \sigma^2} \tag{4.17}$$

$$\widetilde{S}_{i,b}^{mBS,IN}(d_{i,b}) = \left(\mathrm{N_m} - \lambda_\mathrm{m}\pi\left(\mathrm{R_{det}^2} - \frac{1}{2}\mathrm{R_m^2}\right)\right) K \mathrm{P_m} d_{i,b}^{-\alpha} \tag{4.18}$$

$$\widetilde{I}_{i,intra}^{mBS,IN} = \frac{2\pi\lambda_\mathrm{m} K \mathrm{P_m}}{\alpha - 2} \left(\mathrm{R_{det}}\right)^{2-\alpha} \tag{4.19}$$

$$\widetilde{I}_{i,inter}^{MBS,ZF} = \frac{2\pi\widetilde{\lambda}_\mathrm{M} K \mathrm{P_M}}{\alpha - 2} \left(\mathrm{R_M}\right)^{2-\alpha}, \tag{4.20}$$

Proof. Please refer to Appendix E. □

Replacing $\hat{\Upsilon}_{i,b}^{\mathrm{mBS,IN}}(d_{i,b})$ in (4.10) by $\widetilde{\Upsilon}_{i,b}^{\mathrm{mBS,IN}}(d_{i,b})$, we have a closed-form approximation $\widetilde{r}_{i,b}^{\mathrm{mBS,IN}}(d_{i,b})$ for the virtual spectral efficiency $\hat{r}_{i,b}^{\mathrm{mBS,IN}}(d_{i,b})$.

$$\hat{r}_{i,b}^{\mathrm{mBS,IN}}(d_{i,b}) \approx \widetilde{r}_{i,b}^{\mathrm{mBS,IN}}(d_{i,b}) = \log_2\left(1+\widetilde{\Upsilon}_{i,b}^{\mathrm{mBS,IN}}(d_{i,b})\right) \tag{4.21}$$

4.3.2 User Served by MBS

Similar to the analysis of users associated with mBSs, we elaborate on analyzing user j served by MBS p. The received signal at user j is

$$\begin{aligned} y_{j,p} = &\underbrace{\sqrt{\frac{\mathrm{P_M}}{|\mathcal{M}_p|}}\boldsymbol{g}_{j,p}^H\boldsymbol{v}_{j,p}s_{j,p}}_{\text{intended signal}} + \\ &\underbrace{\sum_{k\in\Phi_{\mathrm{M}}^{(\mathrm{a})}\setminus\{p\}}\sum_{m\in\mathcal{M}_k}\sqrt{\frac{\mathrm{P_M}}{|\mathcal{M}_k|}}\boldsymbol{g}_{j,k}^H\boldsymbol{v}_{m,k}s_{m,k}}_{\text{intra-tier interference}} + \underbrace{\sum_{l\in\Phi_{\mathrm{m}}^{(\mathrm{a})}}\sqrt{\mathrm{P_m}}\boldsymbol{f}_{j,l}^H\boldsymbol{w}_l s_l}_{\text{inter-tier interference}} + \underbrace{n_0}_{\text{additive noise}} \end{aligned} \tag{4.22}$$

Correspondingly, the received SINR at user j is

$$\begin{aligned} \Upsilon_{j,p}^{\mathrm{MBS,ZF}} &= \frac{S_{j,p}^{\mathrm{MBS,ZF}}}{I_{j,\mathrm{intra}}^{\mathrm{MBS,ZF}} + I_{j,\mathrm{inter}}^{\mathrm{mBS,IN}} + \sigma^2} \\ &= \frac{\frac{\mathrm{P_M}}{|\mathcal{M}_p|}\beta_{j,p}\left\|\boldsymbol{g}_{j,p}^H\boldsymbol{v}_{j,p}\right\|^2}{\sum_{k\in\Phi_{\mathrm{M}}^{(\mathrm{a})}\setminus\{p\}}\sum_{m\in\mathcal{M}_k}\frac{\mathrm{P_M}}{|\mathcal{M}_k|}\left\|\boldsymbol{g}_{i,k}^H\boldsymbol{v}_{m,k}\right\|^2 + \sum_{l\in\Phi_{\mathrm{m}}^{(\mathrm{a})}}\mathrm{P_m}\left\|\boldsymbol{f}_{j,l}^H\boldsymbol{w}_l\right\|^2 + \sigma^2}, \end{aligned} \tag{4.23}$$

where $S_{i,b}^{\mathrm{MBS,ZF}}$, $I_{i,\mathrm{intra}}^{\mathrm{MBS,ZF}}$ and $I_{i,\mathrm{inter}}^{\mathrm{mBS,IN}}$ correspondingly represent the intended signal, intra-tier interference and inter-tier interference powers received by user j.

Conditioned on the distance between user j and MBS p being $d_{j,p}$, and then $\beta_{j,p} = Kd_{j,p}^{-\alpha}$, the ergodic spectral efficiency of user j associated with MBS p is

$$\bar{r}_{j,p}^{\mathrm{MBS,ZF}}(d_{j,p}) = \mathbb{E}\left[\log_2\left(1+\Upsilon_{j,p}^{\mathrm{MBS,ZF}}\right)\middle|\,\beta_{j,p} = Kd_{j,p}^{-\alpha}\right]. \tag{4.24}$$

Similar to the micro tier analysis, the virtual SINR and virtual spectral efficiency are used to approximate $\bar{r}_{j,p}^{\mathrm{MBS,ZF}}$, which are defined as

$$\hat{\Upsilon}_{j,p}^{\mathrm{MBS,ZF}}\left(d_{j,p}\right)=\frac{\mathbb{E}\left[S_{j,p}^{\mathrm{MBS,ZF}}\middle|\,\beta_{j,p}=Kd_{j,p}^{-\alpha}\right]}{\mathbb{E}\left[I_{j,\mathrm{intra}}^{\mathrm{MBS,ZF}}+I_{j,\mathrm{inter}}^{\mathrm{mBS,IN}}+\sigma^{2}\middle|\,\beta_{j,p}=Kd_{j,p}^{-\alpha}\right]} \tag{4.25}$$

$$\hat{r}_{j,p}^{\mathrm{MBS,ZF}}\left(d_{j,p}\right)=\log_{2}\left(1+\hat{\Upsilon}_{j,p}^{\mathrm{MBS,ZF}}\left(d_{j,p}\right)\right) \tag{4.26}$$

According to Assumption 1, Assumption 2, and Lemma 4.3.1, the following lemma is given to approximate virtual SINR $\hat{\Upsilon}_{j,p}^{\mathrm{MBS,ZF}}\left(d_{j,p}\right)$.

Lemma 4.3.3. *The virtual SINR $\hat{\Upsilon}_{j,p}^{MBS,ZF}\left(d_{j,p}\right)$ is approximated as $\widetilde{\Upsilon}_{j,p}^{MBS,ZF}\left(d_{j,p}\right)$, which can be expressed as*

$$\widetilde{\Upsilon}_{j,p}^{MBS,ZF}\left(d_{j,p}\right)=\frac{\widetilde{S}_{j,p}^{MBS,ZF}\left(d_{j,p}\right)}{\widetilde{I}_{j,intra}^{MBS,ZF}+\widetilde{I}_{j,inter}^{mBS,IN}+\sigma^{2}} \tag{4.27}$$

$$\widetilde{S}_{j,p}^{MBS,ZF}\left(d_{j,p}\right)=\left(\lambda_{\mathrm{m}}\pi\mathrm{R}_{\mathrm{M}}^{2}\left(\mathrm{N}_{\mathrm{m}}-1\right)+1\right)\frac{K\mathrm{P}_{\mathrm{M}}}{\lambda_{\mathrm{m}}\pi\mathrm{R}_{\mathrm{M}}^{2}}d_{j,p}^{-\alpha} \tag{4.28}$$

$$\widetilde{I}_{j,intra}^{MBS,ZF}\left(d_{j,p}\right)=\int_{0}^{2\pi}\frac{K\mathrm{P}_{\mathrm{M}}\widetilde{\lambda}_{\mathrm{M}}}{\alpha-2}\left(d_{j,p}\cos\theta+\sqrt{4\mathrm{R}_{\mathrm{M}}^{2}-d_{j,p}^{2}\sin^{2}\theta}\right)^{2-\alpha}\mathrm{d}\theta \tag{4.29}$$

$$\widetilde{I}_{j,inter}^{mBS,IN}\left(d_{j,p}\right)=\int_{0}^{2\pi}\frac{K\mathrm{P}_{\mathrm{m}}\lambda_{\mathrm{m}}}{\alpha-2}\left(\max\left\{\mathrm{R}_{\mathrm{det}},d_{j,p}\cos\theta+\sqrt{\mathrm{R}_{\mathrm{M}}^{2}-d_{j,p}^{2}\sin^{2}\theta}\right\}\right)^{2-\alpha}\mathrm{d}\theta. \tag{4.30}$$

Proof. Please refer to Appendix F. □

Using Lemma 4.3.3, the macro tier user virtual spectral efficiency in (4.26) is therefore approximated as

$$\hat{r}_{j,p}^{\mathrm{MBS,ZF}}\left(d_{j,p}\right)\approx\widetilde{r}_{j,p}^{\mathrm{MBS,ZF}}\left(d_{j,p}\right)=\log_{2}\left(1+\widetilde{\Upsilon}_{j,p}^{\mathrm{MBS,ZF}}\left(d_{j,p}\right)\right). \tag{4.31}$$

Due to the closed-form expressions derived in Lemma 4.3.2 and Lemma 4.3.3, we will use the virtual spectral efficiency approximations as the user rate metrics in the following analysis. The accuracy verification of the virtual spectral efficiency approximations will be illustrated in Section 4.5.

4.4 Power Saving MBS Deployment

Denote the total power consumption of MBS and mBS as $\mathrm{P}_{\mathrm{MBS}}$ and $\mathrm{P}_{\mathrm{mBS}}$, respectively. Based on the linear model proposed in [79], the base station power consumption can be given as

$$\mathrm{P}_{\mathrm{MBS}} = a_{\mathrm{M}}\mathrm{P}_{\mathrm{M}} + b_{\mathrm{M}}, \tag{4.32}$$

$$\mathrm{P}_{\mathrm{mBS}} = a_{\mathrm{m}}\mathrm{P}_{\mathrm{m}} + b_{\mathrm{m}}, \tag{4.33}$$

where a_{M} and a_{m} account for the power consumption that scales with the transmit power. The terms b_{M} and b_{m} represent the transmission-independent power consumption due to signal processing, battery backup, site cooling, etc. Typically, each MBS consumes more power than single mBS. Therefore, we have $a_{\mathrm{M}} \geq a_{\mathrm{m}}$ and $b_{\mathrm{M}} \geq b_{\mathrm{m}}$ with the values of a_{M}, a_{m}, b_{M} and b_{m} given in [79].

Note that by turning on each MBS, multiple mBSs within the MBS's coverage are muted, and then the total number of active base stations is reduced. Therefore, we are interested in determining the optimal density and coverage areas of the deployed MBSs that minimize the overall network power consumption. To ensure the users' QoS, each user's virtual spectral efficiency has to exceed a threshold value C_0. The power minimization problem is formulated as

$$\underset{\mathrm{R}_{\mathrm{M}},\widetilde{\lambda}_{\mathrm{M}}}{\text{minimize}} \quad f\left(\mathrm{R}_{\mathrm{M}}, \widetilde{\lambda}_{\mathrm{M}}\right) \tag{4.34a}$$

$$\text{subject to} \quad \min_{\forall i}\left\{\log_2\left(1+\widetilde{\Upsilon}_{i,b}^{\mathrm{mBS,IN}}\left(d_{i,b}\right)\right)\right\} \geq C_0 \tag{4.34b}$$

$$\min_{\forall j}\left\{\log_2\left(1+\widetilde{\Upsilon}_{j,p}^{\mathrm{MBS,ZF}}\left(d_{j,p}\right)\right)\right\} \geq C_0 \tag{4.34c}$$

$$\widetilde{\lambda}_{\mathrm{M}} < \frac{1}{\pi \mathrm{R}_{\mathrm{M}}^2} \tag{4.34d}$$

where the objective function $f\left(\mathrm{R}_{\mathrm{M}}, \widetilde{\lambda}_{\mathrm{M}}\right) = \widetilde{\lambda}_{\mathrm{M}}\mathrm{P}_{\mathrm{MBS}} + \left(1-\pi\widetilde{\lambda}_{\mathrm{M}}\mathrm{R}_{\mathrm{M}}^2\right)\lambda_{\mathrm{m}}\mathrm{P}_{\mathrm{mBS}}$ stands for the unit area base station power consumption. Constraint (4.34b) and (4.34c) correspondingly guarantee that all users associated with mBSs and MBSs achieve

the minimum spectral efficiency requirement C_0. Constraint (4.34d) comes from Assumption 1 to ensure that the coverage of the deployed MBSs does not overlap.

The spectral efficiency constraints (4.34b) and (4.34c) can be rewritten as

$$\log_2\left(1+\min_{\forall i}\left\{\widetilde{\Upsilon}_{i,b}^{\mathrm{mBS,IN}}\left(d_{i,b}\right)\right\}\right)\geq C_0 \Leftrightarrow \min_{\forall i}\left\{\widetilde{\Upsilon}_{i,b}^{\mathrm{mBS,IN}}\left(d_{i,b}\right)\right\}\geq 2^{C_0}-1 \tag{4.35}$$

$$\log_2\left(1+\min_{\forall j}\left\{\widetilde{\Upsilon}_{j,p}^{\mathrm{MBS,ZF}}\left(d_{j,p}\right)\right\}\right)\geq C_0 \Leftrightarrow \min_{\forall j}\left\{\widetilde{\Upsilon}_{j,p}^{\mathrm{MBS,ZF}}\left(d_{j,p}\right)\right\}\geq 2^{C_0}-1. \tag{4.36}$$

From Lemma 4.3.2, $\widetilde{\Upsilon}_{i,b}^{\mathrm{mBS,IN}}\left(d_{i,b}\right)$ is a non-increasing function of $d_{i,b}$. The maximum value of $d_{i,b}$ equals to the mBS coverage radius $\mathrm{R_m}$. Thus, $\min_{\forall i}\left\{\widetilde{\Upsilon}_{i,b}^{\mathrm{mBS,IN}}\left(d_{i,b}\right)\right\} = \widetilde{\Upsilon}_{i,b}^{\mathrm{mBS,IN}}\left(\mathrm{R_m}\right)$. For the macro user j, the inter-tier interference term (4.30) in Lemma 4.3.3 can be rewritten as

$$\widetilde{I}_{j,\mathrm{inter}}^{\mathrm{mBS,IN}}\left(d_{j,p}\right)=\frac{2K\mathrm{P_m}\lambda_{\mathrm{m}}}{\alpha-2}\left(\int_{x(d_{j,p})}^{\pi}\left(\mathrm{R_{det}}\right)^{2-\alpha}\mathrm{d}\theta + \int_{0}^{x(d_{j,p})}\left(d_{j,p}\cos\theta+\sqrt{\mathrm{R_M^2}-d_{j,p}^2\sin^2\theta}\right)^{2-\alpha}\mathrm{d}\theta\right), \tag{4.37}$$

where $x(d_{j,p})=\arccos\left(\frac{\mathrm{R_{det}^2}+d_{j,p}^2-\mathrm{R_M^2}}{2\mathrm{R_{det}}d_{j,p}}\right)$. Substituting (4.37) into (4.27), we can show that the first order derivative of $\widetilde{\Upsilon}_{j,p}^{\mathrm{MBS,ZF}}\left(d_{j,p}\right)$ with respect to $d_{j,p}$ is smaller than 0. As a result, the value of $\widetilde{\Upsilon}_{j,p}^{\mathrm{MBS,ZF}}\left(d_{j,p}\right)$ decreases as $d_{j,p}$ increases. Therefore, $\min_{\forall j}\left\{\widetilde{\Upsilon}_{j,p}^{\mathrm{MBS,ZF}}\left(d_{j,p}\right)\right\}=\widetilde{\Upsilon}_{j,p}^{\mathrm{MBS,ZF}}\left(\mathrm{R_M}\right)$.

Using the expressions of $\widetilde{\Upsilon}_{i,b}^{\mathrm{mBS,IN}}\left(\mathrm{R_m}\right)$ and $\widetilde{\Upsilon}_{j,p}^{\mathrm{MBS,ZF}}\left(\mathrm{R_M}\right)$, the following upper bounds of $\widetilde{\lambda}_{\mathrm{M}}$, denoted as $U_{\mathrm{m}}\left(\mathrm{R_M}\right)$ and $U_{\mathrm{M}}\left(\mathrm{R_M}\right)$, can be derived from (4.35) and

(4.36):

$$\begin{aligned}\widetilde{\lambda}_{\mathrm{M}} \leq & U_{\mathrm{m}}\left(\mathrm{R}_{\mathrm{M}}\right) \\ = & \left(\frac{\left(\mathrm{N}_{\mathrm{m}}-\lambda_{\mathrm{m}}\pi\left(\mathrm{R}_{\mathrm{det}}^{2}-\frac{1}{2}\mathrm{R}_{\mathrm{m}}^{2}\right)\right)K\mathrm{P}_{\mathrm{m}}}{\Upsilon_{0}\left(R_{\mathrm{m}}\right)^{\alpha}}-\frac{2\pi\lambda_{\mathrm{m}}K\mathrm{P}_{\mathrm{m}}\left(\mathrm{R}_{\mathrm{det}}\right)^{2-\alpha}}{\alpha-2}-\sigma^{2}\right) \\ & \times\frac{\alpha-2}{2\pi K\mathrm{P}_{\mathrm{M}}\left(\mathrm{R}_{\mathrm{M}}\right)^{2-\alpha}}; \end{aligned} \tag{4.38}$$

$$\begin{aligned}\widetilde{\lambda}_{\mathrm{M}} \leq & U_{\mathrm{M}}\left(\mathrm{R}_{\mathrm{M}}\right) \\ = & \left(\frac{\left(\lambda_{\mathrm{m}}\pi\mathrm{R}_{\mathrm{M}}^{2}\left(\mathrm{N}_{\mathrm{m}}-1\right)+1\right)K\mathrm{P}_{\mathrm{M}}}{\Upsilon_{0}\lambda_{\mathrm{m}}\pi\left(\mathrm{R}_{\mathrm{M}}\right)^{2+\alpha}}-\frac{2K\mathrm{P}_{\mathrm{m}}\lambda_{\mathrm{m}}}{\left(\alpha-2\right)\left(\mathrm{R}_{\mathrm{det}}\right)^{\alpha-2}}\left(\pi-\arccos\left(\frac{\mathrm{R}_{\mathrm{det}}}{2\mathrm{R}_{\mathrm{M}}}\right)\right)\right. \\ & \left.-\frac{2K\mathrm{P}_{\mathrm{m}}\lambda_{\mathrm{m}}\left(\mathrm{R}_{\mathrm{M}}\right)^{2-\alpha}}{\alpha-2}\int_{0}^{\arccos\left(\frac{\mathrm{R}_{\mathrm{det}}}{2\mathrm{R}_{\mathrm{M}}}\right)}\left(2\cos\theta\right)^{2-\alpha}\mathrm{d}\theta-\sigma^{2}\right) \\ & \times\frac{\left(\alpha-2\right)\left(\mathrm{R}_{\mathrm{M}}\right)^{\alpha-2}}{K\mathrm{P}_{\mathrm{M}}\int_{0}^{2\pi}\left(\cos\theta+\sqrt{4-\sin^{2}\theta}\right)^{2-\alpha}\mathrm{d}\theta}, \end{aligned} \tag{4.39}$$

where $\Upsilon_0 = 2^{C_0} - 1$.

Furthermore, constraint (4.34d) gives another $\widetilde{\lambda}_{\mathrm{M}}$ upper bound: $U_{\mathrm{c}}\left(\mathrm{R}_{\mathrm{M}}\right) = \frac{1}{\pi\mathrm{R}_{\mathrm{M}}^{2}}$. Thus, the feasible MBS intensity $\widetilde{\lambda}_{\mathrm{M}}$ that satisfies the constraints (4.34b)–(4.34d) in Problem (4.34) is in the range

$$\widetilde{\lambda}_{\mathrm{M}} \leq \min\left\{U_{\mathrm{m}}\left(\mathrm{R}_{\mathrm{M}}\right), U_{\mathrm{M}}\left(\mathrm{R}_{\mathrm{M}}\right), U_{\mathrm{c}}\left(\mathrm{R}_{\mathrm{M}}\right)\right\}. \tag{4.40}$$

For simplicity, the objective function $f(\mathrm{R}_{\mathrm{M}}, \widetilde{\lambda}_{\mathrm{M}})$ in (4.34a) can be reformulated as $f(\mathrm{R}_{\mathrm{M}}, \widetilde{\lambda}_{\mathrm{M}}) = \left(\mathrm{P}_{\mathrm{MBS}} - \pi\lambda_{\mathrm{m}}\mathrm{R}_{\mathrm{M}}^{2}\mathrm{P}_{\mathrm{mBS}}\right)\widetilde{\lambda}_{\mathrm{M}} + \lambda_{\mathrm{m}}\mathrm{P}_{\mathrm{mBS}}$. Additionally, constraints (4.34b), (4.34c) and (4.34d) can be replaced by (4.40). Therefore, Problem (4.34) can be reformulated as follows

$$\min_{\mathrm{R}_{\mathrm{M}}, \widetilde{\lambda}_{\mathrm{M}}} \quad \left(\mathrm{P}_{\mathrm{MBS}} - \pi\lambda_{\mathrm{m}}\mathrm{R}_{\mathrm{M}}^{2}\mathrm{P}_{\mathrm{mBS}}\right)\widetilde{\lambda}_{\mathrm{M}} + \lambda_{\mathrm{m}}\mathrm{P}_{\mathrm{mBS}} \tag{4.41a}$$

$$\text{s.t.} \quad 0 \leq \widetilde{\lambda}_{\mathrm{M}} \leq \min\left\{U_{\mathrm{c}}\left(\mathrm{R}_{\mathrm{M}}\right), U_{\mathrm{m}}\left(\mathrm{R}_{\mathrm{M}}\right), U_{\mathrm{M}}\left(\mathrm{R}_{\mathrm{M}}\right)\right\} \tag{4.41b}$$

To solve the reformulated Problem (4.41), we first evaluate the term

$\mathrm{P_{MBS}} - \pi\lambda_{\mathrm{m}}\mathrm{R_M^2}\mathrm{P_{mBS}}$ in the objective function (4.41a). Note that when $\mathrm{P_{MBS}} - \pi\lambda_{\mathrm{m}}\mathrm{R_M^2}\mathrm{P_{mBS}} \geq 0$, the minimum value of (4.41a) is achieved at the minimum $\widetilde{\lambda}_{\mathrm{M}}$. In this case, the power consumed by each MBS, i.e. $\mathrm{P_{MBS}}$, is larger than the total power used by the mBSs under a MBS's coverage, i.e. $\pi\lambda_{\mathrm{m}}\mathrm{R_M^2}\mathrm{P_{mBS}}$. As a result, deploying MBSs will not decrease the total network power consumption. On the other hand, when $\mathrm{P_{MBS}} - \pi\lambda_{\mathrm{m}}\mathrm{R_M^2}\mathrm{P_{mBS}} < 0$, the deployment of MBSs can reduce the total network power consumption. In the following analysis, we will elaborate on this case by finding the optimal $\mathrm{R_M}$ and λ_{M} for Problem (4.41) with the additional constraint:

$$\mathrm{P_{MBS}} - \pi\lambda_{\mathrm{m}}\mathrm{R_M^2}\mathrm{P_{mBS}} < 0, \tag{4.42}$$

which is equivalent to $\mathrm{R_M} \geq \sqrt{\frac{a_{\mathrm{M}}\mathrm{P_M}+b_{\mathrm{M}}}{\pi\lambda_{\mathrm{m}}(a_{\mathrm{m}}\mathrm{P_m}+b_{\mathrm{m}})}}$.

With the new constraint (4.42), the objective function (4.41a) decreases with increasing $\widetilde{\lambda}_{\mathrm{M}}$. Thus, the minimum value of (4.41a) is obtained at the maximum achievable $\widetilde{\lambda}_{\mathrm{M}}$, which is $\min\{U_{\mathrm{c}}(\mathrm{R_M}), U_{\mathrm{m}}(\mathrm{R_M}), U_{\mathrm{M}}(\mathrm{R_M})\}$. We will discuss the cases where the maximum $\widetilde{\lambda}_{\mathrm{M}}$ correspondingly equals to $U_{\mathrm{c}}(\mathrm{R_M})$, $U_{\mathrm{m}}(\mathrm{R_M})$, and $U_{\mathrm{M}}(\mathrm{R_M})$ in the following analysis.

Case I: $U_{\mathrm{c}}(\mathrm{R_M}) = \min\{U_{\mathrm{c}}(\mathrm{R_M}), U_{\mathrm{m}}(\mathrm{R_M}), U_{\mathrm{M}}(\mathrm{R_M})\}$

In this case, the objective function in (4.41a) becomes $(\mathrm{P_{MBS}} - \pi\lambda_{\mathrm{m}}\mathrm{R_M^2}\mathrm{P_{mBS}})\frac{1}{\pi\mathrm{R_M^2}}$, which is a decreasing function of $\mathrm{R_M}$. The optimal $\mathrm{R_M}$ in this case equals to its maximum achievable value that satisfies

$$U_{\mathrm{c}}(\mathrm{R_M}) \leq U_{\mathrm{m}}(\mathrm{R_M}) \tag{4.43}$$

$$U_{\mathrm{c}}(\mathrm{R_M}) \leq U_{\mathrm{M}}(\mathrm{R_M}). \tag{4.44}$$

Using the expression of $U_{\mathrm{m}}(\mathrm{R_M})$ in (4.38), we can verify that $U_{\mathrm{m}}(\mathrm{R_M})$ is an increasing function of $\mathrm{R_M}$. Since U_{c} is a decreasing function of $\mathrm{R_M}$, the curves of $U_{\mathrm{c}}(\mathrm{R_M})$ and $U_{\mathrm{m}}(\mathrm{R_M})$ have at most one intersection point $\mathrm{R_1}$. The inequality

in (4.43) therefore gives $\mathrm{R_M} \geq \mathrm{R_1}$. In addition, we use the following lemma to determine the feasible range of $\mathrm{R_M}$ that satisfies (4.44).

Lemma 4.4.1. *The constraint* $U_{\mathrm{c}}(\mathrm{R_M}) \leq U_{\mathrm{M}}(\mathrm{R_M})$ *determines a feasible range of* $\mathrm{R_M}$ *as* $\mathrm{R_M} \leq \mathrm{R_2}$*, where* $\mathrm{R_2}$ *is the single solution of* $U_{\mathrm{c}}(x) = U_{\mathrm{M}}(x)$*.*

Proof. Please see Appendix G. □

Additionally, constraint (4.42) determines $\mathrm{R_M} \geq \sqrt{\frac{a_{\mathrm{M}}\mathrm{P_M}+b_{\mathrm{M}}}{\pi\lambda_{\mathrm{m}}(a_{\mathrm{m}}\mathrm{P_m}+b_{\mathrm{m}})}}$. As a result, the optimal $\mathrm{R_M}$ for Case I equals to $\mathrm{R_2}$ only if $\mathrm{R_2} \geq \max\left\{\mathrm{R_1}, \sqrt{\frac{a_{\mathrm{M}}\mathrm{P_M}+b_{\mathrm{M}}}{\pi\lambda_{\mathrm{m}}(a_{\mathrm{m}}\mathrm{P_m}+b_{\mathrm{m}})}}\right\}$.

Case II: $U_{\mathrm{m}}(\mathrm{R_M}) = \min\{U_{\mathrm{c}}(\mathrm{R_M}), U_{\mathrm{m}}(\mathrm{R_M}), U_{\mathrm{M}}(\mathrm{R_M})\}$

We use $\mathrm{R_3}$ to denote the intersection point of $U_{\mathrm{m}}(\mathrm{R_M})$ and $U_{\mathrm{M}}(\mathrm{R_M})$, i.e. $U_{\mathrm{m}}(\mathrm{R_3}) = U_{\mathrm{M}}(\mathrm{R_3})$. Similar to the analysis in Case I, we have the optimal $\mathrm{R_M}$ equals to $\min\{\mathrm{R_1}, \mathrm{R_3}\}$ under the constraint $\min\{\mathrm{R_1}, \mathrm{R_3}\} \geq \sqrt{\frac{a_{\mathrm{M}}\mathrm{P_M}+b_{\mathrm{M}}}{\pi\lambda_{\mathrm{m}}(a_{\mathrm{m}}\mathrm{P_m}+b_{\mathrm{m}})}}$.

Case III: $U_{\mathrm{M}}(\mathrm{R_M}) = \min\{U_{\mathrm{c}}(\mathrm{R_M}), U_{\mathrm{m}}(\mathrm{R_M}), U_{\mathrm{M}}(\mathrm{R_M})\}$

In this case, the monotonicity of the function $(\mathrm{P_{MBS}} - \pi\lambda_{\mathrm{m}}\mathrm{R_M^2}\mathrm{P_{mBS}})\,U_{\mathrm{M}}(\mathrm{R_M})$ is discussed in the following lemma.

Lemma 4.4.2. *The function* $F(\mathrm{R_M}) = (\mathrm{P_{MBS}} - \pi\lambda_{\mathrm{m}}\mathrm{R_M^2}\mathrm{P_{mBS}})\,U_{\mathrm{M}}(\mathrm{R_M})$ *is an unimodal function with the minimum value achieved at* $\mathrm{R_M} = \mathrm{R_0}$*.*

Proof. Let set $\mathcal{Y} = \left\{y \,\middle|\, \frac{\mathrm{d}F(y)}{\mathrm{d}y} = 0\right\}$. It can be shown that $\left.\frac{\mathrm{d}^2F(y)}{\mathrm{d}y^2}\right|_{y\in\mathcal{Y}} > 0$. Therefore, set $\mathcal{Y}$ has at most one element and function $F(y)$ achieves its minimum value at $y \in \mathcal{Y}$. □

According to Lemma 4.4.2, we can find $\mathrm{R_0}$ though bisection search. Note that the feasible range is $\mathrm{R_M} \geq \max\left\{\mathrm{R_2}, \mathrm{R_3}, \sqrt{\frac{a_{\mathrm{M}}\mathrm{P_M}+b_{\mathrm{M}}}{\pi\lambda_{\mathrm{m}}(a_{\mathrm{m}}\mathrm{P_m}+b_{\mathrm{m}})}}\right\}$ for Case III. As a result, the optimal $\mathrm{R_M}$ for Case III equals to $\max\left\{\mathrm{R_0}, \mathrm{R_2}, \mathrm{R_3}, \sqrt{\frac{a_{\mathrm{M}}\mathrm{P_M}+b_{\mathrm{M}}}{\pi\lambda_{\mathrm{m}}(a_{\mathrm{m}}\mathrm{P_m}+b_{\mathrm{m}})}}\right\}$.

Based on the above discussions, the optimal solutions for Case I, II and III can be determined. Note that the solutions for Case I, II and III include all the local optimal points of Problem (4.34) with constraint (4.42). Therefore, by selecting

the minimum one within the solutions for the three cases, we can obtain the global optimal solution of Problem (4.34) with constraint (4.42).

4.5 Numerical Results

In this section, the analytical results are verified through numerical tests. The transmit power of macro and micro base stations are $\mathrm{P_M} = 10$ W and $\mathrm{P_m} = 1$ W, respectively. We assume each mBS equipped with $\mathrm{N_m} = 4$ transmit antennas. For the small cell network deployment, base station intensity $\lambda_\mathrm{m} = 25$ mBSs/km^2 and the coverage radius of each mBS is $\mathrm{R_m} = 250$ m. Thus, the probability of a user being located within a coverage hole (i.e. $p_{\bar{\mathrm{c}}}$ in (4.1)) is below 0.01. Each mBS can detect the interference nulling request from users within distance $\mathrm{R_{det}} = 250$ m. The PL parameters are $K = 10^{-3}$, $\alpha = 4$. Additive noise variance $\sigma^2 = -134$ dBm.

In Fig. 4.2, we simulate the cell edge user ergodic spectral efficiency on MATLAB platform using Monte Carlo methods and compare the simulated results with the proposed virtual spectral efficiency approximations for users served by mBSs and MBSs, respectively. Two deployments of MBSs with $\widetilde{\lambda}_\mathrm{M} = 0.01\lambda_\mathrm{m}$ and $\widetilde{\lambda}_\mathrm{M} = 0.05\lambda_\mathrm{m}$ are investigated. For each $\widetilde{\lambda}_\mathrm{M}$ setting, MBSs are deployed following the discussion in Section 4.2, where the MBSs are firstly randomly located according to a PPP of intensity $\frac{\widetilde{\lambda}_\mathrm{M}\lambda_\mathrm{p}\pi(2\mathrm{R_M})^2}{1-\exp\left(-\lambda_\mathrm{p}\pi(2\mathrm{R_M})^2\right)}$, and then the MBSs with their coverage areas overlapped with other MBSs are eliminated and the mBSs within the remaining MBSs coverage areas are removed. The radius of each MBS's coverage area satisfies $\pi\mathrm{R_M^2}\widetilde{\lambda}_\mathrm{M} \leq 1$. For each network deployment, the nodes are randomly placed on a 10km x 10km square area. We randomly select the target users on the coverage edges of an MBS and an mBS and calculate the corresponding spectral efficiency based on generated channel matrices. The simulated expected spectral efficiency points in Fig. 4.2 are obtained by averaging over 1000 channel realizations of each of the 1000 random network deployments[2]. Furthermore, the approximated spectral

[2]Similar as in Chapter 3, we run the simulation experiments for 100 times and the 95% confidence interval for the points on the simulated macro user spectral efficiency curve with

efficiency curves are obtained using the virtual SINR expressions derived in Lemma 4.3.2 and Lemma 4.3.3. It is shown in Fig. 4.2 that the approximate micro and macro user virtual spectral efficiencies match well with the simulated results.

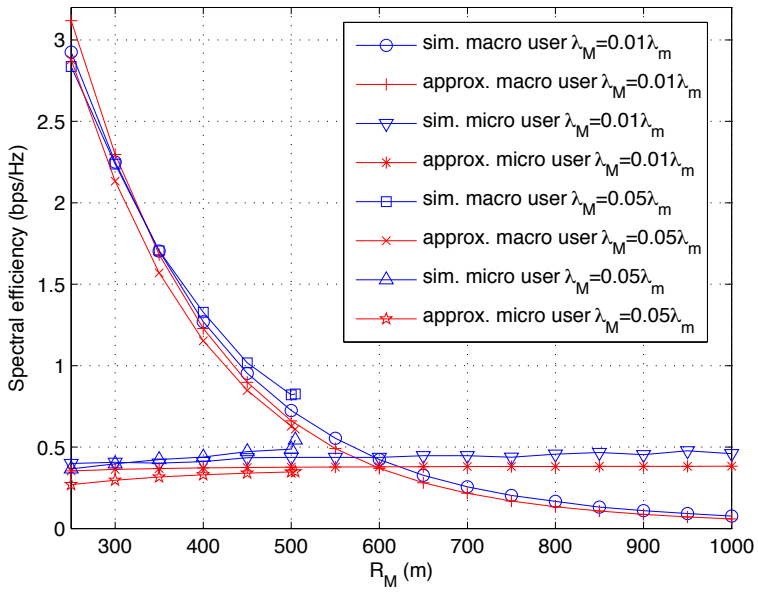


Figure 4.2: The effects of MBS coverage radius on the expected spectral efficiency. Typical user i is located on the cell edge: $r_{i,b} = \mathrm{R_m}$ for micro user, $r_{i,p} = \mathrm{R_M}$ for macro user.

With the spectral efficiency threshold $C_0 = 0.3$, the optimal MBS deployment methodology in Section 4.4 is verified in the following two figures. In Fig. 4.3, we plot the $\widetilde{\lambda}_{\mathrm{M}}$ upper bounds: $U_{\mathrm{m}}(\mathrm{R_M})$ in (4.38), $U_{\mathrm{M}}(\mathrm{R_M})$ in (4.39), and $U_{\mathrm{c}}(\mathrm{R_M}) = \frac{1}{\pi \mathrm{R_M^2}}$ derived from Assumption 1. The intersection points between these three bounds can be obtained. According to the discussion in Section 4.4, the intersection points are denoted as R_1, R_2, and R_3. In Fig. 4.3, $\mathrm{R}_1 = 394$ m, $\mathrm{R}_2 = 626$ m and $\mathrm{R}_3 = 599$ m. For $\mathrm{R_M} \leq \mathrm{R}_1$, $U_{\mathrm{m}}(\mathrm{R_M}) \leq U_{\mathrm{M}}(\mathrm{R_M})$ and $U_{\mathrm{m}}(\mathrm{R_M}) \leq U_{\mathrm{c}}(\mathrm{R_M})$ is satisfied. Consequently, the maximum MBS intensity $\widetilde{\lambda}_{\mathrm{M}}$ is smaller than $U_{\mathrm{m}}(\mathrm{R_M})$ with $\mathrm{R_M} \leq \mathrm{R}_1$, which is the Case II discussed in Section 4.4. For $\mathrm{R}_1 < \mathrm{R_M} \leq \mathrm{R}_2$, Case I is valid where $U_{\mathrm{c}}(\mathrm{R_M}) \leq U_{\mathrm{M}}(\mathrm{R_M})$ and $U_{\mathrm{c}}(\mathrm{R_M}) \leq U_{\mathrm{m}}(\mathrm{R_M})$. When $\mathrm{R}_2 < \mathrm{R_M}$, we have $U_{\mathrm{M}}(\mathrm{R_M}) \leq$

deployed MBS intensity $\widetilde{\lambda}_{\mathrm{M}} = 0.01\lambda_{\mathrm{m}}$ is less than 0.01.

$U_{\mathrm{c}}(\mathrm{R_M})$ and $U_{\mathrm{M}}(\mathrm{R_M}) \leq U_{\mathrm{m}}(\mathrm{R_M})$, and then the analysis for Case III is adopted.

Under the same parameter settings, in Fig. 4.4, we correspondingly plot $f(\mathrm{R_M}, \widetilde{\lambda}_{\mathrm{M}})$ in (4.41a) with $\widetilde{\lambda}_{\mathrm{M}} = U_{\mathrm{c}}(\mathrm{R_M})$ (Case I), $U_{\mathrm{m}}(\mathrm{R_M})$ (Case II) and $U_{\mathrm{M}}(\mathrm{R_M})$ (Case III). It can be seen from Fig. 4.4 that the values of $f(\mathrm{R_M}, U_{\mathrm{c}}(\mathrm{R_M}))$ and $f(\mathrm{R_M}, U_{\mathrm{m}}(\mathrm{R_M}))$ decrease with increasing $\mathrm{R_M}$. In addition, the function $f(\mathrm{R_M}, U_{\mathrm{M}}(\mathrm{R_M}))$ is an unimodal function with its minimum value achieved at $\mathrm{R_M} = \mathrm{R}_0 = 547$ m. Note that the discussion of the three cases in Section 4.4 are all based on the constraint (4.42). Thus, Case I, II and III are feasible only if $\mathrm{P_{MBS}} - \pi\lambda_{\mathrm{m}}\mathrm{R_M^2}\mathrm{P_{mBS}} < 0$ is satisfied. According to Fig. 4.4, the range within which $\mathrm{R_M}$ achieves $\mathrm{P_{MBS}} - \pi\lambda_{\mathrm{m}}\mathrm{R_M^2}\mathrm{P_{mBS}} < 0$ is $[466, \infty)$. As a result, Case II, where $\mathrm{R_M} \leq \mathrm{R}_1$, is infeasible. Additionally, Case I and Case III are valid within the $\mathrm{R_M}$ intervals $[466, \mathrm{R}_2)$ and $[\mathrm{R}_2, \infty)$, respectively. As a result, the optimal MBS coverage radius $\mathrm{R_M^*} = \mathrm{R}_2$ and the corresponding minimum area network power consumption $f(\mathrm{R_M^*}, \widetilde{\lambda}_{\mathrm{M}}^*)$ equals to $f(\mathrm{R}_2, U_{\mathrm{M}}(\mathrm{R}_2)) = f(\mathrm{R}_2, U_{\mathrm{c}}(\mathrm{R}_2)) = 0.52 \times 10^{-3}$ W/m^2.

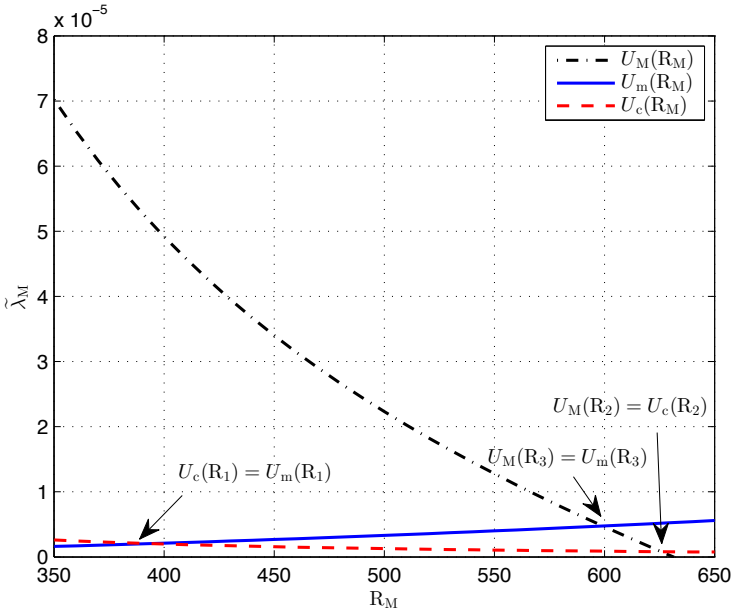


Figure 4.3: The upper bounds of $\widetilde{\lambda}_{\mathrm{M}}$: $U_{\mathrm{M}}(R_M)$, $U_{\mathrm{m}}(R_M)$, $U_{\mathrm{c}}(R_M)$

In Fig. 4.5, we plot the area network power consumption according to the mBS intensity λ_{m}. With spectral efficiency threshold $C_0 = 0.3$ bps/Hz, the optimal

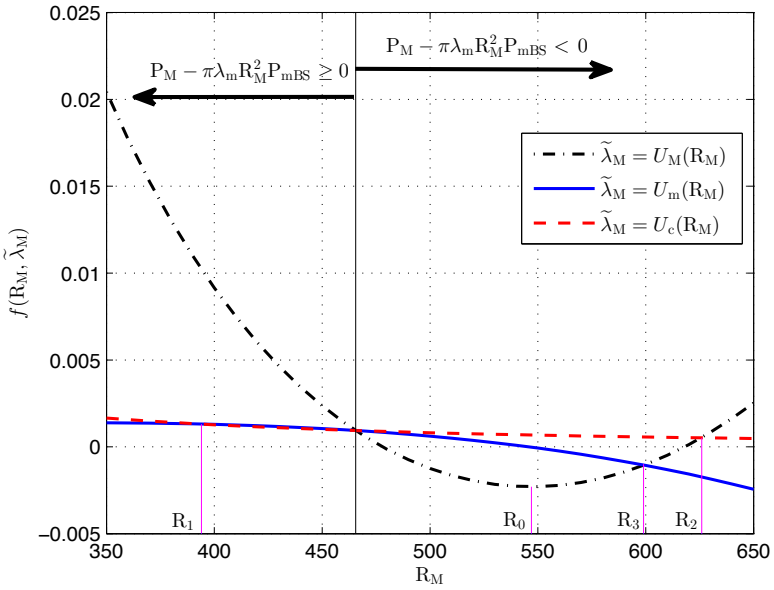


Figure 4.4: The values of objective function $f(\mathrm{R_M}, \widetilde{\lambda}_\mathrm{M})$ in Problem 4.41 when $\widetilde{\lambda}_\mathrm{M}$ correspondingly equals to $U_\mathrm{M}(R_M)$, $U_\mathrm{m}(R_M)$, $U_\mathrm{c}(R_M)$.

deployment of MBSs are determined using the proposed method. We also give the power consumption of the network with only mBSs , denoted as "mBS only" deployment in Fig. 4.5. For the mBS only deployment in Fig. 4.5, the positions of mBSs follow an HPPP with intensity λ_m. As Fig. 4.5 shows, all the curves grow with the mBS intensity. This is due to the increased number of active base stations. However, the power consumption increases slowly in the network using the proposed MBS deployment scheme. This is because the proposed MBS deployment helps reduce the number of active mBSs. Additionally, with the mBS intensity increasing, the gap between the proposed MBS deployment curve and mBS only curve grows, which means that the proposed scheme is efficient to be applied in densely deployed small-cell networks.

In Fig. 4.6, the effects of MBS transmit power $\mathrm{P_M}$ on the network power consumption are investigated. Let $\mathrm{P_M}$ increase from 15 W to 50 W, we can see from Fig. 4.6 that the area power consumption of the network with optimally deployed MBSs decreases. This is because the larger MBS transmit power results in a bigger

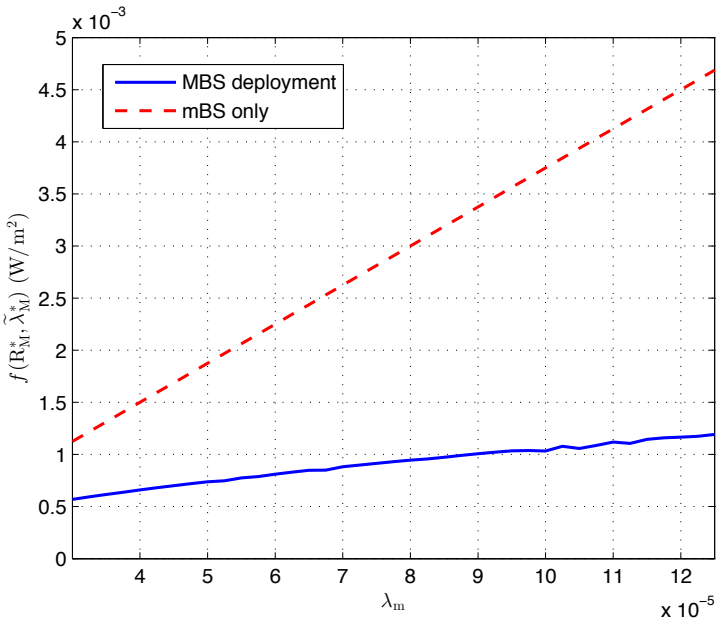


Figure 4.5: The effects of mBS intensity λ_{m} on the network power consumption.

coverage area within which more mBSs can be muted to reduce the total number of active nodes. Additionally, Fig. 4.6 also verified that the proposed deployment scheme outperforms the random deployment scheme with only mBSs in terms of network power saving.

As indicated in [79] and [100], serving all users with small cells is desirable for network throughput optimization but is not optimal for energy efficient network design. The results in Fig. 4.5 and Fig. 4.6 confirm this assertion by showing that deploying MBSs within a small cell network helps in reducing network power consumption.

4.6 Conclusions

In this chapter, we investigated an energy efficient base station deployment problem in a two-tier multi-antenna network by taking geometric dependence into consideration. We proposed a stochastic geometry framework to approximately model and analyse the network performance. Based on the proposed framework, we

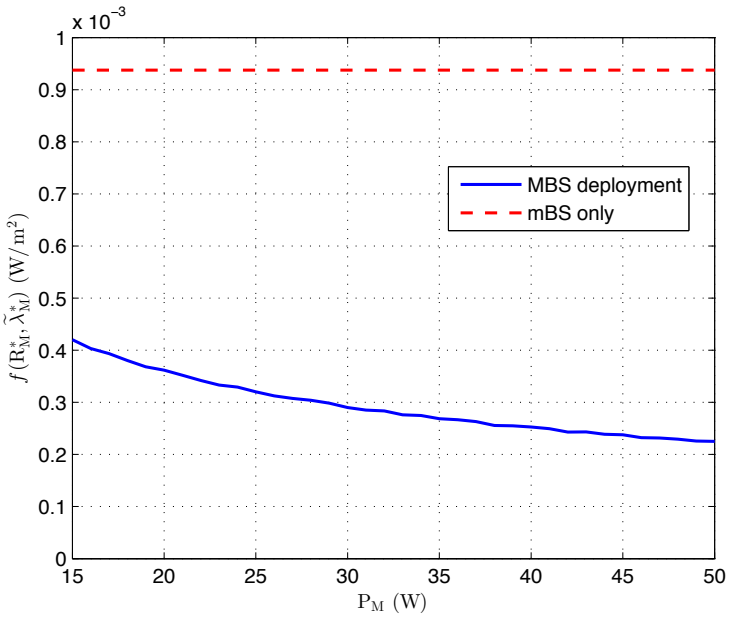


Figure 4.6: The effects of MBS transmit power P_M on the network power consumption.

first derived tractable approximations of the expected spectral efficiency and then determined the optimal density and coverage radius of the deployed MBSs that minimized network power consumption.

Chapter 5

Conclusion and Future Work

5.1 Conclusion

This thesis has comprehensively investigated the optimal power saving strategies in the downlinks of HCNs by applying stochastic geometry analysis. Three novel HCN frameworks that involve joint interference elimination and cell load control were proposed. Then, using stochastic geometry and reasonable approximations, tractable and accurate QoS metrics were derived. Finally, the optimal parameter design strategies that minimize network power consumption were determined based on the proposed HCN frameworks and QoS approximations. The detailed descriptions of the main contributions of Chapter 2–4 are summarized as follows.

- In Chapter 2, we proposed a stochastic geometry framework for a two-tier HCN with sleep-mode base stations that jointly adopted resource partitioning and biased user association to mitigate inter-tier interference of users at the coverage boundaries of different types of base stations. Simple expressions were derived to characterize user throughput. By solving power minimization and coverage maximization problems with throughput constraints, we proposed algorithms to determine the optimal resource partitioning fraction and user association bias. The proposed network design approaches were analytically tractable and provided guidelines for practical HCN design. The results in Chapter 2 revealed that muting small-cells on the partitioned resources

was more efficient in improving network coverage and muting macrocells on the partitioned resources brought about more reduction in network power consumption.

- In Chapter 3, we proposed a novel exclusion zone deployment strategy in femtocell networks to mitigate inter-tier interference and reduce base station transmit power. We derived tractable approximations of success probabilities for users at cell edge, based on which a cell-load dependent MBS transmit power control scheme was determined. By minimizing the average MBS transmit power, the optimal exclusion zone design strategies for open and closed access femtocells were obtained. The results provided insights into energy efficient exclusion zone design for femtocell networks, and verified that using exclusion zones improves cell edge user performance.

- In Chapter 4, we proposed a new base station deployment scheme and applied different multi-antenna beamforming methods at macro and micro base stations to respectively mitigate intra- and inter-cell interference. A novel metric named virtual spectral efficiency was proposed to approximate the exact user ergodic spectral efficiency with high accuracy and tractable expression. The optimal design of the density and coverage of the deployed MBSs was obtained by solving network power minimization problem. Our studies provided promising principles for tractable user spectral efficiency characterization and power saving deployment of multi-antenna base stations in HCNs. Also, the discoveries in this chapter showed the potential of using different types of base stations in improving network energy efficiency.

5.2 Future Work

The work in this thesis may be extended in several directions, as outlined below.

In this thesis, HCNs are designed to minimize the network power consumption under user QoS constraints. It will be interesting to design HCNs to optimize metrics

that reflect the trade-offs between spectral efficiency and energy efficiency [66, 105], which will result in reduced power consumption per transmitted bit. Furthermore, recent works [106, 107] have shown the advantages in terms of network performance improvement of using different user association schemes for the uplink and downlink in HCNs. Therefore, another possible extension of this thesis is to take uplink performance analysis into consideration using the proposed downlink energy efficient design methods. Moreover, some future works are left to be explored in Chapter 2–4.

- In Chapter 2, joint resource partitioning and cell load adaptation is used only to mitigate inter-tier interference for macro users. The user association bias B_2 that determines macro and small-cell users is fixed as constant during the network design. In the future work, it is interesting to determine the optimal value of B_2 that further improves network performance.

- For the transmit power saving design of femtocell exclusion zones discussed in Chapter 3, we only considered using exclusion zones to mitigate inter-tier interference. Thus, it is interesting to study the optimal design of exclusion zones around macro- and femtocells that eliminate both inter- and intra-cell interference. Additionally, it is also promising to jointly reduce the transmit power of macro and femto base stations for networks with multiple macrocells. Moreover, considering power adaptation schemes depending on multiple network parameters (e.g. cell load, SINR, throughput) is also a challenging yet interesting extension of the results in Chapter 3.

- For the energy efficient base station deployment problem studied in Chapter 4, the implementation of other beamforming and power allocation schemes can be considered in the future work. Furthermore, we can include the analysis of resource allocation in the base station deployment discussion to extend our results in Chapter 4.

The stochastic geometry based energy efficient HCN design analysis in this

thesis can be adopted to explore other types of novel networks. One possible extension is to explore millimeter wave cellular networks. Despite its potential to provide high data rates due to the large bandwidths available, millimeter wave transmission suffers from large free-space path loss due to the high carrier frequency [108,109]. In order to implement millimeter wave technique in the cellular networks, cell size should be reduced or multi-hop relay techniques should be used to combat the signal path loss. Thus, we can draw an analogy between millimeter wave cellular system and HCNs with small-cells. Using stochastic geometry analysis, the coverage probability of millimeter wave cellular systems can be characterized [110], which reveals the possibility to implement the energy efficient design schemes proposed in this thesis to mitigate interference and reduce power consumption in millimeter wave cellular systems. Another extension of this thesis is to investigate D2D networks. Since the communications between devices are in an ad hoc manner, stochastic geometry can be applied to characterize D2D network performance [111, 112]. Moreover, in cellular networks overlaid with device to device communications, interference control is an important issue [113]. It is therefore possible to make use of the exclusion zones [114] and beamforming strategies studied in this thesis to help control interference in D2D cellular networks.

Last but not least, the power saving strategies investigated in this thesis can be combined with other green communication strategies such as energy harvesting [115, 116] to further reduce network power consumption. Note that stochastic geometry analysis only provides high level guidelines for network design, and that practical implementation will require further refinements, including real-time adaptations to fluctuations in channels, data throughput requirements, etc. The entire design chain, from the macro view offered by stochastic geometry down to the micro design to match instantaneous user requirements and system state, is yet to be explored. This system design viewpoint offers rich possibilities for future research and development in the Smart Cities/Internet of Things domain.

Appendix A

Proof of Corollary 2.4.1

Using the results of Lemma 2.3.1, Lemma 2.3.2 and (2.16), the values of $\eta_{\mathrm{D1}}(\mathrm{B}_1)$ and $\eta_1(\mathrm{B}_1)$ at $\mathrm{B}_1 \to \mathrm{B}_2$ and $\mathrm{B}_1 \to \infty$ are

$$\lim_{x\to \mathrm{B}_2} \eta_{\mathrm{D1}}(x) = 1 - Z_0 \left(\frac{\lambda_1 \left[1+\rho(\Upsilon,\alpha)\right] + \lambda_2 \left(\frac{\mathrm{B}_2\mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}}}{\lambda_1 + \left(\frac{\mathrm{B}_2\mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}} \lambda_2} \right)^2$$

$$\lim_{x\to \mathrm{B}_2} \eta_1(x) = Z_0 \left(1 + \frac{1.28\lambda_u}{\lambda_1 + \left(\frac{\mathrm{B}_2\mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}} \lambda_2} \right) \frac{\lambda_1 \left[1+\rho(\Upsilon,\alpha)\right] + \lambda_2 \left(\frac{\mathrm{B}_2\mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}} \left[1+\rho\left(\frac{\Upsilon}{\mathrm{B}_2},\alpha\right)\right]}{\lambda_1 + \left(\frac{\mathrm{B}_2\mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}} \lambda_2}$$

$$\lim_{x\to \infty} \eta_{\mathrm{D1}}(x) = 1 - Z_0 \left(1 + \frac{1.28\lambda_u}{\lambda_1 + \left(\frac{\mathrm{B}_2\mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}} \lambda_2} \right) \frac{\lambda_1 \left[1+\rho(\Upsilon,\alpha)\right] + \lambda_2 \left(\frac{\mathrm{B}_2\mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}}}{\lambda_1 + \left(\frac{\mathrm{B}_2\mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}} \lambda_2}$$

$$\lim_{x\to \infty} \eta_1(x) = Z_0,$$

where $Z_0 = \frac{C_1}{W \log_2(1+\Upsilon)}$.

From the above equations, it can be easily obtained that $\lim_{x\to\infty} (\eta_{\mathrm{D1}}(x) - \eta_1(x)) \geq \lim_{x\to \mathrm{B}_2} (\eta_{\mathrm{D1}}(x) - \eta_1(x))$. Hence, we can conclude that $\lim_{x\to\infty} \eta_{\mathrm{D1}}(x) > \lim_{x\to\infty} \eta_1(x)$ is always true if $\lim_{x\to \mathrm{B}_2} \eta_{\mathrm{D1}}(x) > \lim_{x\to \mathrm{B}_2} \eta_1(x)$ holds. In addition, we have $\lim_{x\to \mathrm{B}_2} \eta_{\mathrm{D1}}(x) < \lim_{x\to \mathrm{B}_2} \eta_1(x)$, when $\lim_{x\to\infty} \eta_{\mathrm{D1}}(x) < \lim_{x\to\infty} \eta_1(x)$ is satisfied.

Introducing a new variable $x = \mathrm{B}_1^{\frac{2}{\alpha}}$, the fist order derivative of η_{D1} and η_1 with

respect to B_1 can be derived as

$$\frac{\mathrm{d}\eta_1(\mathrm{B}_1)}{\mathrm{dB}_1} = \frac{C_1}{W\log_2(1+\Upsilon)}\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\bar{\mathrm{N}}_1}{\mathbb{P}_1}\right)\frac{\mathrm{d}x}{\mathrm{dB}_1}$$
$$\frac{\mathrm{d}\eta_{\mathrm{D}1}(\mathrm{B}_1)}{\mathrm{dB}_1} = \frac{C_1}{W\log_2(1+\Upsilon)}\frac{-\mathrm{d}}{\mathrm{d}x}\left(\frac{\bar{\mathrm{N}}_{\mathrm{D}1}}{\mathbb{P}_{\mathrm{D}1}}\right)\frac{\mathrm{d}x}{\mathrm{dB}_1}.$$

The relationship between $\frac{\mathrm{d}\eta_1(\mathrm{B}_1)}{\mathrm{dB}_1}$ and $\frac{\mathrm{d}\eta_{\mathrm{D}1}(\mathrm{B}_1)}{\mathrm{dB}_1}$ can be studied by comparing the term $\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\bar{\mathrm{N}}_1}{\mathbb{P}_1}\right)$ with $-\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\bar{\mathrm{N}}_{\mathrm{D}1}}{\mathbb{P}_{\mathrm{D}1}}\right)$, whose expressions are given below

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\bar{\mathrm{N}}_1}{\mathbb{P}_1}\right) = -\frac{1.28q\lambda_u\lambda_2}{(\lambda_1+q\lambda_2 x)^3}\left(\lambda_1(1+\rho(\Upsilon,\alpha))+q\lambda_2 x\left(1+\rho\left(\frac{\Upsilon}{x^{\frac{\alpha}{2}}},\alpha\right)\right)\right)$$
$$-\left(1+\frac{1.28\lambda_u}{1+q\lambda_2 x}\right)\frac{q\lambda_1\lambda_2\left(\rho(\Upsilon,\alpha)+\rho\left(\frac{\Upsilon}{x^{\frac{\alpha}{2}}},\alpha\right)+\frac{1}{1+\Upsilon x^{\frac{\alpha}{2}}}\right)+q^2\lambda_2^2 x\left(2\rho\left(\frac{\Upsilon}{x^{\frac{\alpha}{2}}},\alpha\right)+\frac{1}{1+\Upsilon x^{\frac{\alpha}{2}}}\right)}{(\lambda_1+q\lambda_2 x)^2} \tag{A.1}$$

$$\frac{-\mathrm{d}}{\mathrm{d}x}\left(\frac{\bar{\mathrm{N}}_{\mathrm{D}1}}{\mathbb{P}_{\mathrm{D}1}}\right) = -\frac{1.28q\lambda_u\lambda_2\left(\lambda_1+\lambda_2 q\mathrm{B}_2{}^{\frac{2}{\alpha}}\right)(\lambda_1(1+\rho(\Upsilon,\alpha))+q\lambda_2 x)}{(\lambda_1+q\lambda_2 x)^3\left(\lambda_1(1+\rho(\Upsilon,\alpha))+\lambda_2 q\mathrm{B}_2{}^{\frac{2}{\alpha}}\right)}$$
$$+\left(1+\frac{1.28\lambda_u}{\lambda_1+\lambda_2 q\mathrm{B}_2{}^{\frac{2}{\alpha}}}-\frac{1.28\lambda_u}{\lambda_1+q\lambda_2 x}\right)\frac{q\lambda_1\lambda_2\left(\lambda_1+\lambda_2 q\mathrm{B}_2{}^{\frac{2}{\alpha}}\right)\rho(\Upsilon,\alpha)}{(\lambda_1+q\lambda_2 x)^2\lambda_1(1+\rho(\Upsilon,\alpha))+\lambda_2 q\mathrm{B}_2{}^{\frac{2}{\alpha}}}, \tag{A.2}$$

$$\frac{\mathrm{d}\mathbb{P}_{\mathrm{c}}}{\mathrm{dB}_1} = \frac{\mathrm{d}\mathbb{P}_{\mathrm{c}}}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{dB}_1}$$
$$= \left(\frac{\lambda_1\lambda_2 q}{(\lambda_1+(1+\rho(\Upsilon,\alpha))\lambda_2 q x)^2}-\frac{\lambda_1\lambda_2 q x^{\frac{\alpha}{2}}/(\Upsilon+x^{\frac{\alpha}{2}})}{\left(\lambda_1(1+\rho(\Upsilon,\alpha))+\lambda_2 q x\left(1+\rho\left(\frac{\Upsilon}{x^{\frac{\alpha}{2}}},\alpha\right)\right)\right)^2}\right)\frac{\mathrm{d}x}{\mathrm{dB}_1}, \tag{A.3}$$

where $q = \left(\frac{\mathrm{S}_2}{\mathrm{S}_1}\right)^{\frac{2}{\alpha}}$.

It can be easily verified from (A.1) and (A.2) that $\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\bar{\mathrm{N}}_1}{\mathbb{P}_1}\right) < -\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\bar{\mathrm{N}}_{\mathrm{D}1}}{\mathbb{P}_{\mathrm{D}1}}\right)$ for all $\mathrm{B}_1 \geq \mathrm{B}_2$.

Given two functions $f(x)$ and $g(x)$, if the first order derivative of $f(x)$ is always smaller than $g(x)$, i.e., $\frac{\mathrm{d}f(x)}{\mathrm{d}x} < \frac{\mathrm{d}g(x)}{\mathrm{d}x}$, then the two functions have at most one intersection point. According to this result, the function $\eta_{\mathrm{D}1}(\mathrm{B}_1) = \eta_1(\mathrm{B}_1)$ has at most one root.

Appendix A. Proof of Corollary 2.4.1

Based on the above analysis, when $\lim_{x\to\infty}\eta_{\mathrm{D1}}(x) < \lim_{x\to\infty}\eta_1(x)$, we have $\eta_{\mathrm{D1}}(\mathrm{B}_1) < \eta_1(\mathrm{B1})$ for all $\mathrm{B}_1 > \mathrm{B}_2$. Therefore, the set $\mathcal{M}_1$ is empty, which can be interpreted as $b_1 = \infty$. Similarly, when $\lim_{x\to\mathrm{B}_2}\eta_{\mathrm{D1}}(x) \geq \lim_{x\to\mathrm{B}_2}\eta_1(x)$, the constraint $\eta_{\mathrm{D1}}(\mathrm{B}_1) \geq \eta_1(\mathrm{B}_1)$ is achieved for all $\mathrm{B}_1 > \mathrm{B}_2$. Thus, $b_1 = \mathrm{B}_2$. The last case in Corollary 2.4.1 is $\lim_{x\to\mathrm{B}_2}\eta_{\mathrm{D1}}(x) \leq \lim_{x\to\mathrm{B}_2}\eta_1(x)$ and $\lim_{x\to\infty}\eta_{\mathrm{D1}}(x) \geq \lim_{x\to\infty}\eta_1(x)$. In this scenario, there exists an intersection point b_1 between $\eta_{\mathrm{D1}}(\mathrm{B}_1)$ and $\eta_1(\mathrm{B}_1)$. Once the value of B_1 exceeds b_1, the constraint $\eta_{\mathrm{D1}}(\mathrm{B}_1) \geq \eta_1(\mathrm{B}_1)$ can be satisfied. Hence, the feasible set $\mathcal{M}_1$ for this case is $\mathcal{M}_1 = \{\mathrm{B}_1 | \, \mathrm{B}_1 \geq b_1\}$. Because there exists only one intersection point between $\eta_{\mathrm{D1}}(\mathrm{B}_1)$ and $\eta_1(\mathrm{B}_1)$ with their expressions known, bisection method can be applied to find the value of b_1.

Appendix B

Proof of Lemma 2.4.1

Firstly, we calculate the first order derivative of $\mathbb{P}_c$ with respect to B_1. For simplicity of notification, a change of variable $x = (B_1)^{\frac{2}{\alpha}}$ is applied. Then (A.3) is derived.

Note that $\frac{dx}{dB_1} > 0$. Thus, $\frac{d\mathbb{P}_c}{dB_1} = 0$ if there exists B_1 that lies within the set $\mathcal{M}_c$, where

$$\mathcal{M}_c = \left\{ B_1 \left| \frac{1}{\lambda_1 + (1+\rho(\Upsilon,\alpha))\,\lambda_2 \left(\frac{S_2 B_1}{S_1}\right)^{\frac{2}{\alpha}}} \right. \right.$$
$$\left. \left. \frac{(B_1/(\Upsilon+B_1))^{\frac{1}{2}}}{\lambda_1(1+\rho(\Upsilon,\alpha)) + \lambda_2 \left(\frac{S_2 B_1}{S_1}\right)^{\frac{2}{\alpha}} \left(1+\rho\left(\frac{\Upsilon}{B_1},\alpha\right)\right)} \right\}.$$

Then we can verify that $\left.\frac{d^2\mathbb{P}_c}{dB_1^2}\right|_{B_1 \in \mathcal{M}_c} < 0$, which means set $\mathcal{M}_c$ has at most one element. If $\mathcal{M}_c \neq \emptyset$, its element is denoted as B_c. From the above analysis, $\mathbb{P}_c$ is unimodal in this case and achieves maximum value at B_c. If $\mathcal{M}_c$ is empty, $\mathbb{P}_c$ is a monotonic function of B_1.

Appendix C

Proof of Lemma 3.3.2

We start the proof of Lemma 3.3.2 by deriving the expressions (3.13)–(3.15). Due to the similarity, we only elaborate on the derivation of $\mathbb{P}_{\mathrm{f}}(\mathrm{N_f}, d_{i,1})$. From (3.8), $\gamma_i^{\mathrm{f}} = \frac{\mathrm{P_f} H_{i,1} g_{\mathrm{f}}(d_{i,1})}{\mathrm{P_M} H_{i,0} g_{\mathrm{f}}^{\mathrm{M}}(d_{i,0}) + I_{\mathrm{f}} + \sigma^2}$, where $I_{\mathrm{f}} = \sum_{j \in \Phi_{\mathrm{f}} \setminus \{1\}} \mathrm{P_f} H_{i,j} g_{\mathrm{f}}^{\mathrm{f}}(d_{i,j})$. We then have $\mathbb{P}_{\mathrm{f}}(\mathrm{N_f}, d_{i,1})$ as

$$
\begin{aligned}
&\mathbb{P}_{\mathrm{f}}(\mathrm{N_f}, d_{i,1}) \\
=&\mathbb{E}_{d_{i,0}} \left[\Pr\left(\gamma_i^{\mathrm{f}} \geq \Upsilon_{\mathrm{f}} \middle| \, \mathrm{N_f}, d_{i,1}, d_{i,0} \right) \right] \\
=&\mathbb{E}_{d_{i,0}} \left[\mathbb{E}_{H_{i,0}, I_{\mathrm{f}}} \left[\Pr\left(H_{i,1} \geq \frac{\Upsilon_{\mathrm{f}}}{\mathrm{P_f} g_{\mathrm{f}}(d_{i,1})} \left(\sigma^2 + \mathrm{P_M} H_{i,0} g_{\mathrm{f}}^{\mathrm{M}}(d_{i,0}) + I_f \right) \right) \right] \right] \\
=&\exp\left(-\frac{\Upsilon_{\mathrm{f}} \sigma^2}{\mathrm{P_f} g_{\mathrm{f}}(d_{i,1})} \right) \mathbb{E}_{H_{i,0}, d_{i,0}} \left[\exp\left(-\frac{\Upsilon_{\mathrm{f}} \mathrm{P_M} H_{i,0} g_{\mathrm{f}}^{\mathrm{M}}(d_{i,0})}{\mathrm{P_f} g_{\mathrm{f}}(d_{i,1})} \right) \right] \mathbb{E}_{I_f} \left[\exp\left(-\frac{\Upsilon_{\mathrm{f}} I_f}{\mathrm{P_f} g_{\mathrm{f}}(d_{i,1})} \right) \right]
\end{aligned}
\tag{C.1}
$$

Note that $H_{i,1}$ is an exponential random variable with expected value 1. Therefore, the term $\mathbb{E}_{H_{i,0}, d_{i,0}} \left[\exp\left(-\frac{\Upsilon_{\mathrm{f}} \mathrm{P_M} H_{i,0} g_{\mathrm{f}}^{\mathrm{M}}(d_{i,0})}{\mathrm{P_f} g_{\mathrm{f}}(d_{i,1})} \right) \right]$ in the last line of (C.1) equals to $\mathbb{E}_{d_{i,0}} \left[\left(1 + \frac{\Upsilon_{\mathrm{f}} \omega \mathrm{P_M} d_{i,0}^{-\alpha_{\mathrm{o}}}}{\mathrm{P_f} d_{i,1}^{-\alpha_{\mathrm{in}}}} \right)^{-1} \right]$.

To find the expectation $\mathbb{E}_{I_f} \left[\exp\left(-\frac{\Upsilon_{\mathrm{f}} I_f}{\mathrm{P_f} g_{\mathrm{f}}(d_{i,1})} \right) \right]$ in (C.1), we use the method in [26] by applying the Laplace transform of the interference. The derivation is

briefly listed as follows,

$$\begin{aligned}
&\mathbb{E}_{I_f}\left[\exp\left(-\frac{\Upsilon_{\mathrm{f}} I_f}{\mathrm{P_f} g_{\mathrm{f}}(d_{i,1})}\right)\right] \\
=&\mathbb{E}_{\Phi_{\mathrm{f}}\setminus\{1\}}\left[\prod_{j\in\Phi_{\mathrm{f}}\setminus\{1\}}\mathbb{E}_{H_{i,j}}\left[\exp\left(-\frac{\Upsilon_{\mathrm{f}} H_{i,j} g_{\mathrm{f}}^{\mathrm{f}}(d_{i,j})}{g_{\mathrm{f}}(d_{i,1})}\right)\right]\right] \\
=&\mathbb{E}_{\Phi_{\mathrm{f}}\setminus\{1\}}\left[\prod_{j\in\Phi_{\mathrm{f}}\setminus\{1\}}\frac{1}{1+\frac{\Upsilon_{\mathrm{f}}\omega^2 d_{i,j}^{-\alpha_{\mathrm{o}}}}{d_{i,1}^{-\alpha_{\mathrm{in}}}}}\right] \\
=&\exp\left(-2\pi\lambda_{\mathrm{f}}\int_{d_{i,1}}^{\infty}\frac{\omega^2\Upsilon_{\mathrm{f}} v}{\frac{v^{\alpha_{\mathrm{o}}}}{d_{i,1}^{\alpha_{\mathrm{in}}}}+\omega^2\Upsilon_{\mathrm{f}}}\mathrm{d}v\right) \\
=&\exp\left(-\pi\lambda_{\mathrm{f}}\left(\omega^2\Upsilon_{\mathrm{f}}\right)^{\frac{2}{\alpha_{\mathrm{o}}}} d_{i,1}^{\frac{2\alpha_{\mathrm{in}}}{\alpha_{\mathrm{o}}}}\int_{\left(\omega^2\Upsilon_{\mathrm{f}}\right)^{\frac{1}{\alpha_{\mathrm{o}}}} d_{i,1}^{\frac{\alpha_{\mathrm{in}}}{\alpha_{\mathrm{o}}}-1}}^{\infty}\frac{1}{1+u^{\frac{\alpha_{\mathrm{o}}}{2}}}\mathrm{d}u\right)
\end{aligned}$$

Substitute the derived results back into (C.1), $\mathbb{P}_{\mathrm{f}}(\mathrm{N_f}, d_{i,1})$ in (3.14) follows. The expressions of $\mathbb{P}_{\mathrm{M}}(\mathrm{N_M}, d_{i,0})$ in (3.13) and $\mathbb{P}_{\mathrm{D1}}(\mathrm{N_{D1}}, d_{i,1})$ in (3.15) can be obtained similarly.

The key to simplify the expression of $\mathbb{P}_{\mathrm{M}}(\mathrm{N_M}, d_{i,0})$ is to approximate the integral in (3.13). The integral in (3.13) can be lower bounded as

$$\int_x^{\infty}\frac{1}{1+u^{\alpha_{\mathrm{o}}/2}}\mathrm{d}u \geq \int_x^{\infty} u^{-\frac{\alpha_{\mathrm{o}}}{2}}\mathrm{d}u = \frac{2}{\alpha_{\mathrm{o}}-2}x^{1-\frac{\alpha_{\mathrm{o}}}{2}}. \tag{C.2}$$

The approximation is tight when $x \gg 1$. According to the assumptions that $\mathrm{P_M} \gg \mathrm{P_f}$ and Υ_{M} takes small value, we therefore use (C.2) to calculate the integral in (3.13) and obtain the accurate approximation of $\mathbb{P}_{\mathrm{M}}(\mathrm{N_M}, d_{i,0})$ in (3.16).

Note that $\left(1+\frac{\Upsilon_{\mathrm{f}}\mathrm{P_M}\omega d_{i,0}^{-\alpha_{\mathrm{o}}}}{\mathrm{P_f} d_{i,1}^{-\alpha_{\mathrm{in}}}}\right)^{-1} > \mathbb{P}_{\mathrm{f}}(\mathrm{N_f}, d_{i,1})$. With the assumption that $\mathbb{P}_{\mathrm{f}}(\mathrm{N_f}, d_{i,1})$ is close to 1, we can therefore conclude that the term $\frac{\Upsilon_{\mathrm{f}}\mathrm{P_M}\omega d_{i,0}^{-\alpha_{\mathrm{o}}}}{\mathrm{P_f} d_{i,1}^{-\alpha_{\mathrm{in}}}}$ is sufficiently small. Consequently, we can apply the Taylor series approximation $e^x \approx 1+x$ with high accuracy to replace $\left(1+\frac{\Upsilon_{\mathrm{f}}\mathrm{P_M}\omega d_{i,0}^{-\alpha_{\mathrm{o}}}}{\mathrm{P_f} d_{i,1}^{-\alpha_{\mathrm{in}}}}\right)^{-1}$ in (3.14) by $\exp\left(-\frac{\Upsilon_{\mathrm{f}}\mathrm{P_M}\omega d_{i,0}^{-\alpha_{\mathrm{o}}}}{\mathrm{P_f} d_{i,1}^{-\alpha_{\mathrm{in}}}}\right)$. Furthermore, approximating the integral in (3.14) using the

lower bound (C.2), we can then have $\widetilde{\mathbb{P}}_{\mathrm{f}}(\mathrm{N_f}, d_{i,1})$ as

$$\widetilde{\mathbb{P}}_{\mathrm{f}}(\mathrm{N_f}, d_{i,1}) = \mathbb{E}_{d_{i,0}}\left[\exp\left(-\frac{2}{\alpha_{\mathrm{o}}-2}\pi\lambda_{\mathrm{f}}\omega^2\Upsilon_{\mathrm{f}}d_{i,1}^{\alpha_{\mathrm{in}}-\alpha_{\mathrm{o}}+2} - \frac{\Upsilon_{\mathrm{f}}\sigma^2 + \Upsilon_{\mathrm{f}}\omega\mathrm{P_M}Kd_{i,0}^{-\alpha_{\mathrm{o}}}}{\mathrm{P_f}Kd_{i,1}^{-\alpha_{\mathrm{in}}}}\right)\right] \tag{C.3}$$

To calculate the expectation in (C.3), the distribution of $d_{i,0}$ is required. Note that user i is assumed to be in set $\mathcal{U}_{\mathrm{f}}$ when calculating $\mathbb{P}_{\mathrm{f}}(\mathrm{N_f}, d_{i,1})$ and the distance between user i and its serving fBS is small (less than $\mathrm{R_f}$). We therefore approximate $d_{i,0}$ by the distance between user i's serving fBS, i.e. fBS 1, and the MBS. Let the distance between fBS 1 and MBS be D. Based on the assumption that fBSs follow a PPP, the PDF of D is

$$f_D(d) = \frac{2d}{\mathrm{R_M^2}} \tag{C.4}$$

with $D \in [0, \mathrm{R_M}]$. Using (C.4), we have (C.3) approximated as

$$\begin{aligned}\widetilde{\mathbb{P}}_{\mathrm{f}}(\mathrm{N_f}, d_{i,1}) =& \int_0^{\mathrm{R_M}} \exp\left(-\frac{2}{\alpha_{\mathrm{o}}-2}\pi\lambda_{\mathrm{f}}\omega^2\Upsilon_{\mathrm{f}}d_{i,1}^{\alpha_{\mathrm{in}}-\alpha_{\mathrm{o}}+2} - \frac{\Upsilon_{\mathrm{f}}\sigma^2 + \Upsilon_{\mathrm{f}}\omega\mathrm{P_M}Ky^{-\alpha_{\mathrm{o}}}}{\mathrm{P_f}Kd_{i,1}^{-\alpha_{\mathrm{in}}}}\right)\frac{2y}{\mathrm{R_M^2}}\mathrm{d}y \\ =& \frac{2}{\alpha}\left(\frac{\Upsilon_{\mathrm{f}}\omega\mathrm{P_M}\mathrm{R_M^{-\alpha_o}}}{\mathrm{P_f}d_{i,1}^{-\alpha_{\mathrm{in}}}}\right)^{\frac{2}{\alpha_{\mathrm{o}}}}\exp\left(-\frac{2}{\alpha_{\mathrm{o}}-2}\pi\lambda_{\mathrm{f}}\omega^2\Upsilon_{\mathrm{f}}d_{i,1}^{\alpha_{\mathrm{in}}-\alpha_{\mathrm{o}}+2} - \frac{\Upsilon_{\mathrm{f}}\sigma^2}{\mathrm{P_f}Kd_{i,1}^{-\alpha_{\mathrm{in}}}}\right) \\ &\times\Gamma\left(-\frac{2}{\alpha_{\mathrm{o}}}, \frac{\Upsilon_{\mathrm{f}}\omega\mathrm{P_M}\mathrm{R_M^{-\alpha_o}}}{\mathrm{P_f}d_{i,1}^{-\alpha_{\mathrm{in}}}}\right)\end{aligned}$$

Replacing the integral in (3.15) by the lower bound (C.2), the expression of $\widetilde{\mathbb{P}}_{\mathrm{D1}}(\mathrm{N_{D1}}, d_{i,1})$ in (3.18) follows.

Appendix D

Proof of Lemma 3.4.1

Note that the MBS serves at most $\mathrm{N_M^{max}}$ users in $\mathcal{U}_\mathrm{M}$, and the MBS cell load $\mathrm{N_M}$ follows a Poisson distribution with mean $\bar{\mathrm{N}}_\mathrm{M}$. Additionally, the expression of $L(\mathrm{N_M}, \mathrm{R_o}, \eta)$ are given in (3.34). Consequently, we have $f(\mathrm{R_o}, \eta)$ in (3.35) as

$$\begin{aligned} f(\mathrm{R_o}, \eta) =& \eta \left(\sum_{n=0}^{\mathrm{N_M^{max}}} \left(2^{\frac{nC_\mathrm{M}}{\eta W}} - 1 \right) \frac{\left(\bar{\mathrm{N}}_\mathrm{M}\right)^n}{n!} e^{-\bar{\mathrm{N}}_\mathrm{M}} + \sum_{n=\mathrm{N_M^{max}}+1}^{\infty} \left(2^{\frac{\mathrm{N_M^{max}} C_\mathrm{M}}{\eta W}} - 1 \right) \frac{\left(\bar{\mathrm{N}}_\mathrm{M}\right)^n}{n!} e^{-\bar{\mathrm{N}}_\mathrm{M}} \right) \\ & \times \left(\frac{2\pi\lambda_\mathrm{f} \mathrm{R_o^{2-\alpha_o}} \omega \mathrm{P_f}}{(\alpha_\mathrm{o}-2)\mathrm{R_M^{-\alpha_o}} \ln\frac{1}{\theta}} + \frac{\sigma^2}{K\mathrm{R_M^{-\alpha_o}} \ln\frac{1}{\theta}} \right). \end{aligned} \tag{D.1}$$

The second summation in (D.1) is upper bounded as

$$\sum_{n=\mathrm{N_M^{max}}+1}^{\infty} \left(2^{\frac{\mathrm{N_M^{max}} C_\mathrm{M}}{\eta W}} - 1 \right) \frac{\left(\bar{\mathrm{N}}_\mathrm{M}\right)^n}{n!} e^{-\bar{\mathrm{N}}_\mathrm{M}} < \sum_{n=\mathrm{N_M^{max}}+1}^{\infty} \left(2^{\frac{nC_\mathrm{M}}{\eta W}} - 1 \right) \frac{\left(\bar{\mathrm{N}}_\mathrm{M}\right)^n}{n!} e^{-\bar{\mathrm{N}}_\mathrm{M}}. \tag{D.2}$$

Under the assumption $\mathrm{N_M^{max}} \gg \bar{\mathrm{N}}_\mathrm{M}$, the probability of cell load $\mathrm{N_M}$ exceeding $\mathrm{N_M^{max}}$ is small. Thus, the approximation (D.2) is tight.

Using the upper bound approximation (D.2), we have $\hat{f}(\mathrm{R_o}, \eta)$ derived as

$$\begin{aligned} \hat{f}(\mathrm{R_o}, \eta) =& \eta \left(\sum_{n=0}^{\infty} \left(2^{\frac{nC_\mathrm{M}}{\eta W}} - 1 \right) \frac{\left(\bar{\mathrm{N}}_\mathrm{M}\right)^n}{n!} e^{-\bar{\mathrm{N}}_\mathrm{M}} \right) \left(\frac{2\pi\lambda_\mathrm{f} \mathrm{R_o^{2-\alpha_o}} \omega \mathrm{P_f}}{(\alpha_\mathrm{o}-2)\mathrm{R_M^{-\alpha_o}} \ln\frac{1}{\theta}} + \frac{\sigma^2}{K\mathrm{R_M^{-\alpha_o}} \ln\frac{1}{\theta}} \right) \\ =& \eta \left(\exp\left(\left(2^{\frac{C_\mathrm{M}}{\eta W}} - 1 \right) \bar{\mathrm{N}}_\mathrm{M} \right) - 1 \right) \left(\frac{2\pi\lambda_\mathrm{f} \mathrm{R_o^{2-\alpha_o}} \omega \mathrm{P_f}}{(\alpha_\mathrm{o}-2)\mathrm{R_M^{-\alpha_o}} \ln\frac{1}{\theta}} + \frac{\sigma^2}{K\mathrm{R_M^{-\alpha_o}} \ln\frac{1}{\theta}} \right) \end{aligned} \tag{D.3}$$

Substituting the expression of $\bar{\mathrm{N}}_\mathrm{M}$ in Lemma 3.3.1 into (D.3), we can get (3.36).

Appendix E

Proof of Lemma 4.3.2

According to the definition of $\hat{\Upsilon}_{i,b}^{\mathrm{mBS,IN}}(d_{i,b})$ in (4.11), the key is to determine the expected values of $\mathbb{E}\left[S_{i,b}^{\mathrm{mBS,IN}}\middle|\,\beta_{i,b}=Kd_{i,b}^{-\alpha}\right]$, $\mathbb{E}\left[I_{i,\mathrm{intra}}^{\mathrm{mBS,IN}}\middle|\,\beta_{i,b}=Kd_{i,b}^{-\alpha}\right]$, and $\mathbb{E}\left[I_{i,\mathrm{inter}}^{\mathrm{MBS,ZF}}\middle|\,\beta_{i,b}=Kd_{i,b}^{-\alpha}\right]$.

Note that $S_{i,b}^{\mathrm{mBS,IN}}=\mathrm{P_m}\beta_{i,b}\left\|\boldsymbol{h}_{i,b}^{H}\boldsymbol{w}_b\right\|^2$. Given the condition that mBS b detects $\mathrm{N_{det}}=n$ user requests, the term $\left\|\boldsymbol{h}_{i,b}^{H}\boldsymbol{w}_b\right\|^2$ in $S_{i,b}^{\mathrm{mBS,IN}}$ follows a Gamma distribution with expected value $\max\{1,\mathrm{N_m}-n\}$. According to Assumption 2, the probability of $\mathrm{N_{det}}>\mathrm{N_m}$ is negligible. As a result, we can further approximate the expected value of $\left\|\boldsymbol{h}_{i,b}^{H}\boldsymbol{w}_b\right\|^2$ conditioned on $\mathrm{N_{det}}=n$ as $\mathrm{N_m}-n$. Then use the result in Lemma 4.3.1 to average over $\mathrm{N_{det}}$, we have

$$\mathbb{E}\left[S_{i,b}^{\mathrm{mBS,IN}}\middle|\,\beta_{i,b}=Kd_{i,b}^{-\alpha}\right]\approx\widetilde{S}_{i,b}^{\mathrm{mBS,IN}}(d_{i,b})=\left(\mathrm{N_m}-\lambda_{\mathrm{m}}\pi\left(\mathrm{R_{det}^2}-\frac{1}{2}\mathrm{R_m^2}\right)\right)K\mathrm{P_m}d_{i,b}^{-\alpha}\tag{E.1}$$

According to the proposed base station deployment scheme proposed in Section 4.2, mBSs within the coverage areas of MBSs are muted. Therefore, the positions of mBSs follow a Poisson hole process (PHP). However, it is challenging to conduct tractable interference analysis with nodes modeled by PHP. Thus, we relaxed the PHP assumption by assuming all mBSs are active, i.e. mBSs follow PPP Φ_{m} with intensity λ_{m}. The expected value of intra-tier interference $\mathbb{E}\left[I_{i,\mathrm{intra}}^{\mathrm{mBS,IN}}\middle|\,\beta_{i,b}=Kd_{i,b}^{-\alpha}\right]$ can be approximated as

$$\mathbb{E}\left[I_{i,\mathrm{intra}}^{\mathrm{mBS,IN}}\middle|\,\beta_{i,b}=Kd_{i,b}^{-\alpha}\right]\approx\widetilde{I}_{i,\mathrm{intra}}^{\mathrm{mBS,IN}}=\mathbb{E}\left[\sum_{l\in\Phi_{\mathrm{m}}\setminus\{b\}}\mathrm{P_m}\left\|\boldsymbol{f}_{i,l}^{H}\boldsymbol{w}_l\right\|^2\right].\tag{E.2}$$

Appendix E. Proof of Lemma 4.3.2

From Assumption 2, mBSs within distance $\mathrm{R}_{\mathrm{det}}$ to user i will completely eliminate their interference through IN-BF. Furthermore, for mBS l doesn't detect user i, $\boldsymbol{f}_{i,l}$ and $\boldsymbol{w}_l$ are mutually independent random variables. Thus, the right hand side of (E.2) is calculated as

$$\mathbb{E}\left[\sum_{l\in\Phi_{\mathrm{m}}\setminus\{b\}} \mathrm{P}_{\mathrm{m}} \left\|\boldsymbol{f}_{i,l}^{H}\boldsymbol{w}_l\right\|^2\right] = \int_{\mathrm{R}_{\mathrm{det}}}^{\infty} 2\pi\lambda_{\mathrm{m}} K\mathrm{P}_{\mathrm{m}} l^{1-\alpha}\mathrm{d}l = \frac{2\pi\lambda_{\mathrm{m}} K\mathrm{P}_{\mathrm{m}}}{\alpha-2}\left(\mathrm{R}_{\mathrm{det}}\right)^{2-\alpha}. \quad \text{(E.3)}$$

In Assumption 1, we assume that MBSs follow a PPP with intensity $\widetilde{\lambda}_{\mathrm{M}}$. The inter-tier interference term $\mathbb{E}\left[I_{i,\mathrm{inter}}^{\mathrm{MBS,ZF}}\middle|\,\beta_{i,b}=Kd_{i,b}^{-\alpha}\right]$ is approximated by $\widetilde{I}_{i,\mathrm{inter}}^{\mathrm{MBS,ZF}}$:

$$\widetilde{I}_{i,\mathrm{inter}}^{\mathrm{MBS,ZF}} = \int_{\mathrm{R}_{\mathrm{M}}}^{\infty} 2\pi\widetilde{\lambda}_{\mathrm{M}} K\mathrm{P}_{\mathrm{M}} l^{1-\alpha}\mathrm{d}l = \frac{2\pi\widetilde{\lambda}_{\mathrm{M}} K\mathrm{P}_{\mathrm{M}}}{\alpha-2}\left(\mathrm{R}_{\mathrm{M}}\right)^{2-\alpha}. \quad \text{(E.4)}$$

Appendix F

Proof of Lemma 4.3.3

Similar as the Proof of Lemma 4.3.2, the virtual SINR for user j associated with MBS p can be obtained by calculating the expected values of signal and interference terms: $\mathbb{E}\left[S_{j,p}^{\mathrm{MBS,ZF}}\middle|\,\beta_{j,p}=Kd_{j,p}^{-\alpha}\right]$, $\mathbb{E}\left[I_{j,\mathrm{intra}}^{\mathrm{MBS,ZF}}\middle|\,\beta_{j,p}=Kd_{j,p}^{-\alpha}\right]$, and $\mathbb{E}\left[I_{j,\mathrm{inter}}^{\mathrm{mBS,IN}}\middle|\,\beta_{j,p}=Kd_{j,p}^{-\alpha}\right]$.

Let the total number of users served by MBS p be n. According to [94], the term $S_{j,p}^{\mathrm{MBS,ZF}}=\frac{\mathrm{P_M}}{n}\beta_{j,p}\left\|\boldsymbol{h}_{j,p}^{H}\boldsymbol{v}_{j,p}\right\|^2$ in (4.23) is a Gamma random variable with mean $\frac{\mathrm{P_M}}{n}\beta_{j,p}\left(\mathrm{N_M}-n+1\right)$. Replacing n by its expected value, which equals to $\lambda_{\mathrm{m}}\pi\mathrm{R_M^2}$, we have

$$\mathbb{E}\left[S_{j,p}^{\mathrm{MBS,ZF}}\middle|\,\beta_{j,p}=Kd_{j,p}^{-\alpha}\right]\approx\widetilde{S}_{j,p}^{\mathrm{MBS,ZF}}\left(d_{j,p}\right)=\left(\lambda_{\mathrm{m}}\pi\mathrm{R_M^2}\left(\mathrm{N_m}-1\right)+1\right)\frac{K\mathrm{P_M}}{\lambda_{\mathrm{m}}\pi\mathrm{R_M^2}}d_{j,p}^{-\alpha}\tag{F.1}$$

Note that the MBSs are separated at least $2\mathrm{R_M}$ away from each other. Therefore, as depicted in Fig. F.1, the interfering MBSs are all located out of the big dashed circle with radius $2\mathrm{R_M}$. Let $x(\theta)$ be the distance from user j to the points on the big dashed circle at angle θ. As shown in Fig. F.1, the value of $x(\theta)$ is a function of $d_{j,p}$, $\mathrm{R_M}$ and θ. Specifically, using the law of cosine, we have [117]

$$x(\theta)=d_{j,p}\cos\theta+\sqrt{4\mathrm{R_M^2}-d_{j,p}^2\sin^2\theta}.\tag{F.2}$$

According to Assumption 1, the positions of MBSs out of the big dashed circle are modelled as HPPP with intensity $\widetilde{\lambda}_{\mathrm{M}}$. Thus, we can approximately calculate

Appendix F. Proof of Lemma 4.3.3

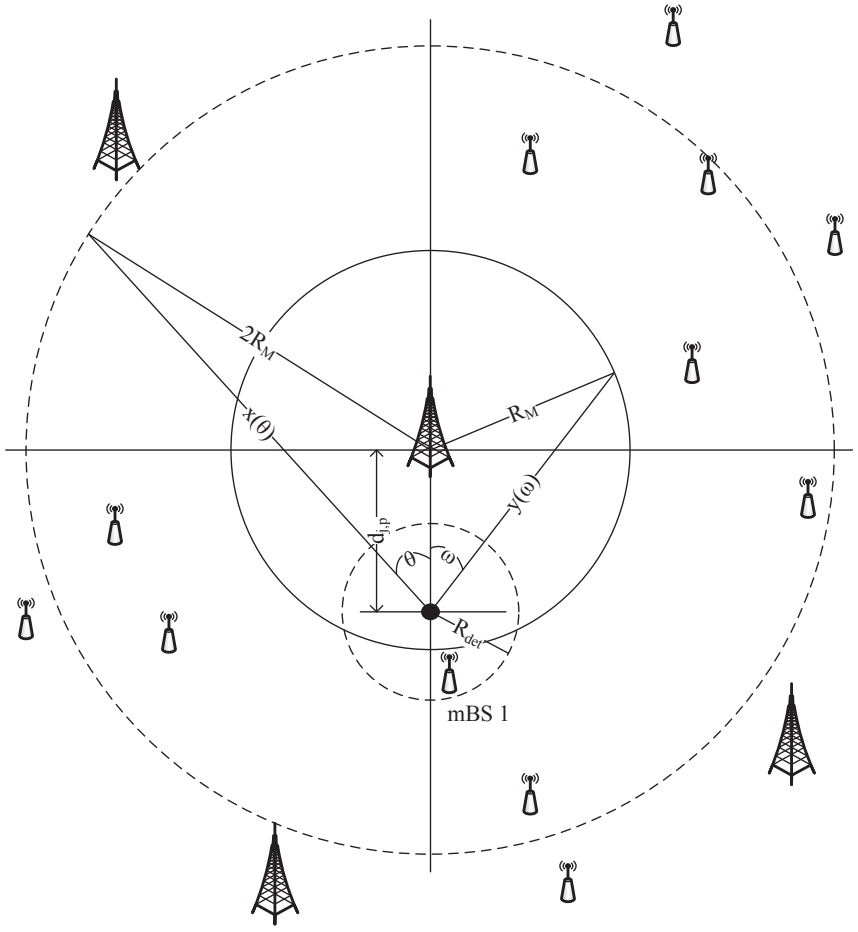


Figure F.1: The network layout. The boundary of the considered macrocell is depicted as a circle with solid line.

$\mathbb{E}\left[I_{j,\text{intra}}^{\text{MBS,ZF}}\middle|\,\beta_{j,p}=Kd_{j,p}^{-\alpha}\right]$ as

$$\mathbb{E}\left[I_{j,\text{intra}}^{\text{MBS,ZF}}\middle|\,\beta_{j,p}=Kd_{j,p}^{-\alpha}\right]\approx\widetilde{I}_{j,\text{intra}}^{\text{MBS,ZF}}\left(d_{j,p}\right)=\int_{0}^{2\pi}\int_{x(\theta)}^{\infty}\widetilde{\lambda}_{\text{M}}K\text{P}_{\text{M}}l^{1-\alpha}\text{d}l\text{d}\theta$$
$$=\int_{0}^{2\pi}\frac{K\text{P}_{\text{M}}\widetilde{\lambda}_{\text{M}}}{\alpha-2}\left(d_{j,p}\cos\theta+\sqrt{4\text{R}_{\text{M}}^{2}-d_{j,p}^{2}\sin^{2}\theta}\right)^{2-\alpha}\text{d}\theta.$$

For the inter-tier interference analysis, we first approximate the locations of mBSs out of the considered MBS' coverage area as a PPP with intensity λ_{m}. Next, we can characterize the distance, denoted as $y(\omega)$, from user j to its associated

MBS's cell edge at angle ω. As shown in Fig. F.1, we have

$$y(\omega) = d_{j,p}\cos\omega + \sqrt{\mathrm{R}_{\mathrm{M}}^2 - d_{j,p}^2\sin^2\omega}. \tag{F.3}$$

In Assumption 2, we have assumed that a mBS eliminates its interference to all users located within a distance $\mathrm{R}_{\mathrm{det}}$. For example, in Fig. F.1, mBS 1 will not cause interference to user j. Correspondingly, we should only consider the mBSs at $y(\omega) \geq \mathrm{R}_{\mathrm{det}}$ when calculating inter-tier interference term $\mathbb{E}\left[I_{j,\mathrm{inter}}^{\mathrm{mBS,IN}}\middle|\,\beta_{j,p} = Kd_{j,p}^{-\alpha}\right]$.

$$\begin{aligned}
&\mathbb{E}\left[I_{j,\mathrm{inter}}^{\mathrm{mBS,IN}}\middle|\,\beta_{j,p} = Kd_{j,p}^{-\alpha}\right]\\
\approx&\widetilde{I}_{j,\mathrm{inter}}^{\mathrm{mBS,IN}}\left(d_{j,p}\right)\\
=&\int_0^{2\pi}\int_{\max\{\mathrm{R}_{\mathrm{det}},y(\omega)\}}^{\infty}\lambda_{\mathrm{m}}K\mathrm{P}_{\mathrm{m}}l^{1-\alpha}\mathrm{d}l\mathrm{d}\omega\\
=&\int_0^{2\pi}\frac{K\mathrm{P}_{\mathrm{m}}\lambda_{\mathrm{m}}}{\alpha-2}\left(\max\left\{\mathrm{R}_{\mathrm{det}}, d_{j,p}\cos\omega + \sqrt{\mathrm{R}_{\mathrm{M}}^2 - d_{j,p}^2\sin^2\omega}\right\}\right)^{2-\alpha}\mathrm{d}\omega.
\end{aligned}$$

Appendix G

Proof of Lemma 4.4.1

We define the set $\mathcal{X} = \{x | U_{\rm c}(x) = U_{\rm M}(x)\}$. To prove Lemma 4.4.1, we resort to calculating the derivative of $g(x) = U_{\rm M}(x) - U_{\rm c}(x)$ at all $x \in \mathcal{X}$.

Using the expressions in (4.39), we have $g'(x) = \frac{{\rm d}g(x)}{{\rm d}x}$ as

$$g'(x) = \left[-\frac{(2\lambda_{\rm m}\pi x^2(\mathrm{N_m}-1)+4)K\mathrm{P_M}}{\Upsilon_0\lambda_{\rm m}\pi x^5} - \frac{2K\mathrm{P_m}\lambda_{\rm m}x^{\alpha-3}}{(\mathrm{R_{det}})^{\alpha-2}}\left(\pi - \arccos\left(\frac{\mathrm{R_{det}}}{2x}\right)\right) - (\alpha-2)\sigma^2 x^{\alpha-3}\right] \times \frac{(\alpha-2)}{K\mathrm{P_M}\int_0^{2\pi}\left(\cos\theta + \sqrt{4-\sin^2\theta}\right)^{2-\alpha}{\rm d}\theta} + \frac{2}{\pi x^3}. \quad \text{(G.1)}$$

From $U_{\rm c}(x) = U_{\rm M}(x)$, we have

$$\frac{(\lambda_{\rm m}\pi x^2(\mathrm{N_m}-1)+1)K\mathrm{P_M}}{\Upsilon_0\lambda_{\rm m}\pi x^4} = \frac{K\mathrm{P_M}\int_0^{2\pi}\left(\cos\theta+\sqrt{4-\sin^2\theta}\right)^{2-\alpha}{\rm d}\theta}{(\alpha-2)} \times \left(\frac{1}{\pi x^2} + \frac{2K\mathrm{P_m}\lambda_{\rm m}x^{\alpha-2}}{(\alpha-2)(\mathrm{R_{det}})^{\alpha-2}}\left(\pi - \arccos\left(\frac{\mathrm{R_{det}}}{2x}\right)\right) + \frac{2K\mathrm{P_m}\lambda_{\rm m}}{\alpha-2}\int_0^{\arccos\left(\frac{\mathrm{R_{det}}}{2x}\right)}(2\cos\theta)^{2-\alpha}{\rm d}\theta + \sigma^2 x^{\alpha-2}\right). \quad \text{(G.2)}$$

Thus, for all $x \in \mathcal{X}$, we have $g'(x) < 0$, which means there is at most one element in $\mathcal{X}$. Denote the single element in $\mathcal{X}$ as $\mathrm{R_2}$. From $g'(\mathrm{R_2}) < 0$, we have $U_{\rm c}(\mathrm{R_M}) \leq U_{\rm M}(\mathrm{R_M})$ for $\mathrm{R_M} \leq \mathrm{R_2}$.

References

[1] E. Oh, B. Krishnamachari, X. Liu, and Z. Niu, "Toward dynamic energy-efficient operation of cellular network infrastructure," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 56–61, Jun. 2011.

[2] Ofcom, "Sitefinder: Mobile phone base station database," Available online at: http://www.sitefinder.ofcom.org.uk/.

[3] M. Alouini and A. Goldsmith, "Area spectral efficiency of cellular mobile radio systems," *IEEE Trans. Veh. Technol.*, vol. 48, no. 4, pp. 1047–1066, Jul. 1999.

[4] V. Chandrasekhar, J. Andrews, and A. Gatherer, "Femtocell networks: a survey," *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 59–67, Sep. 2008.

[5] P. Lin, J. Zhang, Y. Chen, and Q. Zhang, "Macro-femto heterogeneous network deployment and management: from business models to technical solutions," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 64–70, Jun. 2011.

[6] 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 Commun.*, vol. 18, no. 3, pp. 10–21, Jun. 2011.

[7] I. Hwang, B. Song, and S. Soliman, "A holistic view on hyper-dense heterogeneous and small cell networks," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 20–27, Jun. 2013.

[8] P. Demestichas, A. Georgakopoulos, D. Karvounas, K. Tsagkaris, V. Stavroulaki, J. Lu, C. Xiong, and J. Yao, "5G on the horizon: Key challenges for the radio-access network," *IEEE Veh. Technol. Mag.*, vol. 8, no. 3, pp. 47–53, Sep. 2013.

[9] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 122–130, Feb. 2014.

[10] 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 J. Select. Areas in Commun.*, vol. 27, no. 7, pp. 1029–1046, Sep. 2009.

[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 Commun. Surveys & Tutorials*, vol. 15, no. 3, pp. 996–1019, 3rd. 2013.

[12] A. Ghosh, N. Mangalvedhe, R. Ratasuk, B. Mondal, M. Cudak, E. Visotsky, T. Thomas, J. Andrews, P. Xia, H. Jo, H. Dhillon, and T. Novlan, "Heterogeneous cellular networks: From theory to practice," *IEEE Commun. Mag.*, vol. 50, no. 6, pp. 54–64, Jun. 2012.

[13] M. Franceschetti and R. Meester, *Random Networks for Communication: From Statistical Physics to Information Systems*. Cambridge University Press, 2007.

[14] M. Haenggi and R. K. Ganti, *Interference in Large Wireless Networks*. NOW: Foundations and Trends in Networking, 2008.

[15] F. Baccelli and B. Blaszczyszyn, *Stochastic Geometry and Wireless Networks*. NOW: Foundations and Trends in Networking, 2010.

[16] The Climate Group and Global e-Sustainability Initiative (GeSI), "Smart 2020: Enabling the low carbon economy in the information age," Available online at: http://www.smart2020.org, 2008.

[17] J. Hoydis, M. Kobayashi, and M. Debbah, "Green small-cell networks," *IEEE Veh. Technol. Mag.*, vol. 6, no. 1, pp. 37–43, Mar. 2011.

[18] F. Boccardi, R. Heath, A. Lozano, T. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[19] X. Wang, A. V. Vasilakos, M. Chen, and T. T. Kwon, "A survey of green mobile networks: opportunities and challenges," *ACM/Springer J. Mobile Networks and Applications*, vol. 17, no. 1, pp. 4–20, Feb. 2012.

[20] Y. Chen, S. Zhang, S. Xu, and G. Li, "Fundamental trade-offs on green wireless networks," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 30–37, Jun. 2011.

[21] A. Wyner, "Shannon-theoretic approach to a gaussian cellular multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1713–1727, Nov. 1994.

[22] J. Xu, J. Zhang, and J. Andrews, "On the accuracy of the wyner model in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3098–3109, Sep. 2011.

[23] A. Viterbi, A. Viterbi, and E. Zehavi, "Other-cell interference in cellular power-controlled CDMA," *IEEE Trans. Commun.*, vol. 42, no. 234, pp. 1501–1504, Feb. 1994.

[24] C.-B. Chae, I. Hwang, R. W. Heath, and V. Tarokh, "Interference aware-coordinated beamforming system in a two-cell environment," Available online at: http://nrs.harvard.edu/urn-3:HUL.InstRepos:3293263, Sep. 2009.

[25] M. Haenggi, *Stochastic Geometry for Wireless Networks*. Cambridge University Press, 2012.

[26] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122–3134, Nov. 2011.

[27] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and analysis of k-tier downlink heterogeneous cellular networks," *IEEE J. Select. Areas in Commun.*, vol. 30, no. 3, pp. 550–560, Apr. 2012.

[28] H.-S. Jo, Y. J. Sang, P. Xia, and J. Andrews, "Heterogeneous cellular networks with flexible cell association: A comprehensive downlink sinr analysis," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3484–3495, Oct. 2012.

[29] R. Heath, M. Kountouris, and T. Bai, "Modeling heterogeneous network interference using poisson point processes," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4114–4126, Aug. 2013.

[30] S. Mukherjee, "Distribution of downlink SINR in heterogeneous cellular networks," *IEEE J. Select. Areas in Commun.*, vol. 30, no. 3, pp. 575–585, Apr. 2012.

[31] C. de Lima, M. Bennis, and M. Latva-aho, "Coordination mechanisms for self-organizing femtocells in two-tier coexistence scenarios," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2212–2223, Jun. 2012.

[32] K. Huang, V. Lau, and Y. Chen, "Spectrum sharing between cellular and mobile ad hoc networks: transmission-capacity trade-off," *IEEE J. Select. Areas in Commun.*, vol. 27, no. 7, pp. 1256–1267, Sep. 2009.

[33] V. Chandrasekhar and J. Andrews, "Spectrum allocation in tiered cellular networks," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3059–3068, Oct. 2009.

[34] T. D. Novlan, R. K. Ganti, A. Ghosh, and J. G. Andrews, "Analytical evaluation of fractional frequency reuse for OFDMA cellular networks," *IEEE Wireless Commun.*, vol. 10, no. 12, pp. 4294 –4305, Dec. 2011.

[35] S. Z. D. Cao and Z. Niu, "Improving the energy efficiency of two-tier heterogeneous cellular networks through partial spectrum reuse," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 4129 –4141, Aug. 2013.

[36] Y. Lin, W. Bao, W. Yu, and B. Liang, "Optimizing user association and spectrum allocation in HetNets: A utility perspective," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1025–1039, Jun. 2015.

[37] X. Zhang and M. Haenggi, "A stochastic geometry analysis of inter-cell interference coordination and intra-cell diversity," *IEEE Trans. Wireless Commun.*, vol. 13, no. 12, pp. 6655–6669, Dec. 2014.

[38] N. Deng, W. Zhou, and M. Haenggi, "Heterogeneous cellular network models with dependence," to appear in *IEEE J. Select. Areas Commun.*

[39] K. Huang and J. Andrews, "An analytical framework for multicell cooperation via stochastic geometry and large deviations," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2501–2516, Apr. 2013.

[40] N. Lee, D. Morales-Jimenez, A. Lozano, and R. Heath, "Spectral efficiency of dynamic coordinated beamforming: A stochastic geometry approach," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 230–241, Jan. 2015.

[41] Qualcomm Incorporated, "LTE advanced: Heterogeneous networks," 2010.

[42] S. Singh and J. G. Andrews, "Joint resource partitioning and offloading in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 888–901, Feb. 2014.

[43] A. Liu, V. Lau, L. Ruan, J. Chen, and D. Xiao, "Hierarchical radio resource optimization for heterogeneous networks with enhanced inter-cell interference coordination (eICIC)," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1684–1693, Apr. 2014.

[44] K. Pedersen, B. Soret, S. Barcos, G. Pocovi, and H. Wang, "Dynamic enhanced inter-cell interference coordination for realistic networks," to appear in *IEEE Trans. Veh. Technol.*

[45] S. Vasudevan, R. Pupala, and K. Sivanesan, "Dynamic eICIC: A proactive strategy for improving spectral efficiencies of heterogeneous lte cellular networks by leveraging user mobility and traffic dynamics," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 4956–4969, Oct. 2013.

[46] S. Deb, P. Monogioudis, J. Miernik, and J. P. Seymour, "Algorithms for enhanced inter-cell interference coordination (eICIC) in LTE HetNets," *IEEE/ACM Trans. Netw.*, vol. 22, no. 1, pp. 137–150, Feb. 2014.

[47] A. Weber and O. Stanze, "Scheduling strategies for hetnets using eICIC," in *Communications (ICC), 2012 IEEE International Conference on*, Jun. 2012.

[48] O. Arnold, F. Richter, G. P. Fettweis, and O. Blume, "Power consumption modeling of different base station types in heterogeneous cellular networks," in *Proc. 19th Future Network & MobileSummit.*

[49] Y.-F. Liu, Y.-H. Dai, and Z.-Q. Luo, "Joint power and admission control via linear programming deflation," *IEEE Trans. Signal Process.*, vol. 61, no. 6, pp. 1327–1338, Mar. 2013.

[50] C. Xiong, G. Li, S. Zhang, Y. Chen, and S. Xu, "Energy- and spectral-efficiency tradeoff in downlink OFDMA networks," *Wireless Communications, IEEE Transactions on*, vol. 10, no. 11, pp. 3874–3886, Nov. 2011.

[51] L. Xiang, X. Ge, C.-X. Wang, F. Y. Li, and F. Reichert, "Energy efficiency evaluation of cellular networks based on spatial distributions of traffic load and power consumption," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 961–973, Mar. 2013.

[52] 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 J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1525–1536, Sep. 2011.

[53] A. Abdelnasser, E. Hossain, and D. I. Kim, "Tier-aware resource allocation in OFDMA macrocell-small cell networks," to appear in *IEEE Trans. Commun.*

[54] D. Lopez-Perez, X. Chu, A. Vasilakos, and H. Claussen, "Power minimization based resource allocation for interference mitigation in OFDMA femtocell networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 2, pp. 333–344, Feb. 2014.

[55] K.-Y. Wang, N. Jacklin, Z. Ding, and C.-Y. Chi, "Robust MISO transmit optimization under outage-based QoS constraints in two-tier heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1883–1897, Apr. 2013.

[56] Z. Niu, Y. Wu, J. Gong, and Z. Yang, "Cell zooming for cost-efficient green cellular networks," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 74 –79, Nov. 2010.

[57] L. Saker, S. Elayoubi, R. Combes, and T. Chahed, "Optimal control of wake up mechanisms of femtocells in heterogeneous networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 664–672, Apr. 2012.

[58] E. Oh, K. Son, and B. Krishnamachari, "Dynamic base station switching-on/off strategies for green cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2126–2136, May 2013.

[59] D. Cao, S. Zhou, and Z. Niu, "Optimal combination of base station densities for energy-efficient two-tier heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4350–4362, Sep. 2013.

[60] Y. S. Soh, T. Quek, M. Kountouris, and H. Shin, "Energy efficient heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 840–850, May 2013.

[61] W. C. Cheung, T. Quek, and M. Kountouris, "Throughput optimization, spectrum allocation, and access control in two-tier femtocell networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 561–574, Apr. 2012.

[62] S. Luo, R. Zhang, and T. J. Lim, "Optimal power and range adaptation for green broadcasting," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4592–4603, Sep. 2013.

[63] S. Singh, H. S. Dhillon, and J. G. Andrews, "Offloading in heterogeneous networks: Modeling, analysis, and design insights," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2484–2497, May 2013.

[64] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. Andrews, "User association for load balancing in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2706–2716, Jun. 2013.

[65] S. Bashar and Z. Ding, "Admission control and resource allcoation in a heterogeneous ofdma wireless network," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4200 –4210, Aug. 2009.

[66] J. Rao and A. Fapojuwo, "A survey of energy efficient resource management techniques for multicell cellular networks," *IEEE Commun. Surveys & Tutorials*, vol. 16, no. 1, pp. 154–180, First 2014.

[67] 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 Commun. Mag.*, vol. 51, no. 2, pp. 98–105, Feb. 2013.

[68] H. ElSawy, E. Hossain, and D. I. Kim, "HetNets with cognitive small cells: user offloading and distributed channel access techniques," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 28–36, Jun. 2013.

[69] D. Lopez-Perez, I. Guvenc, G. de la Roche, M. Kountouris, T. Quek, and J. Zhang, "Enhanced intercell interference coordination challenges in heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 22–30, Jun. 2011.

[70] I. Guvenc, "Capacity and fairness analysis of heterogeneous networks with range expansion and interference coordination," *IEEE Commun. Letters*, vol. 15, no. 10, pp. 1084–1087, Oct. 2011.

[71] M. Vajapeyam, A. Damnjanovic, J. Montojo, T. Ji, Y. Wei, and D. Malladi, "Downlink ftp performance of heterogeneous networks for LTE-Advanced," in

Communications Workshops (ICC), 2011 IEEE International Conference on, Jun. 2011, pp. 1–5.

[72] Y. Wang and K. Pedersen, "Performance analysis of enhanced inter-cell interference coordination in lte-advanced heterogeneous networks," in *Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th*, May 2012, pp. 1–5.

[73] J. Oh and Y. Han, "Cell selection for range expansion with almost blank subframe in heterogeneous networks," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2012 IEEE 23rd International Symposium on*, Sep. 2012, pp. 653–657.

[74] C. de Lima, M. Bennis, and M. Latva-aho, "Statistical analysis of self-organizing networks with biased cell association and interference avoidance," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 1950–1961, Jun. 2013.

[75] D. Lopez-Perez, X. Chu, and I. Guvenc, "On the expanded region of picocells in heterogeneous networks," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 3, pp. 281–294, Jun. 2012.

[76] M. Cierny, H. Wang, R. Wichman, Z. Ding, and C. Wijting, "On number of almost blank subframes in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 5061–5073, Oct. 2013.

[77] S. Cho and W. Choi, "Energy-efficient repulsive cell activation for heterogeneous cellular networks," *IEEE J. Select. Areas Commun.*, vol. 31, no. 5, pp. 870–882, May 2013.

[78] T. S. Rappaport, *Wireless Communications: Principles and Practice, Second Edition*. Upper Saddle River, N.J.: Prentice Hall, 2002.

[79] A. Fehske, F. Richter, and G. Fettweis, "Energy efficiency improvements through micro sites in cellular mobile radio networks," in *GLOBECOM Workshops, 2009 IEEE*, Nov. 2009, pp. 1–5.

[80] S. Mukherjee and I. Guvenc, "Effects of range expansion and interference coordination on capacity and fairness in heterogeneous networks," in *Proc. 2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2011, pp. 1855–1859.

[81] W. Bao and B. Liang, "Structured spectrum allocation and user association in heterogeneous cellular networks," in *INFOCOM, 2014 Proceedings IEEE*, Apr. 2014, pp. 1069–1077.

[82] Y. Lin and W. Yu, "Optimizing user association and frequency reuse for heterogeneous network under stochastic model," in *Global Communications Conference (GLOBECOM), 2013 IEEE*, Dec. 2013, pp. 2045–2050.

[83] A. Ghosh and R. Ratasuk, *Essentials of LTE and LTE-A*. Cambridge, UK: Cambridge University Press, 2011.

[84] V. Chandrasekhar and J. Andrews, "Uplink capacity and interference avoidance for two-tier femtocell networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3498–3509, Jul. 2009.

[85] S.-M. Cheng, W. C. Ao, F.-M. Tseng, and K.-C. Chen, "Design and analysis of downlink spectrum sharing in two-tier cognitive femto networks," *IEEE Trans. Veh. Technol.*, vol. 61, no. 5, pp. 2194–2207, Jun. 2012.

[86] M. S. Jin, S. A. Chae, and D. I. Kim, "Per cluster based opportunistic power control for heterogeneous networks," in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, May 2011, pp. 1–5.

[87] C. W. Tan, "Optimal power control in rayleigh-fading heterogeneous networks," in *INFOCOM, 2011 Proceedings IEEE*, Apr. 2011, pp. 2552–2560.

[88] Z. Wang, W. Xiong, C. Dong, J. Wang, and S. Li, "A novel downlink power control scheme in lte heterogeneous network," in *Computational Problem-Solving (ICCP), 2011 International Conference on*, Oct. 2011, pp. 241–245.

[89] J. Wu, X. Chu, and D. Lopez-Perez, "Downlink outage probability of co-channel femtocells in hierarchical 3-sector macrocells," *IEEE Commun. Letters*, vol. 16, no. 5, pp. 698–701, May 2012.

[90] G. de la Roche, A. Valcarce, D. Lopez-Perez, and J. Zhang, "Access control mechanisms for femtocells," *IEEE Commun. Mag.*, vol. 48, no. 1, pp. 33–39, Jan. 2010.

[91] A. Golaup, M. Mustapha, and L. Patanapongpibul, "Femtocell access control strategy in UMTS and LTE," *IEEE Commun. Mag.*, vol. 47, no. 9, pp. 117–123, Sep. 2009.

[92] T. Q. S. Quek, G. de la Roche, I. Güvenc, and M. Kountouris, *Small Cell networks: Deployment, PHY Techniques, And Resource Management*. Cambridge University Press, 2013.

[93] X. Chu, J. Wu, and H. Wang, "Outage probability analysis for collocated spectrum-sharing macrocell and femtocells," in *Communications (ICC), 2011 IEEE International Conference on*, Jun. 2011, pp. 1–5.

[94] V. Chandrasekhar, M. Kountouris, and J. Andrews, "Coverage in multi-antenna two-tier networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 5314–5327, Oct. 2009.

[95] D. Calin, H. Claussen, and H. Uzunalioglu, "On femto deployment architectures and macrocell offloading benefits in joint macro-femto deployments," *IEEE Commun. Mag.*, vol. 48, no. 1, pp. 26–32, Jan. 2010.

[96] H. Dhillon and J. Andrews, "Downlink rate distribution in heterogeneous cellular networks under generalized cell selection," *IEEE Wireless Commun. Letters*, vol. 3, no. 1, pp. 42–45, Feb. 2014.

[97] "Guidelines for evaluation of radio transmission technologies for IMT-2000," *ITU Recommendation M.1225*, 1997.

[98] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Courier Dover Publications, 1965.

[99] J. Mietzner, R. Schober, L. Lampe, W. Gerstacker, and P. Hoeher, "Multiple-antenna techniques for wireless communications - a comprehensive literature survey," *IEEE Commun. Surveys & Tutorials*, vol. 11, no. 2, pp. 87–105, Second 2009.

[100] C. Li, J. Zhang, and K. Letaief, "Throughput and energy efficiency analysis of small cell networks with multi-antenna base stations," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2505–2517, May 2014.

[101] C. Li, J. Zhang, M. Haenggi, and K. Letaief, "User-centric intercell interference nulling for downlink small cell networks," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1419–1431, Apr. 2015.

[102] N. Deng, W. Zhou, and M. Haenggi, "A heterogeneous cellular network model with inter-tier dependence," in *Global Communications Conference (GLOBECOM), 2014 IEEE*, Dec. 2014, pp. 1522–1527.

[103] J. Zhang and J. Andrews, "Adaptive spatial intercell interference cancellation in multicell wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1455–1468, Dec. 2010.

[104] M. Haenggi, "Mean interference in hard-core wireless networks," *IEEE Commun. Letters.*, vol. 15, no. 8, pp. 792–794, Aug. 2011.

[105] W. Wang and G. Shen, "Energy efficiency of heterogeneous cellular network," in *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*, Sep. 2010, pp. 1–5.

[106] H. Elshaer, F. Boccardi, M. Dohler, and R. Irmer, "Downlink and uplink decoupling: A disruptive architectural design for 5g networks," in *Global Communications Conference (GLOBECOM), 2014 IEEE*, Dec. 2014, pp. 1798–1803.

[107] K. Smiljkovikj, P. Popovski, and L. Gavrilovska, "Analysis of the decoupled access for downlink and uplink in wireless heterogeneous networks," *IEEE Wireless Commun. Letters*, vol. 4, no. 2, pp. 173–176, Apr. 2015.

[108] S. Rangan, T. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.

[109] T. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. Wong, J. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5g cellular: It will work!" *Access, IEEE*, vol. 1, pp. 335–349, 2013.

[110] T. Bai, A. Alkhateeb, and R. Heath, "Coverage and capacity of millimeter-wave cellular networks," *Communications Magazine, IEEE*, vol. 52, no. 9, pp. 70–77, Sep. 2014.

[111] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Stochastic geometry study on device to device communication as a disaster relief solution," to appear in *IEEE Trans. Veh. Technol.*

[112] N. Lee, X. Lin, J. Andrews, and R. Heath, "Power control for D2D underlaid cellular networks: Modeling, algorithms, and analysis," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 1, pp. 1–13, Jan. 2015.

[113] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Commun. Surveys & Tutorials*, vol. 16, no. 4, pp. 1801–1819, 4th. 2014.

[114] U. Tefek and T. Lim, "Interference management through exclusion zones in two-tier cognitive networks," to appear in *IEEE Trans. Wireless Commun.*

[115] J. Gong, S. Zhou, and Z. Niu, "Optimal power allocation for energy harvesting and power grid coexisting wireless communication systems," *Communications, IEEE Transactions on*, vol. 61, no. 7, pp. 3040–3049, Jul. 2013.

[116] H. Dhillon, Y. Li, P. Nuggehalli, Z. Pi, and J. Andrews, "Fundamentals of heterogeneous cellular networks with energy harvesting," *Wireless Communications, IEEE Transactions on*, vol. 13, no. 5, pp. 2782–2797, May 2014.

[117] Y. Lin and W. Yu, "Downlink spectral efficiency of distributed antenna systems under a stochastic model," *IEEE Trans. Wireless Commun.*, vol. 13, no. 12, pp. 6891–6902, Dec. 2014.

List of Publications

1. C. Jia and T. J. Lim, "Resource partitioning and user association with sleep-mode base stations in heterogeneous cellular networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 3780-3793, Jul. 2015.

2. C. Jia and T. J. Lim, "Power saving design of femtocell exclusion zones," *IEEE Transactions on Communications*, under review.

3. C. Jia and T. J. Lim, "Energy efficient base station deployment in multi-antenna two-tier networks," *IEEE Transactions on Wireless Communications*, under review.

4. C. Jia and T. J. Lim, "Designing femtocell exclusion zones to minimize power in a heterogeneous network," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Austin, USA, 2014.

5. C. Jia and T. J. Lim, "Power minimizing spectrum allocation in a two-tier cellular network," in *Proc. Regional Conference on Computer and Information Engineering (RCCIE)*, Yogyakarta, Indonesia, 2014.

6. C. Jia and T. J. Lim, "Spectrum allocation in heterogeneous networks," in *Proc. IEEE Communications Theory Workshop (CTW)*, Phuket, Thailand, 2013.