Energy Efficiency of Wireless Network using Coordinated Gated Narrow Beams

Jiazhen Zhu

Submitted in total fulfilment of the requirements of the degree of

Doctor of Philosophy

Department of Electrical and Electronic Engineering

THE UNIVERSITY OF MELBOURNE

Australia

July 2019

Abstract

With the ever-increasing demand for wireless service and higher data rate, the wireless network has experienced an unprecedented growth worldwide in the past decade and it is expected to grow continuously. Energy efficiency of the wireless network has become a growing concern for network operators and standardization authorities, not only to reduce the overall electric energy usage but also to reduce its environmental footprint. This has triggered research work to explore future, green wireless technologies and strategies in order to bring energy efficiency improvements in the entire network.

In this thesis, we investigate the opportunity of improving energy efficiency through the use of coordinated gated narrow beams for downlink transmission of data. This is a class of Coordinated Multipoint (CoMP) transmission technique which is originally proposed in LTE standards for enhancing cell-edge throughput. The principle is that multiple base stations are coordinated with each other so that potential interfering source from the adjacent base station can be steered away by appropriate beamforming and scheduling.

The thesis is divided into three parts. In the first part, we develop a realistic coordinated beamforming strategy using gated narrow beams and estimate the network throughput and base station power consumption based on network level simulation. In the simulator, we apply a practical traffic model in which users entering at a rate consistent with time of day dependent traffic level, and a proportional fair resource scheduling scheme to ensure fairness between users. Then, we analyse the required channel state information to support the beam coordination and resource scheduling, and develop a low-overhead signalling and control framework that provides sufficient signalling information. The signalling design and its implementation leverage signalling functionality and protocol in current standards. We develop methodologies for quantifying the signalling overheads

introduced in terms of the percentage of downlink resource occupied by additional coordination reference signals.

Finally, we develop energy consumption models of the functional components and processes in the coordinated network architecture, including backhaul switches and interfaces, and a central coordination unit. We quantify the additional energy costs associated with the coordination and signalling functions, and perform a comprehensive evaluation on the energy efficiency of an LTE network employing the proposed coordinated gated narrow beams. Our results show that significant energy savings can be achieved compared with a conventional network with no coordination.

Declaration

The work presented in this thesis comprises only my original work. No material in this thesis has been previously published and written by another person, except where due reference is made in the text of the thesis. I declare that none of the work presented in this thesis has been submitted for any other degree or diploma at any University and that this thesis is less than 100,000 words in length, excluding figures, tables, bibliographies, appendices and footnotes.

Jiazhen Zhu, July 2019

Acknowledgements

I would like to firstly express my deepest gratitude to my supervisors Prof. Rodney S. Tucker, Dr. Kerry Hinton, Mr. Robert Ayre, and Dr. Jeffrey Cheong. I would not be able to complete this thesis without their continual guidance and generous assistance throughout my PhD candidature. I feel very fortunate to have them as my supervisors and I have benefited tremendously from their wide knowledge, stimulating suggestion, critical thinking and brilliant ideas. I would also like to thank Dr. Qasim Chaudhari for several discussions and his valuable technical advice. I give my special thanks to Prof. Elaine Wong, Prof. Brian Krongold and Prof. Jamie Evans for being in my PhD advisory committee and providing insightful comments.

I also wish to thank Centre for Energy-Efficient Telecommunications (CEET) for providing a great research and work environment. I have met some great colleagues here and many thanks go to them. In addition, I would like to acknowledge the University of Melbourne and CEET for offering me financial support during the course of my study.

Last but not least, I would like to thank my family and friends for their unconditional love, support and patience. I specifically would like to express my gratitude to my parents and my partner. It is their encouragement and understanding that have given me the strength to pursue this PhD study.

Contents

List of Figures

List of Tables

Acronyms

BBU	Baseband Unit
BS	Base Station
CCU	Central Coordination Unit
CoMP	Coordinated Multipoint Transmission
CQI	Channel Quality Indicator
C-RAN	Cloud Radio Access Network
CSI	Channel State Information
CSI-RS	Channel State Information Reference Signal
FLOPS	Floating Point Operations Per Second
FTP	File Transfer Protocol
IMR	Interference Measurement Resource
IP	Internet Protocol
LTE	Long Term Evolution
MCS	Modulation and Coding Scheme
MIMO	Multiple-Input Multiple-Output
PMI	Precoding Matrix Indicator
PRB	Physical Resource Block
RE	Resource Element
RF	Radio Frequency
RI	Rank Indicator
RRH	Remote Radio Head
SGW	Serving Gateway
SINR	Signal to Interference plus Noise Ratio
UE	User Equipment

Chapter 1

Introduction

Wireless devices such as smartphone and tablets have become an integral part of our everyday lives today, with the applications ranging from web browsing, video streaming, social networking, gaming, surveillance, and healthcare monitoring. With the ever-increasing demand for wireless service and higher data rate, the wireless network has experienced an unprecedented growth in the last decade.

Until recently, the primary objectives for the wireless network have been providing high capacity and spectral efficiency, without any consideration of power or energy efficiency. As a result, the wide deployment of wireless communication infrastructure has led to high energy consumption and environmental impact. It is estimated that the mobile communication industry accounts for 15%-20% of the ICT energy footprint [1, 2]. The annual CO_2 emission of base stations in the wireless network is around 20 metric tons, which represents 2% of the worldwide CO_2 emission [3].

The energy cost and carbon footprint of the wireless network will continue to increase as the growth of the wireless network and mobile traffic. Current trends indicate that the annual growth rate of wireless data traffic is 47% or even higher [4]. Under such predictions, the energy efficiency of wireless network is becoming a major concern for the industry and research community [5]. There is an emerging trend of research projects, service providers and regulatory bodies to address the issue of energy consumption of wireless network and work on increasing energy efficiency in wireless technologies. These energy-aware technologies aim at reducing the power consumption while meeting throughput and service quality requirements.

Prior to the commencement of this PhD project, little work had been done to understand the energy gains and costs associated with coordinated multipoint (CoMP) transmission and reception technique, which is proposed in LTE-Advanced standards for the purpose of enhancing spectral efficiency. Therefore, in this thesis, we investigate the energy impact of the introduction of CoMP functions and develop a detailed and realistic estimate of energy efficiency of CoMP enabled LTE network. We determine whether the CoMP operation can bring energy savings and provide insight into the design of an energy efficient coordination strategy. We also develop an understanding of how the energy efficiency will evolve as the number of subscribers and traffic volume increase.

1.1 Focus of the Thesis

IN this thesis, we look to develop a better understanding of the energy efficiency of an LTE network with CoMP. We focus on a practical downlink CoMP solution - coordinated beamforming strategy using gated narrow beams. A large amount of work has been done on the spectral efficiency of CoMP systems [6-9] and CoMP has exhibited significant improvement in terms of spectral efficiency and capacity. However, not much work has been done on the energy aspect of CoMP systems. Their potential for energy efficiency improvement remains to be assessed in detail.

In our work, we look to answer a few key questions:

- Across the network, what extra coordination functions are required for performing CoMP?
- What are the extra energy costs associated with these network functions?
- What is the overall energy efficiency in a CoMP system?

In evaluating the energy efficiency, we focus on energy efficiency metric as measured by the consumed energy divided by the actual data bits delivered. It represents the energy required to carry a unit of data, in other words, the joules per bits carried. It can also be expressed as the ratio of power consumption to the overall bit rate.

A number of estimates of CoMP energy efficiency have already been published [5, 10-12]. However, these estimates were developed using highly theoretical models and simplifying assumptions and there is a lack of complete energy modelling of the CoMP network. Thus, the results don't give a realistic estimate and provide limited insight into

the aforementioned questions. In this thesis, we aim to perform a thorough evaluation of the energy efficiency of a practical CoMP network in the context of LTE.

The estimate of the energy efficiency involves an estimate of CoMP system throughput and an estimate of the energy consumption of the whole network. We look to develop methodologies for determining the effective transmission rate and throughput as well as the energy consumption of the whole network which takes into account the signalling capabilities of LTE standards and signalling overhead required for CoMP, fairness in resource allocation to users, and realistic traffic conditions with varying loads.

The LTE cellular network consists of a radio access network that has base stations to provide radio coverage, aggregation network for backhauling the traffic between the base station and the core network, and a core network for switching of data [13]. A typical base station consists of multiple energy consuming components such as power amplifiers, radio transceivers, baseband processing units, transmit antennas and cooling system [14]. The energy consumption of the LTE cellular network is distributed across all the elements and functions. In addition to understanding the total energy consumption of the network, we seek to also understand the energy consumed in each domain/subsystem and their relative importance on the overall energy efficiency. We also seek to assess the additional energy consumption incurred due to extra CoMP functions and processes across the network.

The CoMP system is expected to involve more functions and processes for supporting the coordination between base stations, compared with a non-CoMP system. In particular, it would require more channel estimation between the coordinating base stations, signal processing, and information sharing between base stations [15]. It may also require architectural changes such as a centralized control unit [16] for processing information from coordinating base stations and establishing an optimum transmission pattern. In this work, we develop energy models of the network components and explicitly quantify the overhead and additional energy costs associated with CoMP.

A question of particular importance is whether CoMP can bring energy savings after taking into account all the extra energy costs in supporting the coordination. To determine the answer, we analyse the comparative energy efficiency performance of CoMP network and a conventional network with no coordination, and provide an understanding of the

trade-off between the energy gain (achieved by enhanced throughput) and the energy costs (for coordination functions).

In addition, we analyse the energy efficiency of both networks as a function of traffic load, so that to understand the energy efficiency performance at different time of day when under different traffic loads. The analysis also enables forecasting future energy efficiency based on the trend of traffic growth. The modelling and analysis in this work are intended to provide a guideline for evaluating the realistic energy efficiency performance of CoMP techniques and other techniques involving base station cooperation.

1.2 Thesis Outline and Original Contributions

In the following, we give an outline of the thesis and state the original contributions of the thesis.

Chapter 2 surveys technologies to improve the energy efficiency of the wireless network, and performs a literature review of the studies on the energy aspect of downlink CoMP schemes, especially the studies pertaining to evaluating the energy consumption and energy efficiency of a CoMP network.

Chapter 3 provides an overview of our approaches to developing a complete and accurate modelling of the energy efficiency of a CoMP coordinated beamforming network. We provide a high-level description of the research problems involved and their solution strategies.

Original contributions in **Chapter 3**:

- Developed a pragmatic approach to estimating the energy efficiency a CoMP coordinated beamforming network taking into account the signalling overhead and additional energy costs incurred in supporting the coordination.

Chapter 4 develops a realistic coordinated beamforming strategy using gated narrow beams in an LTE network, and evaluates the network throughput and energy consumption of coordinated beamforming base stations based on network level simulation. We consider a practical traffic model in which users entering at a rate consistent with time of day dependent traffic level, and a proportional fair resource scheduling scheme to ensure fairness between users. We develop a preliminary estimate of network energy efficiency

using the throughput and base station power consumption estimates. We use it to identify the potential benefit brought by coordination and provide insights into the trade-off between energy gains and costs of CoMP.

Original contributions in **Chapter 4**:

- Developed a distributed coordinated beamforming architecture with multiple gated narrow beams for transmission of data, and a beam coordination and resource scheduling scheme that mitigates interference between the coordinated narrow beams in a coordination cluster.

- Developed a system level simulator of an LTE network of base station with coordinated gated narrow beams, and evaluated various network performance measures including SINR, user download speed, failure rate, cell throughput, base station power consumption and energy efficiency under realistic File Transfer Protocol (FTP) type traffic.

- Used our results to confirm that significant spectral efficiency gain and throughput gain can be achieved through base station coordination.

- Showed that the coordinated network of gated narrow beams has the capacity to carry the entire busy hour traffic load as predicted for the dense urban network in 2020 with negligible failure cases.

- Showed that the base station in the coordinated network of narrow beams consumes less power on average than the base station in a conventional non-coordinated network of wide beams at mid to high traffic levels (50% - 140% of average daily dense urban traffic load in 2020), and estimated that the power savings to be around 140 Watts for one base station.

- Showed that the base station energy savings achieved from the enhanced throughput can potentially lead to up to 48% improvement on energy efficiency, with the greatest improvement delivered at higher traffic loads.

- Showed that the utilisation rate of a base station in the coordinated beamforming network is about 53% at busy hour dense urban traffic level and demonstrated that a commonly used full buffer model in which all base stations transmit 100% of the time can lead to overly optimistic estimate on network energy efficiency.

Chapter 5 analyses the required channel state information (CSI) to serve the beam coordination and resource scheduling scheme, and develops a low-overhead signalling

and control framework that can provide sufficient signalling information to support the coordinated beamforming network modelled in Chapter 4. Our signalling design leverage signalling functionality and protocol in current 3GPP standards. We model and estimate the signalling overhead introduced in terms of the percentage of downlink resource occupied by additional coordination reference signals.

Original contributions in **Chapter 5**:

- Developed a signalling and control framework for the coordinated beamforming network to estimate the signalling overhead, additional signal backhauling and processing effort required for coordination.

- Showed that the coordinated beamforming can be supported entirely based on signalling capabilities of current LTE standards with minimal additional signalling overhead.

- Estimated that the upper bound for the incurred signalling overhead is approximately 3.34% when a minimum periodicity of reference signal measurement (5ms) was assumed.

- Showed that the signalling scheme only incurs a fixed signalling overhead that is not dependent on the number of users or the volume of user traffic.

- Developed a methodology for quantifying the additional signalling overhead, and hence the amount of signalling information exchanged on backhaul and the computation effort for processing the user CSI feedback for CoMP enabled systems.

Chapter 6 develops energy consumption models of the coordinated beamforming network including backhaul switches and interfaces, central coordination unit that processes signalling information and executes the beam coordination and resource scheduling scheme. We use these models to quantitatively estimate the additional energy costs associated with the coordination and signalling functions, and to provide a complete evaluation on the overall energy consumption and energy efficiency of a CoMP network.

Original contributions in **Chapter 6**:

- Developed energy consumption models of the functional components and processes in the coordinated beamforming network architecture.

- Showed that the backhaul power consumption for carrying the user traffic is relatively insensitive to traffic load and adds about 1% to 3% to the base station power consumption.

- Showed that the additional energy consumption for coordination in backhaul and central coordination unit are negligible small across the entire traffic loads.

- Showed that the energy efficiency of the coordinated beamforming network deteriorates by up to 7% by the inclusion of the backhaul energy consumption and coordination signalling overhead etc., and that they should therefore be included in addition to the base station power consumption when evaluating the energy efficiency of CoMP network.

- Showed that the coordinated beamforming network is about 39% and 46% more energy efficient than the conventional non-coordinated network at 100% and 140% traffic load respectively. Demonstrated that significant energy savings can be achieved through CoMP techniques, in particular when the traffic load is high.

- Estimated that the weighted average energy efficiency of the coordinated beamforming network over a complete diurnal cycle for the dense urban environment in 2020 achieves an improvement of about 39% compared with the non-coordinated network.

Chapter 2

Literature Review

This chapter will provide a review of the literature pertaining to the energy consumption of wireless network and potential approaches to reduce energy consumption and improve the energy efficiency of the network. This chapter will also give an introduction to CoMP technology and review previous studies on estimating energy efficiency of CoMP systems.

2.1 Energy Consumption of Wireless Network

IN the past few decades, the Information and Communications Technology (ICT) industry has experienced rapid growth and it is expected to grow continuously. According to [17] The ICT infrastructure is responsible for 10% of the world's total electric energy consumption with 900 billion kWh per year consumption, and it is expected to increase by approximately 16 to 20% per year resulting in an energy consumption rise by a factor of 30 in 23 years. Within the ICT sector, the mobile telecommunication industry is one of the major contributors to energy consumption accounting for 15%-20% of the ICT energy footprint [1, 2], and traffic and energy consumption of mobile networks is growing much faster than ICT on the whole [4, 18].

In addition to the electric energy usage, the rapid growth of ICT also has a deep environmental impact. Since the concern over global warming, reducing CO_2 emission has become a universal goal in the last couple of decades. A number of researches have indicated that the ICT is responsible for about 2%-2.5% percent of worldwide CO_2 emissions, among which mobile radio networks accounts for 0.3%-0.4% [1, 2]. Although it is a low percentage, this number is predicted to double by 2020, particularly associated with the exponential growth of the broadband wireless service [19-21].

Due to the tremendous growth in the number of wireless devices with the availability of low cost notebook, tablets, and smartphones, and the move to higher data rate applications such as video streaming and file transfers, global mobile data traffic is currently increasing at 47% per annum, and it is forecasted that there will be an increase of sevenfold from 2016 to 2021 [4].

Given the current energy and environmental impact, along with the increasing traffic data, there is an urgent need to improve the energy efficiency of wireless networks [20]. Increasing effort is devoted to designing and planning wireless communication network that not only achieves optimal capacity and spectral efficiency but also improves the energy efficiency [22]. Energy Efficiency is becoming a major concern for the industry and research community [5].

Many countries have established commitments to reduce their CO_2 emission levels. From the operator's perspective, energy efficiency not only represents social responsibility but also has an economic benefit as reports have shown that nearly 50% of the overall operating expenses for a mobile telecommunication operator is the energy cost [23]. While previously only concentrating on meeting the traffic demand and Quality of Service, the operator around the world have now realized the importance of managing their networks in an energy efficient way [22, 24, 25].

In the development of 3GPP systems, the energy consumption issues received little attention until recent years. In reaction toward the rising energy costs and carbon footprint, the 3GPP standard has now considered new energy efficient approaches into the design of 4G and beyond 4G mobile network [26] across the entire protocol stack from the physical layer to network layer, and also introducing new energy efficient deployment strategies. Various energy efficiency measurement methods, key performance indicators and control policies are being proposed for 5G networks by 3GPP [27].

Pursuing high energy efficiency (EE) is a trend for the design of future wireless communications, it has also drawn increasing attention from the research community. Several global research collaborations dedicated to energy-efficient wireless communications are being carried out such as Green Radio [13], EARTH [28], GreenTouch [29]. The general message from these projects has been that through adaptation of novel technologies, various system architectures, fundamental trade-offs linked with energy efficiency and the overall performance [20], energy efficiencies can be

significantly improved. In addition, GreenTouch brings major vendors, operators and academia together, and set the goal is to find solutions that can potentially increase network energy efficiency by a factor of 1000 from current levels.

2.2 Energy Efficiency Metric

The development of cellular network has been focusing on the enhancement of throughput and coverage in order to meet increasing demand on the quality of service. The performance metrics and optimization benchmarks have been spectral efficiency and cell-edge user throughput [12]. Spectral efficiency measures the system throughput per unit of bandwidth (bits/s/Hz) and indicates how efficiently a limited frequency spectrum is utilized. The peak value of spectral efficiency is a major performance indicator in the cellular network evolution. For example, the downlink rate in 3GPP standard increases from 0.05 b/s/Hz for GSM to 5 b/s/Hz for LTE [20].

However, this spectral efficiency metric cannot provide insight into how efficiently energy is used. As indicated by [5, 20] spectral efficiency and energy efficiency can be conflicting objectives and there is a trade-off between these two objectives. For instance, spatial multiplexing and diversity using multiple antennas are adopted in LTE to improve spectral efficiency, but the configuration of multiple antennas thus multiple RF chains causes higher power consumption such as higher circuit power [5]. The use of multiple antennas affects the system energy efficiency in a comprehensive manner. Pursuing the goal of spectrum efficiency alone could lead to a design of a system with poor energy efficiency and vice versa.

Therefore, a suitable energy efficiency metric must be defined and taken into consideration in the network optimization process. An energy efficiency metric bit-per-joule (bits/J) was first introduced in [30], and it has now become a commonly used metric for evaluation of power and energy associated behaviour of communication systems [1]. Both bit-per-joule (calculated as $\frac{\text{data bits delivered}}{\text{energy consumption}}$) and joule-per-bit (calculated as $\frac{\text{energy consumption}}{\text{data bits delivered}}$) can be used to represent how efficient the energy is consumed. Optimizing the energy efficiency is essentially maximizing the amount of data transmitted for per unit of energy, or equivalently, minimizing the energy cost to transmit a unit of data.

Some previous studies on evaluating the energy efficiency of wireless network [30, 31] took an information-theoretic approach that only considers transmit power of individual wireless link for energy consumption estimation. As pointed out by the authors in [12, 32], a system-wide energy efficiency measurement should be used in characterizing the energy efficiency of a network. The bit-per-joule energy efficiency metric should take into consideration the overall power consumption of the network such as for data processing and backhauling, rather than the transmit power at the wireless link level only. For correct estimation of throughput, only the successfully transmitted data bits should be accounted for, i.e. exclude dropped and retransmitted packets, overheads incurred in sending signalling packets.

In this thesis, a system-level energy efficiency metric will be adopted for evaluating the energy efficiency of the network, and for comparing the energy efficiency of competing systems. The system-wide approach allows trade-offs to be made between sub-systems in such a way that the energy savings in one part would not be negated by the increased energy consumption in another part.

2.3 Energy Saving Technologies and Deployment Strategies in Wireless Network

As the energy consumption of wireless network has received attention in recent years, a large amount of work has been completed that seeks to reduce the energy consumption [22]. In this section, we review techniques and strategies to reduce energy consumption and improve the energy efficiency in wireless network.

Base station hardware inefficiencies are pointed out by many studies [14, 33, 34]. Work has been undertaken to improve the hardware design and efficiency, such as high-efficiency power amplifiers and digital signal processing, low power circuit design, advanced cooling systems etc. The performance of power amplifiers are of special concern as more than 50% of base station energy is consumed by power amplifiers [35]. Enhanced designs of the power amplifier are proposed [36] to achieve better efficiency and reduce the power dissipated as heat.

Besides improving hardware efficiency, turning off components in the base station and letting it enter sleep mode is a simple and effective solution for energy saving [37]. It is considered one of the most popular energy saving methods when the traffic load is low

[38]. The rationale behind is that, traffic demand fluctuate significantly over time [39, 40] and base station capacity is dimensioned to meet peak load. Thus, base stations are often underutilized and unnecessarily powered on [41, 42]. Powers saving protocols for LTE are proposed with different levels of base station sleeping and the corresponding components and subcarriers to be turned off [43, 44]. A number of studies seek to quantify the energy savings by enabling sleep mode operation [45-47], and energy savings identified ranges from 20% to 50% for different deployment and traffic conditions.

Another energy-saving strategy is resource and spectrum scheduling. The majority of the work in this area is related to developing schemes and algorithms that efficiently utilize network resource in terms of power, time slots and spectrum. It has been shown that these resource scheduling techniques can reduce energy consumption while meeting the requirement of quality of service [48, 49].

Deploying a certain number of relays to exchange information between a base station and a mobile terminal in wireless network can also improve energy efficiency [50, 51]. A relay node is a low power network element that receives the data from the source node, stores, processes and forwards the data to the destination node following some cooperative routing protocols [52]. The relays reduce the transmission distance, thus the transmit power and energy requirement can be lower [53]. In addition, they create multiple connections between the base station and mobile terminal, enabling a diversity gain to be exploited. This diversity gain can potentially lead to extra spectral efficiency and energy efficiency improvements [54].

Smaller size cells were originally proposed in LTE standards for meeting the higher throughput needs in given areas [55]. The placement of small cells can improve spatial reuse and provide higher data rates [56, 57], and recent studies investigate into its potential for saving energy [36, 55, 58, 59]. It is pointed out in [60] that in theory significant network energy efficiency gain can be achieved with the deployment of small cells due to the smaller path loss, penetration loss when deployed indoor, lower transmit power, and thus energy consumption. Studies on small cell look into the optimal number of small cells overlaying per macro base station [61], optimal location and size of small cells to maximize network energy efficiency [45, 62], and hierarchical deployment of different sizes of cells in the network [63].

These various energy savings approaches discussed above can be categorized into two main categories. The first is to reduce the energy consumption of individual base stations, either through the use of more energy efficient hardware component, or selectively switching off components when the traffic is low. The second is to exploit deployment strategies, for example, using energy efficient radio resource allocation and scheduling techniques, or placement of low power small cell and relay nodes to lower the number of base stations while meeting the performance requirement.

2.4 Coordinated Multipoint (CoMP) Transmission and Reception

In this section, we give an introduction to CoMP transmission and reception technology, which is the focus of this work. We investigate the opportunities of energy saving through the use of downlink CoMP in this thesis.

Modern cellular networks such as LTE and LTE-Advanced reuse the same frequency band in every sector and in every cell, with the aim to support the maximum number of concurrent channels in a cell. However, this also brings the problem of inter-cell interference, which is now the main obstacle to achieving spectral efficiency and throughput [64]. As all signals coming from base stations other than the user's serving base station are interference, the frequency reuse dramatically limits the achievable signal to interference ratio and average spectral efficiency in the cell [15]. In particular, users at the cell edge areas suffer from the strongest interference levels from neighbouring cells leading to poor cell-edge throughput and coverage.

3GPP has been working on these issues and a technique named Coordinated Multipoint (CoMP) transmission/reception is proposed as a promising solution for managing inter-cell interference and improving cell-edge throughput [65]. The fundamental principle of CoMP is to coordinate multiple base stations such that the transmission signals of a base station do not cause serious interference into the coverage area of other base stations or can be exploited as a useful signal [16]. The coordination is enabled by the architecture of a cellular network where base stations are connected to a common infrastructure by backhaul [66]. Information such as user location, channel state information and traffic data can be passed on and exchanged between base stations through the backhaul [65],

allowing base stations to have knowledge of neighbouring base stations and be able to perform coordination.

The CoMP studies performed in 3GPP Release 10 LTE-Advanced involves three possible coordination approaches and different approaches have different levels of coordination, scheduling complexity and requirements on the infrastructure such as backhaul links [6]. Figure 2.1 depicts the architectural examples of these approaches. The red lines with arrows represent the signal transmitted from the serving base stations. They also represent the interference incurred by the transmission to the coverage of other adjacent cells.

Coordinated beamforming/scheduling as shown in Figure 2.1 (a) is a relatively simple approach where each user is served by a single serving base station and the potential interfering source from the adjacent base station is steered away by appropriate beamforming and scheduling. In joint processing as shown in Figure 2.1 (b), multiple base stations jointly transmit the same data to a user simultaneously. The transmission data intended for a user is made available at multiple base stations, and these base stations all transmit desired signal rather than interference signal to the user. Dynamic cell selection as shown in Figure 2.1 (c) can be seen as a special form of joint processing, where only one base station transmits data at a time and the user selects the most appropriate dynamically.
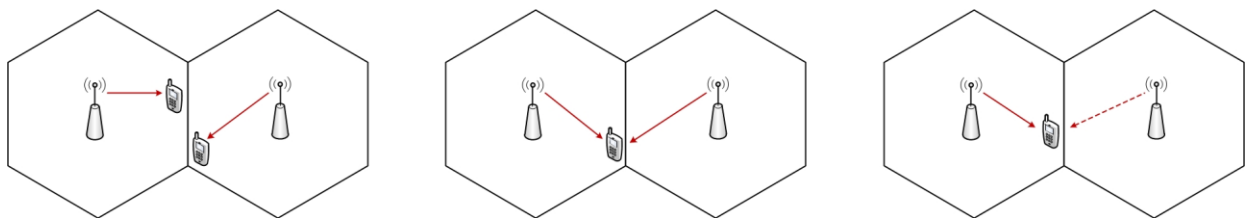
 (a) Coordinated beamforming/scheduling   (b) Joint processing   (c) Dynamic point selection

Figure 2.1 CoMP downlink approaches

The CoMP studies on the coordination techniques in the literature can also be broadly classified into the above three approaches. In this work, we select the approach of coordinated beamforming/scheduling for downlink CoMP solution.

Coordinated beamforming/scheduling jointly combines coordinated beamforming and coordinated scheduling. Under this approach, user traffic data is only available at one

base station, which is its serving base station. The multiple coordinating base stations share only channel state information of multiple user terminals, and the user beamforming and scheduling decisions are made via coordination between the multiple neighbouring base stations. This approach essentially reduces the interference level experienced by the user terminal by appropriately scheduling and selecting the beamforming weights of the interfering base stations.

One group of techniques under this approach is fixed beam pattern coordination. It consists of coordinating the transmit precoders (beamforming matrices) in the coordinating base stations in a pre-defined manner. Both the user terminal and base stations have knowledge of a codebook containing a specific set of quantized precoders. The transmitters are constrained to use only these quantized precoders, and thus a fixed set of beams are generated for coordination. The user terminals also report the perceived channel state information in a reduced form - a restricted or recommended precoding matrix indicator corresponding to a potential precoder in the codebook. Based on the user feedback, a coordinated selection of the transmit precoders in each coordinating base station is made by a central controller and passed to the base stations.

Different to the fixed beam pattern coordination, another group of techniques are user-specific coordinated beamforming that aims at eliminating or suppressing the effect of interference. It allows an arbitrary and adaptive transmit precoder (beamforming matrix) for each user without the constraint of a codebook. The base stations compute a new beamforming matrix for downlink transmission based on the channel state information from the serving and the interfering base stations. Two typical user selection and beamforming strategies are zero-forcing beamforming and joint leakage suppression. Zero-forcing beamforming eliminates the interference by finding a linear precoding matrix in the null space of the transmit channels of the co-scheduled users, and thus forcing the interference from a base station to users served by other base stations to zero. Joint leakage suppression aims to maximize the signal power of the transmission data intended for one user while minimizing the interference leakage to co-scheduled users. It is achieved by maximizing the signal-to-leakage-plus-noise ratio of the user served by the base station.

Compared with the fixed beam pattern coordination, this technique provides more flexibility in conducting the beamforming; however, it has an increased level of

complexity in terms of computation as well as more requirements on the channel station information accuracy and processing. The fixed beam pattern coordination technique has relatively lower level of complexity and also lower overhead in terms of channel state information feedback, backhaul transmission and control. The advantage and limitation of the various different coordinated beamforming/scheduling techniques will be further discussed in Section 4.2.2.

CoMP has attracted a lot of research attention and it has proved to be very effective for improving spectral efficiency, coordinated coverage, cell-edge throughput and user fairness for the cellular network [6-9]. Much of the work on CoMP is related to developing coordination algorithms [67-69] and precoding techniques [70, 71] with the aim to maximize capacity. Studies are also carried out to investigate the impact of practical constraints on the performance of CoMP, including channel estimation errors [72-74], backhaul latency and limited bandwidth [75-77], and system delays and synchronization [74]. The combined benefit of CoMP technique and small cells deployment is also explored in some studies [11]. While the benefit of CoMP transmission and reception in terms of spectral efficiency has been intensively investigated, these work focus on delivering high throughput and they don't usually take into consideration the energy consumption and energy efficiency of CoMP systems.

Although CoMP schemes could provide increased throughput, they come with additional costs such as more channel estimation effort, additional signal processing and extra load on backhaul. The trade-off between gains in cell throughput and increased energy consumption has not been adequately addressed. As energy efficiency is a key performance metric and optimization benchmark in the next generation wireless network [78], it is of great importance to understand the impact of CoMP deployment on the energy efficiency of LTE and LTE-Advanced network.

2.5 Energy Modelling and Energy Efficiency of CoMP Systems

In this section, we review literature pertaining to the energy aspect of CoMP technique. To our knowledge, there are only a few studies in the literature that focus on energy modelling and/or analysis on the energy associated behaviour of CoMP enabled systems, and there is a lack of thorough evaluation on the energy efficiency of CoMP network in

the context of LTE. The following sub-sections review these previous studies with particular focus on their energy consumption models and methodologies for computing energy efficiency metric.

Fehske et al., 2010

One of the earliest studies on the energy efficiency of CoMP network was by Fehske [10]. It was also one of the most comprehensive analyses that considered overhead and energy costs incurred due to CoMP. The study focused on determining the bit per Joule energy efficiency of a joint processing network models where base stations jointly process, send and receive user data. It also provided information on the energy efficiency performance under varying base station inter-site distance and the number of base stations involved in a coordination group.

Fehske et al. identified three categories of overhead that the CoMP induced in the cellular network − additional reference signal, additional backhauling, and additional signalling processing. These three overheads were estimated using linearized models based on the number of BS being coordinated and taken into account when estimating the energy efficiency.

The impact of additional reference signal symbols on reducing the transmission rate was considered in capacity estimation, while the latter two overheads were included as extra energy consumption. The study adopted a base station energy consumption model from [14], which considered the long-term power consumption of cellular base station including power amplifier and baseband signalling processing, cooling loss, and battery backup and power supply loss. They extended the energy consumption models for base stations in [14] to capture energy consumption of additional signalling processing and backhauling links. The estimation of the overheads was however highly theoretical.

It was assumed that the signalling overhead as to resource occupied by reference signal increased linearly as the number of coordinating base stations. Thus, the percentage of signalling overhead in the joint processing network was estimated as the signalling usage in non-coordinating network multiplied by the number of coordinating base stations. The percentage of control overhead associated with the signalling was assumed to be an arbitrary number. The estimation of signalling processing overhead was based on data for

non-coordinating network and theoretical reasoning of additional signalling processing effort for a coordinating network.

A baseline signalling processing power was taken from an LTE testbed with no coordination. It was assumed that, for coordinated network, the signal processing power regarding channel estimations scaled linearly with the number of coordinating base stations, and signalling power regarding MIMO processing scaled quadratically with the number of coordinating base stations, whilst the rest of the signalling power remained the same as in a non-coordinated network. The above estimation of signalling overheads and the additional power for signal processing were made without first conducting a signalling design for the proposed joint processing network using available network signalling protocols and standards at the time.

The additional backhauling was estimated in a similar approach, with the backhaul load assumed to scale linearly with the number of coordinating base stations with some scaling factors. The backhaul network was simply modelled as a collection of wireless microwave links of 100Mbit per second capacity and a power dissipation of 50W each. The power consumption of backhaul was considered to be a pure linear model, and computed as the traffic load divided 100Mbit and multiplied by 50W.

The capacity estimation of the joint processing network was obtained by simulation with LTE-based parameters, where users were dropped into a playground of 57 cells. The channel matrices of each user were computed via a standard path loss model and Rayleigh fading, and then mapped to an upper bound of transmission rate that could be achieved at link level. No specific resource allocation function or consideration for fairness was included in the capacity estimation. Through the simulation results, Fehske et al. found that the energy efficiency in bit per Joule of the joint processing system has a potential gain of 10% and 20% for small and large inter-site distances, respectively.

It was also observed that significant performance improvement occurred when the number of coordinating base stations was increased from two to three; however, for coordination groups larger than three, the additional gains were rather small as the additional energy costs associated with the overheads were larger.

Heliot et al., 2011

Heliot et al. assessed the capacity and energy efficiency of an idealized fully coordinated CoMP system [31]. The study assumed perfect backhaul links between each BS and an idealistic cooperative processing unit, thus the CoMP system was idealized into a distributed antenna or distributed multiple-input multiple-output (DMIMO) system model where antennas from multiple base stations acted as a single antenna array.

This simplified model allowed the channel capacity of CoMP system to be expressed with a closed-form approximation using an information theoretic approach. Heliot et al. extended their prior work on closed-form approximation of single user DMIMO channel capacity [79] to a more generic and accurate approximation for multi-user scenario in this study. It was noted by Heliot et al. that this approximation only represented upper bounds on the channel achievable rate of non-ideal CoMP systems. The assumptions made in the derivation process were not always supported in reality although frequently used in literature for simplification purpose [80].

The energy consumption of the system was calculated by considering baseline base station power consumption and extra CoMP power consumption associated with coordinating multiple base stations. The baseline power consumption of the base station was again taken from [14]. For extra power consumption for CoMP, only the backhauling power was considered, and the estimate was obtained directly from the study by Fehske [10]. An energy efficiency metric 'energy channel capacity' in bit-per-joule was used to evaluate channel energy efficiency, which was defined as the ratio of the channel capacity to the system consumed power.

This study by Heliot et al. also compared the channel energy efficiency of 1-BS system, 2-BS system and 4-BS system for various downlink SNR scenario, where N-BS system refers to coordinating groups of N base stations. The total transmit power remained the same for each multi BS system. It was found that the 2-BS and 4-BS systems outperformed the 1-BS system for very low SNR values. For SNR value above 4 dB, the 1-BS system was the most energy efficient. Heliot et al. concluded that that multi-BS coordination was most likely to be efficient when the link quality between the BSs and user is weak, e.g. cell-edge communication. It was also noted that the potential improvement of CoMP in terms of channel energy efficiency was not as high as in terms of channel capacity due to additional energy cost for coordinating multiple base stations.

Thus, Heliot et al. concluded that the additional energy cost should be kept low for CoMP in order to provide energy efficiency gain.

Huq et al., 2012; Huq et al., 2013

A study by Huq et al. [5] investigated the energy efficiency for different downlink CoMP methods using a system level simulator. Huq et al. considered an LTE network using hexagonal deployment with a central base station surrounded by six base stations in the first tier and twelve base stations in the second tier. Multiple users were randomly and uniformly distributed in the simulated area, and their data rate is determined based on a previous study [81] into the SNR dependent transmission rate for OFDMA channels. The simulation results for performance evaluation were accumulated from the central base station while the other base stations served as interferers.

In estimating the network throughput, Huq et al. considered both joint processing and coordinated beamforming/scheduling for downlink CoMP transmission strategy. However, no further detailed description of the coordination mechanism was provided in this study. The coordination methods such as for determining joint transmitting base stations for a user, or for computing of the beamforming weights were unclear, thus no evaluation of the energy costs associated with specific CoMP functions was possible.

The energy modelling of the network considered base station power consumption only. The power consumption of base stations was estimated based on a simplistic model that consisted of two parts, a transmit power of base station, and a circuit power which accounted for power consumed in the circuitry or dissipated in the form of heat. The base station power consumption was calculated by assuming that both transmit power and circuit power were independent of the base station utilisation rate. In addition, it was assumed that all base stations in the simulator were transmitting 100% of the time according to a full buffer traffic model.

The energy efficiency of the system was estimated by the ratio between downlink throughput and base stations power consumption within a given period. A broad conclusion by Huq et al. was that the energy efficiency of the CoMP improved as the number of users increased, and that CoMP downlink technique could improve the energy efficiency of an LTE network. The energy efficiency was estimated to be around 130 Kbits/Joule when the number of users was 100.

A more recent study by Huq et al. [11] estimated the energy efficiency of a heterogeneous network with coordination between the traditional macro-cell and lower power nodes (such as picocell and relays) using a similar approach to that in [5]. The basic simulation setup was the same as in [5] and again full buffer traffic model was assumed. The base station power consumption was assessed with more detail and discretion. The power model in this study considered the power consumption of power amplifier by scaling the transmit power with the power amplifier efficiency, and a fixed and a dynamic part of circuit power. The dynamic part of circuit power was assumed to scale linearly with base station transmission rate.

Huq et al. compared the energy efficiency of three scenarios where different number and combination of low power nodes were deployed surrounding a macro base station. It was found that 2 pico-cells and 4 relays overlaying with one macro base station provided the best energy efficiency performance with 25% improvement; however, further increase of the number of low power nodes could lead to degradation of energy efficiency.

Hunukumbure et al., 2013

Hunukumbure et al. examined the energy efficiency of cell edges users under CoMP coordinated beamforming [82]. It was estimated by Hunukumbure et al. that the efficiency of energy use by cell edge users improved by a factor of 2 under the proposed coordinated beamforming scheme, when compared to a non-CoMP MIMO network.

For base station energy consumption estimation, a base station power model from EARTH project [28] was adopted, in which the power consumption of base station consists of two parts – an idle power and an incremental power consumption that scales linearly with the number of resource blocks transmitted. Only the incremental power consumption was concerned as this study focused on estimating the incremental energy costs for serving a cell edge user with a certain data rate in CoMP systems. The energy efficiency metric was calculated as 1Mbps data rate over the incremental power consumption required to support this data rate for the user.

The incremental power consumption was estimated based on the cell edge user's link level transmission rate, which determined the number of resource blocks required for the user to achieve the desired data rate. Hunukumbure et al. pointed out the power requirement for extra signalling and control in CoMP scheme, and introduced a 20%

overhead on top of the estimated incremental power consumption. The percentage of CoMP overhead was however determined arbitrarily. The link level transmission rate of cell edge users was estimated by combining adaptations of generic Shannon spectral efficiency limits, realistically measured channels data for cell edge, and SINR values of cell edge user obtained from the system level simulator. The system level simulator involved 3 tier of base stations (19 base stations in total) based on the 4G wireless network model. Full buffer condition was assumed and the users were placed uniformly at random locations.

Hunukumbure et al. estimated the SINR values of cell edge users under CoMP by simply removing the strongest interference source and re-calculate the improved SINR in the simulator. The theory behind was an interference cancellation (or nulling) scheme in which a neighbouring coordinating base station could null the interference on the resource blocks used by a cell edge user by specific precoding. It was noted by Hunukumbure et al. that complete nulling could not be achieved in practical systems like LTE. Reasons were that the precoding vectors in LTE were drawn from a finite codebook, thus they could not completely match the theoretically computed nulling vector. Some residue interference would remain. It was also noted by Hunukumbure et al. that a prerequisite for the interference cancellation to function was the availability of the channel information of all coordinating base stations. For this study, the power required for exchanging channel information between base stations was ignored.

Hunukumbure et al. pointed out that this interference cancellation scheme came at a cost to the in-cell users (users not at or near cell edges) as the precoding vector used to null interference to cell edge users would not normally be the best precoding for in-cell users. The reduced throughput of in-cell users was only evaluated from a spectral efficiency point of view. Its impact on the energy efficiency of the whole network was not evaluated in this study.

Eluwole et al., 2012

A study by Eluwole et al. investigated into the energy efficiency of CoMP joint processing network with varying base station inter-site distance [12]. It was noted by Eluwole et al. that the evaluation of network energy efficiency should consider the various energy consuming components in the system. The power consumption model in this study included backhauling power and power consumption for extra signalling

processing associated with base station coordination, as well as the conventional linear base station power model.

For backhaul power modelling, Eluwole et al. used the same approach and linear model to that used by Fehske [10]. The backhaul power consumption was estimated by assuming the backhaul traffic load to be the theoretical maximally possible peak data rate for LTE downlink. Estimation of power consumption for extra signalling processing was based on a simplifying assumption that the extra signal processing power for cooperation was roughly 50% of the baseline signalling processing power per base station.

The system throughput of the joint processing network was estimated by carrying out an LTE-based simulation with seven hexagonal base stations and each having three randomly located users. Eluwole et al. estimated the energy efficiency of the CoMP system to be 8.5 J/Mbit when the inter-site distance was 1000 meters, while the energy efficiency of non-CoMP system was estimated to be 11.2 J/Mbit for the same inter-site distance. It was also found by Eluwole et al. that the energy efficiency improvement was achieved especially when the inter-site distance was small (less than 1200 meters).

2.6 Summary of Research Gaps

The above sub-sections reviewed the previous studies on the energy aspect of downlink CoMP scheme. In this section, we provide a summary of the research gaps identified in these studies and our approach in addressing these gaps.

In the previous studies reviewed, there were four papers that specifically focused on evaluating the energy efficiency of a CoMP network. The other two papers also included analysis of downlink CoMP scheme but involved slightly different objective - one focused on the energy efficiency of cell edge users, the other evaluated energy efficiency of individual channels. For CoMP downlink scheme, the majority of these studies considered the category of joint processing as opposed to coordinated beamforming/scheduling.

The evaluation of CoMP network energy efficiency involves an estimate of network throughput and an estimate of the energy consumption of the network. In estimating the energy consumption of the network, all previous studies investigated and modelled power consumption of base station. While some studies only considered base station power

consumption, a few studies attempted to extend the power model to cover power consumption of backhaul and additional signal processing for supporting the coordination.

Although the previous studies in the literature provide estimates of the energy efficiency of CoMP network, they exhibit a number of shortcomings. To begin with, energy modelling of the network is not complete and/or practical. The energy consumption estimate should take into account all network components and processes to deliver the information bits, especially those extra process and function introduced by CoMP.

Most of the work only included a power consumption figure for the conventional non-coordinated macro base station. The extra signalling processing energy associated with CoMP was either missing or estimated based on simplistic assumptions. The estimate in the study by Fehske [10] was based on the assumption that the signalling processing power related to channel estimations scaled linearly with the number of base stations in a coordination group. Hunukumbure [82] assumed that the power consumption for extra signal processing in CoMP was 20% of the baseline signal processing power, while Eluwole [12] assumed the extra signalling processing power to be 50% of the baseline signalling processing power. The percentage assumed in both studies was arbitrary with no analysis on the signalling effort required by CoMP.

Fehske's paper [10] was the only study that considered signalling overhead introduced by CoMP; however, the estimation was theoretical, and not based on a practical signalling scheme. The rest of the studies reviewed had no mention of supporting signalling scheme for the joint processing or coordinate beamforming strategy, and no signalling overhead was taken into account. The amount of signalling overhead incurred in CoMP schemes is a practical concern as it will reduce the downlink transmission rate and thus energy efficiency. The exact signalling overhead and signalling processing power can only be ascertained by analysing a practical CoMP signalling scheme.

The impact of the backhaul power consumption should also be analysed in the energy efficiency analysis of CoMP network. In current literature, the backhaul power consumption was either ignored [5, 11, 82], or estimated based on linearized power model [10, 31] without modelling the energy consuming components in the backhaul network. Eluwole [12] assumed the traffic load carried on backhaul to be the theoretical maximum data rate of LTE downlink without estimating the real load of information needs to be exchanged based on the CoMP scheme.

Another limitation of the studies on coordinated beamforming is that the energy efficiency is often estimated based on theoretical upper bounds of the network throughput, which cannot indicate the realistic throughput that could be achieved in a practical LTE network. In the study by Heliot [31], the channel throughput was estimated based on an information theoretic approximation assuming an idealized fully coordinated CoMP system. While the study provides valuable information regarding the theoretically maximum achievable channel rate, the identified throughput and energy efficiency gain is hard to realise in non-ideal CoMP systems when the practical issues are taken into account. The study by Hunukumbure [82] also overestimated the throughput by assuming that the strongest interferer to cell edge users could be completely eliminated by applying a theoretically computed nulling vector to the interference. However, in LTE systems, the precoding vectors can only be selected from a finite predefined codebook and can only approximate the computed nulling vector.

Furthermore, all the previous studies that carried out system level simulation [5, 10, 11, 82] considered a full buffer traffic model in their simulation setup. Under the full buffer traffic model, all base stations run at full power for 100% of the time and users always have data to download. While it simplifies the simulation and provides information on the performance corresponding to full buffer condition, it is not an appropriate traffic model for modelling realistic traffic. In practice, the network is not expected to be fully loaded and run at maximum capacity. The traffic load could be low at night or at non-busy hours; the energy efficiency of the network could potentially be much poorer at those times. Thus, the energy efficiency estimate from the full buffer traffic model cannot characterize a practical network under varying traffic, nor can it provide insight on the dependency of the network energy efficiency on the traffic load.

To sum up, there are a number of shortcomings and limitations in the current literature pertaining to estimating the energy efficiency of a CoMP network. As a result, these studies would give limited insight into the effect on the overall network energy efficiency through the use of a particular downlink CoMP strategy. They are also not particularly useful to the research into the trade-off between capacity gain and energy costs brought by CoMP and the design of more energy efficient CoMP strategies in a practical network.

In this thesis, we perform a thorough evaluation on the energy efficiency of a practical CoMP coordinated beamforming strategy. We take a pragmatic approach and estimate the

system throughput by carrying out a system level simulation of an LTE coordinated beamforming network, while taking into account the throughput reduction due to signalling overhead required for the coordination scheme. The signalling overhead is explicitly quantified based on a practical signalling scheme, and we incorporate 3GPP standards and trends in signalling functionalities in the development of the signalling scheme. In addition, a realistic traffic model with varying traffic load is applied to the simulated coordinated beamforming network for performance evaluation purpose. Our research therefore provides a more realistic assessment of the throughput and energy efficiency of CoMP system, and provides an understanding of the energy efficiency performance at different time of day (i.e., busy hour, night). It can also forecast the energy efficiency in future by incorporating the trends of traffic growth into the traffic model.

Our energy modelling goes beyond the classical base station power consumption model and the theoretical approach of estimating CoMP overhead in previous attempt. We include a base station power consumption that takes into account the increased number of RF chains and circuit power for CoMP. We also include backhaul, additional signal processing, as well as other energy costs associated with coordinating base stations and scheduling resource across the network architecture. The CoMP overheads such as the amount of information exchanged on backhaul and the additional signal processing effort are realistically estimated in the context of an LTE network based on the coordinated beamforming scheme and signalling scheme developed. We therefore present a complete and realistic assessment of the overall energy consumption in a CoMP wireless network. By analysing each of the energy consuming components and processes, our work also provides an understanding of the relative energy consumption in individual domains and subsystem of the network, and their importance to the overall energy efficiency of the network. The inclusion of the additional energy associated CoMP in the energy modelling provides insight into the trade-off between energy gains and costs that CoMP brought to the whole system. It also provides insight into the design consideration towards an energy efficient CoMP system.

Chapter 3

Overview of Research Problems and Tasks

This chapter will provide an overview of the research problems for quantifying the energy efficiency of CoMP beamforming system. It will also present a brief introduction to the research tasks to be carried out in the following chapters.

3.1 Research Problems and Challenges

WE have discussed the benefit of CoMP and pointed out the lack of comprehensive study on the energy efficiency of CoMP system in Chapter 2. This thesis is devoted to the development of methodologies for determining the energy efficiency of a CoMP coordinated beamforming system which takes into account signalling capabilities in LTE standards, effective transmission rate and throughput considering the signalling overhead required for coordination; and the expended energy for processing signalling and coordination information and for backhauling. It also takes into account the fairness of resource allocation, and realistic traffic model with varying traffic load. In addition, it focuses on the development of a coordinated beamforming solution that is scalable, practically realizable, and adheres to LTE standards and development trend. In particular, the coordinated beamforming will be studied in the context of MIMO, which is an integral part of LTE, to ensure that the coordinated beamforming solution supports the benefits provided by MIMO.

Although there are remarkable gains from an idealized fully coordinated CoMP network [31], it is not realistic in implementation due to the high level of signalling, backhaul network traffic load, and computation complexity [6]. Therefore, in practice, only a limited number of base stations can be coordinated in order to keep the overhead

manageable [6]. In this work, we limit the coordination to a cluster of three adjacent base stations to keep a reasonable overhead while exploiting the advantage of coordination inside each cluster. This cluster will be referred to as a coordination cluster. We will take into account the data rate constraints imposed by interference from surrounding clusters and base stations to avoid border effect and ensure the results are valid for larger scale network.

To model a network with coordinated beamforming, the first step is developing a strategy to determine the beamforming pattern to use for transmission and assign resource to users, such that the overall performance of a cluster under coordination is maximized. There are three sub-problems involved here. One is coordination problem – determining which transmit beamforming vectors or precoders to use at each base station. Note there could be many possibilities or combinations in a coordination cluster. Then, another problem is link adaptation problem – determining which modulation and coding schemes (MCS) is most suitable for each transmission. LTE enables channel-adaptive variable-rate transmission to a user and the link adaptation is enabled by channel estimation and feedback. There is also the PRB assignment problem – determining which subset of users should be allocated resources in each subframe, and how much resource should they each be allocated. Although each base station allocates its own resource to its serving users, the decision is made collectively in a cluster through coordination in CoMP system.

As can be seen, the coordinated beamforming strategy is a complex network optimization problem that jointly determines optimal beamforming vectors, the appropriate users to transmit to, the amount of resource allocated and the modulation and coding scheme to be applied. The above coordination and scheduling sub-problems are inter-connected and need to be solved concurrently. In the 3GPP standard, the CoMP strategy and resource scheduler are both implementation dependent, and how they are realised by equipment suppliers will greatly affect the performance of an LTE system [83].

The design of a coordination and scheduling strategy depends on what optimization objective is to be achieved: examples include maximum rate, fairness, and other quality of service (QoS) requirements [84]. There are many studies on resource scheduling in a single cell and some extend to multi-cell scenario [85, 86]. However, finding the optimum solution such as maximising throughput in single base station is already very computational complex and the algorithms are often based on perfect channel knowledge

which is not feasible in real systems. Thus in this work, instead of seeking to determine an overall optimal solution for all the aforementioned sub-problems in CoMP scenario, we will design a computationally efficient coordination and scheduling strategy that dynamically form beam patterns and allocates resources to users in a cluster while keeping a balance between the desired QoS and optimizing the overall system throughput.

An added complexity to the coordination and scheduling is that link adaptation in CoMP is quite different from non-coordination scenario. As mentioned, the link adaptation is enabled by channel estimation and feedback. The feedback sent by the users in the uplink is mainly channel state information (CSI) that contains the user perceived channel properties and information on the signal to interference and noise ratio (SINR) [87]. The knowledge of CSI enables the transmission to adapt to the current channel condition, such as the dynamic selection of an MCS that provides the highest possible data rate while maintaining reliability.

In a non-coordination network, the link adaptation process just involves user equipment measuring the channel from its serving base station, reporting the CSI to the serving base station, and then the base station dynamically chooses an MCS that best suits the channel for data transmission. However, in a CoMP network, it is not sufficient to just obtain the CSI measured from the current serving base station. The functioning of coordinated beamforming would also require knowledge of users' CSI and SINR under multiple coordination scenarios. This knowledge is necessary because that the base stations in a cluster need to collectively determine the beam patterns for transmission. Different combination of beam patterns will produce different signal and interference scenarios to users. Therefore, the knowledge of each user's channel quality or SINR under all beam combinations available for selection is critical in making an optimal decision on beam pattern combination to use for transmission.

As a result, for a user, many CSIs and SINRs would need to be measured or estimated, with each corresponding to a different beam combination available for selection within the coordination cluster. This knowledge is also an important input to the resource scheduler, as it indicates the amount of resource required by users to achieve a certain throughput for different coordination scenarios. Other consideration in resource scheduling also includes throughput in previous transmission and quality of service constraints etc.

The required CSI and SINR under coordination are not supported by the channel estimation and feedback processes used by base stations operating in a non-coordination scenario. Thus, the channel estimation and feedback design must be enhanced in CoMP in order to provide sufficient support for the coordination and scheduling strategy. The base station reference signals for users to listen to and perform channel estimation need to be redesigned to enable the multi-cell channel estimations that are required for CoMP coordination among multiple base stations [15]. In addition, new user capabilities to listen to the CoMP reference signals of the serving base station as well as other coordinating base stations are required. Note that the introduction of reference signals for CoMP will result in a loss of throughput, as the achievable downlink rate will be degraded with the configuration of more reference signals that consume resources originally provided for data. There will be an energy cost associated with the reduced rate.

In this work, we will design a signalling and control scheme that has low overhead while being able to support the coordination and scheduling scheme with the required CSI information. To ensure our system aligns LTE operational principles, the design of this signalling and control scheme will leverage off current standards and trends in LTE network deployments. Compliance to these standards is to ensure the design could be feasibly implemented with minimal or no modifications to LTE equipment and the latest technical advances made in LTE developments are accommodated. In LTE network, the CSI is usually quantized and user only reports back a few indices [87]. We will base our signalling design and overhead estimation on this kind of CSI feedback to ensure that it is of practical relevance.

As discussed in Chapter 2, while the coordination between base stations could provide improved spectral efficiency and throughput, all the coordination schemes come with additional costs such as additional antennas and radio chains, more channel estimation effort and signalling overhead, additional signal processing and computation, extra load on backhaul, and possible changes and new components in system architecture. From an energy efficiency perspective, these additional functions will incur additional energy costs. We will develop energy models and quantify the additional energy costs associated with the designed coordination and scheduling strategy and signalling scheme. The trade-off between attainable gains in cell throughput obtained from coordination, and energy costs associated with supporting the coordination, needs to be investigated in order to evaluate the energy efficiency of networks with coordinated beamforming.

In terms of the changes to system architecture, the information sharing required by CoMP could be supported by a centralized architecture or distributed architecture [15]. We will adopt the centralized architecture for our baseline network architecture, and focus on the Cloud Radio Access Network (C-RAN) architecture [88] when ascertaining overheads and energy costs. The coordination and scheduling schemes, as well as the signalling and control scheme developed in the work will also be applicable to other LTE architectural arrangements. The C-RAN architecture was chosen as current LTE deployment trends to include the centralization and virtualization of baseband processing units (BBU) to improve network efficiency and streamline the provision of advanced LTE functions such as CoMP [88]. In this work, the C-RAN architecture is adopted mainly to facilitate the information exchange in the coordination scheme, the energy savings achieved from BBU pooling is not evaluated in the energy efficiency study in this thesis.

3.2 Research Tasks

The detailed solution to the aforementioned research problems will be developed in the following chapters. Here we provide an overview of the three main research tasks involved in solving the problems and developing an estimate of the energy efficiency of CoMP coordinated beamforming network.

3.2.1 Capacity Model

The first part of work is to build a system level model for a coordinated beamforming network which will be used to ascertain the throughput of the network, and study the performance gains of coordinated beamforming in comparison to a conventional non-coordinated network. This work will be discussed in Chapter 4.

The work involves designing a coordination and resource scheduling strategy with key design criteria that it is suitable and easily implementable for deployment in a realistic network. Accordingly, we aim for a balance between performance and complexity. Based on the system level model, we evaluate the total capacity of the network at system level as well as the user perceived download speed. The total capacity is estimated as the average achieved rate of all users over time based on a traffic model that emulates uniform data download from all locations within a base station coverage.

The features of the model include modelling of clustered base station coordination in groups of three, and modelling of interference from base stations outside the coordination cluster. In addition, the system setup and simulation parameter are from LTE network, and a realistic traffic model with multiple users and varying traffic loads is simulated.

In the capacity model detailed in Chapter 4, we use simplifying assumptions when estimating the throughput and energy consumption of the system. In particular, the SINRs of all users under all coordination scenarios is available through a 'to be designed' signalling scheme which is subsequently addressed in Chapter 5. Another assumption is that base station is the only energy consuming component in the network. Essentially, costs of CoMP are not fully accounted for yet. The estimate will be modified in later chapters when the estimate of signalling overhead and energy consumption of the rest of the network is developed. The preliminary estimate at this stage represents an upper bound to the achievable energy efficiency. It is used to examine the potential gains that coordinated beamforming could provide and study the trade-off between energy gains and costs of CoMP.

3.2.2 Signalling Scheme

The second part of the work is to design a signalling and control framework to provide the coordination and scheduling scheme with the required signalling information and to estimate the signalling overhead introduced. This part of work also presents a methodology for the additional signalling overhead, the amount of signalling information exchanged on backhaul, and the computation effort for processing the user CSI feedback to be quantified for CoMP systems in the context of an LTE network. This work is discussed in Chapter 5.

We made the assumption that the SINRs of all users under all coordination scenarios are available and these SINRs are important inputs to drive the coordination and resource scheduling. In order for the identified performance gain to be valid, we need to justify the information required for signalling can be acquired.

As discussed, the introduction of new CoMP reference signals for user channel estimation of coordination scenarios will result in a loss of resource and throughput. Besides, the user's CSI needs to be constantly measured and fed back to the base station because the users and their channel conditions are not static in reality. Failing to update quickly

enough will lead to the coordination scheme not performing to its best. We will develop strategies to reduce the amount of reference signal and user CSI feedback while still providing sufficient SINRs information and their updates to the coordination scheme.

LTE release 11/12 proposed CoMP and designed new reference signal and signalling functionality to support CoMP in general [89]. However, the configuration specifics are left for implementation as different CoMP schemes require different CSI feedback and signalling configuration [83]. We will focus on making use of and adapting to the procedures and signalling protocols defined in LTE release 11/12 for the design of a signalling and control scheme that supports our coordinated beamforming network.

The signalling scheme not only provides an estimate on the amount of resource occupied by reference signal, it also provides information on the amount of signalling and control information needs to be exchanged on backhaul and further processed. It provides a foundation for estimating the additional energy cost of the CoMP network.

3.2.3 Energy Model

We study the impact of the proposed coordination and signalling schemes on the energy efficiency of the network by estimating the energy cost associated with new functions and architecture enhancements required to support the coordination and signalling. This work is discussed in Chapters 6.

The complete modelling of energy consumption provides a more accurate characterization of the energy efficiency of coordinated beamforming network. In order to quantify the total energy cost of the coordinated beamforming network, we develop a high level architecture of functional and energy consuming components in one coordination cluster of the network model, as shown in Figure 3.1. We analyse the signalling and control information flow, coordination data processing and the energy consumption associated with them based on this architecture.

In Figure 3.1, the BBUs are relocated into a centralized pool and leaving the remote radio heads (RRH) at base station site under the C-RAN architecture [88]. The 3 RRHs in Figure 3.1 represent the three base stations in the same coordination cluster, and the three corresponding BBUs are located in a central site. These RRHs are linked to the centrally located BBU site by CPRI interfaces [90]. The BBUs is connected to a switch, and then

the switch is connected to a serving gateway (SGW). All these components are present in a standard LTE C-RAN deployment, except for the central coordination unit (CCU) placed inside the SGW, at the right-hand bottom of Figure 3.1.

The CCU was added for performing data processing for the coordination and scheduling algorithm, and it will incur an additional computation energy cost. Functionally, such a processing unit is needed in any LTE-based CoMP scheme to support coordination, so its inclusion in our architecture is within the realm of a modern C-RAN implementation. The support of the proposed coordinated beamforming strategy has a minimal impact on the system architecture.
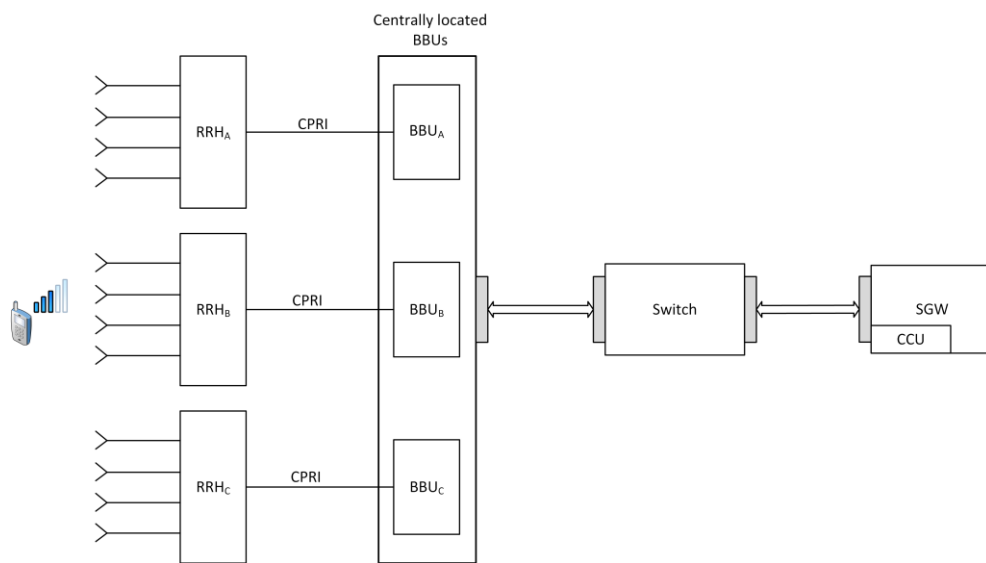
Figure 3.1 High level architecture for one coordination cluster

By adopting this centralized architecture, all the signalling and control information at the base stations such as the user CSI feedback can be passed to the CCU to enable information sharing and thus coordination. In the following, we present a brief description of the information flow on both uplink and downlink in the network architecture.

The key information flow and processing are:

1. User CSI feedback, plus other signalling and control information are sent uplink from users to the centrally located BBUs via CPRI
2. The desired CSIs for coordination and some user information (such as new users and users who have completed their download) are extracted from LTE signalling messages at the BBUs and directed to the CCU via backhaul switches

3. The CSIs along with the user information serve as input to the coordination and scheduling algorithm running at the CCU, to determine the optimum beamforming patterns for transmission and resource allocation

4. The coordination decisions generated at the CCU are sent back to the centrally located BBUs, including the beam pattern decision, the users' resource allocation instruction and expected channel conditions under the chosen beam patterns

5. The BBUs generate the transmission data (IQ data) for the users based on the coordination decision

6. Appropriate beam patterns formed at the (3) RRHs

Information including CSI feedback, beam pattern and scheduling information are passed between the different components thereby generating extra traffic load on the backhaul. The extra traffic will be associated with extra energy consumption. The processing of the extra traffic also requires additional function and energy consumption. Chapter 6 will provide a discussion in more detail and estimate the amount the backhaul load and computation effort.

Chapter 4

Coordinated Beamforming using Gated Narrow Beams

4.1 Introduction

IN Chapter 2, we have identified that in the current literature, there is a lack of thorough evaluation of energy efficiency of CoMP Coordinated Beamforming systems in the context of LTE. An accurate estimation of energy efficiency of such systems will need to take into account some key factors affecting the throughput of the system, such as the interaction between beam coordination strategy and resource scheduling strategy, realistic user traffic modelling, and signalling overhead for coordination purpose. The energy efficiency estimation would also need to fully account for the energy cost associated with functions needed to support the coordination, where a complete energy modelling of the whole network architecture is required. As discussed in Chapter 3, solutions to these sub-problems will be developed by steps in the following chapters. The results will be combined to estimate the overall system energy efficiency.

This Chapter will focus on ascertaining the network throughput of a CoMP Coordinated Beamforming system using gated narrow beam (GNB). A novel GNB architecture is designed which is scalable for large area deployments, and also adaptable to use with existing MIMO antenna systems. A beam combination and resource scheduling scheme is developed to mitigate interference between neighbouring base stations, while ensuring the efficiency and fairness in resource allocation. A key feature of this scheme is managing interference between the coordinated narrow beams by selecting the best narrow beams to be turned "On" in a coordination area.

In addition to the study of throughput, a preliminary estimate of energy efficiency of the GNB system will also be given. The performance of the GNB system is compared against a conventional wide beam system to demonstrate the potential energy savings of GNB system. The preliminary energy efficiency estimation in this chapter is based on several assumptions. Namely, the energy costs and resource usage associated with coordination signalling can be ignored. The energy cost of exchanging and computing coordination information can be ignored. Chapter 5 and Chapter 6 will include the study of these energy costs and signalling resource usage. The impact of these assumptions on energy efficiency will be evaluated and the preliminary energy efficiency estimate will be updated.

The detailed gated narrow beam architecture is proposed in Section 4.2. The section first discusses the potential benefits of a directional narrow beam comparing with a traditional sector wide beam, and then provides an overview of the implementation strategy. Section 4.2 also covers the generation of such narrow beams, coordinated beam gating strategy, and the development of the resource scheduling and gated beam coordination strategies. A system level simulation model is built for quantifying the performance of the gated narrow beam system in Section 4.3. Based on the simulation results, Section 4.4 analyses the performance of a network of gated narrow beams including throughput, energy efficiency and several other metrics. The performance is also compared against a network of wide beams, to demonstrate the achievable improvements. Section 4.5 compares the throughput performance of the gated narrow beam model with other CoMP strategies and schemes under a common reference scenario. Lastly, the performance of gated narrow beam under different traffic models is compared to analyse the impact of traffic model on the network performance.

4.2 Coordinated Gated Narrow Beams

4.2.1 Benefits and Scheduling

This section first discusses the potential benefit of using narrow beams generated through beamforming, and the importance of scheduling strategies (beam scheduling and resource scheduling) when narrow beams are being used. Then, an overview of implementation strategy for the rest of the chapter is provided.

Beamforming is able to provide a more confined antenna radiation pattern with high directional gain and enable reduced interference from other directions. LTE-A includes improved consideration and support for beamforming as a major tool to increase the SINR at the receiver and peak rates [15]. This thesis proposes to generate a pattern of narrow beams using beamforming which can be dynamically switched for transmission of data. The selection of beams to be switched 'On' or 'Off' during any one time epoch is performed by a coordinating and scheduling scheme which will be discussed in Section 4.2.5.

Figure 4.1 (a) shows a central and one ring of base stations; each having three sectors covered by traditional wide beams. In Figure 4.1 (b), wide beams are replaced by a few fixed narrow beams aligned in different directions. For each user, the most suitable narrow beam is selected for transmission during a time epoch. In the simulation developed for this work, the number of narrow beams per sector is adjustable and for this study, 4 narrow beams per sector are assumed. Figure 4.1 (c) gives one possible choice of beams for transmission during a particular time epoch; note that the design only allows at most one narrow beam out of the 4 to be 'On' at one time. The reconfiguration of On/Off state of the beams can be every LTE frame of 10ms or subframe of 1ms. In this study, 1ms reconfiguration interval is assumed.

(a)   (b)   (c)
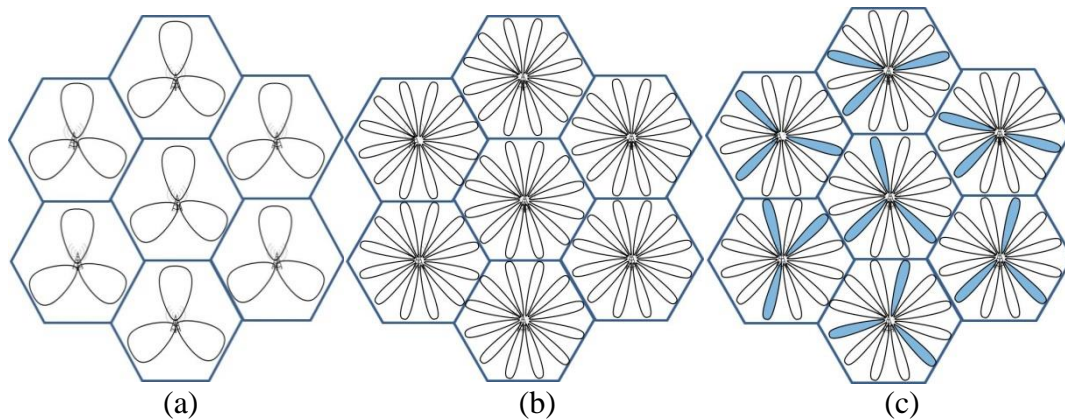
Figure 4.1 Beam pattern for (a) wide beams, (b) for narrow beams, (c) active narrow beams in one time epoch

Traditional base stations in a mobile cellular network radiate three sector-wide beams to cover a service area [15]. Much of the radiated power of the sector-wide beams is wasted as the beams aim to provide coverage and are not directed to the current users.

Supposing the radiated power from the base station in a narrow beam is selected to be the same for a conventional wide beam, the users will receive stronger signal, thus better spectral efficiency. The coverage range is also longer since narrow beam is more directional comparing with spreading the power across a sector-wide area. This is desirable when the narrow beam is from the user's hosting base station; however, the narrow beams from neighbouring sectors may interfere and if so the interference will be stronger as well. Therefore to achieve the potential benefit of narrow beams, a scheduling algorithm for interference management is essential, and the objective is to find out a proper coordination scheme.

A straight forward algorithm is round robin, where each beam will be allocated the same amount of time in a circular order. However in practice, the number of users in each beam coverage zone is not the same at any particular point of time, the users can be clustered, or there could be no user at all in one beam coverage zone. An example of possible users' locations is shown as dots in Figure 4.2. For better throughput and use of energy, beams should be scheduled according to need, in terms of the number of users or number of download requests. A further concern is interference management. Performing round robin separately at each base station (allocating the same amount of time for each beam in a circular order) will not be able to prevent interference between base stations. The base stations need knowledge of the traffic demand and collaboratively determine the optimal beam patterns.
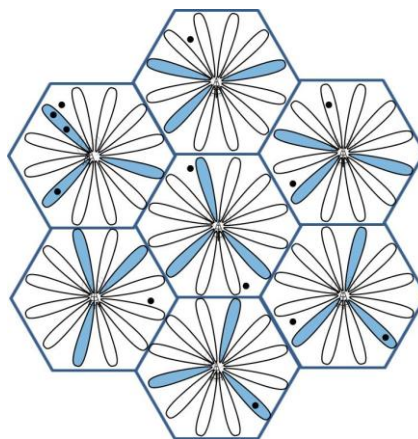
Figure 4.2 A snapshot of narrow beams operating in a playground of 7 base stations

Based on the above discussion, the resource scheduling and beam coordination strategies have to be developed coherently and not independently of each other. The design and implementation of the GNB model include the following key procedures:

1. Firstly, the mechanism for narrow beam generation is investigated. There are different mechanisms available to generate a directional beam. The narrow beam we propose in this thesis is based on digital beamforming where the beam is formed by applying predefined precoding matrices to change the phase and amplitude being output by the radio chains to the multiple antennas. This will help ascertain the associated additional energy in coordinating the narrow beam, including the energy cost at the base station (more antenna and radio chains) and the signalling and control information required to generate them.

2. Next, the gated narrow beam model is proposed with a beam coordination scheme and resource scheduling scheme. Section 4.2.3 first presents the model with the basic principles of how narrow beams within a sector are being turned on and off to serve users. The number of possible beamforming patterns, or narrow beams combinations, within one coordination area (3 neighbouring sectors) is analysed. Section 4.2.4 and 4.2.5 developed beam coordination and resource scheduling scheme which together select the optimum combination of beams to be gated "On/Off" for transmission during a time epoch. In principle, the resource scheduler performs proportional fair scheduling within users in the same sector. The coordination scheme determines the best narrow beam combination to mitigate interference by performing an exhaustive search on all feasible narrow beam combinations with the objective of maximising a defined utility function.

3. In order to evaluate the throughput and energy efficiency of the proposed gated narrow beam model, an LTE network of base stations with narrow beams is simulated. The simulation model adopts parameters from LTE systems, realistic base station power model, and applies realistic FTP download traffic model. Key simulation parameters are summarised in Section 4.3.1. Section 4.3.2 presents the base station power model. The energy consumption of base station is the only energy cost we have included in this chapter, the network has more energy consuming processes and components which will be identified and taken into account in Chapter 5 and 6. A scalable cluster of base station sectors is set as simulation area with interference from outside the cluster also modelled. Section 4.3.3 shows the simulated narrow beam coordination area.

4. As discussed, a network of conventional wide beam base stations is also simulated to demonstrate the advantages of the gated narrow beam model. The key metrics for network performance evaluation is presented in Section 4.4.1. Section 4.4.2

presents and discusses the simulation results, which include throughput, energy consumption, energy efficiency as well as other metrics of gated narrow beam model and wide beam model. Section 4.3.4 analyses the impact of traffic models on network performance by comparing the FTP traffic model with varying traffic levels, and a full buffer model in which all base stations transmit at full power all the time.

4.2.2 Narrow Beam Generation

This section discusses the various methods to generate a narrow beam, and the method we proposed to use in this design. This will set the foundation for ascertaining the signalling and control information needed in generating narrow beams and the associated energy costs.

The proposed narrow beam patterns are developed through the use of multiple antennas and beamforming. There are several methods to achieve beamforming. Firstly, beamforming can be done at the analog domain or digital domain [91]. Analog beamforming is also called passive beamforming where all phasing and amplitudes are controlled by a power divider or phase shifters inside the antenna. For digital beamforming (which is also called active beamforming), the beam is steered and shaped by changing the relative power levels and phases being output by the radio chains to their antennas. In this case, each antenna is fed by a dedicated radio, and a number of antennas collectively form a beam pattern. The technique used in LTE to change the weightings including phase and amplitude is to apply a specific precoding matrix [92].

Both methods could realize the proposed narrow beam but digital beamforming has some advantages over analog beamforming. Digital beamforming adopts reference signal per antenna rather than per beam [93] (the details will be presented in Chapter 5), but analog beamforming will have to include reference signals for each beam in order for channel estimation. As a result, channel estimation cannot be performed for non-scheduled beams. Another advantage is that, the way of applying the beamforming weighting by precoding matrix is consistent with other MIMO techniques such as diversity and spatial multiplexing. This consistency enables the opportunity of combining beamforming and other MIMO techniques [94]. In this design, we will adopt digital beamforming using precoding matrix as defined in LTE.

For digital beamforming solution using a precoding matrix, there are the two choices: codebook based precoding/beamforming and non-codebook based precoding/beamforming [91]. With codebook-based precoding, a codebook with a limited set of possible precoding matrices is predefined and each precoding matrix is associated one to one with the set of narrow beams.

Non-codebook based precoding/beamforming, also known as user-specific beamforming, allows an arbitrary and adaptive precoding matrix for each user without the constraint of a codebook. There are user selection and precoding strategies proposed for user-specific beamforming, such as zero-forcing [95], which eliminates the interference by finding a linear precoding matrix in the null space of the transmit channels of the co-scheduled users. Another method is maximization of signal-to-leakage-plus-noise ratio [96, 97], which aims to maximize the signal power of the transmission data intended for one user while minimizing the interference leakage to co-scheduled users.

However, these user-specific beamforming algorithms have limitations when applying in the real network [95]. The number of users that can be co-scheduled is limited [95]. These algorithms also normally require complete user Channel State Information feedback in order to calculate the optimal beamforming matrix; however, the feedback is quantized in a practical network. Furthermore, these algorithms are already complex for single site and will become more complex when used in CoMP system. As the interference generated from neighbouring base station depends on the user that is scheduled, the base station can only select its own users and precoders assuming certain users are tentatively selected by other base stations. The user selection problem and precoding matrix decision can only be performed in a joint and iterative fashion for the coordinating base stations without achieving optimization [98, 99].

Given the above discussion, codebook-based precoding/ beamforming is chosen as the method to generate the gated narrow beams. This choice also takes into consideration that the nature of the designed gated narrow beam is fixed and each sector includes only a limited set. In this thesis, 8 transmit antennas are used to generate 4 directional beam patterns based on codebook-based beamforming. Previous work on grid of beam [100] has taken a similar approach to generate the fixed narrow beam. Chapter 5 will analyse what user feedback is needed to select the precoding matrix from the codebook. Note that the codebook design is out of the scope of this study.

4.2.3 Gated Narrow Beam Model

Theoretically, CoMP gains increase as the number of coordinating base stations increases [84]. However, a centralized joint multi-cell scheduling optimization problem has been considered impractical due to computational complexity, signalling overhead, user feedback overhead and backhaul traffic in early works [101, 102]. Therefore a global optimization of narrow beams is not practical. In this thesis, we consider a more limited coordination scale and the resulting performance is examined.

As inter-cell interference is the major cause of degradation [64], the coordination scale is set to three sectors facing each other from three neighbouring base stations, rather than the 3 sectors that belong to one base station. The 3 sectors form 3 transmission points, or simply referred to as points as defined in 3GPP as a set of geographically co-located transmit antennas [84]. Note the sectors of the same base station correspond to different points. The design is in accordance with 3GPP Release 11 recommendations that the maximum permissible number of points is 3 to keep the overhead manageable [84, 103].

Figure 4.3 (a) shows the cluster of base stations and hexagonal service areas. Each coloured hexagonal area is a coordination cluster, which consists of three coordinating sectors. The coordination is modularized hence ensures the architecture is scalable for large area deployment. In this work, base stations are identical and in a regular hexagonal pattern. The coordination clusters are fixed and will not change over time. Figure 4.3 (b) is one of the coloured hexagonal areas with representative users and beam pattern.

(a)   (b)
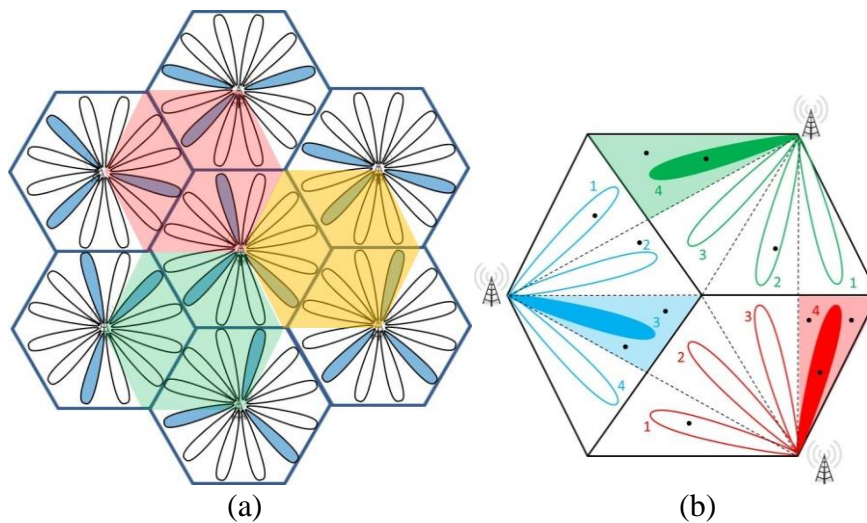
Figure 4.3 Coordination modularization and one possible beam combination in a coordination cluster

In each transmission epoch, one narrow beam out of the four possible beams in each sector may be turned 'On' if there are user(s) in the sector to be served. Alternatively, all beams are turned off if there is no user in this sector, or it is beneficial not to turn on any beam in order to avoid generating interference to users in other sectors within the coordination cluster. Since there are a total of 5 beam choices in each sector (none, or one of the 4 beams), the scheduling of resources within a coordination cluster would need to consider a total number of 5x5x5 =125 beam combinations.

Figure 4.4 illustrates the concept of beam selection to mitigate interference. Note that for each beam combination pattern, each user has a unique signal to noise ratio (SINR). Certain combinations cause worse interference and result in a poor SINR for users. For example in Figure 4.4, users under the red beam will have a worse SINR and spectral efficiency in scenario (a) than in scenario (b) where green and blue beams are shifted away when the red beam is 'On'. From an SINR point of view, the desirable combinations are the ones which avoid direct interference and an aim of the beam coordination and scheduling scheme described in Section 4.2.5 is used to identify such combinations.

(a)   (b)
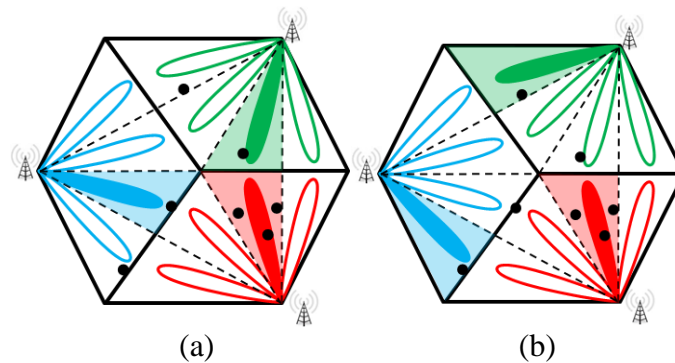
Figure 4.4 Interference level for different beam combinations

The proposed coordinated gated narrow beam approach is applicable for both low and high traffic conditions. The opportunity of scheduling beams to avoid interference improves as traffic load increases. As shown in Figure 4.5, in scenario (a), where there are 5 users in the coordination cluster, the coloured beams have to be turned on to serve the users since they fall within their coverage. In this beam combination, the orientations of the beams show that there will be significant interference. In scenario (b), the

additional 2 users lead to more opportunities for sensible beam combinations to be chosen, and the beam combination shown causes less interference comparing with scenario (a).

(a)   (b)
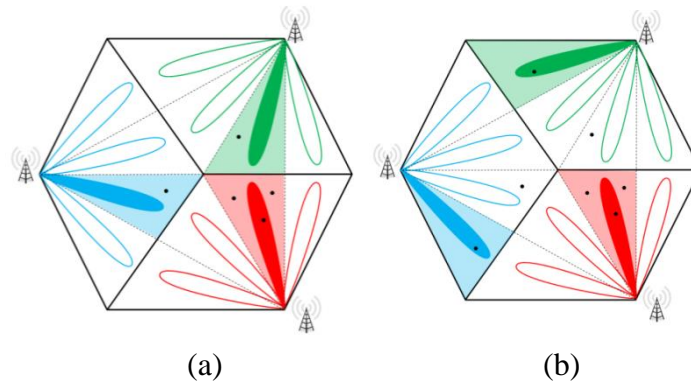
Figure 4.5 Dependency of signal interference on users in the service area

Examples in Figure 4.4 and 4.5 focus on the interference within the coordination cluster, the interference from the outside cluster is less significant but will be included in the modelling work in Section 4.3.4 to capture the total interference. The best beam combination (out of the 125) in the service area should be determined not only based on mitigating interference, but also ensuring fairness and efficiency in resource allocation in terms of throughput to the users.

4.2.4 Resource Scheduling

This section examines how users can be assigned system resources dynamically and efficiently in the proposed gated narrow beam model. As discussed in Chapter 3, the Physical Resource Block (PRB) assignment problem and beam coordination problem are inter-connected and affect the overall system performance. The base station needs to decide which user receives resources and how many resources to allocate. This decision is based on the predicted channel quality in terms of SINR, which in turn depends on the beam combination selected. In this Chapter, all users' SINR under different beam combinations is assumed to be available by channel estimation and the link adaptation problem is ignored. Chapter 5 will discuss how the channel estimation can be achieved.

For coordinated beamforming, each user is served by one sector of a base station. A further rule is made in this design that when a narrow beam is turned 'On' in a sector, the narrow beam will only allocate resource to the users that are within the coverage of this narrow beam. For codebook-based beamforming, part of the users' feedback to base

station includes information of which precoding matrix/beam is preferred (this will be detailed in Chapter 5). This information is used to assign users to their preferred beam. Only when a user's preferred beam is turned 'On', will the user be scheduled for transmission and be allocated resources. In other words, when a certain beam combination is selected, only the users in the coverage of the 'On' beams will be scheduled while the users in the coverage of the 'Off' beams which are very unlikely to have good SINR will not be served in this epoch, hence ensuring the best use of PRB resources.

Although this design decision is heuristic, it ensures that all users are receiving transmission with their best possible beam as in non-coordinating environment. Besides, it also greatly simplifies the overall coordination problem in that user selection and beam selection have become one combined integrated decision. Once a beam is selected, all the users within this beam could be scheduled for transmission and share the frequency-time resource of the transmission epoch.

Since coordination can potentially reduce interference, the user's SINR is improved in a coordinating environment than in a non-coordinating environment and is different under different beam combinations. However, under the assumption in this chapter that the SINRs are known, the resource allocation within one sector is no different to non-coordinating systems. Each sector allocates its own resource to its serving users independently of the other sectors.

The processor sharing (PS) discipline is an appropriate abstraction to model base station downlink [104]. In theory, the PS model assumes infinitely divisible resources, while in an LTE system, the smallest resource that can be allocated is one PRB. However, the scheduled time between update frequency is 1ms which is short compared to the time duration of downloads; hence the PS model is applicable [104].

The choice of the PS principle is critical in meeting Quality of Service requirement. One of the simplest and most well-known scheduling algorithms is Round-robin (RR), which assigns equal frequency/time resource to all users without requiring any user information [84]. RR is only suitable when all users have relatively similar channel qualities; hence it is rarely used in a cellular wireless network as the users' channel quality varies markedly between users near the base station and at the cell edges [84]. An opportunistic scheduler such as maximum rate (MR) utilizes channel state information and allocates resource to

the users with best channel quality with the aim for throughput maximization; however, it is biased against the users with poor channel quality and raises fairness issues [84]. The users at cell edge are rarely allocated resources. Therefore, it is not suitable for cellular network either. To achieve system throughput gain while maintaining a level of fairness among users, proportional fair (PF) [105] can be applied. Instead of making scheduling decisions based on the instantaneous channel quality seen by users, the PF algorithm considers the past and current channel condition. The PF scheduling algorithm is widely used in cellular systems [106-109] including LTE [110] and is adopted in the proposed gated narrow beam model.

4.2.5 Beam Coordination and Scheduling

The past average throughput as used in the standard PF scheduling algorithm is adopted here, and defined as a Weighted Historical Throughput metric *T*. The Weighted Historical Throughput *T* tracks a user's historical performance by means of a weighted sum of a user's past performance and his performance in the current scheduling epoch. The metric dimension is in bits/s.

Studies on dynamic resource scheduling in OFDM systems in a single cell have been commonly based on maximizing the total throughput or utility. In particular, the natural logarithm of the throughput is a typical utility function which is commonly used in network-utility-based maximization problems with diminishing returns [111-113]. This utility function is adopted to select the best beam combination in the gated narrow beam model. Hence, a utility function (*U*) is defined as the natural logarithm of the throughput of user *i*,

$$U_i = \ln(T_i) \tag{4.1}$$

The objective function used in selecting the best beam combination is to maximize the sum of utility of each user. Therefore, the objective function is

$$\sum_{i=1}^{N} U_i = \sum_{i=1}^{N} \ln(T_i) \tag{4.2}$$

where *N* is the total number of users in the coordination cluster. By taking the natural logarithm, priority will be given to users that have a lower average throughput (a throughput increase in the low throughput users will contribute more to the objective function), ensuring a level of fairness across the users.

The PF resource scheduler divides the total available resource in each sector to allocate to its serving users. Transport resources in LTE are allocated in Physical Resource Blocks (PRBs). A pair of PRBs is the smallest element of resource allocation assigned by the base station scheduler and each PRB contains a 200 kHz (f_{PRB}) wide block of frequencies and 0.5 ms (t_{PRB}) in duration [114]. The 200 kHz includes 180 kHz effective usable bandwidth and 20 kHz guard bands. The guard bands have been factored in by a mapping table adopted later for determining the spectral efficiency and data rate. Thus here a PRB is considered to carry data resource over the entire 200 kHz bandwidth.

Each sector of the LTE base station has a spectrum of 20 MHz, hence in a 1 ms time slot (t_{epoch}), 200 PRBs are available in total in each sector. Each PRB carries a different number of payload bits depending on the modulation format, lower SINR leads to fewer payload bits per PRB. The bits are encoded in OFDM symbols and each PRB typically consists of 7 symbols [115].

A user's Weighted Historical Throughput (T) is updated in every scheduling epoch (in the present model this is set to 1ms duration) as,

$$T_i(t) = \begin{cases} \alpha * T_i(t-1) + (1-\alpha) * P_i * S_i * f_{PRB} * t_{PRB}/t_{epoch}, & \textit{if scheduled} \\ \alpha * T_i(t-1), & \textit{if not scheduled} \end{cases}$$

(4.3)

where $T_i(t)$ is the updated weighted historical throughput for user i at time t;

$T_i(t-1)$ is the weighted historical throughput for user i at time $t-1$;

P_i is the number of PRBs allocated to user i at time t;

α is a constant fairness parameter that is used to weigh a user's historical and current throughput performance,

S_i is the spectral efficiency in bits/s/Hz for user i at time t.

The term $S_i * f_{PRB} * \frac{t_{PRB}}{t_{epoch}}$ represents the throughput per allocated PRB for user i at the current epoch. For the values of parameters discussed above this term becomes $S_i * 200000 * \frac{0.5}{1}$.

The S_i is estimated from SINR using a mapping table (will be discussed at end of Section 4.3.1). The SINRs are calculated based on theoretical approach in the simulation. The signal strength of individual narrow beam to a user location is calculated based on path loss, antenna gain, shadowing loss etc., and applied towards the calculation of SINRs. In practice, these SINRs are obtained by implementing appropriate signalling processes that generate specific users' CSI feedback. Chapter 5 will detail the signalling process and show how these SINRs are derived from CSI feedback.

The weighted historical throughput is updated after every epoch, both for users that have been scheduled to receive PRBs and those users who have not been scheduled. To obtain the overall utility for the 125 possible beam combinations in order for maximization, the updated Weighted Historical Throughput (T_i) is required for each user i for each possible beam combination.

Next, for each user, P_i is obtained by applying PF scheduling algorithm in each sector. Suppose that N_1, N_2 and N_3 are the numbers of users being served by beams that have been selected to be switched 'On' within sectors 1, 2 and 3 respectively, in the coordinated playground in a particular epoch. Note that $N_1 + N_2 + N_3 \leq N$, as not every user in the coordination cluster is served simultaneously. For sector 1 (similar for sectors 2 and 3), the number of PRBs $(P)_i$ allocated to user i within the coverage of the beam that is switched on for a given epoch is given by [116]:

$$P_i = \frac{1}{N_1}\left(R + \sum_{i=1}^{N_1} r_i\right) - r_i \qquad (4.4)$$

where $r_i = \frac{\alpha}{1-\alpha} * \frac{T_i(t-1)}{S_i*200000*0.5/1}$, R is the total number of PRBs available in each sector (equal to 200 when the spectrum is 20MHz), and α is the above mentioned constant fairness parameter.

Here we adopt the extended proportional fair algorithm described in [116], which enables direct calculation of the final allocations for each user in an epoch, instead of allocating available resource blocks one by one with the basic proportional fair principle in an iterative process [105]. Both methods give the same resource allocation decisions.

When $\alpha = 0$, the allocation does not take into consideration of past performance, and all users are allocated the same number of PRBs. As α approaches 1 (but should not be equal

to 1), the allocation places greater weighting on the ratio of historical throughput to current spectral efficiency, emphasizing more on fairness.

To sum up, the operation of the beam coordination scheme is based on selecting the best beam combination for the next transmission epoch. For each beam combination, the utility functions in Equation (4.1) for each user and hence the objective function for one beam combination is calculated. The best beam combination for the next transmission epoch, out of the 125 possibilities, corresponds to the choice where the objective function in Equation (4.2) is maximized. The value of a user's utility function depends on the users' Weighted Historical Throughput, which is a weighted sum of a user's past performance and its performance in the current scheduling epoch Equation (4.3). The user's performance in the current scheduling epoch essentially depends on its spectral efficiency under a certain beam combination, and the amount of resource allocated to it based on the PF scheduling algorithm Equation (4.4).

4.3 Coordinated Gated Narrow Beam System Level Simulator

This section develops a MATLAB-based system level simulator of an LTE network of base stations with gated narrow beams, in order to evaluate the throughput and energy efficiency of the proposed gated narrow beam model. A similar simulation is also conducted for base stations of traditional wide beams, as the performance of gated narrow beam model will be compared against that of the wide beam model in Section 4.4.2. A few other performance metrics including SINR and base station power consumption are also studied; a summary of the metrics will be presented later in Section 4.4.1. The focus of this section is on establishing the simulator. In this section, important system simulation assumptions are summarized, including base station and user parameters, radio channel link including path loss model and fading characteristic, traffic model, base station power model, and also base station layout used in the simulator.

4.3.1 Simulation Parameters

The parameters used in the simulation of narrow beams are summarized in Table 4.1 below. The parameters used for traditional wide beams are also shown in comparison. The simulation parameters used in modelling network of narrow beams and wide beams

are based on base station characterises used in the LTE network modelling by GreenTouch Mobile Communication Working Group [117, 118].

Simulation Parameter	Narrow Beam	Wide Beam
Maximum RF Launch Power	46 dBm	46 dBm
Idle RF Power	0 dBm	0 dBm
Number of Transmit Antennas	8	2
Transmission Scheme	Beamforming + SU MIMO	SU-MIMO
Noise Power (for SINR calculations)	-174 dBm/Hz (or -101 dBm for a 20 MHz bandwidth)	-174 dBm/Hz (or -101 dBm for a 20 MHz bandwidth)
Bandwidth	20 MHz	20 MHz
Maximum Antenna Gain	20 dB	14 dB
Horizontal Front-To-Back Antenna Ratio (A_m)	31 dB	25 dB
Vertical Front-To-Back Antenna Ratio (SLA_v)	25 dB	25 dB
Horizontal 3 dB Beam Width (φ_{3dB})	17.5°	70°
Vertical 3 dB Beam Width (θ_{3dB})	10°	10°
Vertical Antenna Tilt (θ_{etilt})	15°	15°
Antenna Height	32 m	32 m
Receiver Height	1.5 m	1.5 m

Table 4.1 Simulation parameters of narrow beam and wide beam base stations

To make a fair comparison, the wide beam model also has a gating function rather than having sector beams continuously on, hence the wide beams are turned 'On' and 'Off' depending on whether there is a user within its coverage, i.e. a user that is best served by the beam. Multiple users within the same wide beam receive equal PRB allocation. The maximum RF launch power of narrow beam base station is assumed to be the same as the wide beam base station, in order to provide the best possible throughput performance.

For both models, the idle RF power at no load is assumed to be zero so that there is minimal RF power transmission when a beam is gated 'Off'. For idle base stations, the time-frequency slots for the remaining signalling such as broadcasting signalling are reserved and will not interfere with user data from other base stations. The use of a separated signalling/control and data architecture proposed in GreenTouch BCG2 (Beyond Cellular Green Generation) [119] could also achieve the same effect.

The narrow beam base stations have 8 transmit antennas; cross-polarization is adopted as it is the current solution for LTE-A standards [15]. Hence the antenna configuration for narrow beam base station is 4 columns with cross-polarization antennas on each column and the columns are closely spaced with half wavelength. The 4 antennas in the same polarization perform beamforming for the purpose of generating narrow beam and the two polarizations enables further exploit of 2 layer spatial multiplexing inside the narrow beam. Equally, for wide beam base station, there are 2 transmit antennas that provide the opportunity for 2 layer spatial multiplexing. The number of receive antennas at the user side is assumed to be 2 for all the users.

As the gated narrow beam model is using four narrow beams to replace one traditional sector wide beam of 70-degree, the antenna beamwidth of narrow beam is assumed to be reduced four-fold over the wide beam antenna model, and the antenna gain and front-to-back ratio are assumed to be increased four-fold. The antenna gain and path loss experienced by signals transmitted from a base station to a given location in the simulation area are calculated as follows [117]:

For horizontal antenna pattern, the antenna gain $A_H\ (\varphi)$ is calculated by

$$A_H\ (\varphi) = -Min\left[12\left(\frac{\varphi}{\varphi_{3dB}}\right)^2, A_m\right] \tag{4.5}$$

where for wide beam the half power beam width $\varphi_{3dB} = 70^0$ and the backward attenuation $A_m = 25$ dB; for narrow beam $\varphi_{3dB} = 17.5^0$ and the backward attenuation $A_m = 31$ dB (to account for the 4x increase (or 6 dB) in power density due to the smaller half power beam width).

For vertical antenna pattern, the antenna gain $A_v(\theta)$ is calculated by

$$A_v(\theta) = -Min\left[12\left(\frac{\theta - \theta_{etilt}}{\theta_{3dB}}\right)^2, SLA_v\right] \tag{4.6}$$

where half power beam width $\theta_{3dB} = 10^0$; backward attenuation $SLA_v = 20\ dB$; downtilt angle $\theta_{etilt} = 15°$.

Combining method in 3D antenna pattern, the overall antenna gain $A(\varphi, \theta)$ is

$$A(\varphi, \theta) = -Min\{-[A_H(\varphi) + A_v(\theta)], A_m\} \tag{4.7}$$

The path loss of signals at a distance R (km) from a base station is given by [117] as

$$L = 128.1 + 37.6 log\ 10(R) \tag{4.8}$$

When calculating the antenna gain and path loss, users are assumed to be at least 35m away from any base station. In terms of fading losses, a shadowing loss is simulated as a Gaussian distribution of zero mean and a 3dB standard deviation. Shadowing loss is experienced at each user location as well as at a base station. Fast fading is not directly simulated (as explained in the next paragraph). In addition, users are presumed to be indoors and experience a building penetration loss of 20dB.

The expectation value of spectral efficiency as a function of the long-term SINR from full simulations in a 3GPP calibrated simulator [117] is derived, using 8x2 Single User MIMO for narrow beam and 2x2 Single User MIMO for wide beam. This is applied as a mapping table to convert the calculated SINR of each user into his experienced user data rate. The table has already accounted for a number of features such as the modulation format and level of error-correction coding appropriate to the SINR, retransmissions of PRBs and coding efficiency, and fast fading statistic for user equipment speed of 3km/h. Transmission of signalling and control PRBs which involves approximately 10-15% of the total available PRBs has also been taken into account in the mapping. This table is also adopted in [118, 120].

4.3.2 Base Station Power Model

The power profiles for the base stations are defined by 2 reference points:

1. Full Load Mode (maximum power): power value for 100% PRB utilization
2. No Load Mode (idle power): power value for 0% PRB utilization

Note that in the No Load Mode, signalling PRBs are still being transmitted, the 0% PRB utilization refers to no user data being transmitted. Other energy saving methods, such as placing the base station into a sleep mode when there is no data traffic, are not considered in this study. The power consumption for base stations of narrow beams and base stations of wide beam are shown in Table 4.2. These are based on the power model given in [117]. The base station of gated narrow beam consumes more power than that of wide beam due to having more antennas and more radio chains and associated signal processing.

Simulation Parameter	Narrow Beam (8x2 MIMO)	Wide Beam (2x2 MIMO)
BS Maximum Power (3 sectors)	665 W	638 W
BS Idle Power (3 sectors)	189 W	132 W

Table 4.2 Base station power model

When a narrow beam is turned 'On' to serve users within its coverage, it will be transmitting at maximum RF power, and transmitting all the available resources (full load mode). The power consumption to serve the sector in this mode is $\frac{665}{3}$ watts. When a narrow beam is turned 'Off', it will be transmitting at idle RF power, and transmitting no data resource (no load mode). The power consumption to serve the sector in this mode is $\frac{189}{3}$ watts. The two state of the beam (On/Off) corresponds to the two states of the base station. The same applies for the wide beam base station with different power consumption.

4.3.3 Playground Configuration

As shown in Figure 4.6, the simulated playground includes a central base station (BS 0) and another 18 base stations arranged in two tiers surrounding the central base station (BS 1-18).

The simulated narrow beam coordination area, where users are dropped and performance evaluation is performed, is essentially only a single coordination cluster consisting of 3 sectors. The simulated coordination area is shown as a pink hexagon in Figure 4.6. Users are dropped at random locations according to a certain rate into the coordination cluster, and trigger the 12 potential 'in-cluster' beams to be turned 'On' and 'Off' according to the coordination scheme.

For narrow beam modelling, only 1 concentric ring surrounding the pink hexagon is considered and interference from beams from BS 11- 17 is ignored. This is because the path loss from these base stations will be significantly greater than those from BS 0-10, and 18, and this simplification helps reduce the computation time of the simulator.

To ascertain the impact of this simplification, we performed a few test simulations with 2 concentric rings surrounding the pink hexagon and compared results. In these test simulations, all the settings remain the same (including simulation parameters for base stations and users, modelling of playground and the traffic applied etc.) except for the deployment of additional base stations outside the simulated coordination cluster. It is found that the difference in the network performance is negligible. The difference in SINR is around 0.4 dB at 10% traffic level, and around 0.3 dB at 50% and 120% traffic levels. The percentage difference in average user throughput (this metric will be defined later in Section 4.4.1) is around 0.5%, 1% and 1.5% at 10%, 50% and 120% traffic levels respectively. The concept of these traffic levels shown in percentage will be discussed in Section 4.3.4. The traffic model applied here is the first traffic scenario - FTP model as will be discussed in Section 4.3.4.

The beams within the cluster are coordinated using the proposed beam coordination and resource scheduling scheme while the outside cluster interference from all surrounding sectors is outside the control of the scheduler. However, all of the narrow beams outside the coordination cluster should be pulsed "On" and "Off" generating interference as the in cluster beams. The pulsing rate depends on the traffic model applied which will be discussed in the next section.

In the case of the wide beam model, users are dropped at random locations anywhere in the extended playground. Beams from all of the base stations in the central location and two concentric rings (drawn in Figure 4.6) were active with full RF power if a user required service from that beam, or idle and emitting limited RF signalling power if no user required service from that beam. The results for users served by the central and 6 base stations in the first ring are recorded for performance evaluation.

Beams omitted in simulation

Beam triggered by user arrival (solid line) and beam pulsed at

predetermined rate (dash line)
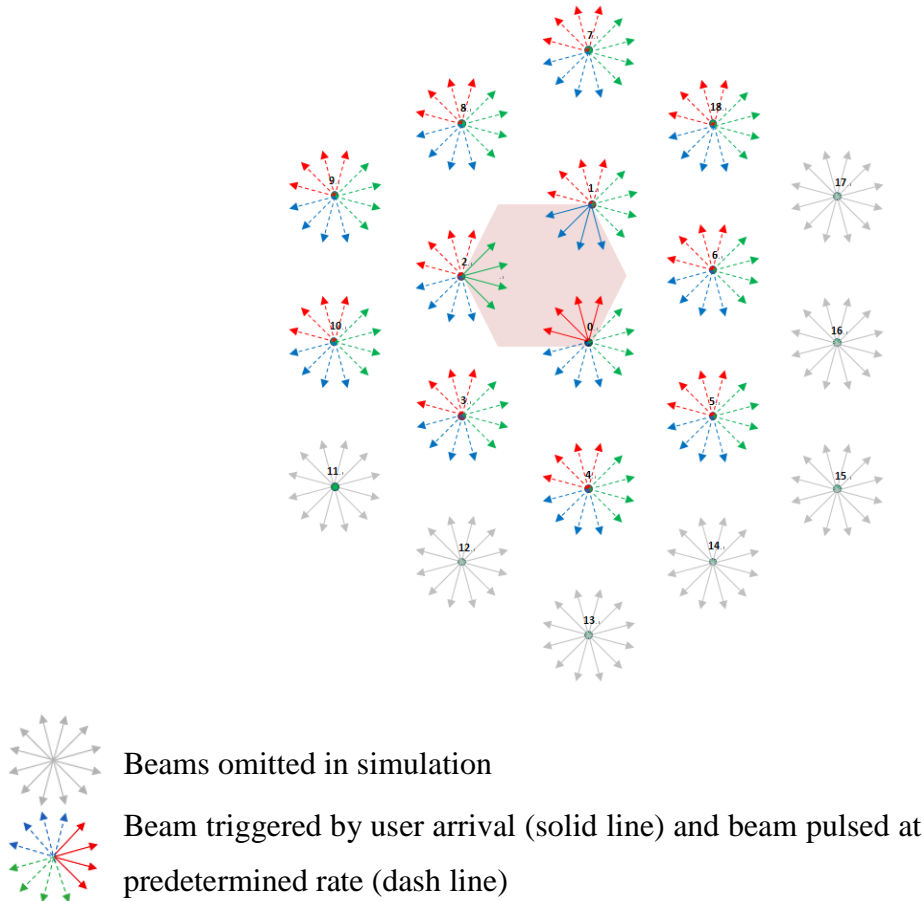
Figure 4.6 Playground configuration of gated narrow beam network simulation

4.3.4 Traffic Model and Outside Cluster Interference

This section discusses various traffic models, and the traffic condition and outside cluster interference scenarios to be applied to this simulation.

A widely used traffic model named full buffer model assumes all base station transmit 100% of the time [121]. This assumption is often used in the literature to simplify the analysis, especially when attempting to derive a closed-form formula for base station performance for various metrics including SINR and throughput [121]. However, base station networks are designed to meet peak hour user requirements and are under-utilized in non-busy hours especially during nights. Hence, the assumption may be valid in congested areas during peak hours but is not true for most of the day. Such a model is valid for determining theoretically achievable gains but is not particularly realistic and will lead to overestimated throughput [121].

In many system level simulations (3GPP [122], GreenTouch [117] and other CoMP scheme [123]), a type of non-full buffer traffic model (FTP traffic model) is used. This

presents a more realistic model for bursty data traffic, especially outside of peak times. Real networks are often running idle between data bursts [121]. In the FTP traffic model [117, 122], users come online at Poisson distributed time intervals with a specified average overall traffic rate, as shown in Figure 4.7. Let the user arrival rate be λ, then the length of the user inter-arrival times D follows the exponential distribution, with mean equal to $1/\lambda$. The probability density function (pdf) for the inter-arrival time, D, of the exponential distribution is [122],

$$pdf: \ f_D = \lambda e^{-\lambda D} \tag{4.9}$$
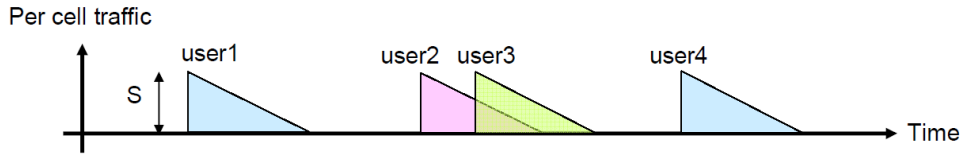
Figure 4.7 Traffic generation of FTP model

It is assumed that each user comes online, requesting a download that contains a fixed file size S of 2 MB. After finishing the download, the user is removed from the playground, with its result marked as successful. If a 2 MB download is not finished after 4 seconds, the download is terminated and marked as a failure. This is a performance metric for determining Quality of Service, and is part of the GreenTouch modelling team's agreed common set of parameters for analysing competing network technologies [117].

In terms of traffic level, predicted traffic load in Dense Urban (DU) area in 2020 as in GreenTouch traffic model [117] is adopted. The user density is 10,000 users/km^2 and each user generates an average of 70 kbps of traffic during the busy hour (BH). This represents the highest traffic load that system needs to support. The layout of the base stations is assumed to take the form of a regular hexagonal pattern, and the base stations' Inter-Site Distance is assumed to be 0.5km for this dense urban area [117].

The drop rate of users corresponding to the busy hour traffic level is then (nos. of users/km^2 = 10,000) x (one coordination cluster area = 0.217 km^2) x (BH Rate/user = 70000 bit/s)/(file size = 16 Mbits) = 9.49 users/s. For the FTP traffic model, this is the value of λ in Equation (4.9). Besides busy hour traffic load, a range of traffic loads ranging from busy hour (140% of the average daily traffic) down to very low load (5% of the average daily traffic) is simulated to evaluate the performance of coordinated gated narrow beam scheme under different traffic loads during a day. Later in this chapter and

throughout the thesis, a traffic level is often nominated in shorthand as a percentage, which represents the percentage of the full-day averaged traffic level. We model all traffic as downlink traffic as the downlink traffic is approximately 90% of the traffic [117] and receiving uplink traffic causes less power consumption than transmitting downlink traffic due to the transmitter amplifiers.

To model the impact of variability of traffic on base station usage, an FTP model will be adopted for evaluating the performance evaluation of gated narrow beams network, as well as the performance comparison over wide beam network. In addition, another 2 traffic scenarios (see 2 and 3 below) are also constructed to estimate the impact of simplifying assumption of a full buffer traffic model. The following describes the various traffic models used in this thesis to simulate beam usage and beam interference within the coordination cluster and outside the cluster:

1. FTP model. FTP download traffic with varying traffic load is applied to the coordination cluster. Beams within the coordination cluster were turned on or off as necessary to serve users according to the proposed coordination and scheduling scheme. Meanwhile, beams outside of the coordination cluster were turned on or off randomly, at a rate equivalent to the rate within the playground. So an equivalent level of interference can be generated without the computational complexity that would have been involved in scheduling over the entire playground.

   This is achieved via an iterative process where a 'seed' pulsing rate for the outside cluster beams/interference is used to run the simulation, and at the end of each iteration, a new pulsing rate is determined based on the 'in cluster' beam pulse rate from the simulation results. The iteration repeats until the pulse rate set for outside cluster beam and the pulse rate of in cluster beams converge.

2. Hybrid FTP and Full Buffer model: The same FTP download traffic as in Scenario 1 is applied to the coordination cluster. However, beams outside of the coordination cluster are switched on continuously in response to a full buffer model.

   Note that this is not a full buffer model. The three sectors in the coordination cluster still operate as before – FTP traffic model with the flexibility of different traffic level. From an individual user's perspective, it still sees pulsed interference

from two major interferers (the other two sectors in the coordination cluster), but all other interferers are always-on.

3. Full Buffer model: All the beams both within and outside the coordination cluster are turned on continuously. This model is simulated based on applying a high FTP traffic load to the coordination cluster of Scenario 2. When the FTP traffic rises to a level that all beams have to be 'On' all the time in order to serve the users, the inside and outside cluster are both under 100% utilization, thus the model is essentially in full buffer condition. It is observed in simulation that when the traffic load is 500% of the average daily traffic, the probability that all 3 sectors have one beam turned on is more than 99.6%. Hence 500% traffic load will be used for generating results for full buffer model.

This is a special case of scenario 2 when the coordination cluster is loaded with FTP traffic until all its BSs' sectors are 100% utilized. This model is an approximation of what is traditionally referred to in the literature as the full buffer model, as the beams within the coordination cluster are still responding to FTP download traffic. The full buffer traffic model essentially represents just one static network traffic level and base station utilization rate, rather than the varying (e.g. time dependent) network traffic model such as the FTP model.

As discussed, Scenario 1 FTP model is adopted in evaluating the performance of gated narrow beam network and wide beam network in Section 4.4.2. Scenarios 1, 2 and 3 are being applied in Section 4.4.3 to compare the impact of different traffic condition on the evaluation of network performance.

4.4 Simulation Results

In this section, the results obtained from the system level simulator described in Section 4.3 are presented. As discussed, the simulation results include both gated narrow beam model and wide beam model. The metrics for performance evaluation is summarized in Section 4.4.1. Section 4.4.2 presents the performance metrics for gated narrow beam model, and the comparison over wide beam model. Section 4.3.4 presents the performance metrics for the gated narrow beam model when the three different traffic scenarios as discussed in Section 4.3.4 are applied to the simulated playground.

4.4.1 Metrics for Performance Evaluation

The metrics used for user-perceived performance are:

- Average SINR (dB). An individual user's SINR is averaged over the time epochs during which it is receiving transmission and being allocated PRB resources towards the download of a 2MB file. The overall average SINR is the average of all individual users' SINR values at the demand traffic level.

- Average user throughput/download speed (Mbps). This is the average download speed across all users. The download speed for each user is calculated by the total data bits received divided by the duration of the download (both the data bits and download duration are observed and recorded from the simulation). Then the download speed is averaged over all the users.

These metrics are applied to both successful user and failed users.

The metrics used for evaluating network performance are:

- Average failure rate (%). This is the percentage of users that fail to complete the download of a 2MB file within 4s.

- Base station power consumption (Watt). This is the average power consumption of a base station of three sectors. Based on the user traffic demand, base stations switch between full load mode when maximum power is consumed, and no load mode when an idle power is consumed. The base station power consumption is the total energy consumed by the base station divided by the elapsed time.

- Cell throughput (Mbps). This measure the sum throughput/capacity of one cell/base station. This is the total number of data bits transmitted by a base station divided by the elapsed time. Note that, for gated narrow beam model, the simulated coordination cluster consists of 3 sectors, which is equivalent to a cell.

- Energy efficiency (Joule/Mbit). This measures the average energy cost (Joules) of downloading a Mbit of user data. It is calculated by dividing base station power consumption by cell throughput.

4.4.2 Performance of Gated Narrow Beam Model and Comparison against Wide Beams

From the simulation, the performance of a network of coordinated gated narrow beams as measured by the performance metrics (SINR, user throughput, failure rate, base station power consumption, cell throughput, and energy efficiency) is obtained. Figure 4.8 shows the behaviour of these metrics as the FTP traffic load is varied. The x-axis represents the traffic load as a percentage of the average daily traffic for dense urban in 2020, as calculated in Section 4.3.4.

The focus of this chapter is on ascertaining the network throughput of a coordinated beamforming system - gated narrow beam system, as well as the discovering the potential energy savings achieved by the throughput gain. The energy efficiency estimate in this chapter is only preliminary as the energy modelling of the network is not complete. In particular, various energy costs associated with supporting the coordination is ignored. Only base station energy consumption is considered in this chapter.

As discussed, the performance of gated narrow beam model will be compared against that of conventional wide beam model to demonstrate potential performance improvements especially potential throughput gain and energy savings. Thus, the corresponding performance of a network of conventional wide beams is plotted alongside gated narrow beam for the same traffic load in Figure 4.8.

Next, we will discuss and compare the performance metrics of the two networks in detail. As shown in Figure 4.8 (a), the user average SINR achieved in the gated narrow beam model is better than in wide beam model across all traffic levels. More improvement is achieved at higher traffic level. This is because the gated narrow beam model is able to mitigate interference by coordinating beams, and the advantage it brings is even greater when the traffic demand is high (a large number of users on the playground). At the busy hour, the SINR of gated narrow beam model is more than 11dB better than wide beam model.

The improved SINR, and hence spectral efficiency, allow the users in gated narrow beam model to receive more bits for a given amount of allocated resource on average. Thus, the user throughput/download speed achieved in gated narrow beam model is improved as well. The gated narrow beam model provides a 4 fold (83.7 Mbps vs. 21.1 Mbps) to 8

fold (69.4 Mbps vs. 7.9 Mbps) improvement in the user download speed/throughput for daily average and busy hour load levels respectively, as shown in Figure 4.8 (b). The MIMO configuration used in the gated narrow beam model also contribute to this improvement of user throughput, as 8x2 MIMO provides a slightly higher spectral efficiency than 2x2 MIMO for the same SINR.

Figure 4.8 (c) shows the percentage of users unable to complete their 2 MB download within a 4s period. This percentage represents users that are failing to meet the quality of service requirement. Due to the poorer user throughput/download speed, a network of wide beam base stations has an unacceptable failure rate even before the load reaches average daily traffic (100%). The failure rate of a network of gated narrow beams is only 0.2% even at busy hour load (140%). It can be concluded that the gated narrow beam model can support dense urban busy hour traffic with comparatively much better failure rates than a wide beam system.

The average base station power consumption of the two models is shown in Figure 4.8 (d). The wide beam model starts with lower power consumption at low traffic load, and a crossover point occurs at around 50% traffic level (reason for the crossover is discussed below). From 50% to 140% traffic levels, the base station in gated narrow beam model consumes less power than the base station in wide beam model on average. The savings is around 140W at average daily traffic (100%) to busy hour traffic level (140%) for one base station.

For both networks, the power consumption of base station is dominated by the idle power at low traffic level, as base station is at no load mode the majority of the time. The idle power for a gated narrow beam base station with 8x2 MIMO configuration is 189W, whilst the idle power for a wide beam base station with 2x2 MIMO configuration is 132W. Thus, the wide beam model has an advantage at low traffic levels because of its lower idle power consumption. However, the user throughput/download speed of gated narrow beam model is better than wide beam model. This allows users to accomplish their download in a shorter period of time on average. Thus, for a given traffic load, the duration a gated narrow beam base station spent in the idle state is comparatively longer than a wide beam base station. This eventually negates the lower idle power advantage that a wide beam base station consumes. As traffic load level increases, it spends more time in its full load state compared to a gated narrow beam base station.

Figure 4.8 (e) shows the overall cell throughput in Mbps. This represents the average throughput of one base station (for gated narrow beam model, the simulated coordination cluster is equivalent to one base station). At low traffic levels when the cell is not capacity limited (all user download request can be completed without exceeding the download time threshold), the cell throughput is mainly determined by the offered user drop rate/traffic level, and less sensitive to SINR. It can be seen from Figure 4.8 (e) that at low traffic level, the cell throughput of both models is very close and increase almost linearly as traffic level. However, as the traffic level increases, there is a growing gap between the two models. The cell throughput of gated narrow beam model continues to increase linearly as traffic level while the cell throughput of wide beam model is less than the gated narrow beam model. Reasons are that, as traffic level increases and wide system approaches saturation, the worsened SINR starts to have an impact with the result that less throughput/bits can be transmitted for a given amount of resource. The large percentage of failure rate of wide beam model shown in Figure 4.8 (d) also suggests that not all users can complete their 2MB download. Those failed users contribute less than 2MB to the cell throughput. Consequently, the cell throughput is less than the offered traffic.

Note that, although from the network perspective, the cell throughput of the gated narrow beam model increases linearly as the traffic level increases (nearly no failure cases) across the entire traffic levels, the user-perceived throughput keeps going down, as shown in Figure 4.8 (b). From an individual user perspective, although he can still finish the download within the 4-second target at higher traffic level, the duration of the download is generally longer due to the poorer SINR and hence spectral efficiency.
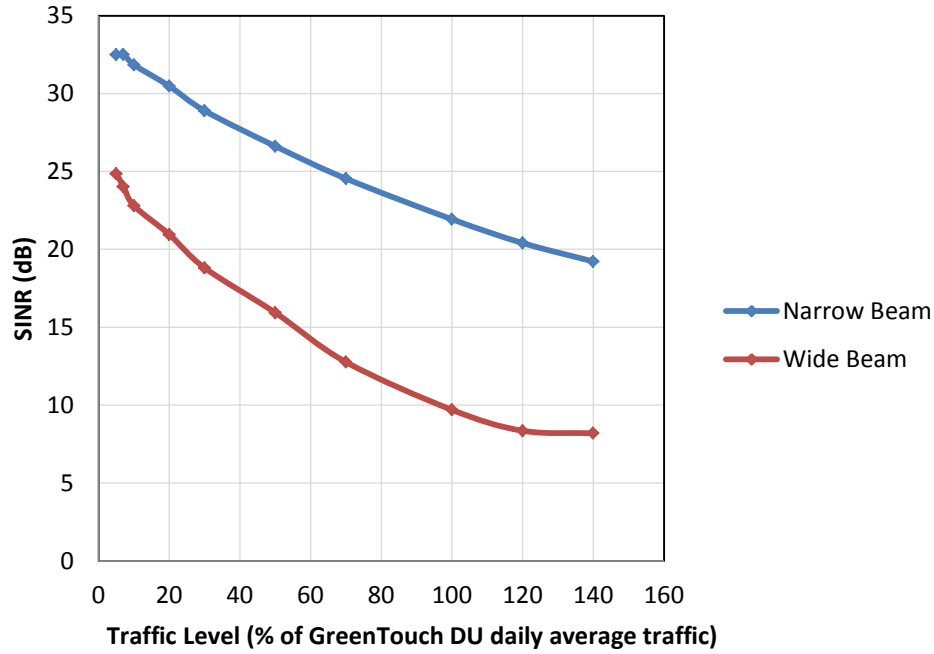
Figure 4.8 (a) SINR vs. traffic level in dense urban environment
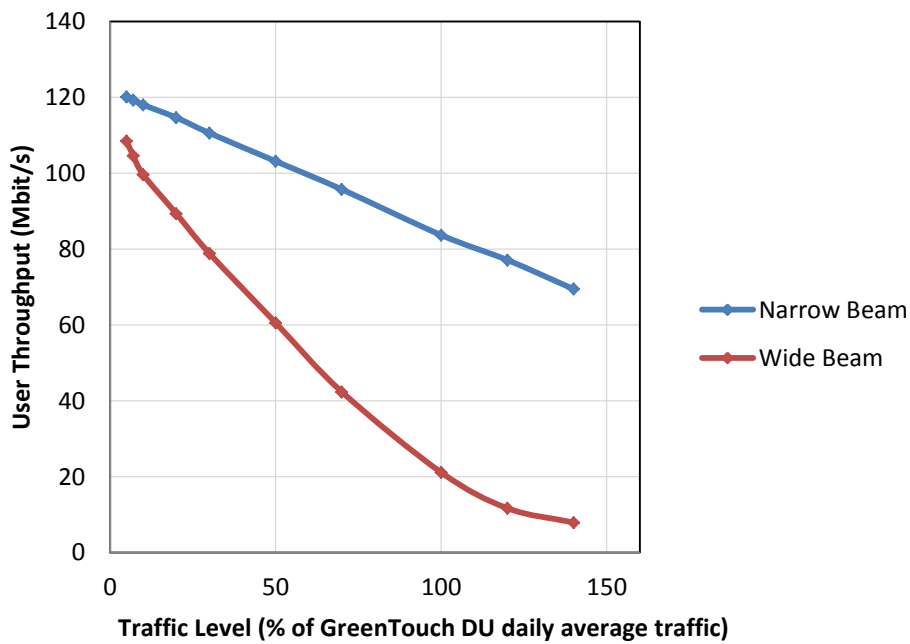
Figure 4.8 (b) User throughput vs. traffic level in dense urban environment
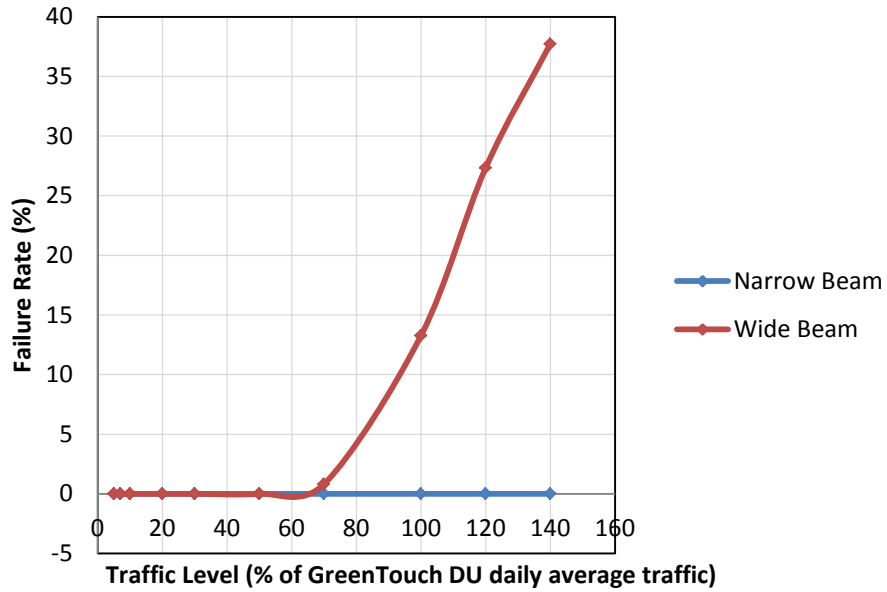
Figure 4.8 (c) Failure rate vs. traffic level in dense urban environment
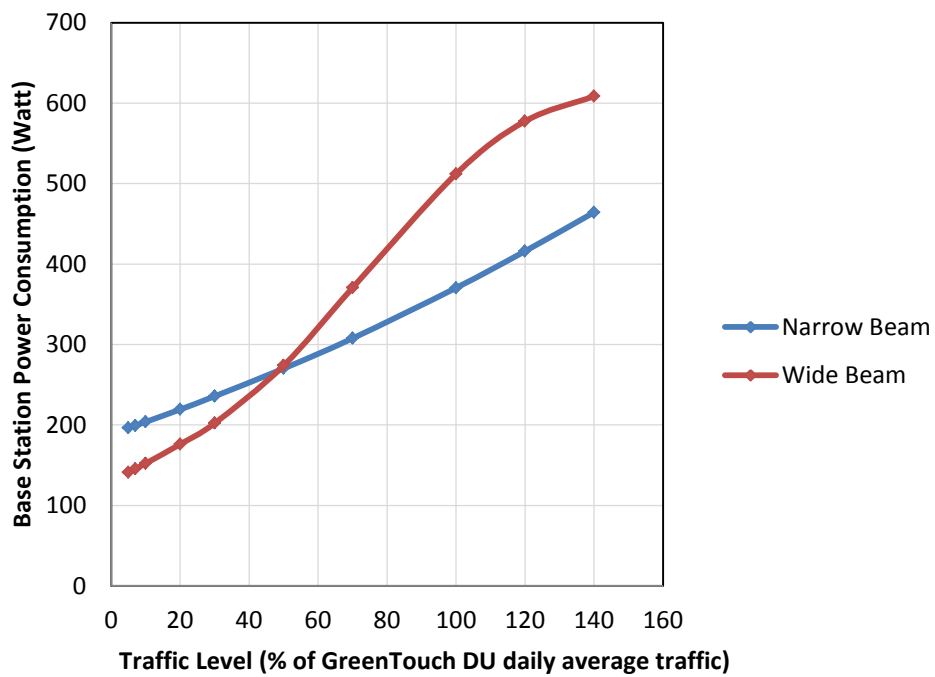
Figure 4.8 (d) Base station power consumption vs. traffic level in dense urban environment
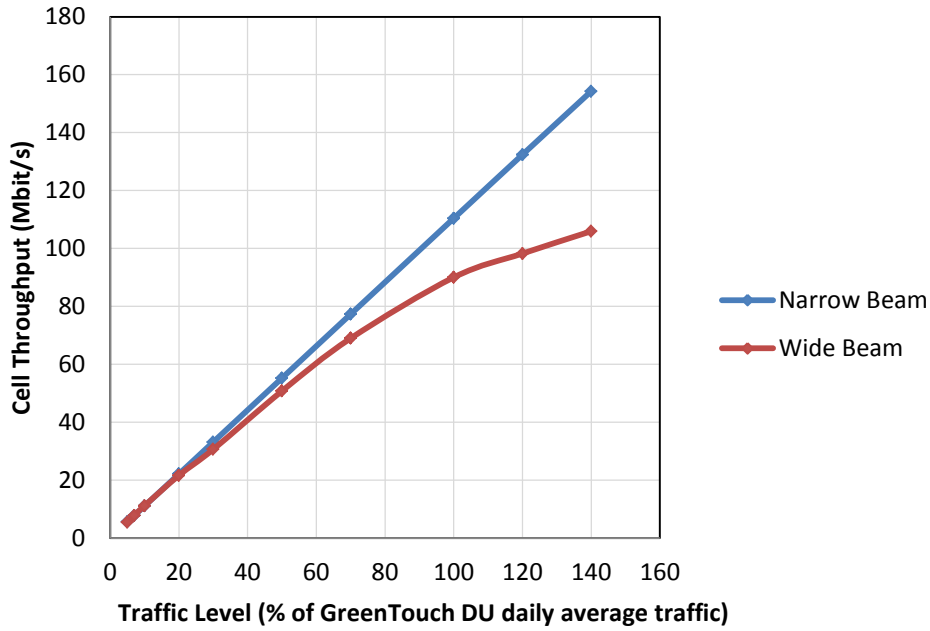
Figure 4.8 (e) Cell throughput vs. traffic level in dense urban environment
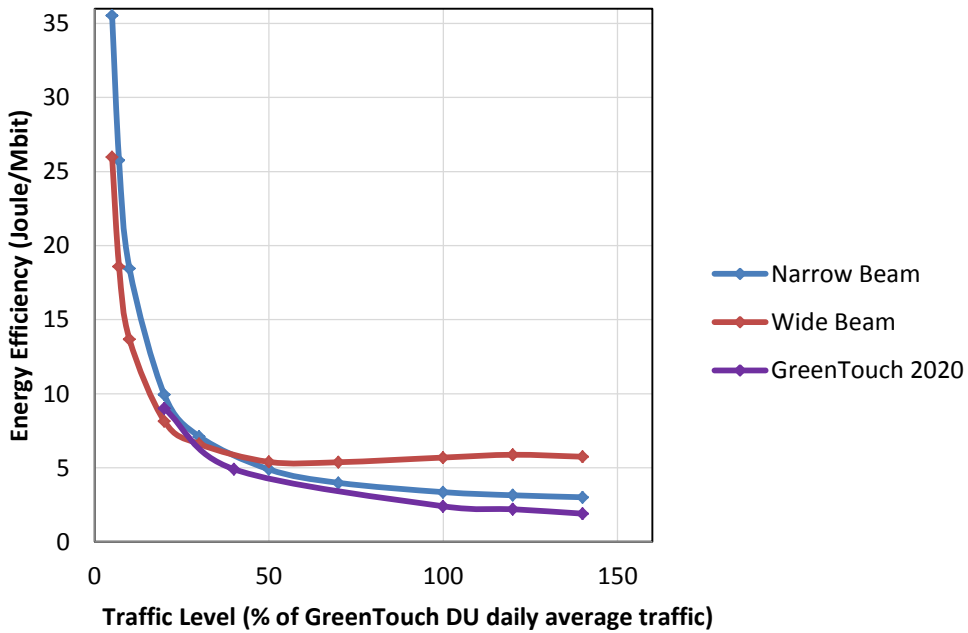
Figure 4.8 (f) Energy efficiency vs. traffic level in dense urban environment

Figure 4.8 (d) and Figure 4.8 (e) have shown that a network of gated narrow beams has the advantage of reduced base station power consumption (50% to 140% traffic level) and improved cell throughput (across all traffic levels). Thus, it can be expected that the energy efficiency of gated narrow beam model is better than that of wide beam model for

a range of traffic levels, from around 45% to 140%. Figure 4.8 (f) presents the energy efficiency comparison between the gated narrow beam model and wide beam model. This is calculated by dividing base station power consumption by cell throughput. The gated narrow beam model is more energy efficient as shown in Figure 4.8 (f) over a large range of DU traffic load (40% to 140%), with the greatest savings of delivered at higher loads. At 100% to 140% load levels, a gated narrow beam model is 41% to 48% more energy efficient than a wide beam model (3.35 Joules/Mbit vs. 5.69 Joules/Mbit to 3.01 Joules/Mbit vs. 5.74 Joules/Mbit). These results demonstrate that a network of coordinated gated narrow beams has great potential for improving energy efficiency. The energy saving is mainly achieved through the throughput gain brought by the coordination of narrow beams. As discussed, several energy costs in supporting this coordination are ignored in this chapter, thus the energy savings are an upper bound of what could be achieved.

Note that there is a 'wobble' in energy efficiency metric of wide beam model at around 120% traffic level in Figure 4.8 (f). The energy efficiency essentially consists of two components – the amortisation of the idle power consumption per bit, and the incremental power consumption per bit. The general trend is that, as traffic level and cell throughput increase, the idle power consumption per bit is decreasing. Meanwhile, the incremental power consumption per bit is increasing due to the worsened SINR. This 'wobble', a minimum followed by a rise, is essentially an artefact of the parameters for wide beam model, when the slope of the decreasing idle power becomes smaller than the slope of the increasing incremental power. The 'wobble' does not show in the behaviour of the gated narrow beam model, because the rate of the SINR deterioration is much slower compared with wide beam model as shown in Figure 4.8 (a), thus the slope of the increasing incremental power is much smaller. Therefore, the energy efficiency metric of the gated narrow beam model behaves as a flat line with no turning point.

In Figure 4.8 (f) the GreenTouch 2020 energy efficiency prediction for a Dense Urban environment from [118] is also included, which includes the benefit of offloading 2/3 of the traffic to highly energy efficient small cells and use of an optical backhaul network. As can be seen, although the gated narrow beam model is more energy efficient than wide beam model (3.01 Joules/Mbit vs. 5.74 Joules/Mbit at 140% traffic load), it is less energy efficient than the GreenTouch 2020 prediction (1.9 Joules/Mbit vs. 3.01 Joules/Mbit at 140% traffic load). However, the higher energy efficiency attained by the GreenTouch

simulation is achieved at the cost of provisioning and managing a large number of small cells, low powered base stations. Deploying macrocells and small cells together could introduce some practical complexities, which is mainly due to the different hardware configurations, fading conditions, inter-tier interference, etc. In addition, the overall energy efficiency of the GreenTouch heterogeneous wireless network deployment is sensitive to the technological mix (e.g. fibre, microwave, VDSL) of backhaul network [124]. An improper choice of backhaul technologies could negate energy savings at the small cell sites. Given the improvement of energy efficiency of gated narrow beam model compared with wide beam model, the gated narrow beam could potentially be used together with small cells in practice to reduce the number of small cells needed. Hence reduce the cost and complexity in the overall small cell deployment. As shown in Figure 4.8 (c), a network of conventional wide beams has an unacceptable throughput failure probability even before the load reaches 100%. This is one of the reasons why small cells are needed in the 2020 scenario in [118], to provide sufficient capacity to avoid failures at the higher load levels. In contrast, the entire Dense Urban busy hour load could be carried by a narrow beam network with negligible failure cases, without the benefit of small cells.

The comparison in Figure 4.8 (a) – (f) demonstrates that the coordinated gated narrow beam network potentially has a significant performance advantage over a conventional wide beam network over these metrics across a large range of traffic load conditions. From the simulation results, the throughput of the gated narrow beam network is ascertained. The signalling traffic for coordination is not taken into account and the throughput will be updated after the overhead is quantified in Chapter 5. The preliminary energy efficiency comparison over wide beam model demonstrates great improvements and confirms the potential of energy savings brought by the proposed gated narrow beam model.

4.4.3 Comparison of Performance under FTP Traffic Model and Full Buffer Model

This section presents the performance metrics for the gated narrow beam model when the three different traffic scenarios as discussed in Section 4.3.4 are applied to the simulated playground. As discussed, the FTP traffic model (Scenario 1) is the traffic model adopted for the purpose of performance evaluation of the network of gated narrow beam base stations. The simulation results in Section 4.4.2 are all based on an FTP traffic model.

The other two scenarios are constructed to estimate the impact of the simplifying assumption of the full buffer condition in the full buffer traffic model, which is commonly used in literature. The difference in performance with full buffer model (Scenario 3) and FTP traffic model (Scenario 1) is explored.

In the following, we will discuss the performance of Hybrid FTP and Full Buffer model (Scenario 2, referred to as hybrid model for short), and compare it with the performance of FTP model (Scenario 1). Then the performance of FTP model (Scenario 1) and full buffer (Scenario 3) is compared. As discussed, Scenario 2 evolves to Scenario 3 as the offered traffic in the coordination cluster is increased to 500% when all sectors are 100% utilized.

Figure 4.9 (a) – (c) presents the metrics of base station power consumption, cell throughput and energy efficiency, respectively. The definitions of these performance metrics have been provided in Section 4.4.1. The behaviour of the metrics for all three traffic scenarios is plotted in the same figure for comparison.

The difference between the full buffer model and the hybrid model is essentially only the outside cluster interference. Because for hybrid model, there are continuously on interferers; and for the FTP model, the interference is less due to the pulsing of outer-ring base stations. For the same FTP traffic level being applied to inside the coordination cluster, the FTP model is expected to outperform the hybrid model in every metric, as the users in the coordination cluster experience less interference from outside the cluster than in hybrid model (thus higher user-perceived SINR), while the traffic and interference condition within the coordination cluster is identical. As shown in Figure 4.9 (a) and Figure 4.9 (b), the base station power consumption of the FTP model is less than that of the hybrid model, because the duration of download is generally shorter due to the higher spectral efficiency and hence comparatively longer time that the base station spent in the idle state. Its cell throughput is also higher in high traffic loads (from 100% to 140% traffic level). For low traffic levels (when the system is not capacity limited and the cell throughput is determined only by the offered traffic), the cell throughput of the two models is very close and both linear to traffic level.

Figure 4.9 (c) compares the energy efficiency behaviour. As expected, the energy efficiency in Joule/Mbit of FTP model is lower than hybrid model over the entire range of traffic level. At 100% to 140% load levels, the hybrid model is 9% to 14% less energy

efficient than the FTP model (3.65 Joules/Mbit vs. 3.35 Joules/Mbit to 3.44 Joules/Mbit vs. 3.01 Joules/Mbit). The energy efficiency difference in the two models confirms the necessity of modelling the outside cluster interferers with the pulsing rate. As discussed in Section 4.3.3 playground configuration, the pulsed interferers crease an equivalent level of interference as inside the cluster. A simplified traffic modelling as in the hybrid model will cause a degraded energy efficiency.

The performance metric of full buffer model (Scenario 3) is shown as a dot in Figure 4.9 (a) – (c) as the full buffer essentially corresponds to just one traffic condition/level. The base stations in full buffer model are running at full load mode 100% of the time, while the base stations in FTP model are running at a much lower utilisation. Thus, the base station power consumption and cell throughput of the full buffer model is greater than FTP model, as shown in Figure 4.9 (a) and Figure 4.9 (b).

The energy efficiency of the full buffer model is 2.86 Joules/Mbit as shown in Figure 4.9 (c), and appears to be lower than the FTP model. However, the energy efficiency of a full buffer network is not attainable in practice. As discussed in Section 4.3.4, the base stations are not provisioned on the basis of running at full capacity (full load mode and 100% PRB utilisation). As observed from the FTP model simulation, even at the busy hour, the utilisation of a base station is only about 53%. Due to the base station idle power consumption (base station still consuming power when no bits/PRBs are transmitted), a base station is less energy efficient at low utilisation, and more energy efficient at higher utilisation.

Thus, for a realistically dimensioned network, the base station would operate at lower energy efficiency than what is achieved in the full buffer model. The results from full buffer model are overly optimistic and misleading when characterising the energy efficiency of real network. On the contrary, the FTP traffic model can more realistically characterise the energy efficiency of a network with varying traffic levels, according to different time of the day (such as night, busy hour etc.). In addition, the energy efficiency figures corresponding to different traffic levels enables the computation of the daily energy consumption a base station.
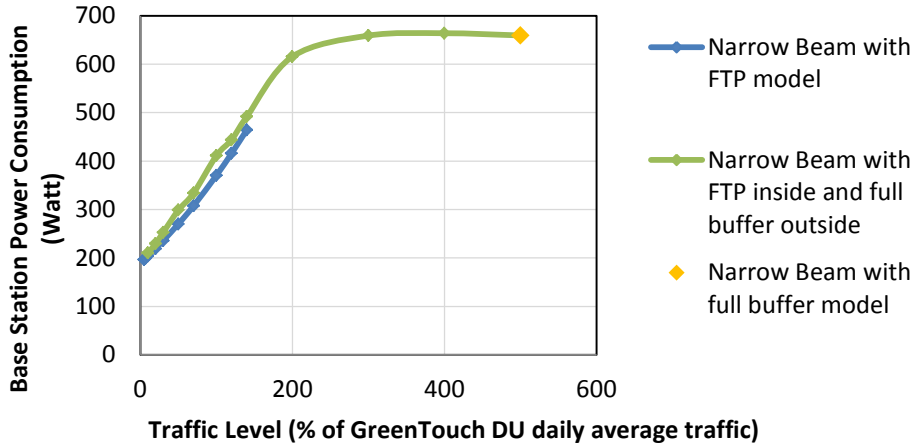
Figure 4.9 (a) Base station power consumption vs. traffic level in dense urban environment under 3 traffic scenarios
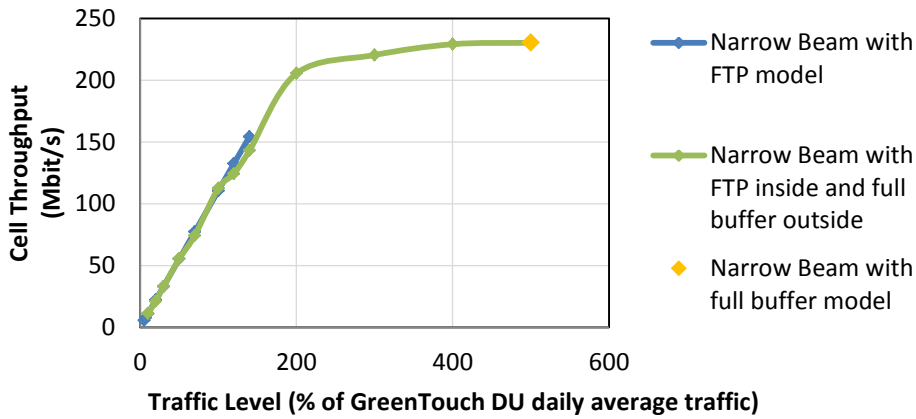
Figure 4.9 (b) Cell throughput vs. traffic level in dense urban environment under 3 traffic scenarios
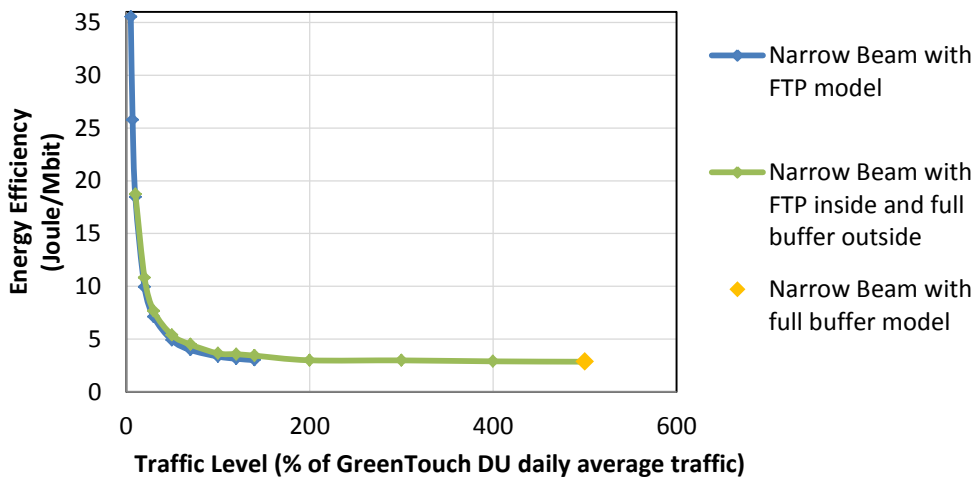
Figure 4.9 (c) Energy efficiency vs. traffic level in dense urban environment under 3 traffic scenarios

4.5 Performance Comparison of Gated Narrow Beam Model and other CoMP strategies

This section presents the performance comparison of the proposed gated narrow beam model and other existing CoMP downlink transmission strategies including coordinated scheduling and beamforming (CS/CB) and joint processing (JP). The performance evaluation in this section is performed through system level simulation under a 3GPP proposed reference scenario – 3GPP Case 1 [125]. The purpose of the 3GPP reference scenarios was to calibrate the simulators from various sources and ensure that they produce comparable results for evaluation. The 3GPP Case 1 simulation scenario and relative parameters are in common use in the literature.

Thus, in this section the gated narrow beam model will be simulated based on specification for this reference scenario in order to make direct comparison with simulation results of other CoMP schemes. 3GPP has proposed the system performance metrics to be evaluated, including cell spectral efficiency, cell-edge spectral efficiency and user throughput distribution [125]. These metrics of GNB model will be estimated from the system level simulation. The details of the metrics will be discussed later.

4.5.1 Simulation Parameters

The 3GPP Case 1 reference scenario considers a homogenous deployment of 3-sector macro base stations with an inter site distance of 500 meters. Users are randomly and uniformly distributed with a fixed density of 10 users per sector. System throughput are to be assessed using full buffer traffic model capturing continuous traffic and non-varying out-tier interference [125].

The detailed simulation parameters and settings are given in Table 4.3 below,

Parameter	Values used for evaluation
Simulation Case	3GPP Case 1 (spatial channel model)
Cellular layout	Hexagonal grid, 19 base stations, 3 sectors per base station
Inter-site distance	500 meter
Carrier frequency	2.0 GHz
System bandwidth	20 MHz
Duplex mode	TDD

Penetration loss	20 dB
User speed	3 km/h
Lognormal shadowing standard deviation	8 dB
Shadowing correlation between base stations	0.5
Distance-dependent path-loss	$128.1 + 37.6\log 10(R)$, R in km
Traffic model	Full buffer traffic model
Number of users per sector	10
Minimum distance between base station and user	35 meter
Thermal noise power	-174 dBm/Hz
CoMP sets	3 coordinated base stations (intra-site or inter-site)
Transmission mode	MU-MIMO and/or SU-MIMO operation in conjunction with CoMP
Number of Antennas (Tx, Rx)	(2, 2), (4, 2), (8, 2) or (4,4)
Total BS Tx Power	46 dBm
Transmit antenna pattern (horizontal)	$A_H(\varphi) = -Min\left[12\left(\dfrac{\varphi}{\varphi_{3dB}}\right)^2, A_m\right], \varphi_{3dB} = 70^0, A_m = 25$ dB. (for narrow beams in GNB model, $\varphi_{3dB} = 17.5^0, A_m = 31$ dB)
Transmit antenna pattern (vertical)	$A_v(\theta) = -Min\left[12\left(\dfrac{\theta - \theta_{etilt}}{\theta_{3dB}}\right)^2, SLA_v\right]$ $\theta_{3dB} = 10^0, SLA_v = 20$ dB, $\theta_{etilt} = 15°$
Combining method in 3D antenna pattern	$A(\varphi, \theta) = -Min\{-[A_H(\varphi) + A_v(\theta)], A_m\}$
Maximum transmit antenna gain	14 dBi (for narrow beams in GNB model, 20 dBi)
Antenna height at base station	32 m
Antenna height at user	1.5 m

Table 4.3 Simulation parameters of 3GPP Case 1 reference scenario

The base station parameters including transmit power, antenna pattern, maximum antenna gain and height are the same as those used in the system level simulator presented in Section 4.3. Base station inter-site distance, pathloss model and penetration loss are also the same. The standard deviation of shadowing loss is increased to 8dB. The shadowing values of a user to the different base stations are correlated. For each user location, a

common shadowing value is computed. The correlation is achieved by adding the common value with the base station shadowing value and dividing the sum by $\sqrt{2}$.

We will continue to use the 8x2 Single-User (SU) MIMO mapping table by GreenTouch [117] to convert the users' SINR into spectral efficiency in GNB network simulation. The MIMO scheduler in the simulator from which the table is derived uses single-user scheduling, thus Multi-User (MU) MIMO gains has not been included in the table nor in the GNB simulation. As discussed, the table has accounts for a number of features such as modulation format and level of error-correction coding appropriate to the SINR, retransmissions of PRBs and coding efficiency, and fast fading statistic for user equipment speed of 3km/h. The table has also taken into account transmission of signalling and control PRBs in non-CoMP scenario, which involves approximately 10-15% of the total available resource.

In the GNB simulation, we have further accounted for the additional signalling overhead incurred due to gated narrow beam coordination by reducing the total number of available PRBs by 3.34%. This percentage is estimated later in Chapter 5 and is based on the signalling and control scheme developed to support the GNB coordination.

For CoMP evaluation, 3GPP set the baseline scenario to 3 coordinated base stations which involves both intra-site coordination and inter-site coordination. Possible transmission modes include MU-MIMO and/or SU-MIMO operation in conjunction with JP or CS/CB. The antenna configurations involves having 2 or 4 or 8 transmit antennas at base station side, and 2 or 4 receive antennas at user side.

4.5.2 Playground Configuration and Traffic Model Analysis

Most of the system level simulations based on 3GPP Case 1 scenario model 19 base stations consisting of a central base station, a first ring of 6 base stations and a second ring of 12 base stations. In some of the studies [126-128] that are considered for performance comparison with GNB, a wrap-around layout is further applied to approximate the outer tiers' interference. An additional third ring of 18 base stations is modelled in GNB simulation to create an equivalent total interference.

The simulated narrow beam coordination area remains to be the single cluster consisting of 3 sectors as described in Section 4.3.3 and shown as a pink hexagon in Figure 4.6. All

of the narrow beams outside of the coordination cluster generate interference continuously in response to the full buffer condition. One of the four narrow beams in each sector is randomly selected to be switched 'On' at each epoch in the simulation.

Full buffer traffic model is used in the 3GPP Case 1 reference scenario for benchmarking the results of different LTE technologies. As previously discussed in Section 4.3.4, the full buffer model assumes all base stations transmit 100% of the time. Although not particularly realistic, it is widely used in the literature for determining theoretically achievable gains.

A user density of 10 users per sector is considered in 3GPP full buffer traffic model. Thus there are a total of 30 simultaneous users in the simulated GNB coordination cluster (pink hexagon). For simulation of the users in this full buffer traffic model, all the 30 random users (10 users per sector) are dropped into the coordination cluster from the beginning of a run. The users remain at their locations throughout the entire simulated time and continuously request for data. Note the user download is continuous and not file based in full buffer traffic model.

After a run is finished, the users will be terminated and their statistics are collected. This process is repeated with the users dropped at new random locations. For a given user drop (in one run), only 30 random locations are evaluated. Thus a large number of drops are simulated to ensure accuracy in system performance estimation. In the following section of simulation results, the 3GPP performance metrics presented for GNB model are assessed from a total of 3000 runs and 90,000 user locations. Each run has a simulated time of 3000 ms (3000 time epochs).

Note that full buffer traffic model is applied in this section for the purpose of performance comparison with other CoMP studies under the 3GPP reference scenario. For energy efficiency evaluation of GNB model across this thesis (and comparison with conventional wide beam model), we adopt the more realistic FTP traffic model with varying dense urban traffic levels predicted for 2020. As have discussed, the throughput and energy efficiency estimations corresponding to full buffer condition can be over optimistic when charactering realistically dimensioned network.

Compared with full buffer traffic model, the base stations under dense urban FTP traffic are running at a much lower utilization. The maximum BS utilization is round 53% at

busy hour traffic. An approximation of full buffer condition of FTP traffic model is modelled in Section 4.4.3, by raising the FTP traffic to a level that all beams have to be 'On' almost all the time in order to serve the users. Note that this scenario is not a full buffer traffic model as the beams within the coordination cluster are still responding to FTP download traffic with users coming on and off. Despite that the base stations are running at nearly 100% of time in this scenario we have modelled in Section 4.4.3, the average number of simultaneous users in the coordination cluster is less than half of that in the 3GPP full buffer traffic model.

The user density of the 3GPP full buffer traffic model is fixed at 10 users per sector. This large number of users can potentially provide GNB model with a better opportunity to perform narrow beam coordination and to mitigate interference compared with the FTP traffic simulated before. Thus an improved cell capacity can potentially be achieved. We discuss this further in the following paragraphs.

As stated in Section 4.2.3, the coordination approach of gated narrow beams is applicable for both low and high traffic conditions. However, the opportunity of scheduling beams to avoid interference improves as traffic load increases. An illustration was shown in Figure 4.5 of how additional users under the coverage of different narrow beams can lead to more opportunities for sensible beam combinations to be chosen.

It is observed from simulation of 3GPP Case 1 full buffer traffic that when 10 users are dropped within a sector of 4 narrow beams randomly and uniformly, the probability that all 4 narrow beams have users to host is more than 77%. The probability that 3 narrow beams have users to host (and only 1 narrow beam doesn't) is more than 21%. These add up to more than 98% probability. This high probability suggests that this traffic scenario can provide the coordination cluster with a large number of sensible beam combinations to be chosen. The chance that a good beam combination (one that avoids direct interference to each other) being selected by the scheduler is greatly increased.

The GNB model has adopted proportional fair scheduling in the beam coordination and resource allocation scheme. The proportional fair algorithm is applied in each individual sector's resource allocation process. In one transmission epoch, all the users being hosted by the narrow beam that is switched 'On' share the total available PRBs according to proportional fair scheduling. The proportional fairness parameter α is also applied to

weigh users' historical and current throughput performance in the process of evaluating beam combinations.

The large number of users starting download simultaneously also provides the proportional fair scheduling with more flexibility. Under low-level FTP traffic, the performance of GNB system is insensitive to the value of alpha as the probability of multiple users being in the same narrow beam is very low. On the contrary, there is more than 75% probability that a narrow beam hosts two or more users at the same time under the 3GPP Case 1 full buffer traffic. We conduct a set of simulations with different α values in the beam coordination and resource scheduling scheme. It is found that the best throughput performance is achieved when alpha is close to 1 (when the resource allocation places greatest weighting on the ratio of historical throughput to current spectral efficiency). An alpha value of 0.99 is set in the simulation to produce results for performance comparison.

4.5.3 Simulation results

First, we discuss the system performance metrics proposed by 3GPP for evaluation with full buffer traffic model. In 3GPP terminology, one sector of macro base station is referred to as 'cell'. The 3GPP terminology is adopted in the following discussion in this section. The metrics used for performance evaluation and comparison are:

- Cell spectral efficiency (bit/sec/Hz). This measures the sum throughput /capacity of one cell. This is the total number of data bits transmitted by a cell divided by the elapsed time and system bandwidth.
- Cell-edge user spectral efficiency (bit/sec/Hz). This measures the performance of cell-edge users. It evaluates a CoMP scheme's ability to manage inter-cell interference and enhance cell-edge throughput. The cell edge throughput is represented by the 5% worst user spectral efficiency. User spectral efficiency is calculated by the total data bits received by a user divided by the duration of download and bandwidth.
- Distribution of user throughput (bit/sec/Hz). This is the cumulative distribution function (CDF) of user spectral efficiency. It shows performance statistics of all simulated users. The cell-edge user spectral efficiency is the 5th percentile in the CDF.

These performance metrics by 3GPP are adopted and we estimate the cell spectral efficiency and user spectral efficiency of our proposed GNB system via simulation. The total data bits received by each user and each cell over the download time are recorded from simulation. The spectral efficiency metrics are then assessed based on the total data bits as per the 3GPP definition.

For the 3GPP Case 1 scenario, 3GPP has set targets for the spectral efficiency metrics corresponding to 2x2 and 4x2 antenna configurations. 3GPP has also reported evaluation results from extensive simulation conducted by a number of companies for 2x2 and 4x2 JP and CS/CB CoMP schemes under 3GPP Case 1 scenario.

The performance metrics of our proposed GNB system is estimated via simulation, and is compared against the 3GPP target and companies' evaluation results. The comparison of cell spectral efficiency is summarized and shown in below Figure 4.10. The round dots represent CS/CB schemes and the diamond-shaped dots represent JP schemes. The 3GPP targets are marked as short solid lines.

The results from different sources are shown in different colors. The evaluation results from 3GPP are colored black. In addition, we have included a number studies in the literature that have conducted system level simulation of various CoMP schemes based on 3GPP Case 1 scenario. The following paragraphs provide a brief explanation on the CoMP schemes considered by those studies.

Sun et al. [126] from Bell Labs conducted simulations of a MU-MIMO Joint Transmission network and a CS/CB network under 2x2 antenna configuration. Their CS/CB scheme considered a modified signal-to-leakage-and-noise ratio precoding algorithm to mitigate the leakage interference under the assumption of imperfect CSI. Sun et al. has also conducted simulation of a coherent Joint Transmission scheme under 8x2 antenna configuration [129]. Both simulations used proportional fair as the scheduling algorithm and assumed a fixed signalling and control overhead of 28.6%. The results are colored blue in Figure 4.10.

Zhang et al. [127] from Nokia Siemens Networks estimated the spectral efficiency metrics for an intra-site Joint Processing network with a layer adaptive user scheduling algorithm. We use their simulation results generated with 8x2 antenna configuration and

singular-value decomposition precoding for performance comparison with GNB model. The result is colored green in Figure 4.10.

The results by Xiong et al. [128] from Bell Labs is for an intra-site CS/CB network under 4x2 antenna configuration. The proposed CS/CB scheme used a new CQI feedback method along with an improved greed search scheduling method. Non-ideal channel estimation and a fixed signalling and control overhead of 30.63% are assumed. The result is colored purple in Figure 4.10.

Zhu et al. [130] from Docomo estimated the spectral efficiency metrics of a MU-MIMO JP network under different feedback scenarios including ideal feedback and limited feedback. We show the result corresponding to practical feedback scenario (subband 8-bits feedback) in Figure 4.10. The simulation uses 4x2 antenna configuration. The result is colored pink.
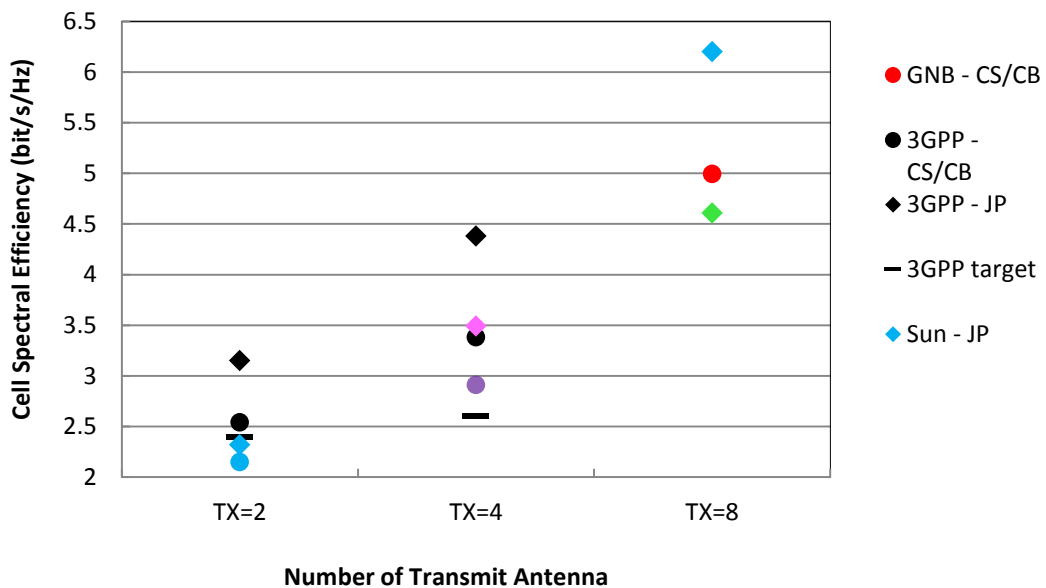
Figure 4.10 Cell spectral efficiency comparison of GNB model and other CoMP schemes

Figure 4.10 presents the cell spectral efficiency comparison of GNB model, other CoMP schemes in the literature and 3GPP targets. The x-axis is the number of transmit antenna configured in the system level simulations. The y-axis is the cell spectral efficiency in bit/sec/Hz. The estimated cell spectral efficiency of GNB network is around 4.99 bit/s/Hz and is shown as a red dot in the figure.

From Figure 4.10, we have the following observations concerning cell spectral efficiency performance of CoMP systems.

- JP scheme can achieve the best performance in terms of average cell spectral efficiency under all the antenna configurations considered. In general, the cell spectral efficiency increases as the number of transmit antenna increases for both categories CoMP-JP and CoMP-CS/CB.

- There is a large performance difference between different coordination schemes (or strategies) under the same CoMP category and under the same antenna configuration. However, note that there are other factors in the simulations that could contribute to the difference, such as the detailed signal processing algorithm at the transmitter and receiver, inclusion of signalling overhead etc.

- With consideration of practical constraints, the performance of JP scheme can be degraded to around the same level as CS/CB scheme (comparing Zhu's JP results in pink and 3GPP's CS/CB result in black under TX=4). Certain CS/CB scheme can even outperform JP scheme (comparing 3GPP's CS/CB result in black and Sun's JP result in blue under TX=2).

It can be seen from Figure 4.10 that the network of GNB achieves better cell spectral efficiency than all other CoMP schemes deploying 2 or 4 transmit antennas. There is no cell spectral efficiency target set for 8x2 antenna configuration by 3GPP. However, the performance of GNB model exceeds the target for 4x2 antenna configuration by about 92%. Comparing the three cell spectral efficiency estimates under 8x2 antenna configuration, the performance of GNB model is 8.2% better than Zhang's JP scheme (4.99 bit/s/Hz vs. 4.61 bit/s/Hz), and 19.5% worse than Sun's JP scheme (4.99 bit/s/Hz vs. 6.20 bit/s/Hz).

As discussed in the literature review, downlink joint transmission requires both users' data and CSI to be shared between coordinating base stations. The higher throughput of joint processing technique is achieved at the cost of much higher backhaul requirement, increased signalling overhead and computational complexity [126]. In contrast, our proposed GNB model under CS/CB category is capable of delivering comparable throughput gain to joint processing schemes with less constraint on the backhaul and signalling and control.

In the following discussion of cell-edge user spectral efficiency and user throughput distribution, we will focus on comparing GNB model with the joint processing schemes that have the same number of transmit antennas as GNB. Figure 4.11 presents the

performance comparison of cell-edge user spectral efficiency. For each CoMP scheme, the cell-edge user spectral efficiency is shown against its cell spectral efficiency. Figure 4.12 depicts the CDFs of user spectral efficiency.
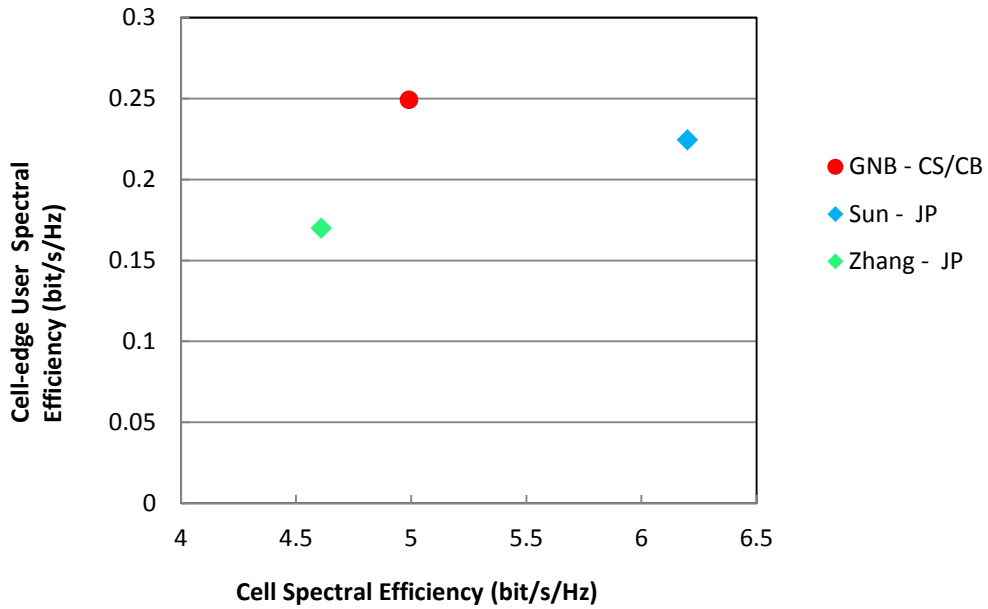
Figure 4.11 Cell-edge user spectral efficiency comparison of GNB model and other CoMP schemes
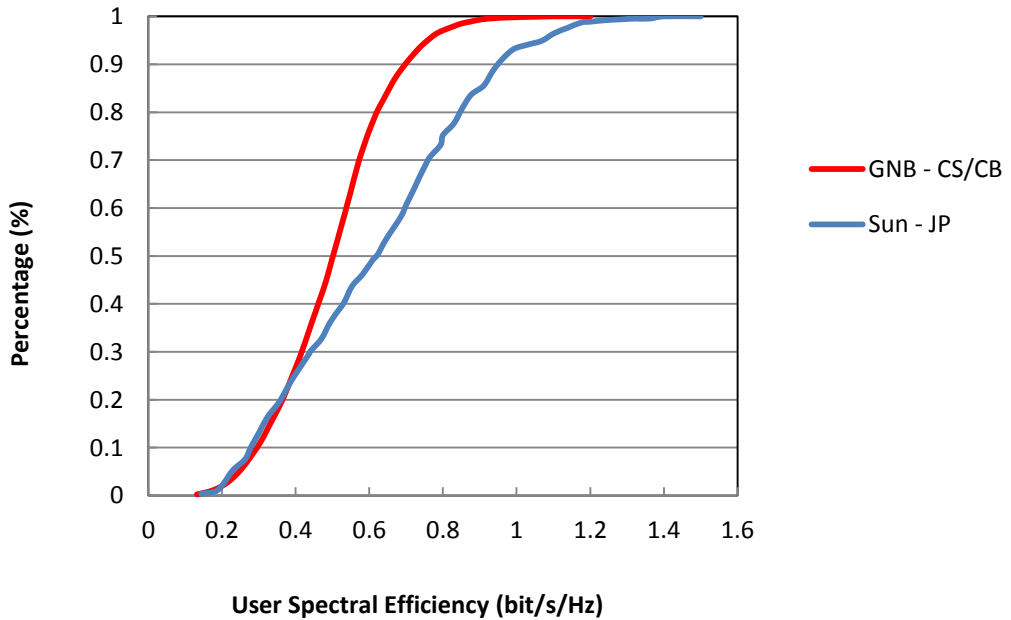
Figure 4.12 Comparison of CDFs of user spectral efficiency

It can be seen that the GNB model (in red) performs better than Zhang's JP scheme (in green) in both cell spectral efficiency and cell-edge user spectral efficiency. While the cell spectral efficiency of Sun's JP scheme (in blue) is better than GNB model, GNB model can provide higher cell-edge user spectral efficiency.

Figure 4.12 shows the CDF of user spectral efficiency of GNB model and Sun's JP scheme. X-axis is the user spectral efficiency in bit/s/Hz and y-axis is the cumulative percentage. It can be seen that the red curve is below the blue curve in the low spectral efficiency regions and on top of the blue curve in the high spectral efficiency region. A crossover point occurs at around 0.4 bit/s/Hz.

The low spectral efficiency corresponds to cell-edge users, while the high spectral efficiency corresponds to cell-center users. The figure suggests that the GNB model can provide better performance for users in the cell edge regions (low spectral efficiency areas). From mid to high spectral efficiency regions, Sun's JP scheme outperforms the GNB model (thus achieving greater total cell throughput).

Cell-edge performance is usually a bottleneck problem in realistic communication scenario [126]. The results in this section show that the proposed GNB model not only provides considerable gain in total cell throughput gain, it is also a good candidate scheme for improving cell edge performance in practical network.

4.6 Conclusion

To summarize, this chapter has demonstrated a scalable coordinated beamforming architecture, and designed a coordination scheme with proportional fair resource scheduler to mitigate inter-cell interference. A system level model and simulator was built to evaluate the throughput and energy efficiency of the proposed gated narrow beam network and the performance is benchmarked against traditional non-coordinating wide beam network. The throughput performance of the gated narrow beam model is also compared with other CoMP schemes under a common reference scenario.

In evaluating the performance, a realistic FTP traffic model with varying traffic loads and thus varying base station utilization rate is applied. The performance of a full buffer traffic model of 100% utilization was also investigated and compared to that of FTP traffic model to analyses the practicability and impact of the full buffer assumption. It

was concluded that the full buffer model is overly optimistic and misleading when characterising the energy efficiency of real network.

The simulation results demonstrated that significant energy efficiency improvement can potentially be achieved through the use of the proposed gated narrow beam model. However, it needs to be noted that the throughput and energy efficiency study of coordinated gated narrow beam model obtained in the chapter is not complete. The signalling overhead in support of the coordination scheme is not included in the throughput estimation, and there are other energy consuming components and procedures besides the base stations in the network that have not been included in the energy consumption. These studies will be conducted in Chapter 5 and Chapter 6 to provide completeness to the energy efficiency modelling.

Chapter 5

Signalling Scheme and Overhead

5.1 Introduction

CHAPTER 4 has developed a distributed and scalable coordinated beamforming scheme with multiple gated narrow beams for transmission of data. It has been demonstrated by simulation that the use of coordinated gated narrow beams in an LTE network has a significant performance advantage over conventional wide beams across the entire range of dense urban traffic load conditions. In particular, the coordinated gated narrow beam network delivers improved SINR and user download speed, higher cell throughput, and better average energy efficiency. The improvement in energy efficiency is mainly achieved by the user download speed, hence reduced download time which allows the base stations to operate in idle mode for a longer time where the associated energy consumption is less.

The energy consumption model in Chapter 4 only considers base station power consumption. The base station of gated narrow beam consumes more power than that of wide beam due to having more antennas and more radio chains and associated signal processing. However, besides the base station power, the realization of CoMP schemes including the coordination scheme proposed in Chapter 4 would also require additional signalling, more channel estimation effort, extra load on backhaul, processing of signalling and coordination information, plus possible changes in system architecture and other requirements such as synchronization and backhaul latency etc.

These additional functions to support the coordination will incur additional energy costs which have not been captured in the simplified energy consumption model in Chapter 4. Hence the study of average energy efficiency is only preliminary, and the improvement

could only represent the upper bound of possible energy efficiency gains. In order to give a realistic estimate of the overall energy efficiency of the network, the energy computation model should fully account for the energy cost associated with functions needed to support the beam coordination. The tradeoff between the attainable gains in cell throughput and costs associated with supporting the coordination will also need to be investigated.

This Chapter will focus on the design of a signalling and control framework to support the coordination scheme proposed in Chapter 4 and estimation of associated signalling overheads. The signalling overhead in terms of the (frequency-time or PRBs/REs) resources dedicated to signalling will reduce the amount of resource for data transmission and hence reduce the downlink rate. It also leads to an increase in energy cost as the base stations idle time will be shortened to compensate for the reduced downlink rate.

The study of signalling and components will also provide the foundation for Chapter 6 to quantify the energy consumption of major components and processes not included in the simplified power consumption model given in Chapter 4. These include the additional energy for processing of extra signalling information, computational energy for coordination and scheduling, and the energy cost of carrying extra signalling and control information in the backhaul network.

For the rest of the chapter, Section 5.2 will discuss the key design consideration of the signalling scheme and analyse the signalling requirement and information to support the coordination and scheduling scheme proposed in Chapter 4. An overview of implementation strategy will also be given in Section 5.2. Section 5.3 will develop an LTE based signalling scheme for GNB and will show that it to provide the required information identified in Section 5.2. Section 5.4 will analyse the signalling overhead associated with the proposed signalling scheme and Section 5.5 will present the signal and control information flow between various entities in the network. Section 5.6 concludes this chapter.

5.2 Signalling Requirements and Implementation Strategy for Gated Narrow Beam Network

5.2.1 Signalling Requirements

In the GNB model in Chapter 4, the narrow beam patterns are generated via codebook-based digital beamforming. A codebook with a limited set of possible precoding matrices is predefined. Each antenna is being fed with the same data signals, and applying the precoding matrix changes the amplitude and phase of the signal directed to each antenna, forming a spatially directed beam. Each precoding matrix is associated one to one with the set of narrow beam patterns. With this codebook-based beamforming approach, the users' Channel State Information (CSI) feedback can be implicit, containing only a few parameters [131]. The details of the parameters of implicit CSI feedback will be provided in Section 5.2.2. In general, the CSI feedback provides information on the state or condition of the communication link, and typically includes information about the SINR being experienced by a user.

The decision of coordination and resource scheduling algorithm developed in Chapter 4 is mainly based on the knowledge of users' SINR. Given the users' SINR, the proposed coordination and scheduling algorithm selects the best beams (or precoding matrix) for operation in the coordination cluster, and determines users to service and the amount of resource to be allocated. Plus, estimation of the users' achievable spectral efficiency under the chosen beam combination so that the link adaptation can function (a proper modulation and coding scheme is selected).

In Chapter 4 simulation, the users' SINRs are calculated based on the theoretical approach. The signal strength of every individual narrow beam is calculated based on path loss, antenna gain, shadowing loss etc., and applied towards the calculation of SINRs. However, in practice, these SINRs can only be obtained by implementing appropriate signalling processes that could generate the required users' CSI feedback. Chapter 4, as a first step to quantifying potential throughput gain, simply assumed that all necessary CSI feedback to support the coordination has been made available without investigating what signalling processes are actually required. This Chapter will design a signalling and CSI feedback scheme that delivers the same SINRs required to serve the coordination and resource scheduling algorithm discussed in Chapter 4.

The amount of CSI feedback provided by the user to support coordination of GNB would be greater than in non-coordination network. This is because a user's SINR under all possible beam combinations will need to be known in order for the selection of the best beam combination for the coordination cluster. In contrast, in non-coordination networks, only the SINR corresponding to the current state of the network would need to be known (instead of under all possible beam combinations). Hence a key design criterion of the signalling scheme is to provide all the required CSI feedback and SINR while maintaining a relatively low signalling overhead. In achieving this, and also to minimize the impact on existing signalling process, the signalling scheme for GNB will be based on the application of 3rd Generation Partnership Project (3GPP)'s existing signalling processes and reference signals that have been standardised in LTE network to support CoMP.

5.2.2 Required SINRs for Gated Narrow Beam Coordination

This section gives a detailed analysis of the set of SINRs required for any given user to serve the coordination and scheduling algorithm. Provision of this set of SINRs forms the target of the signalling scheme to be developed.

Recall that there are 4 possible narrow beams (or 5 including the null beam) in each sector and 125 (=5x5x5) possible beam combinations in a coordination cluster. The coordination and scheduling algorithm examines every beam combination exhaustively by calculating the sum of the users' utility function for each beam combination. The beam combination (out of the 125) that provides the best value of the sum of users' utility function is selected and scheduled to be used for next epoch. During the process of evaluating a certain beam combination, the utility function of each user is updated and the amount of resource allocated to the user is calculated based on the proportional fair scheduler. An accurate SINR of the user is critical to implementing the above scheme.

Although there are a total of 125 beam combinations in the coordination cluster, for any given user, only 25 of these combinations are valid. This is because once a user has selected a hosting beam, it changes only semi-statically, namely, it is assumed that there is no sufficient change in user's channel conditions that would warrant a change of hosting beam within a scheduling epoch. Hence the hosting beam is considered fixed during the dynamic coordination in epochs.

When the user is receiving service, the hosting beam will be at On state, and the other beams within the same sectors are Off. The valid beam combinations for the user are ones which include the hosting beam and any one of the 25 beam On or Off combinations in the two adjacent sectors. This reduces the number of SINR needed for a given user to support GNB. The following example and discussion are also focusing on a single representative user.

Figure 5.1 shows 3 base stations sectors (or referred to as 3 Transmission Points A, B and C where each has 4 beams A1 to A4, B1 to B4, and C1 to C4, respectively. A0, B0 and C0 are used to represent that there is no beam in the corresponding Transmission Point. Together the three sectors served by these beams form a coordination cluster.
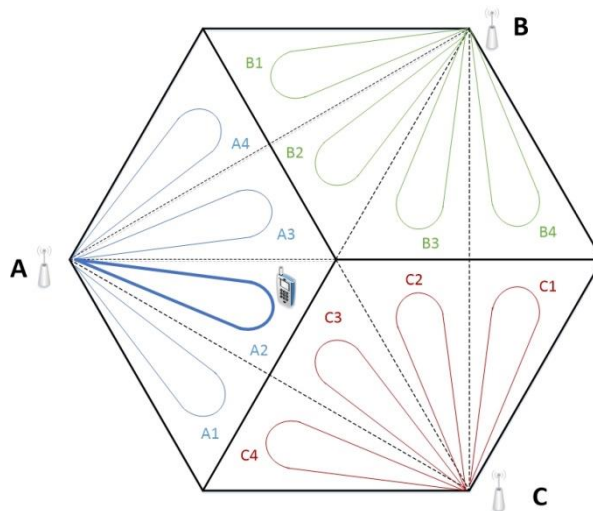
Figure 5.1 Interference scenarios for a representative user

A representative user is shown to be hosted by beam A2. The required SINRs under different interference scenarios are shown in Table 5.1. S_{A1} to S_{A4}, S_{B1} to S_{B4}, S_{C1} to S_{C4} represent the received signal power of the corresponding beam to the user in Figure 5.1. S_{A0}, S_{B0}, S_{C0} have zero value as there is no beam being turned on. I is the interference from outside the coordination cluster and N is the additive white Gaussian noise. As discussed, once the user has chosen a serving beam within a sector, it remains unchanged throughout the beam coordination process. Hence there are only 25 valid beam combinations from this user (1 (from sector A) $\times$ 5 (from sector B) $\times$ 5 (from sector C).

	B0	B1	B2	B3	B4
C0	$\dfrac{S_{A2}}{S_{B0} + S_{C0} + I + N}$	$\dfrac{S_{A2}}{S_{B1} + S_{C0} + I + N}$	$\dfrac{S_{A2}}{S_{B2} + S_{C0} + I + N}$	$\dfrac{S_{A2}}{S_{B3} + S_{C0} + I + N}$	$\dfrac{S_{A2}}{S_{B4} + S_{C0} + I + N}$
C1	$\dfrac{S_{A2}}{S_{B0} + S_{C1} + I + N}$	$\dfrac{S_{A2}}{S_{B1} + S_{C1} + I + N}$	$\dfrac{S_{A2}}{S_{B2} + S_{C1} + I + N}$	$\dfrac{S_{A2}}{S_{B3} + S_{C1} + I + N}$	$\dfrac{S_{A2}}{S_{B4} + S_{C1} + I + N}$
C2	$\dfrac{S_{A2}}{S_{B0} + S_{C2} + I + N}$	$\dfrac{S_{A2}}{S_{B1} + S_{C2} + I + N}$	$\dfrac{S_{A2}}{S_{B2} + S_{C2} + I + N}$	$\dfrac{S_{A2}}{S_{B3} + S_{C2} + I + N}$	$\dfrac{S_{A2}}{S_{B4} + S_{C2} + I + N}$
C3	$\dfrac{S_{A2}}{S_{B0} + S_{C3} + I + N}$	$\dfrac{S_{A2}}{S_{B1} + S_{C3} + I + N}$	$\dfrac{S_{A2}}{S_{B2} + S_{C3} + I + N}$	$\dfrac{S_{A2}}{S_{B3} + S_{C3} + I + N}$	$\dfrac{S_{A2}}{S_{B4} + S_{C3} + I + N}$
C4	$\dfrac{S_{A2}}{S_{B0} + S_{C4} + I + N}$	$\dfrac{S_{A2}}{S_{B1} + S_{C4} + I + N}$	$\dfrac{S_{A2}}{S_{B2} + S_{C4} + I + N}$	$\dfrac{S_{A2}}{S_{B3} + S_{C4} + I + N}$	$\dfrac{S_{A2}}{S_{B4} + S_{C4} + I + N}$

Table 5.1 25 SINRs required for a representative user

In general, for a user hosted by beam Ai of transmission point A (similar equation applies to users hosted by transmission point B and C), the required SINRs under different beam combinations are

$$SINR_{Ai,Bm+Cn} = \frac{S_{Ai}}{S_{Bm} + S_{Cn} + I + N} \quad for\ i \in (1,4), m = (0,4), n = (0,4) \quad (5.1)$$

Note that the 25 SINRs for each user will need to be updated periodically. The user SINR keeps changing as the user is often mobile and suffers from multipath fading even when the user is not moving. Up-to-date SINRs are required to be fed into the coordination and scheduling algorithm to ensure the accuracy of coordination decision. The frequency of updates will affect the overall signalling overhead, and it will be factored into the overhead calculation in Section 5.4. The following discussion in this section and Section 5.3 will focus on obtaining one set of 25 SINRs values for a given periodical update.

5.2.3 Implementation Strategy

As seen from Table 5.1, to fully assist the proposed coordination and scheduling scheme of GNB, the signalling scheme would only need to deliver the 25 SINRs for each user based on its CSI feedback. The design of the signalling scheme for GNB includes the following key concepts and implementation strategy:

1. Implicit CSI feedback is adopted for channel estimation and feedback in the signalling design for GNB. This is a standardized LTE procedure in the user handset to conduct measurements, computing and reporting the CSI feedback. To be able to provide the implicit CSI feedback, the user equipment essentially

measures two items: one concerns the elements in the MIMO channel matrix H which represent the propagation path between each of the transmit and receive antennas. The other is the interference plus noise power. The user's implicit CSI feedback includes a set of parameters - CQI (channel quality indicator), PMI (precoding matrix indicator) and RI (rank indicator). CQI is essentially a quantized form of an SINR value; they are considered to be equivalent in function and used interchangeably in this chapter. PMI and RI together represent the user's preferred precoding matrix selected from the codebook. Details of how these parameters are calculated will be discussed in Section 5.3.1.

2. The channel matrix and interference plus noise power are measured by the user equipment using a 3GPP Reference Signal called Channel State Information Reference Signal (CSI-RS) in the signalling design for GNB. The CSI-RSs are inserted in the downlink time-frequency resource grid for the purpose of channel estimation. The details of CSI-RS in LTE standards will be presented in Section 5.4.1, including how they are coded, where they are located in the PRB resource grid and the frequency of their transmission.

3. For one user, each CSI feedback will provide one SINR (or CQI). It can be expected that directly acquiring all the 25 SINRs (or CQIs) from CSI feedback for every user will introduce excessive signalling overhead. In the signalling design for GNB, a strategy in the 3GPP proposal [84, 132-134] is adopted to reduce the overhead by only acquiring a small number of SINRs (or CQIs) of selected scenarios from CSI feedback. The concept of this strategy is introduced in Section 5.3.2. Based on the small number of SINRs (or CQIs) of the selected scenarios, the required 25 SINRs for GNB can be derived by a series of further calculation.

4. The strategy adopted from the 3GPP proposal is named 'Per Point Feedback'. The selected scenarios under this strategy involve having the user served by each of the transmission points in the coordination cluster, and there is no interference from other transmission points inside the coordination cluster. By configuring a set of CSI-RS for these selected scenarios, the user measures the channel matrix and interference plus noise power, and feedback the parameters - CQI, PMI and RI corresponding to the selected scenarios. The implementation of this strategy on GNB will be discussed in Section 5.3.2. The specific configuration of CSI-RS for the selected scenarios will be discussed in Section 5.4.2.

5. The CQI (or SINR), PMI and RI acquired from CSI feedback for the small set of scenarios forms the basis for deriving the 25 required SINR values for the coordination and scheduling scheme. Conceptually these CQIs correspond to an expected CQI that a user would have experienced if they had been served by a given narrow beam and the interference only includes interference from outside the coordination cluster. The given narrow beam could be one that is hosting the user (e.g. beam A2 in the above example) or one of the beams from other transmission points (namely, B1-B4, or C1-C4) that the user would prefer if they were to be served by that transmission point. These CQIs (or SINRs) are used to calculate the 25 SINRs as in Table 5.1, by exploiting the correlation between precoding matrices to generate narrow beams. The details of the calculation steps will be shown in Section 5.3.3.

Essentially, the above-described signalling scheme for GNB requires measurement of the elements in the channel matrix for the channel from each transmission point to the user, and interference plus noise power from outside the coordination cluster. These measurements are performed on specific configured CSI-RSs and the usage of CSI-RS will impact on available PRB resources that could be deployed for user data transmission. From a power consumption perspective, the base station will spend less time in its idle state as each user will generally take a longer time to download. Consequently, more energy will be consumed. Section 5.4.2 will quantify the amount of signalling overhead incurred. These overheads will be accounted for in the estimation of a base station throughput and its energy efficiency in Chapter 6.

There are other approaches available to obtain the required 25 SINRs besides the proposed solution strategy involving implicit CSI feedback and per point feedback strategy. For example, it is possible to directly acquire all the 25 SINRs (or CQIs) from CSI feedback for every user. Or, a fewer number of SINRs of selected scenarios based on other strategies different from the 'Per Point Feedback' strategy. Generally, the greater number of SINRs being obtained by user CSI feedback, the more signalling overhead and signalling information carried on backhaul, and potentially less further signalling processing. It is also possible to have user feedback the entire channel matrix H including all the elements in it, which is different to the implicit CSI feedback, and then shift all calculations to obtain SINRs to the base station side. This would involve much more user feedback in the uplink, and also more computation effort in signalling processing.

The approach selected for CSI feedback and SINR estimation will affect the amount of PRB resource being taken up by reference signals, as well as overheads in the upstream signalling traffic. The intention of this chapter is to develop a signalling scheme that can support the proposed coordination and scheduling strategy in Chapter 4, and more importantly, present a methodology for quantifying the additional signalling overhead, the amount of signalling information exchanged on backhaul, and the computation effort for processing the user CSI feedback for CoMP systems. This will then enable the estimation of additional energy cost required in signalling and control, which will partly offset the energy savings identified in Chapter 4.

The GNB signalling scheme developed in this chapter leverages the CSI feedback and SINR estimation techniques proposed in 3GPP Technical Specification Group Radio Access Network Working Group 1 for CoMP. In addition, the implementation of this technique is based on the application of 3GPP LTE compliant signalling procedures and reference signals. This provides support that an LTE based GNB signalling scheme can be implemented without significant penalties in signalling overheads and hence throughput performance. However, it is beyond the scope of this thesis to fully examine the accuracy and optimality of the application of the 3GPP proposals in all CoMP systems. This is left as open issues for future research.

5.3 Signalling Scheme Design

5.3.1 Implicit Channel State Information Feedback

This section will first look at the procedure of implicit CSI feedback in LTE standards, especially the calculation of the parameters CQI, PMI, RI feedback. GNB signalling scheme will adopt this standardized procedure at user equipment, and make use of the CSI feedback of selected scenarios to derive SINRs in Table 5.1.

First consider the transmission model of a single user MIMO OFDM downlink transmission with N_T transmit antenna and N_R receive antenna is,

$$y = Hx + z = HWs + z \qquad (5.2)$$

where H is the MIMO channel matrix of dimension $N_T \times N_R$, z is background noise, x is the transmission signal vector, y is the received signal vector. s is the input data stream and has L independent layers, and a linear matrix operation is performed to obtain the

transmit vector $x = Ws$, where x is $N_T \times 1$ vector, W is the precoding matrix (or precoder) of dimension $N_T \times L$ that maps layers of input data to be ready for transmission on the transmit antennas. The number of transmission layers (or referred to as transmission rank) is equal to or smaller than the lower of the number of transmit antennas and the number of receive antennas. The layer (or rank) essentially represents the number of simultaneously transmitted parallel streams in a MIMO channel.

In implicit CSI feedback, a pre-defined codebook C is maintained at both the user and base station. The standard LTE codebook for base station with 2 or more antennas could be found in [93]. In general, the codebook has a set of possible precoding matrices W under each possible rank. The channel matrix H is used by the user to determine which precoding matrix W is preferred. Once the user measures the channel matrix H, the user search the codebook over all possible ranks and precoding matrices in each rank, then selects the preferred precoding matrix by maximizing the channel norm [84],

$$W' = arg \ \underset{W \in C}{max} \ \|HW\| \tag{5.3}$$

W' is the preferred precoding matrix that will maximise the argument $\|HW\|$ over all possible W in the codebook C. The user then feedback the optimal transmission rank called RI (that the preferred precoding matrix falls in) and the index of W' under the optimal transmission rank called PMI. The reported PMI and RI together correspond to a specific precoding matrix entry in the codebook.

As discussed, CQI is a quantized form of SINR value. In practice, a user will first obtain an SINR measurement (typically between -10dB and 25dB) and then map it to a CQI index (value between 1 and 16). The mapping of an SINR to a CQI value at user equipment is implementation dependent and varies between vendors and equipment types. In this chapter, the conversion between SINR and CQI is omitted in for simplicity. The potential inaccuracy that quantization error could introduce to the proposed beam coordination and resource scheduling algorithm is out of the scope of this study. As they are functionally equivalent, the terms CQI and SINR will be used interchangeably.

With the channel matrix measured and the preferred precoding matrix selected as per Equation 5.3, the signal power is calculated under the assumption that the selected precoder W' is applied. The CQI/SINR is calculated by dividing the signal power by the interference plus noise power [84, 135] as shown in Equation 5.4,

$$CQI = \frac{\text{Signal power}}{\text{Interference power +Noise power}} = \frac{\|HW'\|^2}{I_{total}+N} \qquad (5.4)$$

$I_{total} + N$ are measured by the user, I_{total} is the overall interference experienced by the user at the time of measurement (Note I in Equation 5.1 is a particular instance of I_{total}, i.e. the total interference is from base stations outside the coordination cluster). N is the noise.

The above process of determining the parameters PMI, RI and CQI is shown in a diagram in Figure 5.2 (red lines in User Entity). Essentially, the user needs to measure two items, the channel matrix H from the transmission point to the user, and the interference plus noise power (i.e. $I_{total} + N$). The user then searches the codebook with the measured channel matrix and finds the preferred precoding matrix W' (and corresponding indices PMI and RI). The precoding matrix together with the channel matrix allows the estimation of signal power. With the measured interference plus noise power, the CQI (or SINR) is achieved.

The PMI, RI and CQI obtained are fed back over uplink to the base station and be used in the signal processing of transmit signal as shown by the blue lines in Figure 5.2. Each CQI index maps to a specific modulation and coding scheme in the base station. In LTE, the base station can dynamically select appropriate modulation scheme based on the reported CQI in CSI feedback. This functionality is called link adaptation.

In a non-coordination network, the reported CQI would directly be used by the modulation mapper without further processing or calculation. The modulation scheme that corresponds to the CQI will simply be applied. The reported indices PMI and RI that together define the user preferred precoding matrix will be used at the precoding stage. As discussed, the same codebook is maintained at the base station, thus the precoding matrix corresponds to the PMI and RI will be acquired from the codebook, and be applied at the precoding stage.
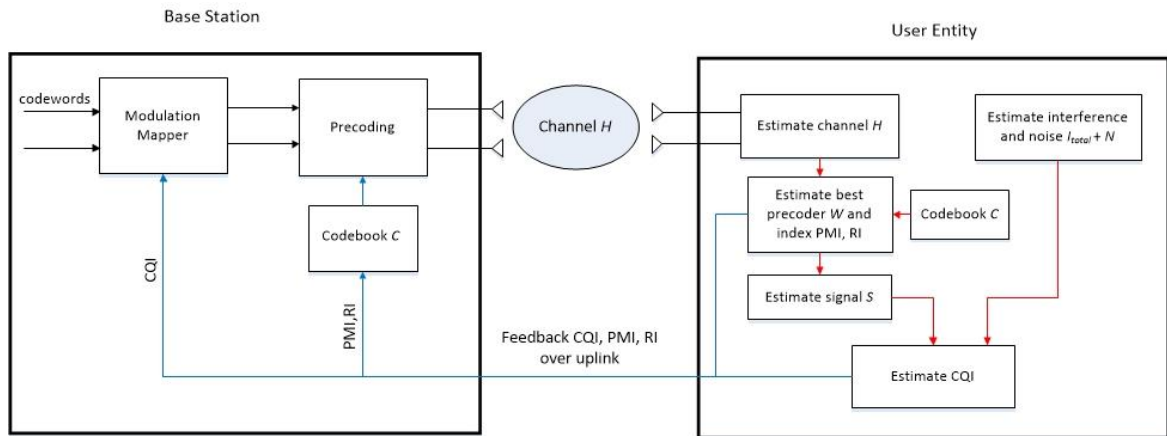
Figure 5.2 Process of standard implicit CSI feedback

Above discussion and figure describes the standard process of implicit CSI feedback. For GNB network, the process of implicit CSI feedback at the user equipment will remain the same as in Figure 5.2 since it is the standard procedure of the handset. The process includes the measurement of channel matrix and interference plus noise power, and the calculation of the parameter CQI, PMI and RI. The CSI feedback will be for certain selected scenarios though.

However, for a GNB network, the base station would not directly make use of the parameters (applying the corresponding modulation scheme and preferred precoding matrix). The CQIs reported for the selected scenarios will go through a series of calculation to derive the required 25 SINRs and then be fed into the coordination and scheduling algorithm. The coordination decision made by the algorithm includes predicted CQI for each user and chosen precoding matrix (or narrow beam) for each transmission point under coordination. This process will be discussed in details in the signalling flow diagram given in Section 5.4.3.

As discussed in Chapter 4, in GNB model, there are a total of 8 antennas where 4 antennas performing beamforming in each transmission layer. A maximum of 2 layers can be supported, to provide opportunity for further exploiting spatial multiplexing or diversity within the narrow beam. The patterns of beamforming (the narrow beams) are limited to 4 directions in each transmission point. Hence under each rank (equals 1 or 2), 4 precoding matrices need to be defined that each corresponds to one narrow beam pattern. Therefore, a total of 8 precoding matrices are needed in the codebook for GNB model. The PMI value ranges from 1 to 4, and the RI value ranges from 1 to 2 for CSI feedback of the GNB codebook.

In the following discussion, the numbering of a narrow beam pattern is the same as the PMI of the precoding matrix that forms the beam, in order to show the correspondence between narrow beams and precoding matrices. Take the example in Figure 5.1, beams A1 to A4 are generated by the precoding matrices of PMI 1 to 4, respectively. The same applies to beams of transmission points B and C. The codebooks used at each transmission point are essentially the same. They contain precoding matrices which change the beam direction relative to the boresight. For the example, in each sector in the coordination cluster, these precoding matrices will form beams which are +/- 15° and +/- 45° to the bore-sight. Designing of codebook is not straightforward and has attracted much research attention [136, 137]. In principle, a precoder should be matched to the eigendirections of the channel correlation matrix, and reported PMI essentially contain information about the eigenvectors of the channel covariance matrix. The codebook design is beyond the scope of this thesis and will not be further investigated.

5.3.2 Per Point Feedback for Gated Narrow Beam Model

As discussed in the solution strategy given in Section 5.2.3, the signalling scheme for GNB adopts a strategy proposed by 3GPP [84, 132-134] that only involves acquiring a small number of SINRs of selected scenarios from CSI feedback from which the 25 SINRs can be derived. The strategy is named 'Per Point Feedback' strategy.

The purpose of the 'Per Point Feedback' strategy when it was proposed by 3GPP is to reduce the amount of signalling and CSI feedback for various CoMP schemes [134]. The CSI feedback for the selected scenarios in the strategy could potentially be used as a basis to estimate SINRs for a range of CoMP transmission conditions including joint transmission, coordinated scheduling and beamforming, dynamic point selection etc. with reasonable accuracy [134].

This section will discuss the concept of 'Per Point Feedback' strategy and adaptation for the GNB model. Section 5.3.3 will demonstrate the derivation process of 25 SINRs shown in Table 5.1 from the limited CSI feedback obtained under the per point feedback strategy.

As discussed in solution strategy in Section 5.2.3, the selected scenarios in the per point strategy involve having the user served by each one of the transmission point under the interference assumption that the interference is only from outside the coordination cluster.

In 3GPP terminology, the channel from the assumed transmission point to the user is referred to as 'signal channel hypothesis'. The assumed interference is referred to as the 'interference hypothesis'.

Refer back to Equation 5.3 and Equation 5.4 for determining preferred precoding matrix (and corresponding indices PMI and RI) and CQI (or SINR). The 'signal channel hypothesis' determines the channel matrix, the 'interference hypothesis' determines the interference plus noise power. The preferred precoding matrix (and corresponding indices PMI and RI) can be obtained by searching the pre-defined codebook with the measured channel matrix. The CQI can also be obtained based on the channel matrix, precoding matrix, and interference plus noise power. Hence, the CSI feedback (CQI, PMI, RI) can be defined by specifying its signal channel hypothesis and interference hypothesis.

A new term 'CSI process' is defined by 3GPP [138] that describes a complete process of specifying of a signal channel hypothesis and an interference hypothesis, configuring specific CSI-RS for user to measure, and the user giving standard implicit CSI feedback via uplink based on the measurements. The 'Per Point Feedback' strategy essentially involves a number of CSI processes for each user. The number of CSI processes equals to the number of coordinating transmission points (which is 3 for GNB model). Each CSI process involves signal channel hypothesis of one transmission point. The interference hypotheses are common across all CSI processes that only the interference from outside the coordination cluster is accounted for.

Specific CSI-RSs are to be configured to enable the signal channel matrix measurement corresponding to the signal channel hypotheses, and interference plus noise power corresponding to the interference hypotheses. Measuring the channel matrix from non-hosting transmission point to a user is not standard in non-coordination network, and measuring interference and noise power from only outside the coordination cluster is not standard either. However, these measurements could be achieved by certain configuration of the CSI-RSs across the transmission points in the coordination cluster. The details of the configuration of the CSI-RS will be discussed in Section 5.4.2.

The following will demonstrate the implementation of the 'Pre Point Feedback' strategy to GNB model, including the specification of signal channel hypothesis and interference hypothesis, and the calculation and implication of parameters in CSI feedback.

In GNB network, every user within a coordination cluster will undertake 3 CSI processes and will report 3 CSI feedbacks (each containing a set of parameters CQI, PMI, RI). The 3 reported CQIs are as shown in Table 5.2, and the specification of signalling channel hypotheses and interference hypotheses in the 3 CSI processes are also shown. The hosting transmission point would instruct the user to apply these CSI processes via radio resource control signalling [84].

	Signal channel hypothesis	Interference hypothesis
CQI_A	Transmission point A	Outside coordination cluster
CQI_B	Transmission point B	Outside coordination cluster
CQI_C	Transmission point C	Outside coordination cluster

Table 5.2 Signal channel and interference hypotheses in Pre Point Feedback for a user

In the 3 CSI processes, the user equipment calculates 3 'Per Point CQI' (namely CQI_A, CQI_B, CQI_C) as well as the preferred precoding matrices according to the 'Per Point Feedback' strategy as follows:

$$CQI_A = CQI_{Ai} = \frac{\|H_A W_i\|^2}{I + N}; where \quad W_i = arg \ \underset{W_l \in C_A}{max} \|H_A W_l\| \qquad (5.5.a)$$

$$CQI_B = CQI_{Bj} = \frac{\|H_B W_j\|^2}{I + N}; where \quad W_j = arg \ \underset{W_m \in C_B}{max} \|H_B W_m\| \qquad (5.5.b)$$

$$CQI_C = CQI_{Ck} = \frac{\|H_C W_k\|^2}{I + N}; where \quad W_k = arg \ \underset{W_n \in C_C}{max} \|H_C W_n\| \qquad (5.5.c)$$

CQI_A, CQI_B, CQI_C are the three 'Per Point CQIs' calculated by the user according to the 'Per Point Feedback' strategy as shown in Equations 5.5 (a)-(c), respectively.

H_A, H_B, H_C are the channel matrices corresponding to the signal channel from transmission point A, B and C respectively, to the user.

W_i, W_j, W_k are the preferred precoding matrices determined by searching through all the precoding matrices of transmission point A, B, C's codebook C_A, C_B and C_C, respectively, as shown in Equations 5.5 (a) – (c).

W_l, W_m, W_n are the precoding matrices that belong to codebooks C_A , C_B and C_C , respectively.

C_A, C_B and C_C are the codebooks of transmission point A, B and C, respectively.

I is the interference from outside the coordination cluster and N is the noise.

Note that the selected preferred precoding matrices $\boldsymbol{W_i, W_j, W_k}$ correspond to narrow beams Ai, Bj and Ck of transmission A, B and C respectively. Recall that a CQI is always calculated based on the assumption that the preferred precoding matrix is being applied. The signal powers in the calculation of CQI_A, CQI_B, CQI_C are actually the potential received signal power of narrow beams Ai, Bj and Ck to the user, respectively. Hence CQI_A, CQI_B, CQI_C are also denoted as $CQI_{Ai}, CQI_{Bj}, CQI_{Ck}$. The 3 'Per Point CQI' are essentially expected CQIs that the user would have experienced if they had been served by the preferred narrow beam of each transmission point under the interference hypothesis.

Figure 5.3 below is an example for a user served by beam A2 of transmission point A (continuing on the example given in Figure 5.1).
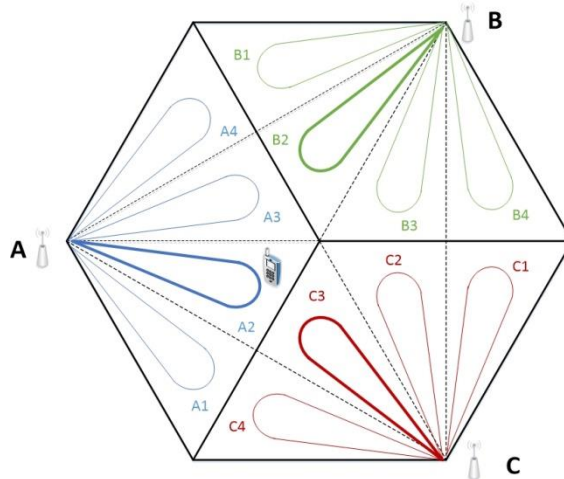
Figure 5.3 A representative user in the coordination cluster

The highlighted narrow beams A2, B2 and C3 in each transmission point represent the narrow beams corresponding to the preferred precoding matrices selected through Equation 5.5 (a) – (c). For this example, the preferred narrow beam indices in Equation 5.5 (a) – (c) are i=2, j=2, k=3. The reported CQIs essentially characterize the channel quality when narrow beam A2, B2 and C3 are to be used for transmission under the assumption that all interference is from outside the coordination cluster.

5.3.3 CQI Computation for Gated Narrow Beam Model

The next step is to derive the 25 required SINRs in Table 5.1 based on the user CSI feedback in the 3 CSI processes according to 'Per Point Feedback' strategy. The 3 reported CSI feedbacks include the 3 'Per Point CQI' as well as the indices (PMI and RI) for the preferred precoding matrices. It will be shown that a user would only need to undertake the 3 CSI processes according to the 'Per Point Feedback' strategy to derive all 25 SINRs.

In order to calculate the SINRs in Table 5.1, the following relationship needs to be recognized first:

$$CQI_A = CQI_{Ai} = \frac{\|\boldsymbol{H_A W_i}\|^2}{I + N} = \frac{S_{Ai}}{I + N} \qquad (5.6.\,a)$$

$$CQI_B = CQI_{Bj} = \frac{\|\boldsymbol{H_B W_j}\|^2}{I + N} = \frac{S_{Bj}}{I + N} \qquad (5.6.\,b)$$

$$CQI_C = CQI_{Ck} = \frac{\|\boldsymbol{H_C W_k}\|^2}{I + N} = \frac{S_{Ck}}{I + N} \qquad (5.6.\,c)$$

where $S_{Ai} = \|\boldsymbol{H_A W_i}\|^2$, $S_{Bj} = \|\boldsymbol{H_B W_j}\|^2$, $S_{Ck} = \|\boldsymbol{H_C W_k}\|^2$ are the power of narrow beams Ai, Bj and Ck to have been received by the user.

As discussed, CQI and SINR are used interchangeably in this chapter while in practice, a user will first obtain an SINR measurement and then map it to a CQI index to report. It is assumed that the Central Coordination Unit (CCU) will keep a same set of mapping between CQI and SINR that the user equipment uses. So that the CCU could convert the reported CQI back to the original SINR value, to assist further calculations. This process is omitted in this chapter; however, Chapter 6 will estimate the computation effort involved in the conversion. The converted SINR may only be an approximation of the original user calculated SINR value, since CQI is only a quantized value with lower granularity. The following discussion assumes that the conversion is done, and CQI and SINR are continued to be used interchangeably.

As have been explained, the 3 reported 'Per Point CQI' $CQI_{Ai}, CQI_{Bj}, CQI_{Ck}$ are the CQIs corresponding the preferred precoding matrices $\boldsymbol{W_i}, \boldsymbol{W_j}, \boldsymbol{W_k}$, respectively. The narrow beams corresponding to these preferred precoding matrices are Ai, Bj, and Ck of transmission point A, B and C, respectively. Equation 5.7 (a)-(c) below show how the 3 reported 'Per Point CQI' can be applied to calculate similar 'Per Point CQI'

corresponding to other precoding matrices/narrow beams of their respective transmission points, namely, CQI_{Al}, CQI_{Bm} and CQI_{Cn}.

$$CQI_{Al} = \frac{S_{Al}}{I + N} = \|(W_l)^H W_i\|^2 CQI_{Ai} \qquad \text{for } l = 1,4 \text{ and } l \neq i \qquad (5.7.a)$$

$$CQI_{Bm} = \frac{S_{Bm}}{I + N} = \|(W_m)^H W_j\|^2 CQI_{Bj} \qquad \text{for } m = 1,4 \text{ and } m \neq j \qquad (5.7.b)$$

$$CQI_{Cn} = \frac{S_{Cn}}{I + N} = \|(W_n)^H W_k\|^2 CQI_{Ck} \qquad \text{for } n = 1,4 \text{ and } n \neq k \qquad (5.7.c)$$

The calculation of these CQIs is achieved by exploiting the correlation between the preceding matrices in the codebook. Conceptually, the correlation between two precoding matrices can indicate the correlation between two signal channels provided by two corresponding narrow beams [132]. This specific derivation process as in the above set of equations involves multiplying the conjugate transposition of the target precoding matrix with the original precoding matrix, and then taking the square of the Frobenious Norm of the product. This approach is adopted from the work in [132, 133]. This formulation will be used for estimation of the computational effort for GNB coordination in Chapter 6.

The above derivation process is accurate and optimal when the precoding matrix is of rank one, as pointed in the discussion in [139, 140]. There will be some inaccuracy when the rank is more than one due to inter-stream interference. Given that the only one 'Per Point CQI' is needed for each narrow beam in this design, and the non-preferred matrices/narrow beams are very unlikely to support two lays of transmission, the choice of W_l, W_m, W_n here are limited to precoding matrices of rank one.

Note, CQI_{Al}, CQI_{Bm} and CQI_{Cn} correspond to the precoding matrices/narrow beams that were not chosen as the preferred precoding matrices/narrow beams by the user equipment in the CSI feedback. In addition to Equation 5.7 (a)-(c), let $CQI_{Al}, CQI_{Bm}, CQI_{Cn}$ be defined to have zero values representing the cases where there is no narrow beam being turned on in the corresponding transmission points.

Continuing on the example shown in Figure 5.3 where A2, B2 and C3 are the preferred narrow beams for transmission point A, B and C. The user is hosted by beam A2 of transmission point A. By applying Equation 5.7 (a)-(c), the 'Per Point CQI' corresponding to other narrow beams can be calculated and are given in column 3 in Table 5.3 below. The calculated 'Per Point CQI' is simply shown as a ratio of the received signal power of this narrow beam and the interference plus noise power. Each

term would need to be calculated by applying Equation 5.7 (a)-(c). For example, $CQI_{B1} = \frac{S_{B1}}{I+N} = \|(W_1)^H W_2\|^2 CQI_{B2}$ or $CQI_{C2} = \frac{S_{C2}}{I+N} = \|(W_2)^H W_3\|^2 CQI_{C3}$, where CQI_{B2} and CQI_{C3} are the 'Reported Per Point CQIs'.

Note that, there is actually no calculation required for the hosting transmission point, as the 'Per Point CQI' of non-preferred precoding matrices/narrow beams of the hosting transmission point will not come into use in this design. This follows from the assumption stated in Section 5.2.2, that a user's hosting beam remains unchanged in each scheduling epoch.

	Reported 'Per Point CQI'	Calculated 'Per Point CQI'
Transmission point A	$CQI_{A2} = \frac{S_{A2}}{I+N}$	$CQI_{A1}, CQI_{A3}, CQI_{A4}$ are not required.
Transmission point B	$CQI_{B2} = \frac{S_{B2}}{I+N}$	$CQI_{B1} = \frac{S_{B1}}{I+N}$, $CQI_{B3} = \frac{S_{B3}}{I+N}$, $CQI_{B4} = \frac{S_{B4}}{I+N}$
Transmission point C	$CQI_{C3} = \frac{S_{C3}}{I+N}$	$CQI_{C1} = \frac{S_{C1}}{I+N}$, $CQI_{C2} = \frac{S_{C2}}{I+N}$, $CQI_{C4} = \frac{S_{C4}}{I+N}$

Table 5.3 Example of the reported 'Per Point CQI' and the calculated 'Per Point CQI'

Hence there will be 3 reported 'Per Point CQI' ($CQI_{A2}, CQI_{B2}, CQI_{C3}$ in above example) and another 6 calculated 'Per Point CQI' (CQI_{B1}, CQI_{B3}, CQI_{B4} , CQI_{C1}, CQI_{C2}, CQI_{C4} in above example) available for computing the SINRs listed in Table 5.1. CQI_{A2} corresponds to the CQI of the hosting narrow beam under the condition that interference is only from beams outside the coordination cluster.

Similarly, CQI_{B1} to CQI_{B4} and CQI_{C1} to CQI_{C4} are the CQIs from other beams from the adjacent sectors in the coordination cluster. These CQIs are again calculated under the condition that interference is only from beams outside the coordination cluster. Note that these beams are essentially the narrow beams from adjacent sectors which could potentially interfere with signal from the hosting beam if they have been selected by the CCU to be switched "On".

It will now be shown how the SINR of a user served by beam Ai, as given in Equation 5.1, can be computed using the above reported and calculated 'Per Point CQIs' (namely CQI_{B1} to CQI_{B4} and CQI_{C1} to CQI_{C4}).

Reproducing Equation 5.1 below that expresses the 25 SINRs required for a user hosted by beam Ai of transmission point A. Similar equations apply to users hosted by transmission point B and C. The signalling scheme needs to deliver these 25 SINRs for every user to support the coordination of GNB.

$$SINR_{Ai,Bm+Cn} = \frac{S_{Ai}}{S_{Bm} + S_{Cn} + I + N} \quad \text{for i} \in (1,4), m = 0, 4, n = 0, 4 \tag{5.1}$$

The above equation can be rearranged as:

$$SINR_{Ai,Bm+Cn} = \frac{\dfrac{S_{Ai}}{I+N}}{\dfrac{S_{Bm}}{I+N} + \dfrac{S_{Cn}}{I+N} + \dfrac{I+N}{I+N}} \quad \text{for i} \in (1,4), m = 0, 4, n = 0, 4 \tag{5.8}$$

Now replacing the terms $\frac{S_{Ai}}{I+N}$ by CQI_{Ai}, $\frac{S_{Bm}}{I+N}$ by CQI_{Bm}, and $\frac{S_{Cn}}{I+N}$ by CQI_{Cn}, Equation 5.8 becomes:

$$SINR_{Ai,Bm+Cn} = \frac{CQI_{Ai}}{CQI_{Bm} + CQI_{Cn} + 1} \quad \text{for i} \in (1,4), m = 0, 4, n = 0, 4 \tag{5.9}$$

With Equation (5.9), all the SINRs required can be derived using the reported and calculated 'Per Point CQI'. Continuing on the example of user served by beam A2 of transmission point A, all entries in Table 5.1 can now be derived. The SINR values can now be rewritten as given in Table 5.4.

	B0	B1	B2	B3	B4
C0	$\dfrac{CQI_{A2}}{CQI_{B0} + CQI_{C0} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B1} + CQI_{C0} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B2} + CQI_{C0} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B3} + CQI_{C0} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B4} + CQI_{C0} + 1}$
C1	$\dfrac{CQI_{A2}}{CQI_{B0} + CQI_{C1} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B1} + CQI_{C1} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B2} + CQI_{C1} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B3} + CQI_{C1} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B4} + CQI_{C1} + 1}$
C2	$\dfrac{CQI_{A2}}{CQI_{B0} + CQI_{C2} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B1} + CQI_{C2} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B2} + CQI_{C2} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B3} + CQI_{C2} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B4} + CQI_{C2} + 1}$
C3	$\dfrac{CQI_{A2}}{CQI_{B0} + CQI_{C3} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B1} + CQI_{C3} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B2} + CQI_{C3} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B3} + CQI_{C3} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B4} + CQI_{C3} + 1}$
C4	$\dfrac{CQI_{A2}}{CQI_{B0} + CQI_{C4} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B1} + CQI_{C4} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B2} + CQI_{C4} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B3} + CQI_{C4} + 1}$	$\dfrac{CQI_{A2}}{CQI_{B4} + CQI_{C4} + 1}$

Table 5.4 25 SINRs required for a representative user expressed using reported 'Per Point CQI' and calculated 'Per Point CQI'

Now it has been shown that, following the above-described signalling scheme, the 25 SINRs can be derived entirely on the 3 CSI feedbacks of 3 CSI processes according to the

'Per Point Strategy'. As discussed, each CSI process involves a signal channel hypothesis of one transmission point and a common interference hypothesis that there is only interference from outside the coordination cluster.

For a user to determine CQI, PMI and RI in its CSI feedback, specific CSI-RSs will need to be configured to enable the signal channel matrix measurement corresponding to the signal channel hypotheses, and interference plus noise power corresponding to the interference hypotheses. The following Section 5.4 will discuss the configuration of CSI-RSs and signalling overhead associated.

5.4 Signal Channel and Interference Measurement Design using Channel State Information Reference Signal

5.4.1 Channel State Information Reference Signal in LTE Standards

This section will first examine a 3GPP Reference Signal called CSI-RS defined in the LTE standard to understand how it could be configured and applied to assist with the measurement of signal channel matrices and interference plus noise power.

LTE Release 9 and prior used Cell Reference Signal (CRS) for channel estimation and interference measurement; however, CRS requires significant usage of PRB resource when the number of antennas is large [84]. Thus, LTE Release 10 introduced CSI-RS of low time/frequency density to support the channel measurement for up to 8 antennas [141]. When LTE Release 11 introduced CoMP, it was decided that CSI-RS would continue to be used with some enhancements for the channel and interference measurement involved in CoMP scenarios [93, 142]. One important enhancement is enabling the user to estimate the CSI-RS for multiple neighbouring transmission points in addition to the user's hosting transmission point. This will be discussed in Section 5.4.1 below.

LTE Release 11 supports transmission of CSI-RS for 1, 2, 4 and 8 antennas [93]. There are a total of 40 Resource Element (REs) that could potentially be used for the reference signal symbols of CSI-RS in a PRB pair. A subset of the 40 REs is used for CSI-RS transmission for a given transmission point, and all the remaining REs can be used for

data transmission. The exact CSI-RS structure, in terms of the exact set of REs used for CSI-RS in the LTE resource grid, depends on the number of antennas in use at a transmission point, and varies for different transmission points. The CSI-RS pattern for 4 and 8 antenna as defined in Release 11 [93] is shown in Figure 5.4(a) and Figure 5.4(b).

CSI-RS pattern for 4 antennas in a PRB pair

									C8				
									C3				
						C5			C7			C9	
						C0			C2			C4	
									C6				
									C1				
									C8				
									C3				
						C5			C7			C9	
						C0			C2			C4	
									C6				
									C1				

Figure 5.4 (a) CSI-RS pattern for 4 antenna system in LTE Release 11

CSI-RS pattern for 8 antennas in a PRB pair

									C3				
									C3				
						C0			C2			C4	
						C0			C2			C4	
									C1				
									C1				
									C3				
									C3				
						C0			C2			C4	
						C0			C2			C4	
									C1				
									C1				

Figure 5.4 (b) CSI-RS pattern for 8 antenna system in LTE Release 11

In the case of a transmission point configured with 4 antennas, the CSI-RS consists of 4 REs (2 RE pairs to be specific) per PRB pair. Thus, there can be 10 different CSI-RS configuration (or called reuse patterns) based on the 40 possible REs. These different reuse patterns can be used by different transmission points, especially neighbouring transmission points, to avoid CSI-RS collisions (i.e. two or more transmission points are

transmitting their reference signals on the same frequency-time RE). The 10 reuse patterns are labelled C0-C9 as in Figure 5.4 (a). For a transmission point of 8 antennas, the CSI-RS consists of 8 REs (4 RE pairs), and there are only 5 reuse patterns, C0-C4, as shown in Figure 5.4 (b).

Note that the CSI-RSs of different antennas are orthogonally multiplexed by applying mutually Orthogonal Cover Codes (OCC), to enable physical separation of reference signals (from different transmit antennas) by a receive antenna. From the resource grid map perspective shown in Figure 5.4, all antennas will use the same selected reuse pattern for transmission CSI-RS. OCCs will be used to code the reference signal symbols to distinguish their point of origin, namely their transmission antenna.

A transmission point would typically configure one reuse pattern from Figure 5.4 to let the user perform signal channel matrix measurement for this transmission point. It is recommended the data transmissions in neighbouring transmission points that collide with the CSI-RS be muted to improve the accuracy of CSI-RS measurement by the user [143, 144]. For this purpose, the neighbouring transmission point will configure muted CSI-RS, which is essentially the same pattern as normal CSI-RS but with zero power. This feature will be implemented for GNB, and it will generate additional overhead by reducing the number of REs for data transmission.

The interference measurement could also benefit from the use of muted CSI-RS, especially CoMP systems. By configuring muted CSI-RS in some transmission points and normal CSI-RS in other transmission points, the user is then able to measure interference power for certain selected scenarios [145]. For example, in the 'Per Point Feedback' strategy, the interference hypothesis is that there is only interference from outside the coordination cluster, and the other two non-hosting transmission points do not generate interference. In this case, the two non-hosting transmission points could configure muted CSI-RS to enable the interference measurement. The muted CSI-RS for interference measurement purpose is named Interference Measurement Resource (IMR). Interference measurements are much simpler to conduct. The transmission points only have to agree on which RE(s) to mute. In 3GPP, it was recommended that the IMR be taken over 4 REs and configured as one CSI-RS reuse pattern of Figure 5.4 (a) [142].

The periodicity of the CSI-RS is between 5ms and 80ms [89]. The configured CSI-RS pattern (for signal channel matrices measurement) and IMR pattern (for interference

measurement) are repeated with this periodicity. The signalling overheads are dependent on (a) the number of antenna in use at the transmission points – which defines what reuse patterns are available; (b) the selected reuse pattern for CSI-RS and IMR to support the measurements; and (c) the periodicity of the measurements. The specific configuration of CSI-RS and IMR of GNB signalling scheme is given the following Section 5.4.2.

5.4.2 CSI-RS and IMR Configurations plus Overhead Analysis

Under the 'Per Point Feedback' strategy, the measurements essentially include the 3 MIMO channel matrices H_A, H_B, H_C corresponding to signal channel from transmission points A, B and C respectively, to the user. In addition, the measurements also include $I + N$ which is the interference power received by the user from outside the coordination cluster plus noise. Note there is only one measurement for interference plus noise power as the interference hypothesis in 3 CSI processes are the same. The user could simply reuse this measurement three times. With these measurements, the user will be able to determine CQI, PMI and RI in its CSI feedback as discussed in Section 5.3.2. The CSI feedback is reported to the user's hosting transmission point and then passed to the CCU to derive the 25 SINRs of its valid beam combinations.

Recall that 8 antennas are configured at the transmission points in GNB model in Chapter 4. Thus the CSI-RS pattern used here is the 8 antenna pattern as given in Figure 5.4 (b), so that the signalling overhead will align with the capacity modelling in Chapter 4.

5.4.2.1 CSI-RS Configuration and Associated Overhead

The selected CSI-RS patterns to support the signal channel matrices measurement at transmission points A, B and C are shown in Figure 5.5. The colour coded REs (CSI-RS-A, CSI-RS-B, CSI-RS-C) represents the set of CSI-RSs configured at transmission points A, B and C, respectively, at a given measurement epoch. Out of the 5 possible reuse patterns, it is assumed that the CSI-RS of transmission points A, B and C will use the reuse patterns as denoted by C0, C1, C2 in Figure 5.4.

Transmission point A
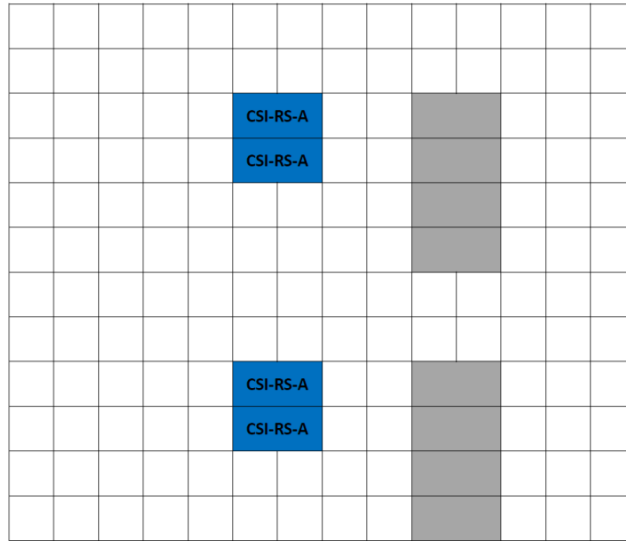
Transmission point B
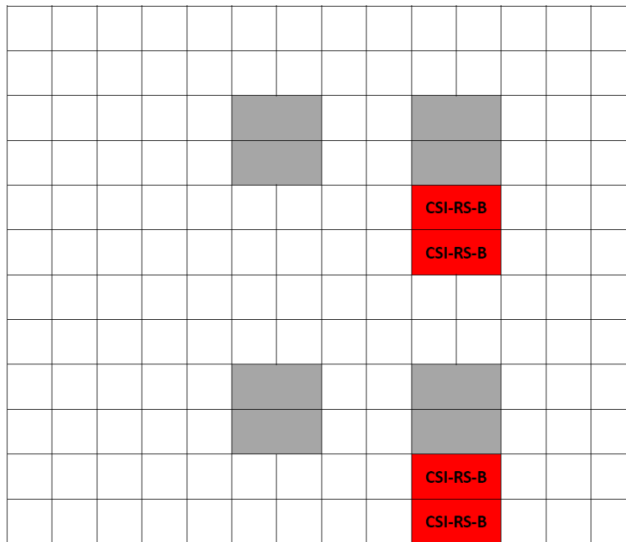
Transmission point C

Figure 5.5 CSI-RS patterns for 3 transmission points

The grey shaded REs represent CSI-RSs that are muted. As discussed in Section 5.4.1, these REs are muted to ensure that they do not interfere with the measurement of CSI-RS of neighbouring transmission point.

Hence the following subset of reuse patterns of 8 antennas will be configured by transmission point A (similar for transmission point B and C):

C0: configured as CSI-RS to support the measurement of the signal channel matrix $\boldsymbol{H_A}$;

C1: configured as muted CSI-RS so that transmission from transmission point A will not interfere with the CSI-RS in transmission point B;

C2: configured as muted CSI-RS so that transmission from transmission point A will not interfere with the CSI-RS in transmission point C;

The number of REs that need to be reserved at each transmission point to support channel measurements is further discussed below.

In the above design, the users are instructed by their hosting transmission point to measure the signal channel matrices of all 3 transmission points, namely its hosting transmission point and the other two non-hosting transmission points. As discussed in Section 5.4.1, this is an important adaptation of CSI-RS to support CoMP. This enables a user to estimate the CSI-RS for multiple neighbouring transmission points in addition to the user's hosting transmission point.

The reference signal sequence is initialized with a scrambling seed that contains a cell ID [84]. Traditionally (in Release 10 and prior), only the hosting base station's cell ID can be used to initialize its reference signals. A user only knows the cell ID of its hosting base station, thus a user can only descramble the reference signal sequences that are initialized by its hosting base station. To support CoMP systems, enhancement in Release 11 relaxed this requirement by allowing the reference signal sequence of CSI-RS to be initialized with any value ranging from all possible cell IDs [84]. The choice of proper cell IDs is left for specific implementation of CoMP.

One possible approach for GNB signalling design is to have the non-hosting transmission points transmit a CSI-RS sequence that is initialized with cell ID of the user's hosting transmission point. Then, each transmission point will need to transmit 2 additional CSI-RS patterns initialized with the other two transmission point's cell IDs, so that the users

hosted by the other two transmission points can measure its CSI-RS. This approach essentially triples the number of CSI-RS reuse patterns as shown in Figure 5.5.

Alternatively, if the users in a coordination cluster can be made aware of a 'cluster ID' that is specifically designed for GNB signalling, the CSI-RS sequence from all transmission points could potentially be initialized with this 'cluster ID'. In this approach, users will be able to descramble the CSI-RS sequences from non-hosting transmission points. The CSI-RSs configured as in Figure 5.5 can be measured by all users in the coordination cluster, and no additional CSI-RS is required than those shown in Figure 5.5. This approach of 'cluster ID' is adopted for signal channel measurement of non-hosting transmission points in GNB signalling.

The number of REs consumed to support the measurements of signal channel matrices can be deduced by inspecting Figure 5.5. For each transmission point, 8 REs are configured for CSI-RS (coloured) and another 16 REs are muted (grey). Hence this will contribute to (8 + 16) REs to the signalling overhead.

The configured CSI-RS reuse patterns along with the muted REs could be transmitted with the periodicity of 5, 10, 20, 40, or 80 ms [89]. If a periodicity of 5 ms is selected, this would correspond to the highest frequency of transmission of CSI-RSs and result in the greatest signalling overhead. Since there are 84 REs per PRB and 10 PRBs are transmitted in a cycle time of 5 ms, the percentage of signalling overheads to support the measurements of signal channel matrices would be:

$$\frac{(8+16)\,RE}{84\,RE\,per\,PRB\,\times\,10\,PRB} \approx 2.86\,\%$$

Note that it is not necessary that CSI-RS-A, CSI-RS-B and CSI-RS-C are all sent in the same subframe or measurement epoch. The example of them all being in one subframe is only for illustrative purpose. They could be in different subframes but the overall signalling overhead will be the same.

5.4.2.2 IMR Configuration and Associated Overhead

The selected IMR patterns to support measurement of interference plus noise power are shown in Figure 5.6. The REs coloured orange (IMR-A, IMR-B, IMR-C) represent the set of IMRs configured at transmission A, B and C, respectively, at a given measurement

epoch. As discussed, the IMR is essentially muted CSI-RS, and is configured as one CSI-RS reuse pattern of 4 antennas. Out of the 10 possible reuse patterns, it is assumed that the IMRs of all transmission points will use the reuse pattern as denoted by C9 in Figure 5.4 (a).

Transmission point A
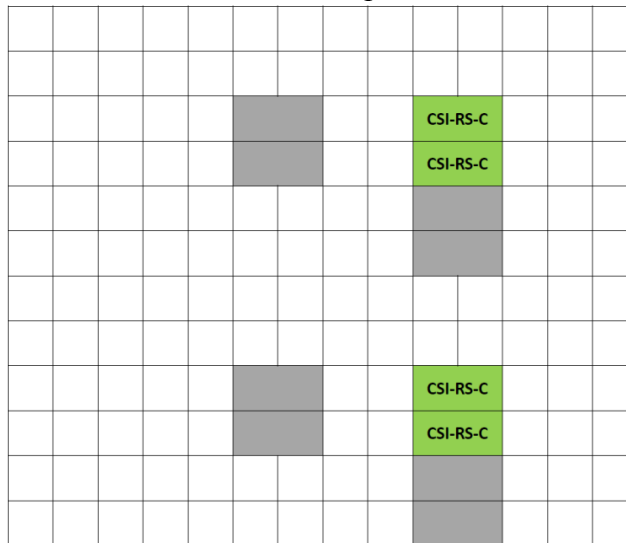

Transmission point B
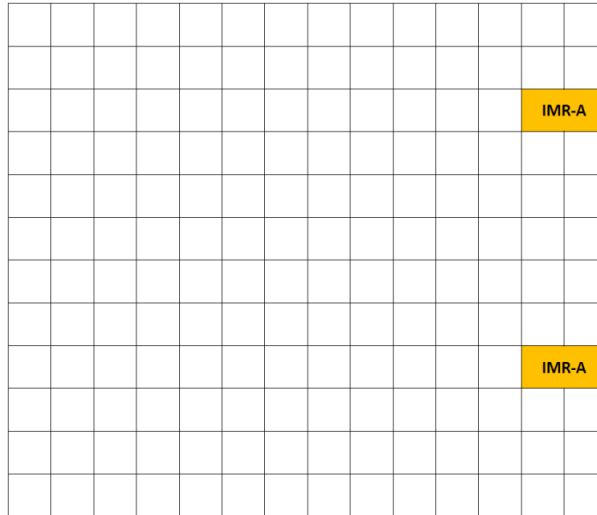
Transmission point C
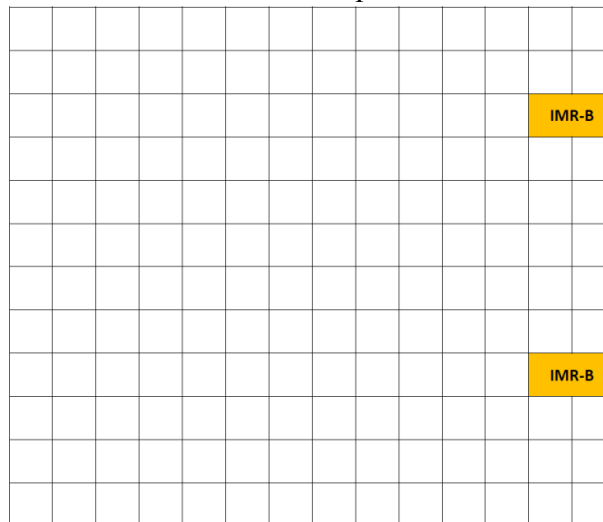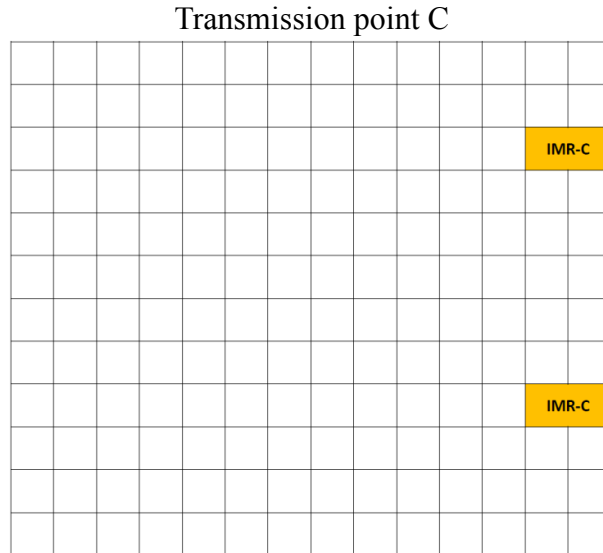
Figure 5.6 IMR patterns for 3 transmission points for interference plus noise power measurement

Note that it is essential that all transmission points use the same reuse pattern. Because all transmission points need to be muted at the same time and frequency (same RES) so that the received power measured by the users would be the interference power from only outside the coordination cluster plus noise power. Thus, the measured power would correspond to the interference hypothesis in the 'Per Point Feedback' strategy.

As shown in Figure 5.6, 4 REs are configured as IMR for each transmission point. Hence the percentage of signalling overhead to support the measurement of interference plus noise power would be:

$$\frac{4\ RE}{84\ RE\ per\ PRB \times 10\ PRB} \approx\ 0.48\ \%$$

5.4.2.3 Total Signalling Overhead

Hence the total signalling overhead associated with the proposed GNB signalling scheme is the sum of the overheads calculated in Section 5.4.2.1 and Section 5.4.2.2,

$$2.86\% + 0.48\ \% = 3.34\%$$

This overhead is calculated for our proposed coordination scheme which involves 3 transmission points each having 8 antennas in one cluster, and is based on the chosen appropriate parameter values and configuration patterns provided in the LTE standard. In specific, this percentage of overhead is estimated based on: configuration of CSI-RS

using 8 antenna pattern; configuration of IMR using 4 antenna pattern; and periodicity of CSI-RS and IMR transmission of 5ms.

Note that, under the 'cluster ID' approach, the configured CSI-RSs can be measured by all users. The configured IMRs can be measured by all users as well. Hence, the overhead introduced by the signalling scheme does not scale with the number of users or user traffic volume. However, the periodicity of measurement will affect the amount of signalling overhead. The estimated figure of 3.34% is the upper bound of the signalling overhead as a minimum periodicity of measurement (5ms) was assumed. The periodicity of measurement is dependent on the frequency with which the CSI feedbacks (and the derived 25 SINRs) are needed to serve the coordination and scheduling scheme. For example, if a user's signal channel conditions are changing quickly, its CSI feedbacks need to be reported more frequently and hence the measurements of CSI-RS and IMR need to be conducted more frequently too.

The coordination and scheduling scheme proposed in Chapter 4 works on a subframe basis (1ms) or on a longer term depending on the tradeoffs between signalling overhead and the potential performance degradation of using outdated CSI in the coordination. The determination of this tradeoff is beyond the scope of this thesis.

5.5 Signalling and Control Information Flow

This section describes the logical flow of signalling and control information of GNB network based on the proposed signalling scheme. The processes include the measurement and calculation performed at the user equipment, how the user CSI feedback is passed to Central Coordination Unit (CCU) to support the coordination and scheduling algorithm, and the coordination decisions from CCU to the BBUs.

Figure 5.8 shows the logical signalling and control information flow for 3 users (labelled UE1, UE2 and UE3) in one coordination cluster as shown in Figure 5.7. Each user will perform 3 CSI processes according to the 'Per Point Feedback' strategy. In each CSI process, the user measures the signal channel matrix of one transmission point on its CSI-RS, selects the preferred precoding matrix by searching the codebook, and calculates the signal power. The user also measures the interference plus noise power from outside the coordination cluster on the configured IMRs. Then the user calculates CQI. The user report 3 set of CQI, PMI and RI in its 3 CSI feedback.

To facilitate discussion, the subscript '1', '2' etc. are added to previous notations to identify UE1, UE2 etc. For UE1 performing a CSI process for transmission point A, it measures the channel H_{1A} on CSI-RS-A, selects the preferred precoder W_{1A} and index PMI_{1A}, and calculates the signal power S_{1A} assuming W_{1A} is applied. It also measures the interference plus noise power I+N on the configured IMR-A, and calculates CQI_{1A}. By repeating 3 CSI processes corresponding to 3 transmission points, it feedbacks the 3 'Per Point CQI' namely $CQI_{1A}, CQI_{1B}, CQI_{1C}$, and the indexes for the preferred precoding matrices $PMI_{1A}, PMI_{1B}, PMI_{1C}$ and $RI_{1A}, RI_{1B}, RI_{1C}$. The CQIs, PMIs and RIs are forwarded to the CCU and then applied to calculate the 25 SINRs according to the CQI computation process as explained in Section 5.3.3. Note that there is a conversion step from CQI to SINR prior to the computations which is omitted in Figure 5.8. This above process is the same for all users in the coordination cluster, including UE2 and UE3.

Once every user's 25 SINRs of valid beam combination are obtained, the main coordination and scheduling algorithm (as shown inside the CCU in Figure 5.8) will perform the coordination and scheduling scheme proposed in Chapter 4. The output from the CCU is the coordination decision for the 3 BBUs (each corresponds to one transmission point in the coordination cluster) including 3 PMI indices (PMI_A, PMI_B, PMI_C) for each BBU and one predicted CQI for each user in the coordination cluster.

In the example given in Figure 5.7, there are 3 users, thus 3 CQIs are produced, CQI_1, CQI_2, CQI_3 for UE1, UE2, UE3, respectively. Note that the coordination decision would also include the number of PRBs allocated to each user, which was determined by proportional fair scheduling as given in Chapter 4. For clarity, this plus certain overheads (user ID and protocol headers) have again been omitted from Figure 5.8. They will be included in calculation of the additional traffic flow in the backhaul in Chapter 6.

The PMI_A, PMI_B, PMI_C in the coordination decision represent the precoding matrices chosen by the CCU for narrow beam generation at the 3 transmission points. Recall that each precoding matrix when applied at the precoding stage forms a specific narrow beam. PMI_A, PMI_B, PMI_C collectively define a beam combination. The predicted user CQIs are under the assumption that the beam combination defined by PMI_A, PMI_B, PMI_C is applied. These predicted CQIs will be used in selecting proper modulation and coding scheme at the modulation mapper.
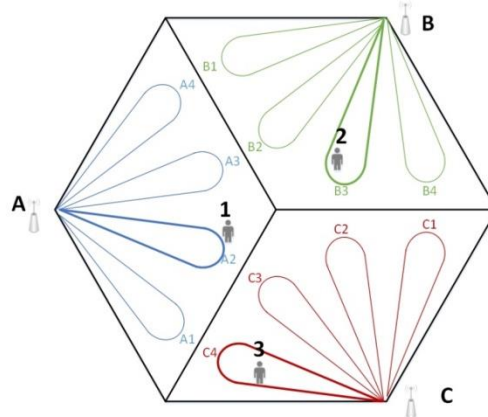
Figure 5.7 Representative users in the coordination cluster
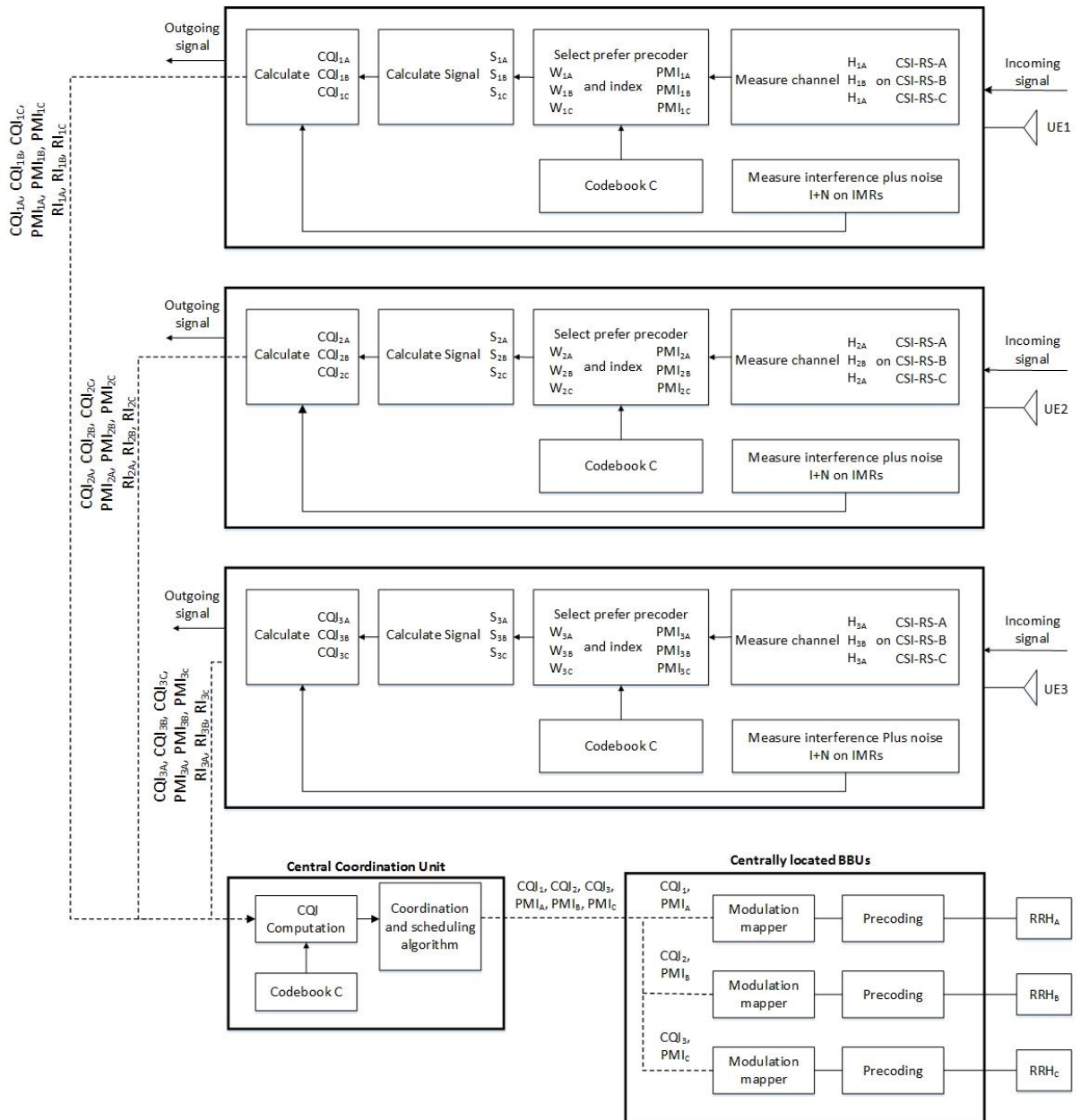
Figure 5.8 Logical signalling and control information flow for gated narrow beam model

5.6 Conclusion

To summarize, this chapter has developed a detailed signalling scheme for GNB that could fully support the proposed coordination and scheduling scheme in Chapter 4. The design of the proposed signalling scheme, where possible, is in compliance with the LTE standards. In particular, it makes use of CSI-RS signals and signalling procedures defined in Release 11 to support CoMP reference signal, CSI-RSs for user measurement of signal channel performance (characterized by channel matrices) and interference plus noise power from outside the coordination cluster. It also makes use of the standard LTE implicit CSI feedback of user equipment. A 'Per Point Feedback' strategy based on 3GPP proposal is adopted to help derive the SINRs required by the coordination and scheduling scheme from the user implicit CSI feedbacks. The main contribution of this chapter is presenting a methodology for quantifying the additional signalling overhead, the amount of signalling information exchanged on backhaul, and the computation effort for processing the user CSI feedback for CoMP systems in the context of an LTE network.

CSI-RSs of specific configuration are inserted into the resource grid, taking the place of the resource that could be otherwise occupied by user data transmission. The upper bound for the incurred signalling overhead is estimated to be approximately 3.34% when a minimum periodicity of measurement (5ms) was assumed. The percentage of the overhead is only dependent on the periodicity of measurement, and is not dependent on the number of users in the coordination cluster or the volume of user traffic.

Due to the signalling overhead, there are less available resources for data transmission resulting in a degraded cell throughput. To compensate for the reduced throughput, the base station will spend less time in idle state hence more energy will be consumed. The performance evaluation of GNB in Chapter 4 has not factored these in, but the signalling overhead will need to be accounted for to give a true estimate of energy efficiency of GNB network. The updated performance will be presented in Chapter 6.

In addition, the signalling and control framework presented in this chapter will also provide the foundation for quantifying the energy consumption of major components and processes not included in the simplified power consumption model given in Chapter 4. These include the additional energy for processing of extra signalling information, computational energy for coordination and scheduling algorithm, and the energy cost of

carrying extra signalling and control information in the backhaul network. Their calculations will also be presented in Chapter 6.

Chapter 6

Backhaul and Computation Energy

6.1 Introduction

A coordinated beamforming scheme was developed in Chapter 4 and its potential for energy efficiency improvements examined. A signalling and control framework was designed in Chapter 5 to provide the channel estimation information required for the beam coordination and resource scheduling.

The next step was to study the impact of the proposed coordination scheme and signalling scheme on energy efficiency by estimating the energy cost associated with the new functions and architecture enhancements needed to support the coordination and signalling. As discussed in Chapter 3, the other two energy consuming components and processes besides base stations are the backhaul information exchange and Central Coordination Unit (CCU) computation. In general, the signalling information needs to be exchanged through backhaul. The CCU uses this signalling information to decide the optimal beam coordination and resource scheduling based on the coordination and scheduling algorithm. This chapter will estimate the energy consumption associated with the backhaul network and with functions supporting coordination and scheduling.

For the rest of the chapter, Section 6.2 describes the proposed mobile network architecture and the procedures used to quantify its energy consumption. The energy consumption of the network includes energy consumption associated with information exchanged in the backhaul, signalling information processing and the main coordination

and scheduling algorithm at CCU. Section 6.3 summarizes the identified key information flow in the Upstream (or uplink) and Downstream (or downlink). This section also presents the estimate of the extra information volume generated for the purpose of coordination in each section of the flow, based on the signalling scheme proposed in Chapter 5. Furthermore, the additional energy consumption associated with the extra signalling information volume is also estimated. Section 6.4 investigates the computation energy consumption of the CCU that performs the critical calculations for processing of extra signalling information, and running the coordination and scheduling algorithm. An overall energy efficiency study that takes into account the backhaul and CCU energy as well as signalling overhead is presented in Section 6.5. The tradeoff of the energy gains and costs of the gated narrow beam model compared with the traditional wide beam model is discussed. In addition, a preliminary energy efficiency comparison against other CoMP schemes under the 3GPP reference scenario is performed. Finally, a conclusion of this chapter is given in Section 6.6.

6.2 Architecture of the Network and Procedures for Energy Consumption Estimation

Chapter 3 developed a high level architecture of functional and energy consuming components in one coordination cluster of network model. Here, the mobile network layout at a larger scale is provided including backhauling component and core control in Figure 6.1.
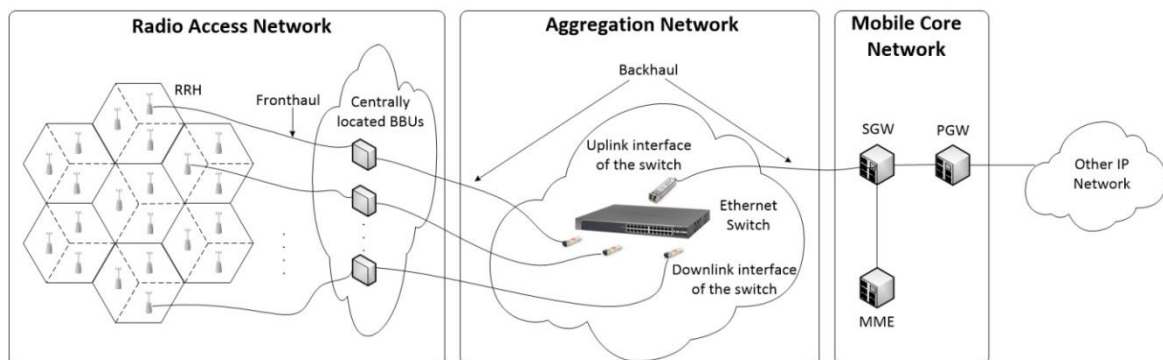
Figure 6.1 Network Architecture (7 base stations each serving 3 sectors with one RRH per sector)

As shown in Figure 6.1, the LTE cellular system consists of a radio access network, aggregation network and mobile core network. In the radio access network, a physical separation of the baseband processing unit and the radio unit of the base station is introduced following the current trend of LTE network [146]. The signal processing unit

is called Baseband Unit (BBU) and the radio unit is called Remote Radio Head (RRH). BBUs from a number of base stations are co-located in a central office leaving only the RRHs at the site to provide radio access. The separation between BBU and RRH create a connectivity segment named 'fronthaul'.

The fronthaul is based on digital radio over fibre technology, such as the widely used Common Public Radio Interface (CPRI) [90], where the baseband transmit and receive radio signals, i.e., IQ samples are digitized and carried over optical fibre. These fronthaul optical transport links are of low latency and high bandwidth supporting up to 10Gbit/s [90], hence the fronthaul could easily carry the additional traffic load due to coordination, and any additional energy consumption incurred would be minimal because of the low data rate (see later in Section 6.3.2 and Section 6.3.3).

The latency constraint could potentially limit the scale of BBU centralization. The fronthaul between RRH and BBU needs to be included in the total latency that radio signal can tolerate. The allowed time for round trip propagation delay is 400 μs for LTE-Advanced [90], and this corresponds to 40km permitted reach between RRH and BBU. The typical scale of the centralization for a medium-sized urban network is between 100 base stations for a coverage area of 5km x 5km and 1000 base station for 15km x 15km [88].

In the modelling architecture in this chapter, a service area covered by 5 concentric rings of base stations plus a central base station (a total of 91 BSs) is considered. Note that Figure 6.1 only shows a subset of those base stations (1 ring of BS and a central BS). As discussed in Chapter 4, the modelling is performed for the Dense Urban environment adopted by GreenTouch Mobile Communication Working Group [117]. The base stations are deployed in a hexagonal pattern with an inter-site distance of 0.5km. For the service area of 5 rings, the maximum distance between any 2 BS is thus 5km. Thus it is not unreasonable to assume that the distance between the centrally located BBUs and any RRH is well within the 40km allowed distance and associated latency.

The backhaul connects the BBUs with the mobile core network. There are different topology and technology options to implement backhauling [147] and the power consumption of the backhaul will vary for different deployment selections. The solution adopted in this modelling includes Ethernet backhauling with the medium of fibre optic, which is readily available and commonly used for LTE network [147]. Ethernet switches

can be flexibly located to have several levels of aggregation [148]. Figure 6.1 only shows one level of aggregation for the sake of simplicity. Optical fibre is the used for backhauling links from the BBUs to the aggregation switches. The traffic from all 5 rings of base stations plus a central base station is accumulated and collected at one central aggregation node with one or more aggregation switches (depending on the need).

The mobile core network consists of three entities: Serving GateWay (SGW), Mobility Management Entity (MME), and Packet data network Gateway (PGW). The backhaul network is connected to Serving GateWay (SGW) in the mobile core network. The main purpose of the SGW is to route and forward user data packets among LTE transmission points, and to manage handover among LTE and other 3GPP technologies [149]. For CoMP transmission, SGW coordinates the base stations within the coordination clusters, and is responsible for exchanging of channel station information, generating scheduling decisions [150]. Hence, the Central Coordination Unit (CCU) in the gated narrow beam model that operates the coordination and scheduling scheme is assumed to be a process running inside the SGW, operating along with other existing LTE scheduling protocols.

A complete energy modelling of the elements in the whole network architecture is required to give a true estimate of overall energy efficiency. In Figure 6.1, the energy consumption in the radio access network (including BBUs, fronthaul, RRHs and antennas) has already been factored into the base station energy consumption modelling in Chapter 4. The energy consumption in the backhauling of aggregation network and CCU computation within the SGW in mobile core network has not yet been studied, and needs to be accounted for in order to make a complete energy analysis.

The estimation of the energy consumption in the backhaul and CCU includes the following key procedures:

1. The energy consuming components in the backhaul are mainly the switches and their uplink and downlink interfaces. The power profile of these components needs to be identified to quantify their energy consumption, and most importantly, how it scales with the carried traffic. This profiling is necessary so that a relationship between the amount of traffic passing through the backhaul and the associated energy consumption can be established. Based on the understanding of the energy consuming components, a baseline backhaul power consumption that only considers IP/user traffic load will be estimated for the service area of 5 rings

of base stations plus a central base station. The IP/user traffic volume is common to both coordinated and non-coordinated network, thus this baseline backhaul power consumption is the same in both wide beam model and narrow beam model. The detail is discussed in Section 6.3.1.

2. Quantify the additional traffic passing through the backhaul to facilitate coordination for gated narrow beam model, in order to quantify the additional energy consumption in backhauling. The aggregation switch traffic has 3 components: IP/user traffic, the upstream user Channel State Information (CSI) feedback for coordination, and the downstream coordination decisions. The latter two components form the additional traffic exchanged through backhaul to support the gated narrow beam coordination. The bit rate of the upstream and downstream signalling and control data as a function of user traffic loads are studied in Section 6.3.2 and Section 6.3.3 respectively. The bit rate is initially estimated for a single coordination cluster and then scaled to the service area of 5 rings of base stations plus a central base station. The additional energy consumed in the backhaul is estimated based on the additional traffic passing through the backhaul energy consuming components.

3. Next is the estimation of the computation energy consumption of performing the calculations in the CCU for gated narrow beam model. The CCU is assumed to run on the existing processors in SGW along with other existing scheduling protocols. The calculations mainly involve two parts. One is processing of extra signalling information which is essentially the uplink user CSI feedback for coordination as in the design outlined in Chapter 5. The other is performing the core coordination and scheduling algorithm to generate the coordination decisions, including beam pattern decision and user scheduling decision, as in the design outlined in Chapter 6. By estimating the computation effort required, the incremental energy consumed by the processor can be estimated based on the energy efficiency of the processor. This is discussed in Section 6.4.

4. Finally, for gated narrow beam model, the complete energy consumption in the network can be obtained by adding the energy consumption in backhauling and CCU to the base station energy consumption computed in Chapter 4. In comparison with the preliminary energy efficiency evaluation in Chapter 4, the energy consumption increases. In addition to the increased energy consumption, the energy efficiency of the network is also impacted by the reduction of delivered

data bits due to signalling overheads as discussed in Chapter 5. The overall energy efficiency figure can then be derived by dividing the increased energy consumption by the reduced data bits. The final step is to compare the overall energy efficiency of a network using wide beams and a network using gated narrow beams to determine whether there is still advantage of the gated narrow beam. Note that the wide beam model also needs to factor in the baseline backhaul power consumption. This is discussed in Section 6.5.

6.3 Estimation of Baseline Backhaul Energy and Additional Backhaul Energy for Gated Narrow Beam Network

6.3.1 Baseline Backhaul Power Consumption

This section estimates the baseline backhaul power consumption of mobile network by modelling of the energy consuming components in backhaul. Only IP/user traffic is considered in this section, which is common to both wide beam model and narrow beam model.

The backhaul architecture considered in this modelling has been presented in Figure 6.1 - fibre optic network based on point-to-point Ethernet. One downlink interface and switch port for each BBU, and one BBU per RRH. The backhaul power consumption model is calculated based on the model presented in [147, 148, 151].

The backhaul power is shown in Equation 6.1 and includes 3 components: the power consumed at the aggregation switches (1st term); the downlink interface power consumption from the BBU to the aggregation switch (2nd term); and the uplink interface power consumption from the aggregation switch to the core mobile network ((3rd term).

The detail expression of backhaul power consumption model is as follows,

$$P_{bh} = \left\lceil \frac{N_{BS}}{max_{dl}} \right\rceil P_s + N_{BS}P_{dl} + \left\lceil \frac{Tot_{Ag}}{U_{max}} \right\rceil P_{ul} \tag{6.1}$$

where P_{bh} is the total energy of the backhaul (includes all switches and interfaces). N_{BS} is the number of base stations; max_{dl} and U_{max} are constants that denote the number of downlink interfaces of the switch, and the maximum transmission rate of an uplink interface of the switch, respectively. P_s represents the power consumed by the switch and

Tot_{Ag} is the total aggregate traffic collected at the switches; P_{dl} and P_{ul} are the power consumption of the downlink interface of an aggregation switch.

The number of downlink interfaces available at one switch (max_{dl}) is used to compute the total number of switches needed. It is assumed that each base station uses a dedicated downlink interface, hence $\frac{N_{BS}}{max_{dl}}$ is the number of switches required (round up if not integer) to collect the backhaul traffic from N_{BS} base stations. $\frac{N_{BS}}{max_{dl}} P_S$ is the total power consumption of the switches. As the number of downlink interfaces needed is the same as the number of base stations, $N_{BS}P_{dl}$ is the total power consumption of downlink interfaces. The number of uplink interfaces needed is calculated by $\frac{Tot_{Ag}}{U_{max}}$ (round up if not integer), and $\frac{Tot_{Ag}}{U_{max}} P_{ul}$ is the total power consumption of the uplink interfaces.

The power consumption of the switch (P_S) includes two main components. The first one is the traffic-independent idle power, which is mainly the power consumption of the backplane of the switch. The second one is the power consumption that is proportional to the amount of traffic passing through the switch(Ag_{switch}). The expression of the switch power consumption is as shown below,

$$P_S = P_{idle} + (P_{max} - P_{idle}) \frac{Ag_{switch}}{Ag_{max}} \tag{6.2}$$

P_{idle} is the idle power consumption of the switch, Ag_{max} is the maximum amount of traffic that the switch can handle. P_{max} is the maximum power consumption of the switch, corresponding to when the maximum traffic is passing through. With this equation, the switch power consumption at any load can be estimated. A more detailed explanation of these parameters can be found in [147, 148, 151].

In this modelling, the Cisco Catalyst 3750-X Switch (model C3KX-NM-10G) [152] is chosen as a representative aggregation switch. It could support 24 1Gbit/s small-form factor pluggable (SFP) ports. SFP is a device that is plugged into the ports of the aggregation switch with the other end connecting optical fibre cables from the BBUs, and converts between the serial electrical signals and serial optical signals. As each switch supports 24 BBU connections, for the service area of 5 rings plus a central base stations (which is 91 base station in total), 4 switches are required each supporting 23 or 22 base stations (91=23+23+23+22). For simplicity, the following modelling will assume that the

traffic passing through each switch (Ag_{switch}) is the same, and is estimated for 23 base stations. According to the datasheet provided by Cisco [152], the values for the parameters described in the backhaul power consumption model are: max_{dl} =24, U_{max}=10 Gbit/s, P_{dl}=2W, P_{ul}=3.5W, Ag_{max}=24Gbit/s, P_{idle}=85.7W, P_{max}=90.1W. For the designed service area, N_{BS}=91.

In order to quantify the power consumption of the switch (P_S), the amount of traffic passing through the switch needs to be estimated. The IP/user traffic volume in one base according to GreenTouch traffic load is shown in the 1st and 2nd row in Table 6.1. The traffic load (%) is the percentage of the predicted daily average traffic density in dense e urban environment, as used in GreenTouch modelling [117]. The traffic per base station is multiplied by 23 to obtain the traffic passing through each switch as shown in the 3rd row, and the power consumption of one switch as shown in the 4th row is calculated based on Equation (6.2). Finally, the total power consumption of the switches is given in the last row.

Traffic load (%)	20	30	50	70	100	120	140
Traffic per BS in Mbit/s	21.72	32.58	54.3	76.02	108.6	130.31	152.03
Traffic in each Switch in Gbit/s (Ag_{switch})	0.500 (=21.72x23/1000)	0.749	1.249	1.748	2.498	2.997	3.497
Power consumption of 1 switch in W (P_S)	85.79 (=85.7+4.4x0.500/24)	85.84	85.93	86.02	86.16	86.25	86.34
Total power consumption of (4) switches in W	343.17 (=85.79x4)	343.35	343.72	344.08	344.63	345.00	345.36

Table 6.1 Switch power consumption as a function of traffic load

Then, the baseline backhaul power consumption for the service area of 91 base stations can be calculated using Equation (6.1), the results are as shown in Table 6.2. As discussed, the number of uplink interfaces needed is calculated by $\lceil \frac{Tot_{Ag}}{U_{max}} \rceil$. In this modelling, the uplink interface configuration should be able to support the maximum total aggregated traffic, thus for each switch, the traffic is 3.497Gbit/s corresponding to the maximum traffic load of 140% as shown in Table 6.1.

The Cisco Catalyst 3750-X series allows stacking of switches to form a single logical unit [152]. Therefore, the 4 switches can be internally connected, and share common uplink

interfaces connecting with SGW. The number of uplink interfaces is calculated as $\frac{3.497\ Gbit/s\ per\ switch\ \times 4\ siwtches}{10\ Gbit/s} = 1.4$ and then round up to 2.

Given the stacking configuration, only one level of aggregation is required in this modelling, as shown in Figure 6.1. Note that, the number of downlink and uplink interfaces do not scale with traffic load. The backhaul architecture has been configured with one downlink interface per BBU, which provides more capacity than needed. The number of uplink interfaces is dimensioned for the maximum traffic. The power consumption of downlink and uplink interfaces will stay the same across different traffic loads as well.

Traffic load (%)	20	30	50	70	100	120	140
Total power consumption of (4) switches in W	343.17	343.35	343.72	344.08	344.63	345.00	345.36
Total power consumption of downlink Interfaces in W	182 (=91x2)	182	182	182	182	182	182
Total power consumption of uplink interfaces in W	7 (=2x3.5)	7	7	7	7	7	7
Baseline backhaul power consumption in W	532.17 (=343.17+182+7)	532.35	532.72	533.08	533.63	534.00	534.36

Table 6.2 Baseline backhaul power consumption as a function of traffic load

Thus, the baseline backhaul power consumption is obtained, and it is common to both wide beam model and narrow beam model. To implement the coordination scheme in the network, there will be extra signalling and control information on the backhaul. However, it will be shown in Section 6.3.2 and Section 6.3.3 that the increased signalling and control information is not going to be large enough to have an impact on the topology of connections in backhaul. The number of switches needed, and the number of downlink and uplink interfaces, for the backhauling of 91 base stations of which BBUs are centrally located, will remain the same. Therefore, the only component that will incur additional energy is the switch itself, as its energy consumption scales with traffic load (as seen in Equation 6.2).

As discussed, the additional traffic load will be estimated for one single coordination cluster initially and scale to the service area of 91 base stations. The additional backhaul power consumption of gated narrow beam model is estimated based on the additional

aggregated traffic passing through the switches. Chapter 3 has presented a high level architecture of functional and energy consuming components in one coordination cluster of network model. To facilitate discussion, the Figure 3.1 from Chapter 3 is reproduced below in Figure 6.2.
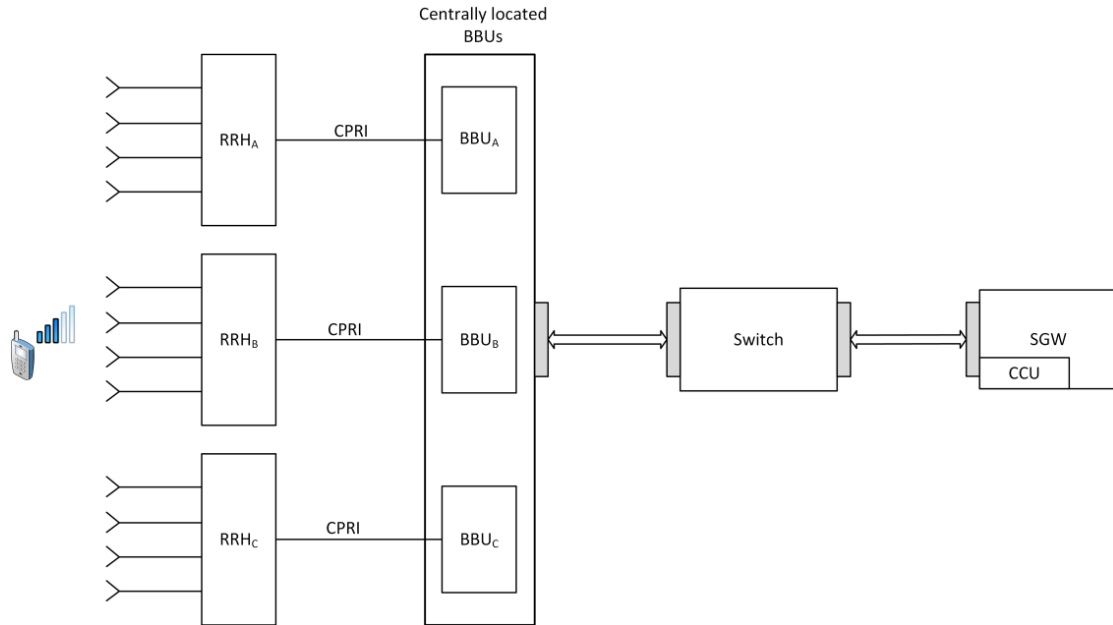
Figure 6.2 High level architecture for one coordination cluster

This architecture is focusing on one coordination cluster. The 3 RRHs represent the 3 sectors from 3 base stations in the coordination cluster, and the 3 corresponding BBUs are located in a central site. Note that physically each BBU is connected to the switch individually, and Figure 6.2 only shows one link for conceptual demonstration. Also, note that the one switch shown in Figure 6.2 is not dedicated to serve only one coordination cluster, the 4 internally connected switches support the entire service area of 91 base stations, which comprises 91 coordination clusters. The calculation in the next two sections starts with quantifying the additional information generated by one single coordination cluster.

6.3.2 Extra Channel State Information in the Upstream

This section studies the information flows and volume in the Upstream (or uplink) directions, focusing on the user CSI feedback required for coordination. The user CSI feedback for coordination plus other existing signalling and control information are generated by the user equipment and sent uplink towards the centrally located BBUs via fronthaul; this CSI feedback together with user connection and hosting information are

extracted from the centrally located BBUs and sent to the CCU via the aggregation switch. Figure 6.3 below shows the upstream information flow. For completeness, user data traffic in the uplink direction is also shown.
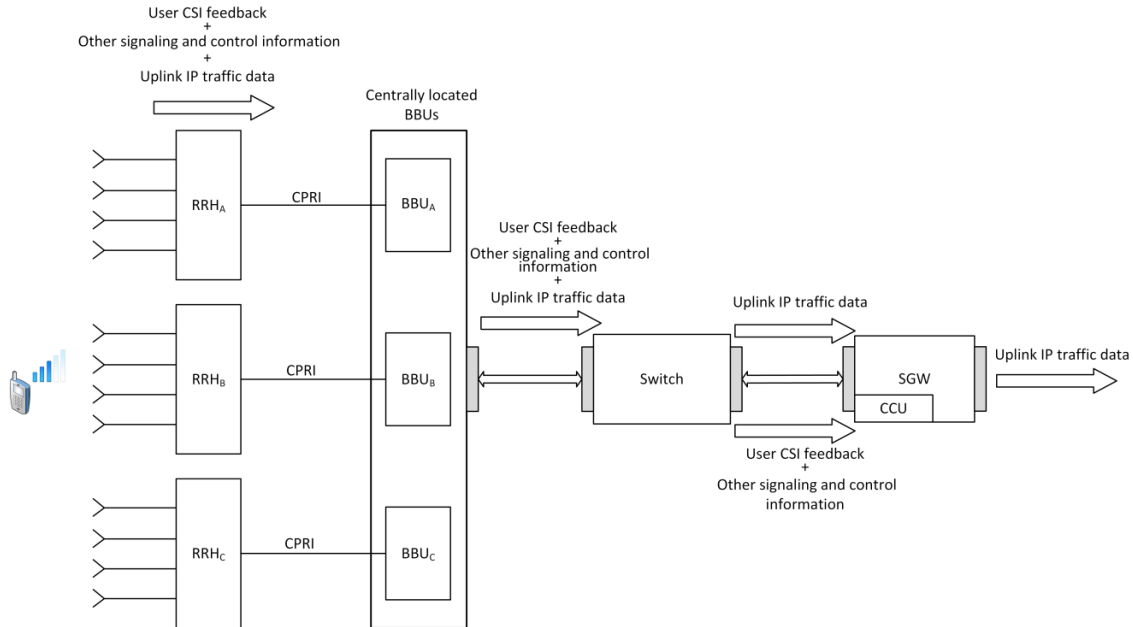
Figure 6.3 Uplink stream signalling and control information flow

Besides the user CSI feedback, another element needed as input to the coordination and scheduling algorithm is the user connection and hosting information, including users requesting connection, their hosting beam selection, and user ceasing connection. This information is controlled by higher signalling – Radio Resource Control (RRC) protocol [153, 154]. It is already part of LTE signalling and not extra in our gated narrow beam model. The user connection and hosting information is already available at the SGW and thus would be available to the CCU. It is beyond the scope of this study to quantify the amount of information transfer between these two entities.

The authors in [155] discussed in detail the size of parameters, CQI (channel quality indicator), PMI (precoding matrix indicator) and RI(rank indicator), in the CSI feedback for codebook-based implicit feedback scheme (which is the feedback scheme adopted in Chapter 5 signalling design). The size of the CQI feedback depends on the reporting type as well as reporting mode, such as wideband/sub-band, periodic/aperiodic. In general, wideband CQI feedback is configured periodically to provide basic and essential information about the downlink channel information. Narrowband feedback can be configured as needed when frequency selective scheduling is to be performed. Therefore, in this modelling, periodic wideband CQI feedback is assumed, which consists of 4 bits

per reporting interval [155]. The size of PMI feedback depends on the number of transmit antenna ports and the number of transmission layers. As in the design outlined in Chapter 4 and 5, at each sector, there are 8 antennas in total (4 antennas per layer performing beamforming to provide narrow beams pointing to 4 possible directions), and the transmission layer is restricted to no more than 2. Hence there are 8 entries in the codebook, the PMI feedback can be represented by 3 bits. RI with the value of 1 or 2, can be represented by 1 bit. Hence a complete (CQI, PMI, RI) feedback is 4+3+1=8 bits.

In addition to the CSI parameters, the feedback needs to be accompanied by the user ID in the form of Globally Unique Temporary Identifier (GUTI) which consists of up to 80 bits [156]. Hence the payload is a total of 88 bits (=11 bytes). The total traffic volume passing through the switch includes payload as well as the Ethernet header. The size of the Ethernet header is 38 bytes for an application payload of 11 bytes according to the study in [157]. Therefore, the total traffic passing through the switch per CSI feedback is 49 bytes.

For the (CQI, PMI, RI) feedback frequency, 5msec periodicity is assumed which represents the worst case provided for in the LTE standard [89]. The same periodicity was assumed in Chapter 5's signalling overhead calculation. For every 5msec, each user in the coordination cluster needs to generate and feedback in the upstream 3 CSI feedbacks, one per transmission point/sector, as designed in the per point feedback scheme in Chapter 5. The bit rate generated by one user can be calculated as follows,

$$the \ number \ of \ bits \ per \ report \times \frac{the \ number \ of \ report \ per \ user}{periodicity} = \frac{49 \times 8 \ bit \ \times 3}{5ms} \qquad (6.3)$$
$$= 235.2 \ kbit/s$$

Then, the total bit rate in one coordination cluster can be calculated by the following equation,

$$bit \ rate \ per \ user \ \times average \ number \ of \ users \ in \ the \ coordination \ cluster \qquad (6.4)$$

The average number of users in one coordination cluster (i.e. 3 sectors and hence equivalent to 1 base station) is obtained in Chapter 4 simulation and shown in Table 6.3 below,

Traffic load (%)	20	30	50	70	100	120	140
Average number of users	0.197	0.312	0.580	0.919	1.700	2.340	3.630

Table 6.3 Average number of users in one coordination cluster as a function of traffic load

By adopting Equation (6.4), the average bit rate associated with user CSI feedback in one coordination cluster can be obtained, as shown in Table 6.4 below,

Traffic load (%)	20	30	50	70	100	120	140
Average bit rate in kbit/s	46.33 (=235.3x0.197)	73.38	136.42	216.15	399.84	550.37	853.78

Table 6.4 Average bit rate associated with user CSI feedback in one coordination cluster

The additional power consumption associated with switching the additional information/traffic at the aggregation switch can be calculated using Equation (6.5), which expresses the traffic dependent component in Equation (6.2). Recall the values of the parameters, P_{idle} is 85.7W, P_{max} is 90.1W, Ag_{max} is 24Gbit/s. $Ag_{additional}$ represents the additional traffic which is the average bit rate in Table 6.4,

$$P_{additional} = (P_{max} - P_{idle}) \frac{Ag_{additional}}{Ag_{max}}$$ (6.5)

The additional power consumption of individual switch (which support 23 base stations) and the total consumption of the (4power) switches are shown in Table 6.5 below,

Traffic load (%)	20	30	50	70	100	120	140
Additional power consumption of 1 switch in mW	0.20 (=(90.1-85.7)x$\frac{46.33\times10^3\times23}{24\times10^9}x10^3$)	0.31	0.58	0.91	1.69	2.32	3.60
Additional power consumption of the (4) switches in mW	0.80 (=0.20x4)	1.24	2.32	3.64	6.76	9.28	14.40

Table 6.5 Additional power consumption at aggregation switch due to additional uplink traffic

As discussed, the additional upstream signalling traffic shown in Table 6.4 causes an increase in power consumption in the aggregation switches, and consequently an increase in backhaul power consumption. The additional power consumption of the 4 switches in Table 6.5 is the additional backhaul power consumption in upstream in gated narrow beam model.

6.3.3 Coordination Decision in the Downstream

This section studies the information flow and volume in the Downstream (or downlink) directions, focusing on the coordination decision. The coordination decision includes the beam pattern decision along with the users' resource allocation instructions and expected channel conditions under the chosen beam patterns. This information is generated at the CCU and sent back to the centrally located BBUs; then it is used to help generate baseband I/Q data for the scheduled users at the centrally located BBUs and sent to RRH. Figure 6.4 below shows the downstream information flow. For completeness, user data traffic in the downlink direction is also shown.
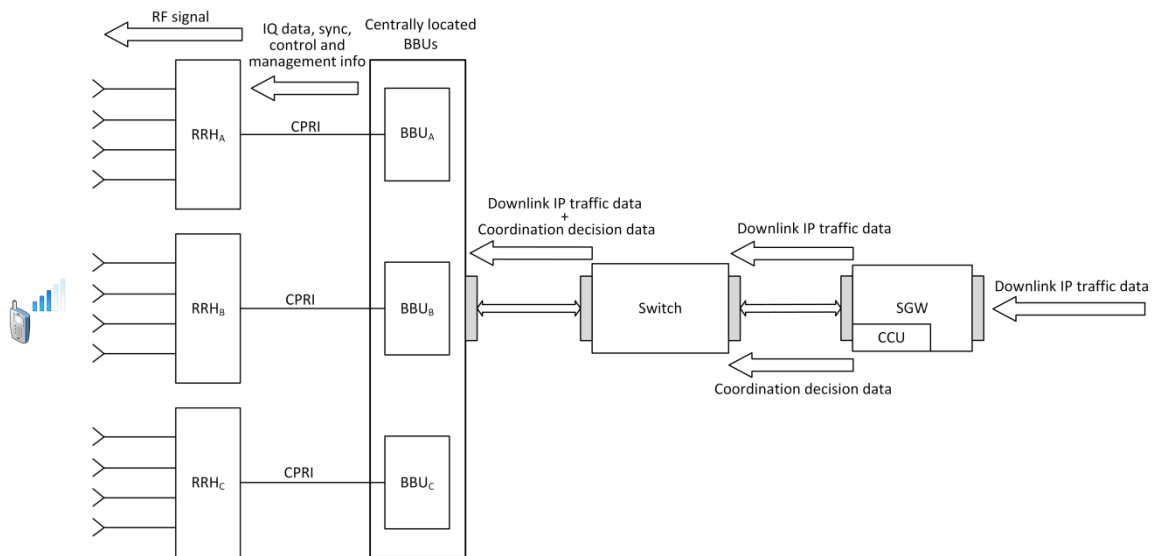
Figure 6.4 Downlink stream signalling and control information flow

The downlink IP traffic data is common in coordination and non-coordination scenarios, and has been accounted for in Section 6.3.1 when quantifying the baseline backhaul power. There are three additional information items in the coordination decision that are needed at the BBUs from the CCU:

1. The beamforming decision for each transmission point/sector
2. The estimated CQIs under the selected beam pattern for each user
3. The number of PRBs allocated for each user

In the BBUs, the additional information of coordination decisions will be used for signal processing. As discussed in Chapter 5, for each transmission point/sector, the beamforming decision is essentially the index of one precoding matrix that is to be used to form the desired narrow beam pattern at the RRH. The precoding matrix in the

beamforming decision will be applied to all scheduled users within the sector at precoding stage. The user's estimated CQI will be used to choose the modulation format.

In the following, the number of bits in one set of coordination decisions containing these 3 additional information items is estimated, and also the corresponding bit rate. The estimation is done for one coordination cluster, and for varying traffic load.

As discussed earlier, there are 8 entries in the codebook (8 precoding matrices). Therefore, for each sector, the beamforming decision can be selected from 9 options (8 precoding matrices selections plus turning off a beam). The beamforming decision of one sector can be represented by 3 bits (2^3= 16). The 3 coordination sectors together use 9 bits (= 3 sectors x 3 bits/sector) for a complete beamforming decision. The number of bits for beamforming decision in one coordination cluster is a constant 9 bits as shown in 1^{st}row in Table 6.6.

For each user, the estimated CQI is assumed to be 4 bits, same as the size of CQI in the CSI feedback. The maximum number of PRBs a user can be allocated at each epoch is 200, thus the number of PRBs allocated to a user can be represented by 8 bits. Therefore, for one user, the CQI and PRB information together is 12 bits. The CQI and PRB information volume for one coordination cluster depends on the number of users in the coordination cluster (the information is required for every user, if some users are not scheduled in one epoch, their CQI and PRBs are simply set to zero). The simulation-monitored average number of users as shown in Table 6.3 is used to obtain the total information volume as shown in 2nd row in Table 6.6.

Then, adding up the beam pattern information volume and the user related CQI and PRB information volume gives the total payload bits as shown in 3rd row in Table 6.6, further adding the Ethernet header of 38 bytes [157] gives the total traffic passing through the switch for one set of information as shown in 4th row in Table 6.6. Finally, these additional items of downlink information are generated and passed every epoch of 1msec. Hence the bit rate can be obtained by dividing the information bits by the periodicity, results shown in the last row in Table 6.6.

Traffic load (%)	20	30	50	70	100	120	140
Number of bits for beam pattern decision	9	9	9	9	9	9	9
Average number of bits for user CQI and PRB information	2.36 (=12x0.197)	3.74	6.96	11.03	20.40	28.08	43.56
Average total number of payload bits	11.36 (=9+2.36)	12.74	15.96	20.03	29.40	37.08	52.56
Average total number of bits including Ethernet header	315.36 (=11.36+38x8)	316.74	319.96	324.03	333.40	341.08	356.56
Average bit rate in kbit/s	315.36 ($=315.36/10^{-3}/10^3$)	316.74	319.96	324.03	333.40	341.08	356.56

Table 6.6 Average bit rate associated with coordination decision in one coordination cluster

These additional coordination decision information needs to be sent back to the centrally located BBUs through the aggregation switch. The additional information volume on any one switch is essentially 23 times the information volume per coordination cluster. Similarly to the previous calculation of additional switch power consumption for uplink, Equation (6.5) is adopted to obtain the additional power consumption of individual switch and then the total power consumption of the (4) switches can be derived. These calculations are shown in Table 6.7 below,

Traffic load (%)	20	30	50	70	100	120	140
Additional power consumption of 1 switch in mW	1.33 $(=(90.1\text{-}85.7)\text{x}\frac{315.36\,\text{x}10^3\text{x}23}{24\text{x}10^9}\text{x}10^3)$	1.33	1.33	1.33	1.33	1.33	1.33
Additional power consumption of the (4) switches in mW	5.32 (=1.33x4)	5.32	5.32	5.32	5.32	5.32	5.32

Table 6.7 Additional power consumption at aggregation switch due to additional downlink traffic

As discussed, the additional downstream signalling and control traffic shown in Table 6.6 causes an increase in power consumption in the aggregation switches, and consequently an increase in backhaul power consumption. The additional power consumption of the 4 switches in Table 6.7 is the additional backhaul power consumption in downstream in gated narrow beam model.

6.3.4 Additional Backhaul Power Consumption for Gated Narrow Beam Network

Section 6.3.2 and Section 6.3.3 estimated the additional power consumption of the switches due to additional upstream and downlink traffic to support the coordination in gated narrow beam model, respectively. As discussed, the switch is the only component that incurs additional power consumption as its power consumption scales with traffic passing through. Therefore, the additional backhaul power consumption in a gated narrow beam network for the service area consisting of 91 base stations can be obtained by adding up the last rows in each of Table 6.5 (upstream) and Table 6.7 (downstream) and is shown in Table 6.8 below.

Traffic load (%)	20	30	50	70	100	120	140
Additional backhaul power consumption in mW	6.12 (=0.80+5.32)	6.56	7.64	8.96	12.08	14.60	19.72

Table 6.8 Additional backhaul power consumption for GNB as a function of traffic load

6.4 Estimation of Computation Energy for Gated Narrow Beam Network

In this section, the energy consumption at the CCU is quantified. As discussed, the CCU is responsible for processing of extra signalling information and performing the coordination and resource scheduling. The CCU function is a process presumed to operate inside the SGW and run on the existing processor of SGW.

The amount of computation effort is quantified by counting the FLoating-point Operations Per Second (FLOPS) performed in the simulations reported in Section 4. For real floating point numbers, an addition, subtraction, or multiplication of them are counted as one floating point operation. There are difficulties obtaining an accurate count of flop for other advanced operations such as division, logarithm etc. It is assumed in this modelling that division and logarithm are 4 times and 50 times computationally expensive, consuming an average of 4 and 50 floating point operations respectively.

There are two major processes running on the CCU. One is the process of SINR estimation described in the signalling scheme in Chapter 5; the other is the main coordination and scheduling scheme described in Chapter 4. A complete algorithm for the processes is shown below to assist the counting of operations.

Algorithm: Gated Narrow Beam Coordination and Scheduling Algorithm

Part One: SINR estimation algorithm at the CCU

1: Each User Equipment (UE) estimates the channel from the all 3 transmission points within the coordination cluster, and also estimates the interference from sources outside of the cluster plus noise power as instructed by radio resource control signalling.

2: Based on the measurements, each UE feeds back the implicit Channel State Information (CSI) in uplink according to the 'Per Point Feedback' scheme. The implicit CSI feedback contains (CQI, PMI and RI) which are derived based on Equation (5.5).

3: The user CSI feedback from each user, along with other signalling and control information are directed to the Central Coordination Unit (CCU) via backhaul.

4: At the CCU, the user CSI feedback goes through a series of calculations steps.

 4.1: Each UE's 3 'Per Point CQI' is mapped back to (estimates of) the original SINR values.

 4.2: All the SINRs required to support coordination are calculated according to Equation (5.7) and Equation (5.9).

 4.3: These SINRs are further mapped to spectral efficiency values (same process as used in Chapter 4) to serve the coordination and scheduling algorithm.

Part Two: Coordination and scheduling algorithm at the CCU

 5.1: Update the list of users with newly added users and finished users according to the control information in radio resource control signalling.

 5.2: Select one potential beam combination (i.e. one turned-on narrow beam or no beam for each sector).

  5.2.1: Test whether the beam combination is feasible. The feasible beam combinations are those that have at least one user in any turned on beams.

  5.2.2: Perform scheduling within each sector in the coordination cluster.

- Obtain an estimate of the SINR for each user under the selected beam combination, and thus the spectral efficiency the user would achieve under the selected beam combination.

- For those users covered by active beams in the selected beam combination, determine the number of allocated PRB to each user according to the proportional fair scheduling algorithm according to Equation (4.4).

- For both the users under the active beams and turned off beams, determine the predicted weighted throughput for each user based on current spectral efficiency, the number of allocated PRB and historical throughput according to Equation (4.3).

5.2.3: For all users, the utility function is calculated based on the predicted weighted throughput according to Equation (4.1).

5.2.4: Compute sum of utility function of all users within the coordination cluster according to Equation (4.2).

5.3: Continue until all feasible beam combinations are scanned and tested. The potential beam combination that has the highest sum utility function is selected as the best beam combination.

5.4: Transmit with the best beam combination.

5.4.1: Each transmission point receives the PMI corresponding to the selected beam.

5.4.2: Each user receives the predicted CQI when the selected beam combination applies.

5.4.3: Update the user's historical throughput based on the throughput achieved under the best beam combination in the scheduling epoch.

6.4.1 Operations in SINR Estimation Algorithm

First, the calculation steps and the number of floating-point operations for the part one SINR estimation algorithm are estimated. Out of steps 1 to 4, steps 1 and 2 happen at the user equipment; hence only the calculation of step 4 happens at the CCU.

The calculation effort of step 4 is related to the number of users (hence traffic load) and the frequency of their CSI feedback. As a first step, the counting is focusing on obtaining one set of spectral efficiency estimations of one user in one reporting cycle.

Step 4.1: CQI to SINR conversion

The 3 'Per Point CQI' need to be converted back to SINR for the following calculations. CQI values range from 1 to 15. It is assumed that the CCU keeps a table of Modulation and Coding Scheme which is the same at user equipment, and then the corresponding SINR value of the CQI can be obtained by examining the table. This is a simple fetch step, and no flop is required. As there are 3 'Per Point CQI', there are three fetches involved in this step.

Step 4.2: SINR estimation

This step enables estimation of 25 SINRs based on the 3 SINRs obtained in Step 4.1.

Equation (5.7) in Chapter 5 is reproduced below.

$$CQI_{Al} = \frac{S_{Al}}{I+N} = \|(W_l)^H W_i\|^2 CQI_{Ai} \qquad \text{for l} = 1,4 \text{ and l} \neq \text{i} \qquad (5.7.a)$$

$$CQI_{Bm} = \frac{S_{Bm}}{I+N} = \|(W_m)^H W_j\|^2 CQI_{Bj} \qquad \text{for m} = 1,4 \text{ and m} \neq \text{j} \qquad (5.7.b)$$

$$CQI_{Cn} = \frac{S_{Cn}}{I+N} = \|(W_n)^H W_k\|^2 CQI_{Ck} \qquad \text{for n} = 1,4 \text{ and n} \neq \text{k} \qquad (5.7.c)$$

The same operation is executed for each of the non-hosting transmission points (2) and each precoding matrices that are not preferred in user feedback (3), which is a total of 6 times. Take Equation (5.7.a) as an example to illustrate the calculation process. W_i is an 8 × 1 or 8 × 2 matrix of complex numbers depending on whether the transmission rank is 1 or 2. W_l is 8 × 1 matrix of complex numbers.

The conjugate transposition of W_l denoted by $(W_l)^H$ is obtained by taking the transpose and then taking the complex conjugate (negating their imaginary parts but not their real parts) of each entry. This process involves array-element-copy operations and no floating point operation.

$(W_l)^H$ is a 1×8 matrix, and W_i is assumed to be 8 × 2 matrix to represent the worst case. Then the matrix multiplication of $(W_l)^H$ and W_i requires 16 complex number multiplication and 14 complex number addition.

The results of the multiplication of $(W_l)^H$ and W_i is 1×2 matrix of complex numbers. The Frobenious Norm operation ($\| \quad \|$) is the square root of the sum of the absolute squares of its elements. The outer square operation cancels out the square root, hence the operation ($\| \quad \|^2$) is simply the sum of the absolute squares of its elements. The absolute of one complex number can be calculated by the sum of the square of the real part and the square of the imaginary part, which consists of 2 real multiplications and 1 real addition. As $(W_l)^H W_i$ contains 2 elements, the operations are 2 times - 4 real multiplications and 2 real additions, plus 1 real additions to sum up the absolutes. Finally, $\|(W_l)^H W_i\|^2$ is multiplied with CQI_{Ai}, which is 1 real multiplication.

Equation (5.9) in Chapter 5 is reproduced as below.

$$SINR_{Ai,Bm+Cn} = \frac{CQI_{Ai}}{CQI_{Bm} + CQI_{Cn} + 1} \quad \text{for i} \in (1,4), \text{m} = 0,4, \text{n} = 0,4 \qquad (5.9)$$

It is executed 25 times to obtain the total 25 required SINRs. In Equation (5.9), there are two real additions and one real division. To add up, at the worst case, step 4.2 includes $6 \times 16 = 96$ complex number multiplications, $6 \times 14 = 84$ complex number additions, $6 \times (4+1) = 30$ real multiplications, $6 \times (2+1) + 25 \times 2 = 68$ real additions, and 25 real divisions.

Step 4.3: SINR to spectral efficiency mapping

This step maps the 25 SINRs to 25 spectral efficiencies which will be used in the scheduling algorithm. Consider the predefined mapping table used by GreenTouch, where SINR values (from -7.5dB to 39.5dB) are mapped to spectral efficiencies in bit/Hz/second. The granularity is every 1 dB gap, hence there are 48 entries in the table. A binary search is performed through the 48 entries. Let the SINR value first be compared against the middle element in the mapping table; if they are not equal, the lower or upper half of the table is eliminated depending on the results. Continue this half-interval search for 4 times, and there are only 3 entries left. The nearest SINR value can be obtained by comparing these 3 entries one by one. Thus, this process involves 4+3=7 comparisons at the worst case, plus one fetch. Hence a total of $25 \times 7 = 175$ comparisons and 25 fetches are required in this step.

To sum up, to obtain one set of spectral efficiency estimations of one user in one reporting cycle, the operations needed are summarized in the following Table 6.9,

Operations	Floating point operation required per operation	Number of operations	Total number of floating point operations
Fetch	N/A	28	N/A
Complex number multiplication	6	96	576
Complex number addition	2	84	168
Real multiplication	1	30	30
Real addition	1	68	68
Real division	4	25	100
Comparison	1	175	175
			Total: 1117

Table 6.9 Table of operation summary for one set of spectral efficiency estimations

Note that in complex number arithmetic, addition is counted as 2 flops (adding the real and imaginary part separately); multiplication is counted as 6 flops. Consider the following operation $(a + bi) \times (c + di)$. This result of this multiplication is real part = ac - bd, imaginary part = cb + ad, and the process consists of 4 real multiplications, 1 real subtraction, and 1 real addition.

From the table, it can be seen that there are a total of 1117 flops in obtaining one set of spectral efficiency estimations of one user in one reporting cycle. The number of estimation sets needed depends on the number of users and frequency of channel state information reporting. To be consistent with Chapter 5, a minimum of 5ms periodicity of CSI reporting is assumed (which corresponds to the maximum reporting rate).

The total FLOPS for the part one SINR estimation algorithm for the total of 91 base stations can be calculated by,

$$\frac{\begin{array}{c} number\ of\ floating\ point \\ operation\ per\ estimation \end{array} \times \begin{array}{c} average\ number\ of\ users \\ in\ coordition\ cluster \end{array} \times 91\ clusters}{periodicity} \quad (6.6)$$

The number of flops per estimation is 1117. The average number of users in one coordination cluster has been presented in Table 6.3. Hence the FLOPS for various traffic loads can be calculated as shown in Table 6.10 below,

Traffic load (%)	20	30	50	70	100	120	140
Computat-ion rate in FLOPS	4.00×10^6 $(= \frac{1117 \times 0.197 \times 91}{5 \times 10^{-3}})$	6.34×10^6	1.18×10^7	1.87×10^7	3.46×10^7	4.76×10^7	7.38×10^7

Table 6.10 Average FLOPS for SINR estimation algorithm as a function of traffic load

6.4.2 Operations in Coordination and Scheduling Algorithm

Next, the calculation steps and the number of floating-point operations for part two, the coordination and scheduling algorithm, are estimated. This is achieved by putting in various counters in the simulation and monitoring these counters. The Matlab simulation in Chapter 4 was designed to run as efficiently as possible, and that the number of computation steps it requires at each stage is likely to be representative of what the CCU software would involve. This simulation is run for various traffic loads and the simulation is only performing the coordination and scheduling algorithm for one coordination cluster.

For a particular traffic load of 50% of the GreenTouch dense urban daily average traffic, as shown in the following table, a breakdown of different operations (addition, subtraction, multiplication, division, logarithm etc.) and the value of their counters in major functions of the algorithm are presented. The beam combination function is responsible for selection of the best beam pattern; the resource scheduling function is mainly performing the proportional fair scheduling within each sector (which is essentially a sub-function of the beam combination function); program control deals with all necessary variables that control the running of the algorithm in the simulation such as loop functions etc.; the data management function is responsible for user metric tracking, beam comparison intermediate metric tracking etc. Each function's corresponding numbered algorithm steps are presented in the second left column in Table 6.11.

The counters are monitored for a large number of epochs, for this specific run, the time elapsed in the simulation is 581,269 ms (581,269 epochs). The computation rate in FLOPS is calculated by dividing the total number of floating point operations with the time elapsed, as shown in the rightmost column in Table 6.11. Then, the total computation rate in FLOPS is obtained by summing up for all operations in all functions as shown in the bottom right corner in Table 6.11.

Function	Main steps in algorithm	Operation type	Floating point operation required per operation	Operation Counter	Total number of floating point operations	Computation rate in FLOPS
Beam combination	5.2.1 5.2.3 5.2.4 5.3	Addition	1	1,092,022	1,092,022	1,879
		Comparison	1	247,131,756	247,131,756	425,159
		Logarithm	50	1,092,022	54,601,100	93,934
		Logical	N/A	184,128,966	0	0
		Multiplication	1	1,092,022	1,092,022	1879
Resource scheduling	5.2.2	Addition	1	1,494,558	1,494,558	2,571
		Division	4	1,494,558	5,978,232	10,285
		Multiplication	1	2,490,930	2,490,930	4,285
		Subtraction	1	1,494,558	1,494,558	2,571
Program control	Across Algorithm	Assign	N/A	1,166,516	0	0
		Compare	0.25	1,165,666	291,417	501
		Increase	0.25	62,468,630	15,617,158	26,867
Data management	5.1 5.4.3	Assign	N/A	30,933,813	0	0
		Decrease	0.25	23,820	5,955	10
		Increase	0.25	1,990	498	1
						Total: 569,942

Table 6.11 Operation breakdown for coordination and scheduling algorithm

The integer operations in program control are assumed to be 4 times more efficient then floating point operations, and hence are considered to consume 0.25 flops per operation. Some operations like logical operations and assign operations do not require floating point operations. However, the final FLOPS should be doubled to represent the effort of memory access such as read and write [158]. Hence, for traffic load of 50%, the final computation rate for supporting one coordination cluster is 1,139,884 FLOPS.

The computation rates for other traffic loads are estimated using the same method, and the final FLOPS for all traffic loads are shown in Table 6.12 below. The 2nd row shows the FLOPS for supporting one coordination cluster, and the 3rd row represents the FLOPS for the whole area of 91 base stations taking into account the doubling for memory access.

Traffic load (%)	20	30	50	70	100	120	140
Computation rate in FLOPS for one cluster	2.20×10^5	3.30×10^5	5.70×10^5	8.15×10^5	1.52×10^6	2.29×10^6	4.08×10^6
Computation rate in FLOPS for the area of 91 base stations	4.00×10^7	6.01×10^7	1.04×10^8	1.48×10^8	2.77×10^8	4.17×10^8	7.43×10^8

Table 6.12 Computation effort for performing the coordination and scheduling algorithm as a function of traffic load

6.4.3 Computation Power Consumption for Gated Narrow Beam Network

Finally, the computation rate in FLOPS needs to be translated to energy consumption, which requires the energy efficiency of the processor. In this modelling, the performance of power consumption of a modern server using general purpose processor is used to estimate the efficiency of a processor in the SGW. This is likely to be conservative as some specialised processors that might be used in the SGW can often be more efficient than the general purpose processors.

In this modelling, Dell Inc. PowerEdge R630 (Intel Xeon E5-2699 v3 2.30 GHz) is adopted as a representative server. The average active power (W) in relation to the target load (%) can be found in [159]. At 100% load, the power consumption is 287W. The floating point operation performance of various common algorithms varies between 70 Gflops-240 Gflops [160], and an average rate of 100 Gflops is assumed in this study. Hence at 100% load, the energy efficiency of the server is 0.35 GFLOPS per Watt (= 100 GFLOP/287W). Note that the power consumption of 287W includes both idle and incremental power consumption of the CPU, and the power consumption of machine mainboard and I/O. Hence, it is a conservative estimate of the additional energy consumption of a shared processor in the SGW.

The power consumption of CCU computation for supporting the gated narrow beam coordination can be obtained by dividing the sum computation rate in FLOPS by the server energy efficiency in GFLOP/Watt. The sum computation rate is obtained by adding up the FLOPS for SINR estimation algorithm in Table 6.10 and the coordination

and scheduling algorithm in Table 6.12. The calculations are shown in the Table 6.13 below,

Traffic load (%)	20	30	50	70	100	120	140
Sum computation rate in GFLOPS	$(= \dfrac{0.044}{\dfrac{4.00 \times 10^6 + 4.00 \times 10^7}{10^9}})$	0.066	0.116	0.167	0.312	0.465	0.817
Computation power consumption in mW	126 $(=0.044/0.35x10^3)$	189	330	477	891	1329	2334

Table 6.13 Computation power consumption for GNB as a function of traffic load

6.5 Overall Energy Efficiency Analysis of Gated Narrow Beam Model and Tradeoffs

Previous sections in this chapter have studied the mobile network architecture and the energy consuming components in the network. Those sections have also estimated the energy costs in backhauling and computation in core network based on the necessary functions and processes to support the coordination of gated narrow beams. Specifically, Section 6.3.1 estimated baseline backhaul power consumption which is common to coordinated and non-coordinated network. Section 6.3.4 provided the additional backhaul power consumption for GNB. Section 6.4.3 provided the computation power consumption for GNB.

This section summarizes this chapter and brings together the preliminary study in Chapter 4. Chapter 4 modelled the GNB network in a system level simulation and estimated the preliminary network performance metrics including cell throughput, power consumption and energy efficiency. The results in Chapter 4 were preliminary because the energy consumption modelling only considered the power consumption of the base stations. Besides, they assumed the same amount of available resource for user traffic as non-coordinated scenario. In other words, it did not account for the signalling overhead incurred due to coordination of gated narrow beams.

As discussed, to give a realistic estimate of the overall energy efficiency of a wireless network, a complete energy modelling is needed that factors in the energy cost across all components and processes across the network for delivering the bit rate. Therefore, the overall energy efficiency of GNB network needs to be obtained by taking into account the

backhaul and computation energy (which is studied in this chapter) and the signalling overhead according to the signalling scheme designed in Chapter 5.

6.5.1 Overall Energy Efficiency of Gated Narrow Beam network

Firstly, base station power consumption and all other power consumption for GNB model are summarized in Table 6.14. The average base station power consumption of GNB network for various traffic loads has been obtained in Chapter 4. The power consumption figures are reproduced in the 2nd row in Table 6.14. They are for one single base station. The backhaul (3rd, 4th rows) and computation energy (5th row) are from Table 6.2, Table 6.8 and Table 6.13 respectively. The power consumption figures in these tables are estimated for the service area of 91 base stations, and the figures are scaled to represent one single base station by diving 91.

Traffic load (%)	20	30	50	70	100	120	140
Average base station power consumption in W	219.274	235.696	270.158	307.667	370.356	415.957	464.176
Baseline backhaul power consumption in W	5.848 (=532.17/91)	5.850	5.854	5.858	5.864	5.868	5.872
Additional backhaul power consumption in mW	0.067 (=6.12/91)	0.072	0.084	0.098	0.133	0.160	0.217
Computation power consumption in mW	1.38 (=126/91)	2.08	3.63	5.24	9.79	14.60	25.65
Sum power consumption	225.123 (=219.274+5.848+0.0 67x10^{-3}+1.38x10^{-3})	241.548	276.016	313.530	376.230	421.840	470.074

Table 6.14 Complete power consumption of one base station in GNB network

The components in this table form a complete power consumption model of the mobile network and the sum of them provides the complete energy cost for supporting the various load levels. It can be seen that the additional energy power consumption incurred for GNB coordination are negligible compared with the base station power consumption. At maximum traffic load, the base station power consumption is around 464W, while the

computation power consumption is only 25.65 mW and the additional backhaul power consumption even smaller, 0.217 mW. It could be concluded that the additional backhaul and computation energy consumption to support coordination of GNB is marginal.

The baseline backhaul power consumption adds a few percentages to the previously calculated preliminary base station power consumption (about an additional 2.7% at 20% traffic load, and an additional 1.3% at 140% traffic load). Note that the baseline backhaul power consumption is relatively insensitive to the traffic load, as the power consumption of uplink and downlink interfaces does not scale with the traffic load. Whilst the switch power consumption is dependent on traffic load; however, the idle power consumes more than 95% of the total power, and the traffic dependent component consumes only less than 5%.

The following plots Figure 6.5-6.8 compare the performance metrics of the preliminary results given in Chapter 4, and those of the overall results obtained in this chapter that consider complete energy consumption and reduced throughput due to signalling overhead.

Figure 6.5 compares the power consumption, the preliminary power consumption considering only base station power consumption (2nd row in Table 6.14) is shown in the blue curve named 'Narrow Beam Chapter 4 results'. The complete power consumption by adding backhauling and computation power (6th row in Table 6.14) is shown in the red curve named 'Narrow Beam Overall results'.
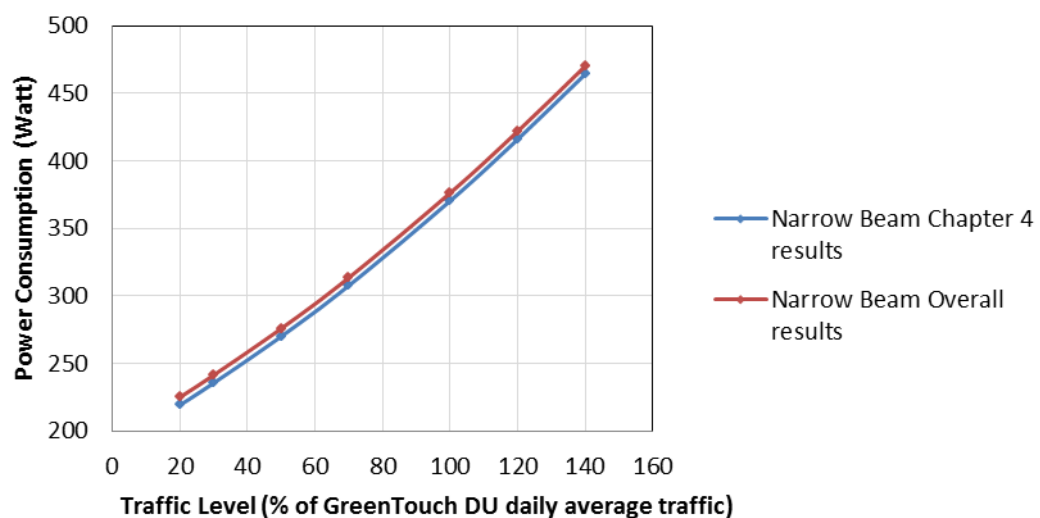

Figure 6.5 Power consumption comparison between Chapter 4 results and overall results

It can be seen that the gap between the two curves is almost constant. This is because the baseline power consumption is insensitive to traffic loads, and the other two additional power consumptions for GNB coordination are negligible compared with base station power consumption.

Figure 6.6 compares the cell throughput. To obtain the overall cell throughput, the cell throughput in Chapter 4's preliminary results needs to be corrected to reflect the signalling overhead due to coordination. Chapter 5 has estimated that the amount of signalling overhead incurred in the proposed signalling scheme is a fixed percentage of 3.34%. In other words, 3.34% of the overall resources in terms of (frequency-time or PRBs/REs) which could carry user data traffic are taken up by reference signals. Thus the overall cell throughput shown as the red curve is 100%-3.34%=96.66% of the preliminary cell throughput shown as the blue curve. Note that the 3.34% overhead estimate is based on the highest periodicity of the sending of reference signals to facilitate CQI measurements. Hence the overall cell throughput shown as the red curve is a conservative estimate.
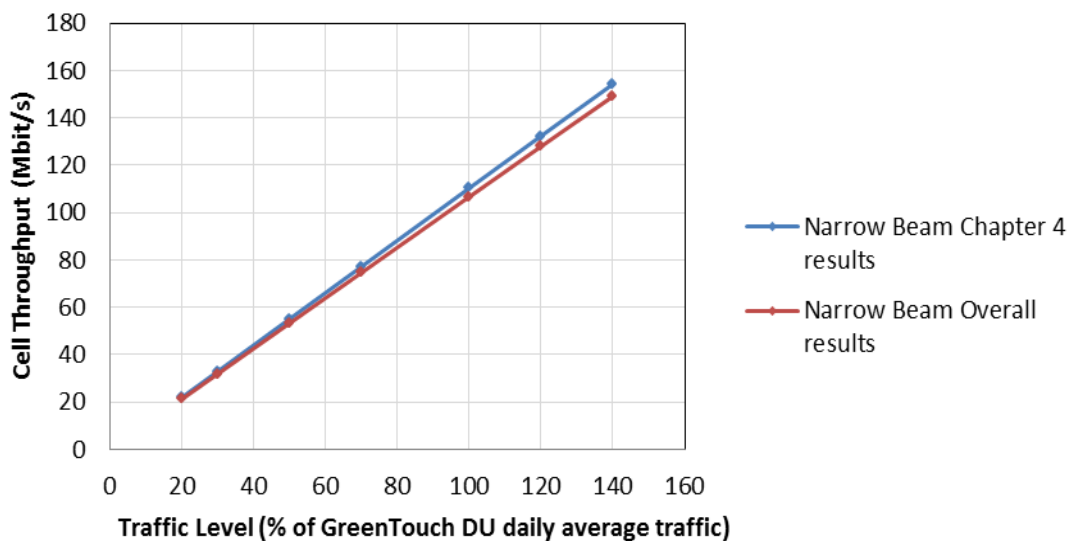
Figure 6.6 Cell throughput comparison between Chapter 4 results and overall results

Based on the complete power consumption in Joule/s in Figure 6.5 and the reduced cell throughput in Mbit/s in Figure 6.6, the overall energy efficiency in Joule/Mbit can be derived by dividing the complete power consumption by reduced cell throughput. The overall energy efficiency as well as the preliminary results from Chapter 4 are shown in Figure 6.7,
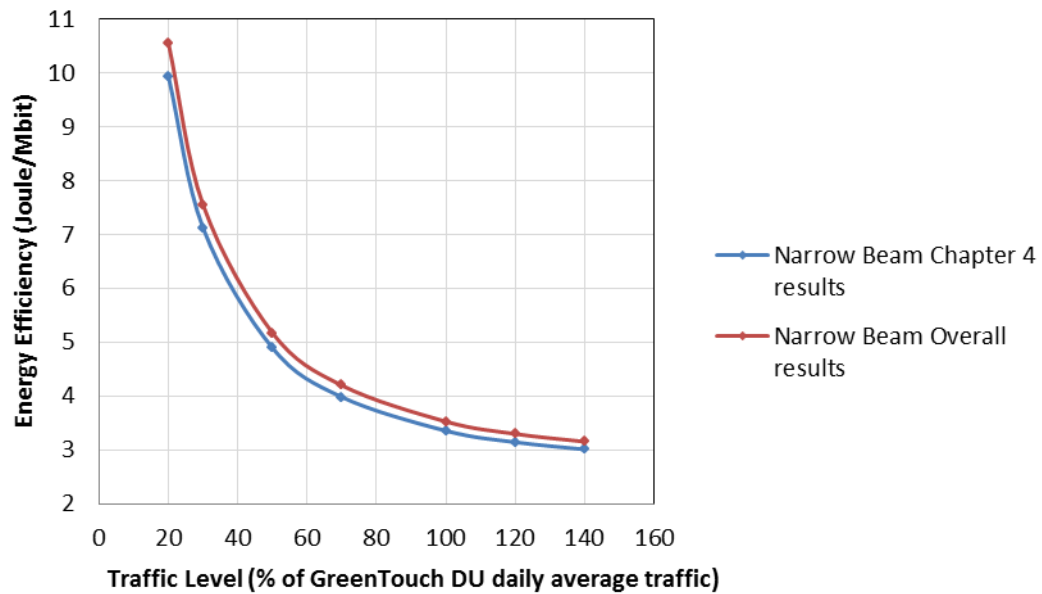
Figure 6.7 Energy efficiency comparison between Chapter 4 results and overall results

After taking into account the backhaul and computation energy as well as the reduced throughput due to signalling, the energy efficiency of GNB is deteriorated (shown in red). It now represents the true energy efficiency of the network rather than a preliminary result that Chapter 4 provided.

Comparing with Chapter 4 results (which is essentially an upper bound), the overall results is less energy efficient across the entire range of traffic load (20% to 140%). In specific, it is 4.7% less energy efficient at 140% traffic load (3.15 Joule/Mbit vs. 3.01 Joule/Mbit), and 6.1% less energy efficient at 20% traffic load (10.54 Joule/Mbit vs. 9.93 Joule/Mbit). The difference in percentage is less at the high end of traffic load, and the greater at the low end of traffic load. This is because the difference between the overall energy efficiency and the preliminary energy efficiency is largely due to the baseline backhaul power and the reduced throughput. It is largely unaffected by the additional backhaul power (the incremental backhaul power over the baseline backhaul power) and computation power. The reduced throughput is a fixed percentage thus will not cause this difference. The baseline backhaul power can be considered simply as a constant power overhead being added, thus as at low traffic load (fewer bits transmitted), it affects the energy efficiency more than at high traffic load.

It can be concluded that the overall energy efficiency of GNB network is affected (by 4% to 7% depending on the traffic load) by the inclusion of the energy consumption of baseline plus additional backhaul, CCU computation, and throughput reduction due to

signalling overhead. Most of the correction is due to the baseline backhaul energy and throughput reduction, the additional backhaul energy for carrying the additional traffic for coordination and the CCU computation energy is very marginal.

This difference in the two set of results confirms the importance of including signalling overhead and other energy costs in addition to base station power consumption when ascertaining the true energy efficiency of the GNB network. The corrected overall results have addressed the deficiencies (identified in the literature review) in prior works' energy efficiency modelling concerning CoMP systems, and provided a complete analysis of energy efficiency of a CoMP system. The preliminary results in Chapter 4 are obtained using the same simplified models as many previous works, where the costs of performing coordination are ignored (signalling overhead is neglected and only base station power consumption is accounted for in terms of energy consumption). Hence, the comparison between the preliminary results and the corrected overall results also indicates the amount of error that the incomplete modelling used by previous works could introduce in estimating the energy efficiency of CoMP systems. The relative importance of all the extra overhead and cost incurred by coordination for GNB system is also shown in the above analysis.

This thesis focuses on analysing the gated narrow beam network, which is one design of CoMP under the category of coordinated beamforming. For other CoMP designs, the exact channel state information feedback needed, signalling and control information exchange and processing, and thus costs of coordination would depend on the specific coordination and scheduling algorithm that a certain CoMP scheme chooses to adopt. As have shown, for GNB system, baseline backhaul energy and throughput reduction due to signalling overhead are two major factors for the energy efficiency deterioration. The baseline backhaul energy is dependent on the architecture of the network, and not dependent on the specifics of the CoMP scheme. On the contrary, the signalling overhead is entirely dependent on the coordination strategy and signalling scheme designed to support the coordination strategy. Thus the signalling overhead has the greatest impact on the energy efficiency and throughput of CoMP schemes, especially if the periodicity of sending reference signals to facilitate CQI/SINR measurements is high.

For CoMP schemes whose coordination and scheduling algorithm can be sufficiently driven by the 'per point CQIs' (the low-overhead channel state information feedback

148

strategy used in GNB signalling design) and/or CQIs derived from the 'per point CQIs', their signalling overhead can be obtained directly using the results presented in Chapter 5. For other CoMP schemes, the approach given in Chapter 5 can be used as a template for quantifying the impact of signalling on the system throughput in the context of LTE network once the signalling and CQI estimation technique are known. The percentage of energy efficiency deterioration of GNB system after accounting for all the overheads and costs is expected to be good approximation only to that of coordinated beamforming systems that uses 'per point feedback' for signalling and coordination algorithm of similar complexity.

Nevertheless, this work provided a template for accurately quantifying the energy efficiency of network integrating CoMP techniques. The quantitative analysis in this thesis is based on the proposed gated narrow beam model. The methodology developed for energy efficiency estimation accounting for the impact of signalling overheads on throughput and the added energy costs due to incremental increase in backhaul traffic and computational processing in the CCU, is applicable to other CoMP designs as well.

In this thesis, the energy efficiency of GNB network is estimated in the context of LTE network. However, the modellings and estimation methods can provide some insight for the fifth generation mobile network (5G) as well. Same to LTE networks, both access and backhaul network power consumption should be accounted for in estimating the total energy efficiency of 5G systems. In 5G, the radio access technologies will be deployed at much higher frequency bands to take advantage of the larger spectrum availability. The new technologies such as massive-MIMO and millimeter-wave can bring orders of magnitude increase in spectral efficiency [161]. In massive-MIMO systems, a large number of antennas (e.g. 256) are employed at each base station. Beamforming will be implemented with the large antenna array in order to mitigate the high propagation loss at high frequencies and establish links with satisfying signal-to-noise ratio [162].

The signalling overhead estimations performed in Chapter 5 can serve as an example for analysing signalling costs associated with 5G beamforming. The estimation methods developed are more applicable to systems designed with quantized CSI feedback and codebook-based beamforming precoding. The 5G standards [163-165] allow a high level of flexibility with CSI-RS configurations, in terms of periodicity and the number of ports a resource can be configured with. The exact amount of overhead associated with a

massive-MIMO beamforming strategy will need to be estimated based on the specific signalling information required and the configured CSI-RS resources.

In terms of access network architectures, a number of C-RAN architectures are being proposed for 5G with different level of centralization of baseband processing functions [162]. The C-RAN architecture is discussed in Chapter 3. The modelling in this thesis adopted the C-RAN architecture to facilitate the information exchange in GNB coordination scheme. The analysis of the signalling and control information flow over C-RAN architecture in a coordinated network can be useful for similar analysis for 5G networks. The energy savings that can be achieved from BBU pooling is not taken into account in this work. It is expected in the literature that a fully centralized C-RAN architecture can bring up to 38% of power savings compared with the distributed architecture (all functions remain at BS site) in 5G systems [162].

There are also a variety of backhauling solutions in 5G network to backhaul the traffic generated in access network to the core network [166]. The solutions include wired connection (e.g., fiber optic cable), wireless connection (e.g., high frequency millimeter wave) or a mixed architecture [166]. In this thesis, we constructed models for energy consumption estimation of a fiber optic backhaul network based on point-to-point Ethernet. These models could be used as a template for estimating 5G backhaul energy consumption if a similar backhaul technology is considered.

6.5.2 Energy Efficiency Comparison of Gated Narrow Beam Model and other CoMP strategies

In Chapter 4, we performed performance evaluation of GNB model in terms of network throughput under the 3GPP Case 1 reference scenario. The throughput performance of GNB model is compared against other CoMP strategies and schemes in the literature as well as 3GPP targets. In the following, we extend the throughput performance comparison to a preliminary energy efficiency comparison.

The intra-site coherent joint processing scheme under 8x2 antenna configuration by Sun et al. [129] is included for energy efficiency comparison with GNB model. It is the CoMP scheme that delivers the best cell spectral efficiency among the studies reviewed in Section 4.5. Their CS/CB scheme and JP schemes under 2x2 antenna configuration [126] are also included in the comparison.

As discussed in last section, the CoMP signalling overheads and additional energy costs due to coordination is scheme specific. Additional energy costs include additional backhaul energy for carrying the additional traffic for coordination and the CCU computation energy. An accurate estimation of these overheads and costs for other CoMP schemes can be obtained by performing detailed analysis following the approaches presented in this thesis.

In this section, the preliminary energy efficiency estimation will only account for base station power consumption and baseline backhaul power consumption. As discussed, the baseline backhaul power consumption is dependent on the architecture of the network, but not dependent on the specifics of the CoMP scheme. It is also found to be one of the two major factors for energy efficiency deterioration when comparing preliminary and overall energy efficiency results of GNB model. In this estimation, the majority of the power consuming components of the network are included while the scheme-specific and relatively small power consumptions are not. The throughput reduction due to CoMP signalling overhead is not taken into account either.

The backhaul power consumption of the CoMP schemes under the 3GPP Case 1 reference scenario is estimated using the same approach as was used earlier in this chapter. The backhaul estimations are based on the network architecture presented in Section 6.2 and a service area of 91 base stations is considered. The same backhaul power consumption model and aggregation switch parameters as discussed in Section 6.3.1 are used. Table 6.15 of switch power consumption and Table 6.16 of baseline backhaul power consumption below are in equivalent forms of Table 6.1 and Table 6.2 respectively.

	GNB - CS/CB	Sun intra-site JP (8x2)	Sun inter-site CS/CB (2x2)	Sun inter-site JP (2x2)
Cell spectral efficiency in bit/s/Hz	4.9912	6.2004	2.1475	2.3192
Traffic per BS in Mbit/s	299.47 $(=4.9912 \times 3 \times 20 \times 10^6/10^6)$	372.02	128.85	139.15
Traffic in each Switch in Gbit/s (Ag_{switch})	6.888 $(=299.47 \times 23/1000)$	8.556	2.964	9.601
Power consumption of 1 switch in W (P_S)	86.96 $(=85.7+4.4 \times 6.888/24)$	87.27	86.24	87.46
Total power consumption of (4) switches in W	347.84 $(=86.96 \times 4)$	349.08	344.96	349.84

Table 6.15 Switch power consumption of CoMP schemes under 3GPP Case 1 scenario

	GNB - CS/CB	Sun intra-site JP (8x2)	Sun inter-site CS/CB (2x2)	Sun inter-site JP (2x2)
Total power consumption of (4) switches in W	347.84	349.08	344.96	349.84
Total power consumption of downlink Interfaces in W	182 $(=91 \times 2)$	182	182	182
Total power consumption of uplink interfaces in W	10.5 $(= \lceil 6.888 \times 4/10 \rceil \times 3.5)$	14	7	14
Baseline backhaul power consumption in W	540.34 $(=347.84+182+10.5)$	545.08	533.96	545.84

Table 6.16 Baseline backhaul power consumption of CoMP schemes under 3GPP Case 1 scenario

The IP/user traffic volume in one base station (shown in the 2ndst row in Table 6.15) is obtained from the cell spectral efficiency figures in Section 4.5 (reproduced in the 1st row). Note that the cell spectral efficiency in 3GPP terms corresponds to one sector of the 3-sector macro base station, thus it is multiplied by 3. System bandwidth of 20MHz is assumed for all schemes.

The traffic per base station is multiplied by 23 to obtain the traffic passing through each switch as shown in the 3rd row, and the power consumption of one switch as shown in the

4th row is calculated based on Equation (6.2). For the inter-site joint processing scheme by Sun et al's, the traffic passing through switch is multiplied by an additional factor of 3 as each base station needs to have user traffic data for all 3 coordination clusters that it is involved in. The total power consumption of the switches is given in the last row in Table 6.15.

The backhaul architecture has been configured with one downlink interface per BBU that provides enough capacity. Thus the number of downlink interfaces is the same for all schemes. The number of uplink interface is dimensioned to support the aggregated traffic of the 4 switches. Thus the number of uplink interface varies for different scheme. The baseline backhaul power consumption in Table 6.16 is obtained by adding up the three power consuming components in backhaul, including power consumption of switches (1st row in Table 6.16), power consumption of downlink interfaces (2nd row in Table 6.16) and power consumption of uplink interfaces (3rd row in Table 6.16).

Table 6.17 below presents the preliminary energy efficiency estimates. The base station power consumption figures in Table 4.2 are reproduced in the 1st rows in Table 6.17. In Chapter 4 simulation, the GNB base station (8x2 MIMO) and wide beam base station (2X2 MIMO) consume 665 Watts and 638 Watts respectively when transmitting at full load mode. The same base station power consumptions are assumed for base stations in Sun et al's schemes (665 Watts for schemes under 8x2 antenna configuration and 638 Watts for schemes under 2x2 antenna configuration). As full buffer traffic model is applied in the 3GPP Case 1 reference scenario, the base stations operates at full load mode all the time.

The backhaul power consumption figures in Table 6.16 are estimated for the service area of 91 base stations. These figures are scales to represent one single base station by diving 91 (shown in the 2nd row in Table 6.17). The sum power consumption in the 3rd row is obtained by adding the base station power consumption and baseline backhaul power consumption. The preliminary energy efficiency in Joule/Mbit in the last row is calculated by dividing the sum power consumption by the base station throughput (2nd row in Table 6.15).

	GNB - CS/CB	Sun intra-site JP (8x2)	Sun inter-site CS/CB (2x2)	Sun inter-site JP (2x2)
Traffic per BS in Mbit/s	299.47	372.02	128.85	139.15
Base station power consumption in W	665	665	638	638
Baseline backhaul power consumption in W	5.938 (=540.34/91)	5.99	5.868	5.998
Sum power consumption in W	670.938 (=665+5.938)	670.99	643.868	643.998
Energy efficiency in Joules/Mbit	2.24 (=670.938/299.47)	1.8	5	4.63

Table 6.17 Preliminary energy efficiency estimation of CoMP schemes under 3GPP Case 1 scenario

It can be seen from Table 6.17 that the baseline backhaul power consumption is almost the same for all schemes and it adds less than 1% to the base station power consumption. As discussed, the baseline backhaul power consumption is relatively insensitive to the traffic load. Whilst the switch power consumption is dependent on traffic load, the idle power consumes more than 95% of the total power, and the traffic dependent component consumes only less than 5%.

Under the 3GPP Case 1 full buffer traffic, the percentage of contribution of baseline backhaul power consumption to the total power consumption is lower than when dense urban FTP traffic is applied. For GNB model under dense urban FTP traffic, baseline backhaul power consumption adds around 1.3% to 2.7% to the base station power consumption over different traffic levels. This is because the base stations are 100% utilized under full buffer traffic while the utilization rate is only 53% at busy hour FTP traffic. The base station power consumption under full buffer traffic is much higher than the average base station power consumption under dense urban FTP traffic.

For energy efficiency comparison, Table 6.17 shows that the GNB model is 24.4% less energy efficient than the coherent joint processing scheme with 8x2 antenna configuration (2.24 Joule/Mbit vs. 1.8 Joule/Mbit). The two schemes have the same base station power consumption and about the same sum power consumption. The difference in energy efficiency is caused by the difference in throughput. The base station throughput of GNB

model is around 19.5% less than that of the coherent joint processing scheme with 8x2 antenna configuration (299.47 Mbit/s vs. 372.02 Mbit/s).

The energy efficiency estimates of the two schemes with 2x2 antenna configuration are about twice as high compared to the GNB model. This is because their base station power consumption is only 27 W less while the throughput is more than halved. Comparing the CS/CB scheme and JP scheme with 2x2 antenna configuration, JP scheme has slightly higher sum power consumption due to slightly higher backhaul power consumption; however it achieves a better energy efficiency as its throughput is higher. Note that the throughput reduction due to CoMP signalling overhead and additional computational energy for coordination are not accounted for in above estimations. These preliminary energy efficiency estimates of the CoMP schemes will deteriorate once these overheads and costs are included. The joint processing schemes are likely to involve more signalling overhead and additional energy costs than CS/CB schemes [126].

6.5.3 Comparison of Gated Narrow Beam Model and Wide Beam Model

Chapter 4 has compared the base station power consumption, cell throughput and energy efficiency of GNB model and wide beam model. However, as discussed, all results were preliminary since the energy modelling was incomplete and signalling overhead was not accounted for. Under this condition, the comparison in Chapter 4 gave an unfair advantage to the GNB model. This is because GNB network would consume slightly more backhaul energy when compared with a wide beam network for exchanging additional information to facilitate coordination, as has shown in the chapter. In addition, a GNB network also needs a CCU to process extra signalling information and execute the coordination and scheduling algorithm, which does not exist for the wide beam network. The resource (or PRB) available will also be reduced for a GNB network once the signalling overhead is taken into account, but not for wide beam network. The energy costs associated with these additional functions and processes for supporting the coordination must be included to fully model a network with the coordination scheme.

The preliminary results in Chapter 4 have shown that GNB network is significantly more energy efficient than a network of conventional wide beam network over a large range of traffic conditions, in particular at higher loads. However, the question remains whether this is still the case when the energy modelling is complete and all energy costs have been

accounted for. Now that a complete modelling of the network is possible based on the study of Chapter 5 and Chapter 6, the comparison of the GNB network and wide beam network is revisited here.

The GNB results have been updated in Section 6.5.1. The equivalent overall results including power consumption, cell throughput and energy efficiency for wide beam network are presented in Figure 6.8, Figure 6.9, and Figure 6.10, respectively. Note that the additional power added to the Chapter 4 preliminary results for the wide beam model is only that of the baseline backhaul power consumption. The cell throughput remains unchanged for wide beam model. Its updated energy efficiency is derived the same way as in the last section for GNB model, by including the baseline backhaul power consumption (as per 3rd row in Table 6.14).

Figure 6.10 shows the behaviour of the wide beam energy efficiency metric, which is largely unchanged when compared to the preliminary results shown in Chapter 4. The 'wobble' (the increases in Joules/Mbit to about 120% traffic load then followed by a decrease) still exists for the same reasons discussed in Chapter 4. The updated energy efficiency has only deteriorated slightly due to the almost fixed baseline backhaul power consumption added to all traffic levels.
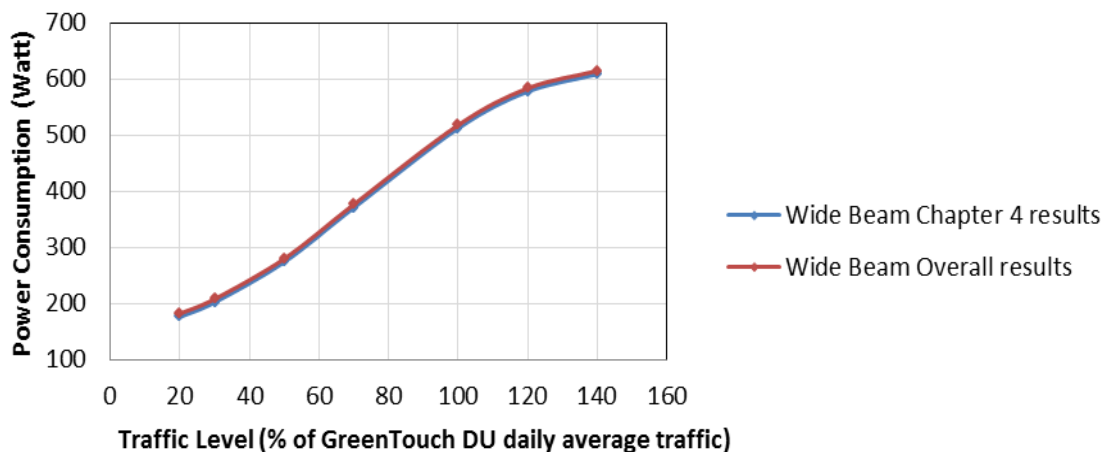

Figure 6.8 Power consumption comparison between Chapter 4 results and overall results
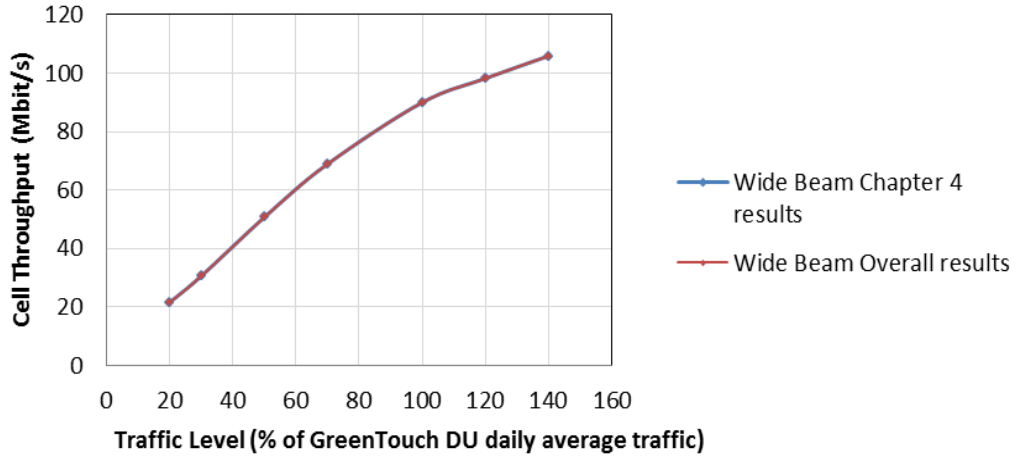
Figure 6.9 Cell throughput comparison between Chapter 4 results and overall results (Note, both the throughputs are the same and hence the graphs are overlapping)
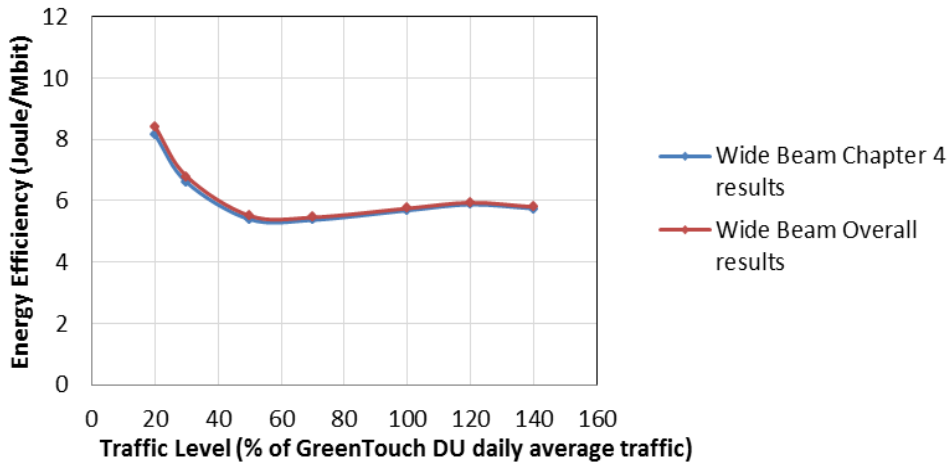
Figure 6.10 Energy efficiency comparison between Chapter 4 results and overall results

Next, the performance metrics of narrow beam model and wide beam model are plotted alongside each other for comparison. Recall that in the system level simulation in Chapter 4, the coordinated GNB network is configured with base stations with 8 antennas at each sector using 8x2 MIMO. The wide beam network is configured with base stations with 2 antennas at each sector using 2x2 MIMO. The 'Narrow Beam Overall Results' and 'Wide Beam Overall Results' refer to the performance of GNB network and wide beam network, respectively. The comparisons of power consumption, cell throughput and energy efficiency are presented in Figure 6.11, Figure 6.12, and Figure 6.13, respectively.
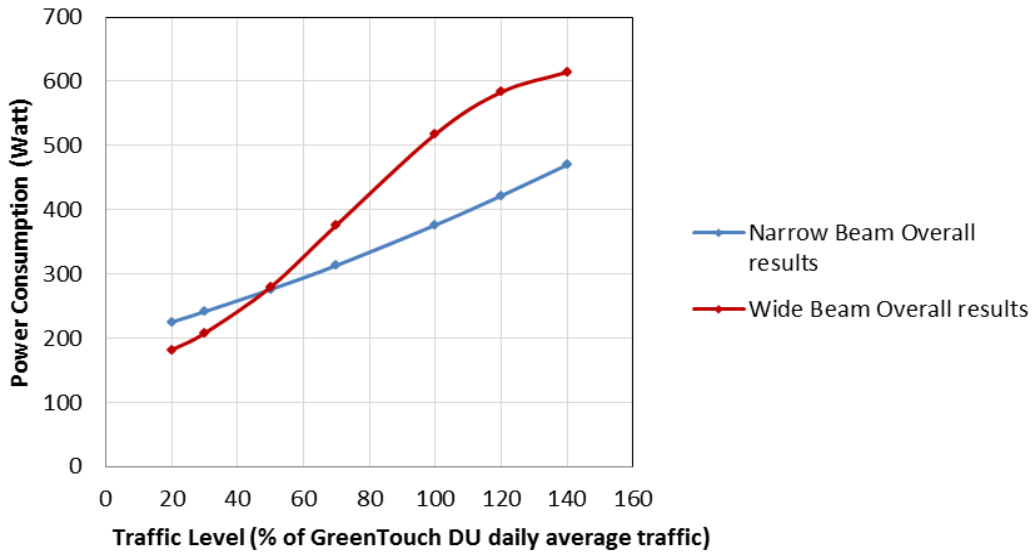
Figure 6.11 Power consumption comparisons between wide beam and narrow beam

Comparing Figure 6.11 with the preliminary comparison in Chapter 4 given Figure 4.8 (f), the power consumption difference between the two models is marginally less in high traffic levels (on the right side of the crossover point). This is because narrow beam has included more energy costs (baseline plus additional backhaul power and CCU computation power) while wide beam only adds the baseline backhaul power.

For Figure 6.12 cell throughput comparison, the difference between the two models is reduced than preliminary comparison in Chapter 4. Because the cell throughput of narrow beam is reduced while wide beam remains no change. The cell throughput advantage of narrow beam has only slightly reduced.

Figure 6.12 Cell throughput comparisons between wide beam and narrow beam

Figure 6.13 presents the overall energy efficiency comparison between a coordinated network of gated narrow beams and a non-coordinated network of conventional wide beams. The GreenTouch 2020 prediction figure is again added for comparison. As mentioned in Chapter 4, this energy efficiency prediction includes the benefit of offloading 2/3 of the traffic to highly efficient small cells. The energy efficiency improvement of GNB model is achieved without deployment of any new cells and the complexity of managing them.
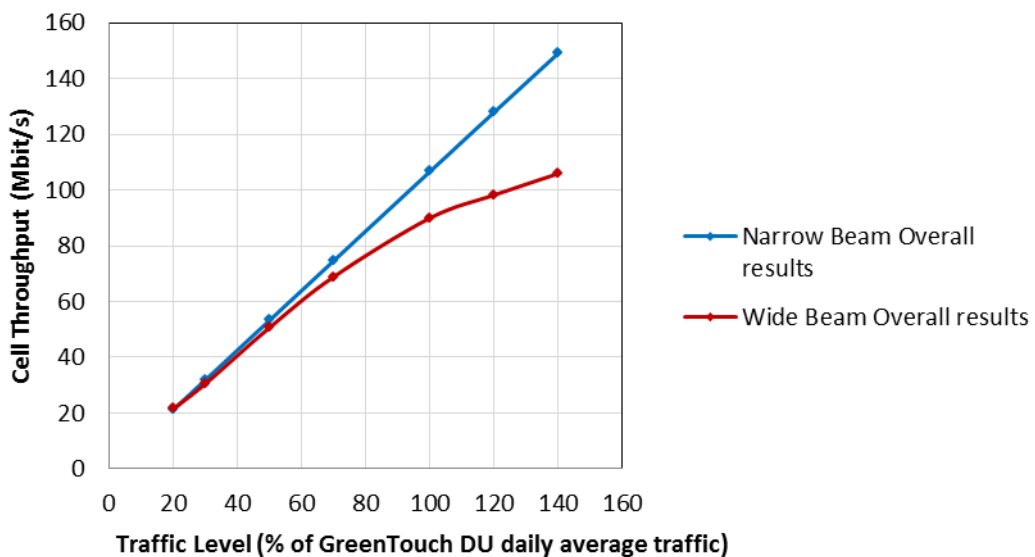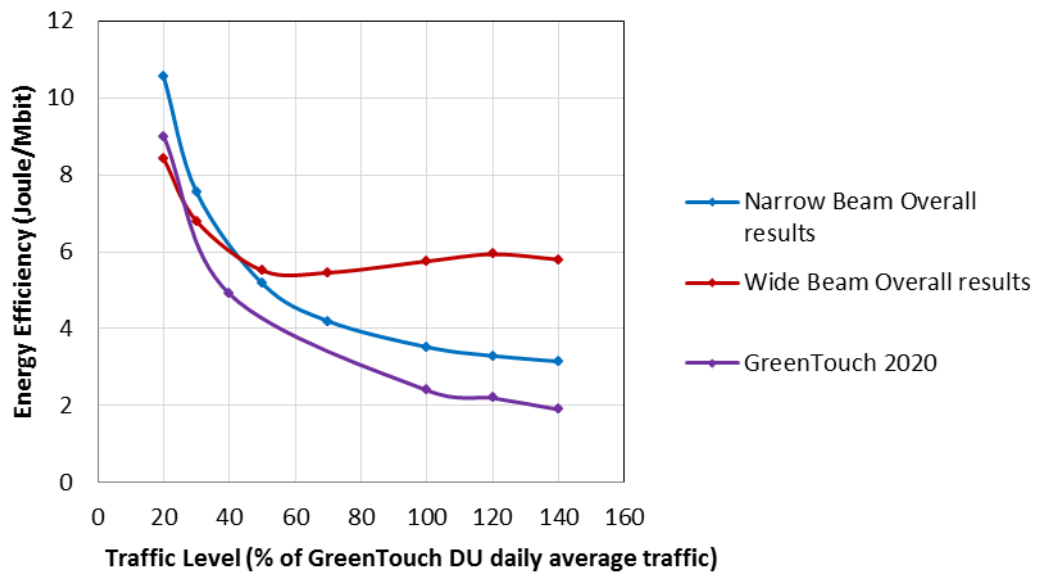

Figure 6.13 Energy efficiency comparisons between wide beam and narrow beam

In general, the main trend is very much similar to the preliminary comparison in Chapter 4. A wide beam network is more energy efficient at low traffic load, because base station idle power dominates the energy consumption and a wide beam base station using 2x2 MIMO consumes less idle power (132W) than narrow beam base station using 8x2 MIMO (189W). This is the case when only base station power consumption is considered, and is still true when backhaul power and computational power are included.

The maximum power consumption of narrow beam base station and wide beam base station is comparable, 665W vs. 638W. However, when base stations are both transmitting at maximum power, a base station employing narrow beam is able to provide much more capacity (given better SINR and spectral efficiency), thus much more energy efficient. Therefore, although wide beam model is more energy efficient at very low traffic, the advantage of narrow beam model being more energy efficient in full load

mode, starts to take effect and cause the crossover. The GNB network is more energy efficient in a large range of higher traffic levels.

In practice, the traffic demand fluctuates during each time of the day. Thus the average energy efficiency during a complete diurnal cycle is an important consideration when determining which deployment (coordinated gated narrow beam vs. wide beam) to implement. To estimate this, the daily traffic profile of the area needs to be known in addition to the energy efficiency as a function of traffic levels. The GreenTouch 2020 prediction reported that, for the dense urban environment, the occurrences of discretised traffic load levels 20%, 40%, 100%, 120%, 140% are 2, 4, 4, 8, 6 hours respectively per day [117]. The relative traffic volume can be obtained by multiplying the traffic load levels in percentage with the duration of these traffic load levels. Then the weighted average energy efficiency for the whole diurnal cycle can be calculated by combining the energy efficiency in Joule/Mbit for various traffic levels and weighting factors corresponding to the relative traffic volume,

For coordinated GNB network,

$$\frac{10.54 \times 20\% \times 2 + 6.15 \times 40\% \times 4 + 3.53 \times 100\% \times 4 + 3.30 \times 120\% \times 8 + 3.15 \times 140\% \times 6}{20\% \times 2 + 40\% \times 4 + 100\% \times 4 + 120\% \times 8 + 140\% \times 6}$$

$$= 3.60 \text{ Joule per Mbit} \tag{6.7}$$

For wide beam network,

$$\frac{8.41 \times 20\% \times 2 + 6.01 \times 40\% \times 4 + 5.75 \times 100\% \times 4 + 5.94 \times 120\% \times 8 + 5.80 \times 140\% \times 6}{20\% \times 2 + 40\% \times 4 + 100\% \times 4 + 120\% \times 8 + 140\% \times 6}$$

$$= 5.91 \text{ Joule per Mbit} \tag{6.8}$$

For GreenTouch dense urban scenario for 2020, the daily average energy efficiency of coordinated GNB network is 3.60 Joule/Mbit. It is 39.1% more energy efficient compared with that of wide beam network (5.91 Joule/Mbit). More improvement can be achieved when the traffic demand continues to increase beyond 2020.

For both narrow beam network and wide beam network, the absolute values of the overall energy efficiency measured in Joules/Mbit have increased when compared with Chapter 4 results (which means less energy efficient). At 140% traffic load, the overall energy efficiency of narrow beam and wide beam are 3.15 Joules/Mbit vs. 5.80 Joules/Mbit (the different is 2.65 Joules/Mbit). In Chapter 4 results, they were 3.01 Joules/Mbit vs. 5.74

Joules/Mbit (the difference is 2.73 Joules/Mbit). At 20% traffic load, the overall energy efficiency of narrow beam and wide beam are 10.54 Joules/Mbit vs. 8.41 Joules/Mbit (the different is 2.13 Joules/Mbit). In Chapter 4 results, they were 9.93 Joules/Mbit vs. 8.14 Joules/Mbit (the difference is 1.79 Joules/Mbit).

The energy efficiency advantage of narrow beam network over wide beam network slightly shrinks. This is because the cell throughput is reduced for narrow beam network but not for wide beam network, and narrow beam network has included more energy costs comparing with wide beam network. This also leads to a small right shift of the crossover point. In Chapter 4 comparison, the crossover is around 40% traffic load. While in Figure 6.13, the crossover point is around 45% traffic load. This means that due to the slighter greater energy costs and, more importantly, the reduction in cell throughput, it takes more traffic before the capacity advantage of narrow beam can negate its higher idle energy consumption.

Nevertheless, the comparison demonstrates that coordinated GNB network is still significantly more energy efficient than wide beam network over a large range of traffic levels, especially higher traffic levels. In particular, it is 38.6% more energy efficient (3.53 Joules/Mbit vs. 5.75 Joules/Mbit) than wide beam at 100% traffic load. It is 45.7% more energy efficient (3.15 Joules/Mbit vs. 5.80 Joules/Mbit) than wide beam at 140% traffic load. Beyond 140%, the energy efficiency of narrow beam will continue to go down. The overall added energy costs associated with the coordination can only slightly offset the throughput gain and energy saving of coordinated GNB network identified in Chapter 4. The entire Dense Urban busy hour load could be carried by the coordinated GNB network with high energy efficiency and negligible failure cases, without the benefit of small cells included in GreenTouch 2020 prediction.

6.6 Conclusion

This chapter presented the mobile network architecture and identified the key information flow in the upstream and downstream in each section of the coordinated gated narrow beam network. Then it quantified the baseline backhaul energy common to coordinated and non-coordinated network, and additional backhaul energy and central coordination unit computation energy for the GNB network according to the proposed coordination and scheduling scheme in Chapter 4 and signalling scheme in Chapter 5. A complete

energy modelling of the GNB network is achieved by adding these energy costs to the simplified energy model in Chapter 4 which only considers base station power consumption. The overall energy efficiency is calculated taking account of the complete energy modelling and the reduced throughput due to signalling overhead. Besides, we also extend the throughput comparison with other CoMP strategies in Chapter 4 to a preliminary energy efficiency comparison. Finally, the overall energy efficiency of GNB is compared against the overall energy efficiency of wide beam. The comparison demonstrated that the additional energy costs associated with GNB coordination is marginal, and the throughput gain brought by narrow beam coordination leads to significant energy efficiency gain.

Chapter 7

Conclusion

This chapter summarizes the key contribution of this thesis and also discuss some future directions for research in the area of energy efficiency of wireless network.

7.1 Summary of Contribution

THE focus of this thesis is on estimating the energy efficiency of a coordinated mobile network of base stations using a gated narrow beam (GNB) scheme. This is a class of Coordinated Multipoint (CoMP) transmission where base stations are coordinated with each other for providing services to user equipment. The thesis develops a methodology on how to ascertain the network throughput and overall energy consumption of such coordinated networks. The study focuses on the development of a scalable CoMP beamforming architecture based on gated narrow beams and a methodology for quantifying the energy efficiency of CoMP systems, using GNB as a specific example of its application. It also provides insights into the trade-off between the energy savings and energy costs associated with the coordination. The results indicated significant energy savings can be achieved from a CoMP coordinated beamforming network compared to non-coordinated wide beam network.

In Chapter 2 we surveyed technologies to improve the energy efficiency of wireless network. The coordinated beamforming network, where beams among different base stations are coordinated to mitigate inter-cell interference, is a promising technology in LTE for increasing spectral efficiency and system capacity. However, we found that the energy aspect of this technology has not been evaluated thoroughly. In Chapter 3, we

163

provided a high level description of the research problems and solution strategies in developing a more complete and accurate modelling of energy efficiency of a coordinated beamforming network. Our modelling overcomes many of the shortcomings of approaches to estimating energy efficiency of wireless network reviewed in the literature (such as incomplete energy modelling, lack of realistic user traffic simulation and consideration for signalling overhead in throughput estimation). This modelling takes into account fairness of resource allocation, the signalling capabilities of LTE, signalling overhead, the extra energy for processing coordination and signalling information and for backhauling. It also involves simulation of a realistic LTE network and a practical traffic model with varying loads is used to simulate user traffic.

The contribution of this thesis can be summarised into three studies (a) "Developing a realistic coordinated beamforming strategy", (b) "Developing an LTE compliant signalling scheme", (c) "Developing energy modelling for mobile network architecture". The three research studies are connected and solutions to them are developed separately for tractability reasons as well as clarity, to illustrate the steps for developing a methodology for quantifying the energy efficiency of CoMP systems. The results are then combined to estimate the overall system energy efficiency.

The contribution of each study towards determining the energy efficiency of a coordinated beamforming base station network is outlined as follows. The contribution in the first study, "Developing a realistic coordinated beamforming strategy" is the estimation of the throughput and energy consumption of a network of base stations employing gated narrow beams. The associated signalling overhead is estimated from a proposed signalling scheme that provides necessary channel state information to support the coordination. The signalling overhead referred to the signalling resource usage in terms of PRBs, and it leads to a reduction of the throughput gain identified in the first subject. The signalling studies are covered in the second subject, "Developing an LTE compliant signalling scheme". A complete energy modelling of the network is developed to quantify the energy consumption of major components and processes which have not been captured in the first subject. These additional energy consumptions are added on to the base station energy consumption for an overall energy estimate. The third subject, "Developing energy modelling for mobile network architecture", combines these results and provides the final energy efficiency estimate of the coordinated beamforming network.

In the following sections, we outline the key contributions of this thesis in each of the three subjects described above.

7.1.1 Developing a realistic coordinated beamforming strategy

In Chapter 4, we developed a distributed and scalable coordinated beamforming architecture with multiple gated narrow beams for transmission of data. We also developed a beam coordination and resource scheduling scheme which can mitigate interference between sectors within the same coordination cluster, and ensure efficiency and fairness in resource allocation among users. The key feature of this scheme is managing interference between the coordinated narrow beams by selecting the best narrow beams to be turned "On" in a coordination cluster at any time. We developed a system level simulator of an LTE network of base stations with gated narrow beams. The simulation setup involves realistic FTP traffic model in which user entering at Poisson distributed time intervals with an average rate consistent with the time of day dependent traffic level. Through simulation, we evaluated several performance metrics including SINR, user download speed, failure rate (percentage of users that fail to complete download within time limit), cell throughput, base station power consumption and energy efficiency.

We compared the performance of a coordinated network of gated narrow beam (narrow beam model) to the performance of a conventional non-coordinated network of wide beam (wide beam model), when the same traffic load is applied. The narrow beam model has an improved SINR, user download speed and cell throughput across the entire traffic levels (20% - 140% of average daily dense urban traffic load). We found that the wide beam model has an unacceptable failure rate before the load reaches average daily traffic (100%), while the failure rate of a network of gated narrow beam is only 0.2% even at busy hour load (140%). We concluded that the narrow beam model has significant throughput gain compared with wide beam model and can carry the entire dense urban busy hour load with negligible failure cases.

The base station power consumption is dominated by idle power at low traffic level, and a base station in a narrow beam network model consumes more power than a base station in wide beam network model. As a result, at low traffic levels, the base station power consumption in the narrow beam model is greater than in the wide beam model. From 50%

to 140% traffic levels, the base station in narrow beam model consumes less power than in wide beam model. The improved user throughput of gated narrow beam model allows the users to accomplish their download in a shorter period, and thus the duration of base station being at idle mode is extended. We estimated that the savings is around 140 Watts at average daily traffic (100%) to busy hour traffic level (140%) for one base station.

In Chapter 4, we also showed a preliminary energy efficiency estimate of the network of coordinated gated narrow beam. This preliminary estimate only takes into account the throughput gain (thus energy savings) brought by coordination, but not the additional energy costs. Ignoring these additional costs for the moment, the narrow beam model is shown to be more energy efficient over a large range of traffic load (40% to 140%), with the greatest savings delivered at higher loads. The narrow beam model is 41% to 48% more energy efficient than a wide beam model at 100% to 140% load levels.

We have also compared the performance metrics of narrow beam model under the FTP traffic model and under a full buffer model, in which all base stations transmit 100% of the time. The full buffer model is commonly used in literature to simplify analysis, and our results indicated that the energy efficiency estimate from full buffer model is inaccurate and misleading. We observed from the FTP traffic model simulation that the utilisation of a base station is only about 53% even at busy hour traffic level. Thus, for a realistically dimensioned network, the base station would operate at lower energy efficiency than what is achieved in the full buffer model. Our simulation results based on FTP traffic model contribute to the existing literature in that it can determine the practical performance and energy efficiency of the network under varying traffic.

In addition, we compared throughput performance of coordinated gated narrow beam model against other CoMP strategies in the literature under a 3GPP proposed reference scenario. We found that the cell spectral efficiency of the network of coordinated gated narrow beam exceeds the 3GPP target for 4x2 antenna configuration by about 92%. The coordinated gated narrow beam model is capable of delivering comparable throughput gain to joint processing schemes with less constraint on the backhaul and signalling and control. Moreover, it delivers a significant throughput gain to cell edge regions.

7.1.2 Developing an LTE compliant signalling scheme

The coordination needs to be supported by a signalling scheme that can supply the channel conditions. In Chapter 5, we analysed the required channel state information to serve the beam coordination and resource scheduling scheme used in Chapter 4, and developed a low-overhead signalling and control framework that can provide sufficient signalling information to enable the coordinated gated narrow beam model to function. The developed signalling solution consists of a framework covering reference signal configuration in the frequency-time resource grid, user equipment measurements and channel estimation, user implicit CSI feedback via uplink, and further processing of the CSI information at a central coordination unit. We have also developed a methodology for quantifying the signalling overheads and hence throughput reduction for supporting a CoMP coordinated beamforming system.

Our signalling scheme leverages reference signal and signalling functionality in current 3GPP LTE standards and provides insights into the signalling design of CoMP coordinated beamforming networks in the context of LTE. In the designed signalling scheme, reference signals of specific configuration are inserted into the resource grid and taking the place of portions of the resource that could be otherwise occupied by user data transmission. We noted that the signalling scheme only incurs a fixed overhead that is not dependent on the number of users in the coordination cluster or the volume of user traffic. The key determinant for determining the exact overhead is the periodicity of measurements. We estimated that the upper bound for the incurred signalling overhead is approximately 3.34% when the shortest periodicity of measurement (5ms) specified by 3GPP was assumed.

The signalling overhead reduces the amount of resource for data transmission and thus reduces the throughput gain identified in the Chapter 4 simulation results. It also leads to an increase in energy cost as the base stations idle time will be shortened to compensate for the reduced throughput. We concluded that the proposed coordinated gated narrow beam architecture model can be fully supported with less than 4% of signalling overhead and thus throughput reduction. The capacity loss due to signalling is markedly smaller than the capacity gain from the use of coordinated gated narrow beams. The estimation of signalling overhead enabled our modelling to account for the costs in signalling thus enabled an accurate and practical estimate of network throughput and energy efficiency.

The signalling and control framework also provides the foundation for Chapter 6 to quantify the computational energy for processing of signalling and control information, and the energy costs of carrying this information in the backhaul network.

7.1.3 Developing energy modelling of mobile network architecture

In Chapter 6, we developed a complete energy modelling of the network architecture and quantified the energy consumption of major components and processes not captured in Chapter 4 in a coordinated network of gated narrow beams. Three key energy consuming areas were analysed: the switches and interfaces power consumption for backhauling both uplink and downlink traffic; the computation energy of the Central Coordination Unit (CCU) for processing of the signalling information, and for executing the beam coordination and resource scheduling algorithm. These additional energy costs are associated with the coordination and signalling functions, they must be estimated and accounted for in order to achieve a more complete and accurate estimate of the energy efficiency of a coordinated network. This estimation of these energy costs also provides insights into the relative energy consumption of base station and other network elements, and the trade-off between the energy gain and energy costs brought by coordination.

The backhaul power consumption is divided into two parts: a baseline backhaul power consumption for carrying the IP/user traffic volume, and the additional backhaul power consumption for exchanging the signalling and control information specific to the coordinated gated narrow beam model. We estimated that the baseline backhaul power consumption adds a few percentages to the previously calculated base station power consumption (about an additional 2.7% at 20% traffic load, and an additional 1.3% at 140% traffic load). We noted that the baseline backhaul power consumption is relatively insensitive to the traffic load, as the power consumption of uplink and downlink interfaces does not scale with the traffic load. Whilst the switch power consumption is dependent on traffic load, the switch idle power consumption represents more than 95% of the total power, and the traffic dependent component consumes only less than 5%. We estimated that the additional backhaul power consumption and the CCU computation power consumption are several orders of magnitude less than that of the base station power consumption. The contribution of them to the total power consumption of the network is less than 0.06‰ across traffic load levels. It could be concluded that the

additional backhaul and computation energy consumption in supporting the coordination of gated narrow beam is marginal. Based on the backhaul power consumption analysis, we also extend the throughput comparison against other CoMP strategies in Chapter 4 to a preliminary energy efficiency comparison.

We found that, the energy efficiency of the network of coordinated gated narrow beam is affected (by 4% to 7% depending on the traffic load) by the inclusion of the energy consumption of baseline plus additional backhaul, CCU computation, and throughput reduction due to signalling overhead. It is 4.7% less energy efficient at 140% traffic load, and 6.1% less energy efficient at 20% traffic load compared with Chapter 4 preliminary results. We identified that most of the correction is due to the baseline backhaul energy and throughput reduction. The difference in the two sets of results confirms the importance of including signalling overhead and other energy costs in addition to base station power consumption when evaluating the energy efficiency of coordinated network.

We have addressed the deficiencies in prior works' incomplete modelling for energy efficiency concerning CoMP systems, and provided a complete analysis of energy efficiency of a CoMP system. As the preliminary results in Chapter 4 are obtained using the same simplified models as many previous works, the comparison between the two sets of results also indicates the amount of error that the incomplete modelling used by prior works could introduce in energy efficiency estimation for networks integrated with CoMP techniques.

We again compared the performance of the narrow beam model and the wide beam model based on the updated energy consumption results. For both models, the absolute values of the overall energy efficiency measured in Joules/Mbit have increased when compared with Chapter 4 results, as expected, which mean less energy efficient. Our results show that the energy efficiency advantage of narrow beam model over wide beam model identified in Chapter 4 is slightly reduced. Nevertheless, the narrow beam model is still significantly more energy efficient than wide beam model over a large range of traffic levels, especially higher traffic levels. We estimated the narrow beam model to be about 39% more energy efficient than wide beam model at the 100% traffic load (daily average), and about 46% more energy efficient at the 140% traffic load (busy hour). Based on the predicted daily traffic profile by GreenTouch for 2020, the weighted average energy efficiency over a diurnal cycle of narrow beam model can be expected to achieve an

improvement of about 39% compared with wide beam model in the dense urban environment.

Our results demonstrated that significant energy savings can be achieved in the coordinated network of gated narrow beams compared with conventional wide beam network under medium to high traffic level. We also concluded from the results that the overall added energy costs associated with coordination can only slightly offset the energy saving from the improved throughput identified in Chapter 4.

7.2 Future Directions

7.2.1 Long Term Assessment

In our analysis, we focused on ascertaining the energy consumption and energy efficiency of CoMP network under certain traffic conditions, and we developed relation between traffic throughput and energy. To obtain an understanding of the long term average operational energy consumption and energy efficiency of wireless network for an area, we need to perform assessment over a typical weekly cycle. The traffic fluctuates significantly during a day and it may also be quite different between weekdays and weekends in an area. Our model developed a detailed estimate of the energy consumption and energy efficiency of CoMP network as a function of its traffic throughout. A long term assessment could be performed by combining our energy estimate with the realistically measured time-varying traffic profile by network operators. We could also substitute the traffic data with prediction of future to develop an understanding of how the energy consumption and energy efficiency will evolve over time as the traffic demand increases. The estimate of long term average energy consumption would also enable understanding of the environmental impact of wireless network. We could combine the estimate of long term average energy consumption with estimates of the typical carbon footprint of electricity generation to estimate the carbon footprint.

7.2.2 Heterogeneous Network

To cope with the traffic growth, LTE-Advanced standardization has also proposed the implementation of heterogeneous network which involves a mixed deployment of macro base station and small cells of various sizes and capacities. The macro base stations remain to provide broad coverage and small cells such as micro, pico and femto base

stations provide coverage extension and extra capacity to higher traffic demand areas. The small cells consume less power compared to macro base station and they have the potential for energy savings. Our model focused on a homogeneous network of macro base stations and coordination between the macro base stations. A more complex coordination framework could be developed that is capable of coordinating both macro base stations and smaller base stations overlaying the macro base stations in a heterogeneous network. The objective would be to explore the combined benefit of coordinated multipoint transmission and small cells in improving energy efficiency of wireless network. Such study would also be useful for developing energy efficient small cell deployment strategies.

7.2.3 Uplink CoMP reception

The CoMP implementation can be made on the uplink as well as downlink for improving spectral efficiency. For uplink CoMP, a number of receivers from multiple base stations can jointly receive and process uplink signals from one mobile terminal, and the received signal is combined using advanced algorithms to improve quality. For this to happen, it is required that the received signals need to be exchanged among the receiving points. Another simpler approach is that, multiple users coordinate and schedule their uplink transmissions such that the interference among them is minimized. Although the uplink capacity and energy consumption is not the bottleneck in today's wireless network, improved uplink spectral efficiency, especially at cell edge areas, can lead to reduced transmit power and energy consumption of smartphone, and thus an extended battery life. We have quantified the energy gain of downlink CoMP and investigated into the trade-off between the increased throughput and the additional energy costs in performing the coordination. The understanding of the energy behaviour of CoMP enabled network could be more complete with a similar analysis on uplink CoMP.

References

[1]   T. Chen, H. Kim, and Y. Yang, "Energy efficiency metrics for green wireless communications," in *Wireless Communications and Signal Processing (WCSP), 2010 International Conference on*, 2010, pp. 1-6.

[2]   R. Tafazolli, "eMobility Mobile and Wireless Communications Technology Platform: Strategic Applications Research Agenda," *NetWorks European Technology Platform,* 2010.

[3]   M. Ismail, K. Qaraqe, and E. Serpedin, "Cooperation incentives and downlink radio resource allocation for green communications in a heterogeneous wireless environment," *IEEE Transactions on Vehicular Technology,* vol. 65, pp. 1627-1638, 2016.

[4]   Cisco Visual Networking Index, "Global Mobile Data Traffic Forecast Update, 2016-2021," ed: White Paper, March 2017.

[5]   K. M. S. Huq, S. Mumtaz, J. Rodriguez, and R. L. Aguiar, "Comparison of energy-efficiency in bits per joule on different downlink CoMP techniques," in *2012 IEEE International Conference on Communications (ICC)*, 2012, pp. 5716-5720.

[6]   R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck*, et al.*, "Coordinated multipoint: Concepts, performance, and field trial results," *IEEE Communications Magazine,* vol. 49, pp. 102-111, 2011.

[7]   C. Yang, S. Han, X. Hou, and A. F. Molisch, "How do we design CoMP to achieve its promised potential?," *IEEE Wireless Communications,* vol. 20, pp. 67-74, 2013.

[8]   D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata*, et al.*, "Coordinated multipoint transmission and reception in LTE-advanced: deployment scenarios and operational challenges," *IEEE Communications Magazine,* vol. 50, pp. 148-155, 2012.

[9]   R. Mahapatra, A. De Domenico, R. Gupta, and E. C. Strinati, "Green framework for future heterogeneous wireless networks," *Computer Networks,* vol. 57, pp. 1518-1528, 2013.

[10]   A. J. Fehske, P. Marsch, and G. P. Fettweis, "Bit per joule efficiency of cooperating base stations in cellular networks," in *GLOBECOM Workshops (GC Wkshps), 2010 IEEE*, 2010, pp. 1406-1411.

[11]   K. M. S. Huq, S. Mumtaz, M. Alam, J. Rodriguez, and R. L. Aguiar, "Frequency allocation for hetnet comp: Energy efficiency analysis," in *Wireless Communication Systems (ISWCS 2013), Proceedings of the Tenth International Symposium on*, 2013, pp. 1-5.

[12]   O. T. Eluwole and M. Lohi, "Coordinated multipoint power consumption modeling for energy efficiency assessment in LTE/LTE-advanced cellular networks," in *2012 19th International Conference on Telecommunications (ICT)*, 2012, pp. 1-6.

[13]   C. Han, T. Harrold, S. Armour, I. Krikidis, S. Videv, P. M. Grant*, et al.*, "Green radio: radio techniques to enable energy-efficient wireless networks," *IEEE communications magazine,* vol. 49, 2011.

[14]   O. Arnold, F. Richter, G. Fettweis, and O. Blume, "Power consumption modeling of different base station types in heterogeneous cellular networks," in *Future Network and Mobile Summit, 2010*, 2010, pp. 1-8.

[15]   I. F. Akyildiz, D. M. Gutierrez-Estevez, and E. C. Reyes, "The evolution to 4G cellular systems: LTE-Advanced," *Physical Communication,* vol. 3, pp. 217-244, 2010.

[16]   P. Chand, R. Mahapatra, and R. Prakash, "Energy Efficient Coordinated Multipoint Transmission and Reception Techniques-A Survey," *power,* vol. 45, p. 30, 2013.

[17]   G. Fettweis and E. Zimmermann, "ICT energy consumption-trends and challenges," in *Proceedings of the 11th International Symposium on Wireless Personal Multimedia Communications*, 2008, p. 6.

[18]   E. C. Strinati and L. Hérault, "Holistic approach for future energy efficient cellular networks," *e & i Elektrotechnik und Informationstechnik,* vol. 127, pp. 314-320, 2010.

[19] M. Nahas, S. Abdul-Nabi, L. Bouchnak, and F. Sabeh, "Reducing energy consumption in cellular networks by adjusting transmitted power of base stations," in *Broadband Networks and Fast Internet (RELABIRA), 2012 Symposium on*, 2012, pp. 39-44.

[20] Y. Chen, S. Zhang, S. Xu, and G. Y. Li, "Fundamental trade-offs on green wireless networks," *Communications Magazine, IEEE,* vol. 49, pp. 30-37, 2011.

[21] G. Koutitas and P. Demestichas, "A review of energy efficiency in telecommunication networks," *Telfor journal,* vol. 2, pp. 2-7, 2010.

[22] Z. Hasan, H. Boostanimehr, and V. K. Bhargava, "Green cellular networks: A survey, some research issues and challenges," *Communications Surveys & Tutorials, IEEE,* vol. 13, pp. 524-540, 2011.

[23] S. Fletcher, "Core 5-Green radio: programme objectives and overview," ed, 2008.

[24] A. Fehske, G. Fettweis, J. Malmodin, and G. Biczók, "The global footprint of mobile communications: The ecological and economic perspective," *Communications Magazine, IEEE,* vol. 49, pp. 55-62, 2011.

[25] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson*, et al.*, "How much energy is needed to run a wireless network?," *Wireless Communications, IEEE,* vol. 18, pp. 40-49, 2011.

[26] 3rd Generation Partnership Project, "Evolved Universal Terrestrial Radio Access (E-UTRA); Potential solutions for energy saving for E-UTRAN " TR 36.927 v10.1.0, October 2011.

[27] 3rd Generation Partnership Project, "Telecommunication management; Study on system and functional aspects of energy efficiency in 5G networks," TR 32.972 Release 15, April 2018.

[28] G. Auer, O. Blume, V. Giannini, I. Godor, M. Imran, Y. Jading*, et al.*, "D2. 3: Energy efficiency analysis of the reference systems, areas of improvements and target breakdown," *EARTH,* pp. 1-69, 2010.

[29] GreenTouch. [Online]. Available: http://www.greentouch.org

[30] H. M. Kwon and T. G. Birdsall, "Channel capacity in bits per joule," *Oceanic Engineering, IEEE Journal of,* vol. 11, pp. 97-99, 1986.

[31] F. Heliot, M. A. Imran, and R. Tafazolli, "Energy Efficiency Analysis of Idealized Coordinated Multi-Point Communication System," in *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*, 2011, pp. 1-5.

[32] D. Feng, C. Jiang, G. Lim, L. J. Cimini, G. Feng, and G. Y. Li, "A survey of energy-efficient wireless communications," *Communications Surveys & Tutorials, IEEE,* vol. 15, pp. 167-178, 2013.

[33] V. Mancuso and S. Alouf, "Reducing costs and pollution in cellular networks," *Communications Magazine, IEEE,* vol. 49, pp. 63-71, 2011.

[34] P. Serrano, A. De La Oliva, P. Patras, V. Mancuso, and A. Banchs, "Greening wireless communications: Status and future directions," *Computer Communications,* vol. 35, pp. 1651-1661, 2012.

[35] A. Amanna, "Green Communications: Annotated Review and Research Vision," *Virginia Tech,* p. 87, 2010.

[36] H. Claussen, L. T. Ho, and F. Pivit, "Effects of joint macrocell and residential picocell deployment on the network energy efficiency," in *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*, 2008, pp. 1-6.

[37] J. Wu, Y. Zhang, M. Zukerman, and E. K. N. Yung, "Energy-Efficient Base-Stations Sleep-Mode Techniques in Green Cellular Networks: A Survey," *IEEE Communications Surveys & Tutorials,* vol. 17, pp. 803-826, 2015.

[38] J. Wu, S. Zhou, and Z. Niu, "Traffic-Aware Base Station Sleeping Control and Power Matching for Energy-Delay Tradeoffs in Green Cellular Networks," *IEEE Transactions on Wireless Communications,* vol. 12, pp. 4196-4209, 2013.

[39] E. Oh and B. Krishnamachari, "Energy savings through dynamic base station switching in cellular wireless access networks," in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, 2010, pp. 1-5.

[40] J. Gong, S. Zhou, Z. Niu, and P. Yang, "Traffic-aware base station sleeping in dense cellular networks," in *Quality of Service (IWQoS), 2010 18th International Workshop on*, 2010, pp. 1-2.

[41] B. Rengarajan, G. Rizzo, and M. Ajmone Marsan, "Bounds on qos-constrained energy savings in cellular access networks with sleep mode," 2011.

[42] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Switch-off transients in cellular access networks with sleep modes," in *Communications Workshops (ICC), 2011 IEEE International Conference on*, 2011, pp. 1-6.

[43] L. M. Correia, D. Zeller, O. Blume, D. Ferling, Y. Jading, I. Gódor*, et al.*, "Challenges and enabling technologies for energy aware mobile radio networks," *Communications Magazine, IEEE,* vol. 48, pp. 66-72, 2010.

[44] P. Frenger, P. Moberg, J. Malmodin, Y. Jading, and I. Gódor, "Reducing energy consumption in LTE with cell DTX," in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, 2011, pp. 1-5.

[45] I. Ashraf, F. Boccardi, and L. Ho, "Sleep mode techniques for small cell deployments," *Communications Magazine, IEEE,* vol. 49, pp. 72-79, 2011.

[46] K. Heimerl, S. Hasan, K. Ali, T. Parikh, and E. Brewer, "An experiment in reducing cellular base station power draw with virtual coverage," presented at the Proceedings of the 4th Annual Symposium on Computing for Development, Cape Town, South Africa, 2013.

[47] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Optimal Energy Savings in Cellular Access Networks," in *2009 IEEE International Conference on Communications Workshops*, 2009, pp. 1-5.

[48] A. He, S. Srikanteswara, K. K. Bae, T. R. Newman, J. H. Reed, W. H. Tranter*, et al.*, "System power consumption minimization for multichannel communications using cognitive radio," in *Microwaves, Communications, Antennas and Electronics Systems, 2009. COMCAS 2009. IEEE International Conference on*, 2009, pp. 1-5.

[49] A. He, S. Srikanteswara, J. H. Reed, X. Chen, W. H. Tranter, K. K. Bae*, et al.*, "Minimizing energy consumption using cognitive radio," in *Performance, Computing and Communications Conference, 2008. IPCCC 2008. IEEE International*, 2008, pp. 372-377.

[50] J. Xu and L. Qiu, "Area power consumption in a single cell assisted by relays," in *Green Computing and Communications (GreenCom), 2010 IEEE/ACM Int'l Conference on & Int'l Conference on Cyber, Physical and Social Computing (CPSCom)*, 2010, pp. 460-465.

[51] G. Chandwani, S. N. Datta, and S. Chakrabarti, "Relay assisted cellular system for energy minimization," in *India Conference (INDICON), 2010 Annual IEEE*, 2010, pp. 1-4.

[52] Q. Li, R. Q. Hu, Y. Qian, and G. Wu, "Cooperative communications for wireless networks: techniques and applications in LTE-advanced systems," *Wireless Communications, IEEE,* vol. 19, 2012.

[53] X. J. Li, B.-C. Seet, and P. H. J. Chong, "Multihop cellular networks: Technology and economics," *Computer Networks,* vol. 52, pp. 1825-1837, 2008.

[54] J.-y. Song, H. J. Lee, and D.-H. Cho, "Power consumption reduction by multi-hop transmission in cellular networks," in *Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th*, 2004, pp. 3120-3124.

[55] C. Li, J. Zhang, and K. B. Letaief, "Energy efficiency analysis of small cell networks," in *2013 IEEE International Conference on Communications (ICC)*, 2013, pp. 4404-4408.

[56] W. C. Cheung, T. Q. Quek, and M. Kountouris, "Throughput optimization, spectrum allocation, and access control in two-tier femtocell networks," *Selected Areas in Communications, IEEE Journal on,* vol. 30, pp. 561-574, 2012.

[57]   A.-H. Tsai, L.-C. Wang, J.-H. Huang, and R.-B. Hwang, "High-capacity OFDMA femtocells by directional antennas and location awareness," *Systems Journal, IEEE,* vol. 6, pp. 329-340, 2012.

[58]   W. Li, W. Zheng, Y. Xie, and X. Wen, "Clustering based power saving algorithm for self-organized sleep mode in femtocell networks," in *The 15th International Symposium on Wireless Personal Multimedia Communications*, 2012, pp. 379-383.

[59]   M. Jada, M. M. A. Hossain, J. Hämäläinen, and R. Jäntti, "Impact of femtocells to the WCDMA network energy efficiency," in *2010 3rd IEEE International Conference on Broadband Network and Multimedia Technology (IC-BNMT)*, 2010, pp. 305-310.

[60]   Y. Hou and D. Laurenson, "Energy efficiency of high QoS heterogeneous wireless communication network," in *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*, 2010, pp. 1-5.

[61]   F. Richter, A. J. Fehske, P. Marsch, and G. P. Fettweis, "Traffic demand and energy efficiency in heterogeneous cellular mobile radio networks," in *Vehicular Technology Conference (VTC 2010-Spring), 2010 IEEE 71st*, 2010, pp. 1-6.

[62]   J. Hoydis, M. Kobayashi, and M. Debbah, "Green small-cell networks," *Vehicular Technology Magazine, IEEE,* vol. 6, pp. 37-43, 2011.

[63]   F. Richter and G. Fettweis, "Cellular mobile network densification utilizing micro base stations," in *Communications (ICC), 2010 IEEE International Conference on*, 2010, pp. 1-6.

[64]   D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *Selected Areas in Communications, IEEE Journal on,* vol. 28, pp. 1380-1408, 2010.

[65]   E. Hossain, D. I. Kim, and V. K. Bhargava, *Cooperative cellular wireless networks*: Cambridge University Press, 2011.

[66]   J. Segel and M. Weldon, "Lightradio portfolio-technical overview," *Technology White Paper,* vol. 1, 2011.

[67]   K. Michail, V. Tatsis, D. Panagiotis, C. Georgios, N. Nomikos, D. N. Skoutas*, et al.*, "A load and channel aware dynamic point selection algorithm for LTE-A CoMP networks," in *Telecommunications and Multimedia (TEMU), 2016 International Conference on*, 2016, pp. 1-5.

[68]   A. S. Ahmad, M. J. Huque, and M. F. Hossain, "A novel CoMP transmission mechanism for the downlink of LTE-A cellular networks," in *Informatics, Electronics and Vision (ICIEV), 2016 5th International Conference on*, 2016, pp. 875-880.

[69]   R. Annavajjala, "Low-complexity distributed algorithms for uplink CoMP in heterogeneous LTE networks," *Journal of Communications and Networks,* vol. 18, pp. 150-161, 2016.

[70]   J. Zhang, R. Chen, J. G. Andrews, A. Ghosh, and R. W. Heath, "Networked MIMO with clustered linear precoding," *IEEE Transactions on Wireless Communications,* vol. 8, 2009.

[71]   M. Stojnic, H. Vikalo, and B. Hassibi, "Rate maximization in multi-antenna broadcast channels with linear preprocessing," *IEEE Transactions on Wireless Communications,* vol. 5, pp. 2338-2342, 2006.

[72]   S. Han, C. Yang, M. Bengtsson, and A. Pérez-Neira, "Channel norm-based user scheduler in coordinated multi-point systems," in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, 2009, pp. 1-5.

[73]   G. Cili, H. Yanikomeroglu, and F. R. Yu, "Coordinated Multi-Point (CoMP) adaptive estimation and prediction schemes using superimposed and decomposed channel tracking," in *Communications Workshops (ICC), 2013 IEEE International Conference on*, 2013, pp. 116-121.

[74]   G. Cili, H. Yanikomeroglu, and F. R. Yu, "Energy efficiency and capacity evaluation of LTE-Advanced downlink CoMP schemes subject to channel estimation errors and system delay," in *Vehicular Technology Conference (VTC Fall), 2013 IEEE 78th*, 2013, pp. 1-5.

[75]   R. Zakhour and D. Gesbert, "Optimized data sharing in multicell MIMO with finite backhaul capacity," *IEEE Transactions on Signal Processing,* vol. 59, pp. 6102-6111, 2011.

[76] N. Seifi, M. Viberg, R. W. Heath, J. Zhang, and M. Coldrey, "Coordinated single-cell vs multi-cell transmission with limited-capacity backhaul," in *Signals, Systems and Computers (ASILOMAR), 2010 Conference Record of the Forty Fourth Asilomar Conference on*, 2010, pp. 1217-1221.

[77] A. Marotta, K. Kondepu, F. Giannone, S. Doddikrinda, D. Cassioli, C. Antonelli*, et al.*, "Performance evaluation of CoMP coordinated scheduling over different backhaul infrastructures: A real use case scenario," in *Science of Electrical Engineering (ICSEE), IEEE International Conference on the*, 2016, pp. 1-5.

[78] G. Y. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen*, et al.*, "Energy-efficient wireless communications: tutorial, survey, and open issues," *IEEE Wireless Communications,* vol. 18, 2011.

[79] F. Héliot, R. Hoshyar, and R. Tafazolli, "An accurate closed-form approximation of the distributed MIMO outage probability," *IEEE Transactions on Wireless Communications,* vol. 10, pp. 5-11, 2011.

[80] H. Ozcelik, M. Herdin, W. Weichselberger, J. Wallace, and E. Bonek, "Deficiencies of'Kronecker'MIMO radio channel model," *Electronics Letters,* vol. 39, pp. 1209-1210, 2003.

[81] G. Miao, N. Himayat, and G. Y. Li, "Energy-efficient link adaptation in frequency-selective channels," *IEEE Transactions on communications,* vol. 58, 2010.

[82] M. Hunukumbure and S. Vadgama, "Improving cell edge energy efficiency through CoMP beam-forming," in *Vehicular Technology Conference (VTC Spring), 2013 IEEE 77th*, 2013, pp. 1-5.

[83] H.-L. Määttänen, K. Hämäläinen, J. Venäläinen, K. Schober, M. Enescu, and M. Valkama, "System-level performance of LTE-Advanced with joint transmission and dynamic point selection schemes," *EURASIP Journal on Advances in Signal Processing,* vol. 2012, p. 247, 2012.

[84] J. Acharya, L. Gao, and S. Gaur, *Heterogeneous Networks in LTE-advanced*: John Wiley & Sons, 2014.

[85] D. Gesbert, S. G. Kiani, A. Gjendemsjo, and G. E. Oien, "Adaptation, coordination, and distributed resource allocation in interference-limited wireless networks," *Proceedings of the IEEE,* vol. 95, pp. 2393-2409, 2007.

[86] M. Rahman and H. Yanikomeroglu, "Interference avoidance through dynamic downlink OFDMA subchannel allocation using intercell coordination," in *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*, 2008, pp. 1630-1635.

[87] M. Boujelben, S. Benrejeb, and S. Tabbane, "Interference coordination schemes for wireless mobile advanced systems: a survey," *arXiv preprint arXiv:1403.3818,* 2014.

[88] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger*, et al.*, "Cloud RAN for mobile networks—a technology overview," *Communications Surveys & Tutorials, IEEE,* vol. 17, pp. 405-426, 2015.

[89] Y.-H. Nam, Y. Akimoto, Y. Kim, M.-i. Lee, K. Bhattad, and A. Ekpenyong, "Evolution of reference signals for LTE-advanced systems," *Communications Magazine, IEEE,* vol. 50, pp. 132-138, 2012.

[90] P. Chanclou, A. Pizzinat, F. Le Clech, T.-L. Reedeker, Y. Lagadec, F. Saliou*, et al.*, "Optical fiber solution for mobile fronthaul to achieve cloud radio access network," in *Future Network and Mobile Summit (FutureNetworkSummit), 2013*, 2013, pp. 1-11.

[91] C. Powell, "Technical Analysis: Beamforming vs MIMO Antennas," Radio Frequency SystemsMarch 2014.

[92] B. Schulz, "LTE Transmission Modes and Beamforming," *Rohde and Schwarz White Paper,* 2011.

[93] 3rd Generation Partnership Project, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation," TS 36.211 v11.0.0, September 2012.

[94] M. Litzenburger, M. Ohm, and H. Haining, "Multi User MIMO with Fixed Beamforming for FDD Systems."

[95] U. Jang, K. Y. Lee, K. S. Cho, and W. Ryu, "Transmit beamforming based inter-cell interference alignment and user selection with CoMP," in *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*, 2010, pp. 1-5.

[96] M. Sadek, A. Tarighat, and A. H. Sayed, "A leakage-based precoding scheme for downlink multi-user MIMO channels," *Wireless Communications, IEEE Transactions on,* vol. 6, pp. 1711-1721, 2007.

[97] Z. Huang, B. Li, and M. Liu, "Coordinated Beamforming of CoMP with limited feedback," in *Network Computing and Information Security (NCIS), 2011 International Conference on*, 2011, pp. 138-141.

[98] NTT DOCOMO, "System Performance for CS/CB-CoMP in Homogeneous Networks with High Tx Power RRHs," 3GPP TSG RAN WG1; R1-110866; February 2011.

[99] Hitachi Ltd., "Preliminary Evaluation of DL Homogeneous CoMP," 3GPP TSG RAN WG1; R1-111283; May 2011.

[100] A. Osseiran, V. Stankovic, E. Jorswieck, T. Wild, M. Fuchs, and M. Olsson, "A MIMO framework for 4G systems: WINNER concept and results," in *Signal Processing Advances in Wireless Communications, 2007. SPAWC 2007. IEEE 8th Workshop on*, 2007, pp. 1-5.

[101] W. Choi and J. G. Andrews, "The capacity gain from intercell scheduling in multi-antenna systems," *Wireless Communications, IEEE Transactions on,* vol. 7, pp. 714-725, 2008.

[102] S. G. Kiani and D. Gesbert, "Optimal and distributed scheduling for multicell capacity maximization," *Wireless Communications, IEEE Transactions on,* vol. 7, pp. 288-297, 2008.

[103] ZTE, "Discussion on aggregated PMI feedback," 3GPP TSG RAN WG1; R1-122366; May 2012.

[104] L. Lei, C. Lin, J. Cai, and X. S. Shen, "Flow-level performance of opportunistic OFDM-TDMA and OFDMA networks," *Wireless Communications, IEEE Transactions on,* vol. 7, pp. 5461-5472, 2008.

[105] R. K. Almatarneh, M. H. Ahmed, and O. Dobre, "Performance analysis of proportional fair scheduling in OFDMA wireless systems," in *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*, 2010, pp. 1-5.

[106] H. Kim, K. Kim, Y. Han, and S. Yun, "A proportional fair scheduling for multicarrier transmission systems," in *Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th*, 2004, pp. 409-413.

[107] C. Wengerter, J. Ohlhorst, V. Elbwart, and A. G. Edler, "Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA," in *Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st*, 2005, pp. 1903-1907.

[108] T. B. L. E. L. R. Ramjee, "Generalized proportional fair scheduling in third generation wireless data networks," 2006.

[109] H. J. Kushner and P. Whiting, "Convergence of proportional-fair sharing algorithms under general conditions," *Wireless Communications, IEEE Transactions on,* vol. 3, pp. 1250-1259, 2004.

[110] P. Kela, J. Puttonen, N. Kolehmainen, T. Ristaniemi, T. Henttonen, and M. Moisio, "Dynamic packet scheduling performance in UTRA long term evolution downlink," in *Wireless Pervasive Computing, 2008. ISWPC 2008. 3rd International Symposium on*, 2008, pp. 308-313.

[111] M. Chiang and J. Bell, "Balancing supply and demand of bandwidth in wireless cellular networks: utility maximization over powers and rates," in *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, 2004, pp. 2800-2811.

[112] D. Dolev, A. Zymnis, S. P. Boyd, D. Bickson, and Y. Tock, "Distributed large scale network utility maximization," in *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*, 2009, pp. 829-833.

[113] J.-W. Lee and M. Chiang, "Jointly optimal congestion and contention control based on network utility maximization," *Communications Letters, IEEE,* vol. 10, pp. 216-218, 2006.

[114] J. Zyren and W. McCoy, "Overview of the 3GPP long term evolution physical layer," *Freescale Semiconductor, Inc., white paper,* 2007.

[115] 3rd Generation Partnership Project, "Requirements for further advancements forevolved universal terrestrial radio access (E-UTRA) LTE-Advanced," TR 36.913 v10.0.0, March 2011.

[116] P. Hosein, "On the optimal allocation of downlink resources in OFDM-based wireless networks," in *Wired/Wireless Internet Communications*, ed: Springer, 2006, pp. 202-213.

[117] GreenTouch Mobile Communications Working Group, "Architecture Doc2: Reference Scenarios," May 2013.

[118] O. Blume, A. Ambrosy, M. Wilhelm, and U. Barth, "Energy Efficiency of LTE networks under traffic loads of 2020," in *Wireless Communication Systems (ISWCS 2013), Proceedings of the Tenth International Symposium on*, 2013, pp. 1-5.

[119] "Green Meter Study: Reducing the Net Energy Consumption in Communications Networks by up to 90% by 2020," GreenTouch White Paper June 2013.

[120] C. Desset, "Quantitative analysis of energy saving potential in future cellular base stations and networks," *IEEE JSAC Series on Green Communications and Networking*.

[121] C. Mueller, "On the importance of realistic traffic models for wireless network evaluations," *COST 2100 12th MCM,* 2010.

[122] 3rd Generation Partnership Project, "Annex A of: Evolved Universal Terrestrial Radio Access (EUTRA); further advancements for E-UTRA physical layer aspects," TR 36.814 v9.0.0, March 2010.

[123] G. Morozov, A. Davydov, and I. Bolotin, "Performance evaluation of dynamic point selection CoMP scheme in heterogeneous networks with FTP traffic model," in *Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2012 4th International Congress on*, 2012, pp. 922-926.

[124] S. Tombaz, P. Monti, F. Farias, M. Fiorani, L. Wosinska, and J. Zander, "Is backhaul becoming a bottleneck for green wireless access networks?," in *Communications (ICC), 2014 IEEE International Conference on*, 2014, pp. 4029-4035.

[125] 3rd Generation Partnership Project, "Evolved Universal Terrestrial Radio Access (EUTRA); further advancements for E-UTRA physical layer aspects," TR 36.814 v9.0.0, March 2010.

[126] H. Sun, W. Fang, J. Liu, and Y. Meng, "Performance evaluation of CS/CB for coordinated multipoint transmission in LTE-A downlink," in *2012 IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications-(PIMRC)*, 2012, pp. 1061-1065.

[127] L. Zhang, S. Peter, and C. Wang, "Performance evaluation of layer adaptive multi-user scheduling in LTE-A downlink," in *7th International Conference on Communications and Networking in China*, 2012, pp. 787-791.

[128] Z. Xiong, H. Yang, M. Zhang, Y. Meng, and Y. Pan, "Feedback and scheduling for coordinated beamforming of CoMP in LTE-advanced system," in *2013 IEEE 78th Vehicular Technology Conference (VTC Fall)*, 2013, pp. 1-5.

[129] H. Sun and T. Yang, "Performance evaluation of distributed scheduling for downlink coherent joint transmission," in *2015 IEEE 82nd Vehicular Technology Conference (VTC2015-Fall)*, 2015, pp. 1-5.

[130] J. Zhu, X. She, X. Yun, L. Chen, and H. Otsuka, "A practical design of downlink coordinated multi-point transmission for LTE-Advanced," in *2010 IEEE 71st Vehicular Technology Conference*, 2010, pp. 1-6.

[131] 3GPP TSG RAN WG1, "On CSI feedback for ITU requirement fulfilling CoMP schemes," San Francisco, USA, no. R1-092024, May 2009.

[132] 3GPP TSG RAN WG1, "CQI definition for CoMP," Prague, Czech Republic, no. R1-122132, May 2012.

[133]   3GPP TSG RAN WG1, "CQI calculation for CoMP," Zhuhai, China, no. R1-113276, October 2011.

[134]   M. Ding and H. Luo, *Multi-point Cooperative Communication Systems: Theory and Applications*: Springer, 2013.

[135]   3GPP TSG RAN WG1, "CQI feedback to support downlink CoMP," Prague, Czech Republic, no. R1-122158, May 2012.

[136]   D. J. Love, R. W. Heath Jr, V. K. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *Selected Areas in Communications, IEEE Journal on,* vol. 26, pp. 1341-1365, 2008.

[137]   D. J. Love, R. W. Heath Jr, and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *Information Theory, IEEE Transactions on,* vol. 49, pp. 2735-2747, 2003.

[138]   3GPP TSG RAN WG1, "CSI feedback modes for CoMP," Prague, Czech Republic, no. R1-121946, May 2012.

[139]   3GPP TSG RAN WG1, "Discussion on CQI definition and Rank reporting for DL CoMP," Prague, Czech Republic, no. R1-122076, May 2012.

[140]   3GPP TSG RAN WG1, "CQI definition for DL CoMP," Prague, Czech Republic, no. R1-122292, May 2012.

[141]   W. Xi, X. Yun, S. Nagata, Y. Kishiyama, and L. Chen, "An enhanced interference measurement scheme for CoMP in LTE-Advanced downlink," in *Communications (ICC), 2013 IEEE International Conference on*, 2013, pp. 4870-4874.

[142]   3rd Generation Partnership Project, "Coordinated multi-point operation for LTE physical layer aspect " TR 36.819 v11.1.0, December 2011.

[143]   Y. Ohwatari, N. Miki, T. Abe, S. Nagata, and Y. Okumura, "Investigation on improvement in channel estimation accuracy using data signal muting in downlink coordinated multiple-point transmission and reception in LTE-Advanced," in *Wireless Communications and Networking Conference (WCNC), 2011 IEEE*, 2011, pp. 1288-1293.

[144]   3GPP TSG RAN WG1, "Impact of Downlink CoMP on the Air Interface," Ljubljana, Slovenia, no. R1-090366, Jan 2009.

[145]   3GPP TSG RAN WG1, "Interference measurement resource configuration and CQI calculation," Prague, Czech Republic, no. R1-121947, May 2012.

[146]   C. F. Lanzani, G. Kardaras, and D. Boppana, "Remote Radio Heads and the evolution towards 4G networks," *Technical white paper from Radiocomp and Altera,* vol. 8, 2009.

[147]   P. Monti, S. Tombaz, L. Wosinska, and J. Zander, "Mobile backhaul in heterogeneous network deployments: Technology options and power consumption," in *Transparent Optical Networks (ICTON), 2012 14th International Conference on*, 2012, pp. 1-7.

[148]   S. Tombaz, P. Monti, K. Wang, A. Västberg, M. Forzati, and J. Zander, "Impact of backhauling power consumption on the deployment of heterogeneous mobile networks," in *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*, 2011, pp. 1-5.

[149]   F. Capozzi, G. Piro, L. A. Grieco, G. Boggia, and P. Camarda, "Downlink packet scheduling in LTE cellular networks: Key design issues and a survey," *Communications Surveys & Tutorials, IEEE,* vol. 15, pp. 678-700, 2013.

[150]   C. Khirallah, D. Vukobratović, and J. Thompson, "On energy efficiency of joint transmission coordinated multi-point in LTE-advanced," in *Smart Antennas (WSA), 2012 International ITG Workshop on*, 2012, pp. 54-61.

[151]   S. Huq, K. Mohammed, S. Mumtaz, J. Bachmatiuk, J. Rodriguez, X. Wang*, et al.*, "Green hetnet comp: energy efficiency analysis and optimization," *Vehicular Technology, IEEE Transactions on,* vol. 64, pp. 4670-4683, 2015.

[152]   Cisco Catalyst 3750-X and 3560-X Series Switches Data Sheet.   [Online]. Available: http://www.cisco.com/c/en/us/products/collateral/switches/catalyst-3750-x-series-switches/data_sheet_c78-584733.html

[153]   F. Qian, Z. Wang, A. Gerber, Z. M. Mao, S. Sen, and O. Spatscheck, "Characterizing radio resource allocation for 3G networks," in *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, 2010, pp. 137-150.

[154]   3rd Generation Partnership Project, "Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification," TS 36.331 v11.4.0, July 2013.

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

[156]   Netmanias Techinal Document, "LTE Identification I: UE and ME Identifiers," August, 2013.

[157]   K. Hinton, R. Ayre, A. Vishwanath, C. Zhang, and M. Feng, "Energy efficiency of optical IP protocol suites," in *National Fiber Optic Engineers Conference*, 2012, p. NTu1E. 3.

[158]   H. Yang and T. L. Marzetta, "Total energy efficiency of cellular large scale antenna system multiple access mobile networks," in *Online Conference on Green Communications (GreenCom), 2013 IEEE*, 2013, pp. 27-32.

[159]   Standard Performance Evaluation Corporation (SPEC) Data Sheet.   [Online]. Available: https://www.spec.org/power_ssj2008/results/res2015q2/power_ssj2008-20150317-00691.html

[160]   Dell Inc. PowerEdge R630 (Intel Xeon E5-2699) Performance Data Sheet.   [Online]. Available: http://browser.primatelabs.com/geekbench3/2021143

[161]   X. Gao, L. Dai, S. Han, I. Chih-Lin, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays," *IEEE Journal on Selected Areas in Communications,* vol. 34, pp. 998-1009, 2016.

[162]   M. Fiorani, S. Tombaz, J. Martensson, B. Skubic, L. Wosinska, and P. Monti, "Modeling energy performance of C-RAN with optical transport in 5G network scenarios," *IEEE/OSA Journal of Optical Communications and Networking,* vol. 8, pp. B21-B34, 2016.

[163]   3rd Generation Partnership Project, "Technical Specification Group Radio Access Network; Physical Channels and Modulation," TS 38.211 v15.6.0, June 2019.

[164]   3rd Generation Partnership Project, "Technical Specification Group Radio Access Network; Multiplexing and Channel Coding," TS 38.212 v15.6.0, June 2019.

[165]   3rd Generation Partnership Project, "Technical Specification Group Radio Access Network; Physical layer procedures for control," TS 38.213 v15.6.0, June 2019.

[166]   M. M. Mowla, I. Ahmad, D. Habibi, and Q. V. Phung, "A green communication model for 5G systems," *IEEE Transactions on Green Communications and Networking,* vol. 1, pp. 264-280, 2017.

Author/s:
Zhu, Jiazhen

Title:
Energy efficiency of wireless network using coordinated gated narrow beams

Date:
2019

Persistent Link:
http://hdl.handle.net/11343/226878

File Description:
Energy efficiency of wireless network using coordinated gated narrow beams