

Cross-Layer Resource Allocation and Scheduling in Bidirectional Cooperative Multichannel Relaying Networks

by

Ahmed Zainaldin, M. A. Sc.

A dissertation submitted to the
Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electrical Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering

Department of Systems and Computer Engineering

Carleton University

Ottawa, Ontario

December, 2015

Abstract

Efficient allocation of the scarce resource spectrum is a fundamental problem in mobile communications. Cooperative Relaying (CoR) and Network Coding (NC) are two promising techniques for improving the performance of next-generation wireless networks. In this thesis, cross-layer resource allocation is studied for relay-assisted bidirectional multichannel wireless networks. We model the network in which two nodes, User Equipment (UE) and eNodeB (eNB), exchange information with the assistance of a Relay Station (RS). The UEs and eNB can choose between different transmission schemes: direct transmission, pure CoR, or via the combination of NC and CoR (NC/CoR). Novel optimization frameworks are proposed for the resource allocation problem for such networks. First, the achievable rate regions are characterized, and the joint optimization problem of power allocation and transmission scheme selection for maximizing the bidirectional stability region is defined. A hybrid transmission scheme with adaptive resource allocation is proposed to dynamically select the best transmission strategy and the optimal resource allocation at each bidirectional transmission frame. The simulation results show that the hybrid scheme with joint power control and channel allocation improves the system performance considerably. Second, the joint resource allocation and relay selection problem is studied for bidirectional Orthogonal Frequency Division Multiple Access (OFDMA) relay networks. The problem of joint channel allocation/pairing, relay selection, and

transmission scheme selection is formulated as a combinatorial optimization problem. First, we adopt a graph-based approach by establishing an equivalent Maximum Weighted Clique Problem (MWCP). In addition, the joint problem is mapped into a three-dimensional assignment problem in which a metaheuristic algorithm based on Ant Colony Optimization (ACO) is employed to find the solution in polynomial time. The results confirm that while the proposed solutions provide similar results, the ACO algorithm is faster and more reliable due to its lower complexity.

key words— Cooperative Communications, Network Coding, Two-Way Relaying, OFDMA, Ant Colony Optimization, Maximum Weighted Clique Problem

I dedicate this thesis to my family, especially my father and mother, for their constant support and motivation.

Acknowledgements

First of all, I would like to thank my advisor, Professor Ioannis Lambadaris, for his continuous inspiration, guidance and support on my thesis work. In the past few years, he constantly encouraged me to focus on the most important and challenging problems, and instructed me how to find a way toward innovative ideas and creative solutions. He also gave me a lot of freedom to choose the favorite research topics to strengthen my interest and enthusiasm in research.

I am also grateful to all the colleagues and students in the Broadband Networks Lab at the Department of Systems and Computer Engineering for their enjoyable discussions and shared research work with me on communications concepts and interesting ideas. I would like to thank Dr. Hassan Halabian for his collaborative work and invaluable comments, which helped improve the quality of my research.

Lastly, I would like to express my deepest gratitude to my parents for their continuous support and motivation. Special gratitude to all my family, who keep encouraging me to pursue my PhD degree. To them all I dedicate this thesis.

Table of Contents

Abstract	ii
Acknowledgements	v
Table of Contents	vi
List of Tables	xi
List of Figures	xii
List of Abbreviations	xiv
1 Introduction	1
1.1 Overview	1
1.2 Research Background	2
1.2.1 Wireless Relay Networks	3
1.2.2 Cooperative Diversity	4
1.2.3 One-Way Relaying	6
1.2.4 Two-Way Relaying	7
1.3 Motivation	9
1.4 Objective	11
1.5 Contributions and Organization of the Thesis	11

1.6	List of Publications and Patent Filings	13
2	Background and Related Work	15
2.1	Two-Way Multichannel Relaying	15
2.1.1	Opportunistic Scheduling in Two-Way Relaying	15
2.1.2	Cross-layer Resource Allocation and Stability Analysis in Two-Way Relaying	18
2.2	Resource Allocation in Two-Way Relaying OFDMA Networks	20
2.2.1	Joint Channel Assignment and Pairing in Two-Way OFDMA Relaying	21
2.2.2	Optimization vs. Heuristic Approaches	24
2.3	Ant Colony Optimization	25
2.3.1	ACO Algorithms	27
3	Optimal Joint Resource Allocation and Power Control in Bidirectional Relaying Networks	29
3.1	System Description and Assumptions	29
3.1.1	Channel Model	33
3.1.2	Queuing Model	35
3.2	Achievable Rate Regions for Direct, CoR and NC/CoR Schemes	37
3.2.1	Direct Transmission	37
3.2.2	CoR Transmission	39
3.2.3	NC/CoR Transmission	40
3.3	Achievable Rate Region for PLNC/CoR	43
3.4	Hybrid Transmission Scheme	45
3.4.1	Hybrid Scheme for the Continuous-flow System	50
3.4.2	Hybrid Scheme with Discrete Packet Arrivals	54

3.4.3	Hybrid Scheme with Power Control	56
3.4.4	Complexity Analysis	57
3.5	Simulations and Discussion	58
3.5.1	Average System Queue Backlog (Delay) and Capacity Evaluation	59
3.5.2	Comparison with Physical Layer Network Coding	64
3.5.3	Performance Evaluation of the Hybrid Scheme with Power Control	67
3.6	Chapter Summary	67
4	Joint Resource Allocation and Relay Selection in OFDMA Net-	
	works	71
4.1	System Description and Assumptions	72
4.1.1	Proposed Frame Architecture	73
4.1.2	Channel Model	74
4.1.3	Queuing Model	75
4.2	Optimization Framework	76
4.2.1	Achievable Bidirectional Rate Pairs	76
4.2.2	Optimization Problem Formulation	82
4.2.3	Modified Optimization Problem	84
4.3	Graph-Based Optimization	85
4.3.1	Maximum Weighted Clique Problem (MWCP)	86
4.3.2	Basic Bron-Kerbosch Algorithm	87
4.3.3	Modified Bron-Kerbosch Algorithm for MWCP with Pivoting	88
4.3.4	MWCP Graph Model	89
4.3.5	Complexity	90
4.3.6	Sequential Greedy Allocation for MWCP problem	91
4.4	ACO-Based Optimization	91

4.4.1	Three-Dimensional Assignment Problem ACO (3DAP-ACO)	
	Algorithm	92
4.4.2	Enhanced Hybrid 3DAP-ACO Algorithm	94
4.4.3	Complexity	100
4.4.4	Fast-3DAP-ACO	101
4.5	Simulations and Discussion	102
4.5.1	Average System Queue Backlog Evaluation	104
4.5.2	Impact of Number of RNs on Average System Capacity	107
4.5.3	Optimization Algorithms Comparison	107
4.5.4	Optimal vs Suboptimal Algorithm Simulation Time	109
4.6	Chapter Summary	110
5	Conclusion and Future Work	112
5.1	Conclusion and Discussion	112
5.2	Future Work	114
	List of References	116
	Appendix A Long Term Evolution (LTE) Background	126
A.1	LTE Network Architecture	127
A.2	Relay Technologies in LTE-Advanced Cellular Networks	129
A.3	Frequency-Division Duplexing (FDD) versus Time-Division Duplexing (TDD)	133
A.4	Spectrum Allocation and Resource Blocks	134
	Appendix B CoR Rate Region	136
	Appendix C NC/CoR Rate Region	140

List of Tables

3.1	Simulation Parameters for System Model	59
4.1	ACO System Parameters	101
4.2	Simulation Parameters in the System	103
A.1	Transmission Bandwidth Configuration in LTE Channel	135

List of Figures

1.1	Relay Deployment Scenarios in LTE-Advanced Networks [1]	4
1.2	Convention Relaying vs. Network Coding	9
2.1	Ants Facing an Obstacle	26
3.1	Two-Way Relaying Schemes	31
3.2	Uplink Model	32
3.3	Two-Way Relay Queuing Model	35
3.4	Achievable Rate Regions for Direct, CoR and NC/CoR Schemes in Example 1	38
3.5	Achievable Rate Regions for Direct, CoR and NC/CoR Schemes in Example 2	48
3.6	Delay and Throughput Performance of Direct, CoR, NC/CoR and Hybrid schemes (SISO)	60
3.7	Capacity and Delay Performance of Different Transmission Schemes under SISO and MIMO Configurations	62
3.8	Average Queue Backlog vs. Number of RSs	63
3.9	Average System Capacity vs. Number of RSs	65
3.10	Network Capacity and Delay Performance of the Hybrid Scheme with Power Control	68

3.11 Average System Capacity vs. Number of RSs - Hybrid Scheme with Power Control	69
4.1 OFDMA Frame Architectures	73
4.2 Two-Way Relay Queuing Model	75
4.3 Downlink Subframe with \mathcal{N} Subchannels (one time-slot)	77
4.4 Direct Transmission Scheme	78
4.5 CoR Transmission Scheme	80
4.6 NC/CoR Transmission Scheme	81
4.7 MWCP graph with $N = 2$ where $i \in N, j \in N$ and $k \in N$	90
4.8 3DAP Cost Matrix	93
4.9 Network capacity and delay performance of different transmission schemes (SISO)	105
4.10 Network capacity and delay performance of different transmission schemes (MIMO)	106
4.11 ACO vs. Greedy for $N = 11$ for SISO	108
4.12 Average Simulation Time per Cycle	110
A.1 Evolution of Radio Access Technologies [2]	127
A.2 LTE-Advanced Relay Architecture	131
A.3 MBSFN Subframe with Relay Support	132
A.4 FDD vs. TDD LTE Frame Structure [3]	133
A.5 LTE Resource Grid [3]	135

List of Abbreviations

3GPP	3rd Generation Partnership Project
3G	3rd Generation
4G	4th Generation
AF	Amplify-and-Forward
DF	Decode-and-Forward
BS	Base Station
eNB	eNodeB
UE	User Equipment
RN	Relay Node
AWGN	Additive White Gaussian Noise
OFDMA	Orthogonal Frequency Division Multiple Access
CSI	Channel State Information
TWRC	Two-Way Relay Channel
DT	Direct Transmission

CoR	Cooperative Relaying
NC	Network Coding
XOR	Exclusive-OR
FDD	Frequency-Division Duplexing
TDD	Time-Division Duplexing
LTE	Long Term Evolution
LTE-A	Long Term Evolution-Advanced
WiMAX	Worldwide Interoperability for Microwave Access
MAC	Multiple Access
SISO	Single-Input Single-Output
MIMO	Multiple-Input Multiple-Output
QoS	Quality of Service
QAP	Quadratic Assignment Problem
3DAP	Three-Dimensional Assignment Problem
MWCP	Maximum Weighted Clique Problem
ACO	Ant Colony Optimization

Chapter 1

Introduction

This chapter provides a brief background and highlights the importance of resource allocation in multichannel relay-enhanced networks. An overview of resource background and author's contributions are also presented.

1.1 Overview

Future wireless systems will face a growing demand for broadband and multimedia services. In the past decade, we have witnessed a tremendous growth in wireless data traffic. The key drivers for this rapid growth include the continuous progress and roll-out of next-generation wireless networks such as Long Term Evolution (LTE) as well as the development of smart phones and other mobile computing devices. The roll-out of smart phones and connected mobile devices has resulted in substantial and continued growth in wireless data traffic. Numerous research efforts have been cast into this actively evolving field, and generated cutting-edge technologies to support the growing usage trends. However, the scarcity of radio resources, such as bandwidth and power has become a major limitation in the development of high-speed wireless communications. The accommodation of user demands with limited system resources

arises as a big challenge, especially when the number of mobile users is large. As such, efficient utilization of radio resources is of critical importance for next-generation wireless systems. In addition, wireless environments, such as channel conditions, user locations, etc., are usually changing rapidly. To maintain satisfactory Quality of Service (QoS), the allocation of resources has to be frequently adapted to the current status of the network, which results in higher implementation complexity.

In this thesis, we build upon the motivation of balancing the user demands and the scarcity of radio resources. The work focuses on the design of optimal radio resource allocation schemes which are dynamically adapted to the environmental changes. Moreover, the stochastic nature in wireless networks is exploited to improve the system performance and user QoS satisfaction.

1.2 Research Background

Traditional wireless communications are based on point-to-point communication in which two end-nodes are involved in the communication network. These two nodes are: the Base Station (BS) or eNodeB (eNB) and Mobile Station (MS) or User Equipment (UE) in cellular networks, access point and laptop in Wireless Local Area Networks (WLANs), or two MSs in device-to-device communications. One of the most severe impairments to wireless communications is channel fading. Fading results in a significant loss in transmission power. Substantial research has been done to reduce the influence of fading. Despite the recent progress in wireless communication techniques, the need to improve wireless capacity and coverage remains a challenge. Relay-based cooperative communication has attracted great attention due to its ability to improve the spectral efficiency, extend the coverage area, mitigate channel impairments and save deployment costs. As such, next-generation wireless technologies

such as LTE-Advanced have adopted relays as one of their key features.

1.2.1 Wireless Relay Networks

Next-generation wireless networks are designed to provide high data rate services and wide coverage with high traffic requirements. In such large coverage areas, users across the cell experience different signal qualities. For example, users at the cell edge often suffer from bad channel conditions. In addition, shadowing by various obstacles can degrade the signal quality significantly. Increasing the transmission power often does not improve the signal quality as it also increases the inter-cell interference. Therefore, next-generation cellular networks face significant capacity and coverage challenges due to the rapidly growing demands for mobile broadband and the limited spectrum available. Next-generation networks including fourth-generation (4G) networks and beyond are increasingly migrating from flat to hierarchical architectures in order to address capacity and coverage issues. Most second-generation (2G) and third-generation (3G) cellular networks have relatively flat architectures, consisting mostly of macrocells with higher power cellular base stations. While such large cells offer the benefits of fast deployment and wide area coverage, they do not provide high capacity due to their non-aggressive frequency reuse and inter-cell interference. Reducing the cell size and deploying more macrocell base stations improve the network capacity, but this comes with high operating costs. The use of small cells including relays is a promising cost-effective solution to tackle this problem. A User Equipment (UE) associates with a macrocell base station if the received signal strength from the macrocell base station on the downlink is at least as high as the received signal strength from a small cell base station. On the uplink, the channel gain from a UE to a small cell is typically large. Hence, the UEs can achieve their desired performance with a lower transmit power to the small cell base station, resulting

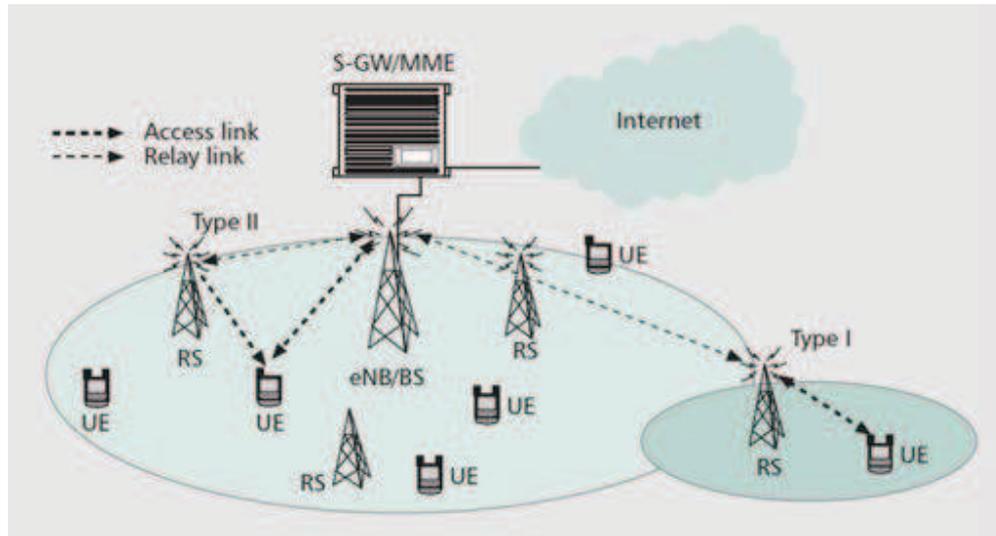


Figure 1.1: Relay Deployment Scenarios in LTE-Advanced Networks [1]

in minimal interference to macrocell. Moreover, offloading users to small cells leads to higher throughput for all users, including those associated with macrocell, since more resources are made available at macrocell and the resources may be used more efficiently. Deploying a set of Relay Nodes (RNs) between base stations or eNodeBs (eNB) and UEs is a cost-effective approach for improving system performance, such as coverage extension, power saving and cell-edge throughput enhancement. In the emerging cellular standards such as 3rd Generation Partnership Project (3GPP) Long Term Evolution (LTE) [4,5], relays have been introduced to provide ubiquitous high-data-rate coverage. Figure 1.1 shows an illustration of relay deployment scenarios in next-generation LTE-Advanced Networks [1].

1.2.2 Cooperative Diversity

When a signal is transmitted through a wireless channel, it undergoes attenuation, fading, distortion, delays, and phase shifts. Due to these factors, system performance can be severely degraded. Diversity is a means employed to diminish the effects of

fading thereby improving the performance of the signal transmission. The idea is to transmit multiple copies of the signal through different channels with independent fading characteristics, thereby reducing the probability of all channels being in deep fade simultaneously. Different types of diversity techniques have been developed [6]. Frequency diversity uses multiple frequencies to transmit the same information. In time diversity, the information is transmitted repeatedly in time at regular intervals to ensure that temporally orthogonal channels exhibit statistically independent fading characteristics. In spatial diversity, multiple transmit antenna (transmit diversity) or multiple receive antenna (receive diversity) are used to transmit and/or receive signals providing multiple independent paths. Such spatial diversity schemes improve the network performance using available channel and power resources.

Traditional wireless networks have been predominantly used for direct transmission in a one-hop fashion. In contrast to direct communications, cooperative communication allows the destination to listen to both the source node and a relay node which overhears and repeats the transmission due to the broadcast nature of the wireless medium. It is a communication paradigm that promises significant capacity and multiplexing gain in wireless networks [7–9]. Since the paths are statistically independent, cooperation results in spatial diversity. The source and relayed information are combined at the receiver for detection.

Recently, a significant amount of research has been done in the field of cooperative diversity. In [7], the concept of user cooperation diversity was introduced. Typically, in a wireless network where a relay or a group of relays are located between the source and destination, the relay can facilitate the transmission by employing different relaying schemes [8, 9]. Cooperation diversity can be divided into two classes: repetition coding and parallel coding. In repetition coding, the relay encodes the source information and uses the same encoder as the source node. On the other hand, in parallel

coding, the relay can encode the source message using a codeword that is generally correlated with, but not necessarily identical to, the source codeword [10].

The combination of the directly transmitted and relayed information offers a form of cooperative diversity, which improves the throughput and power efficiency. Diversity combining is the technique applied to combine the multiple received signals of a diversity reception device into a single improved signal. Diversity combining techniques can be classified into selection combining (SC), maximum-ratio combining (MRC), and equal gain combining (EGC) [11]. In SC, the strongest signal is selected from the multiple received signals. In MRC, each signal branch is multiplied by a weight factor that is proportional to the signal amplitude. That is, branches with strong signal are further amplified, while weak signals are attenuated. The received signals are weighted with respect to their signal-to-noise ratio (SNR) and then summed. In EGC, each signal branch weighted with the same factor, irrespective of the signal amplitude. However, co-phasing of all signals is needed to avoid signal cancellation. Therefore, this is considered a more complex combining technique. The thesis assumes MRC combining technique. However, the diversity combining technique does not limit our proposed hybrid transmission scheme. To fully exploit multi-user diversity, improve spectrum efficiency and utilize the advantages of relay-based cooperative wireless networks, efficient resource allocation is significant. Specifically, the problem formulation may differ remarkably in terms of optimization objectives, relay strategies, channel assignments, transmit power constraints, and system frameworks.

1.2.3 One-Way Relaying

Conventional relaying is typically done in one direction, i.e., one-way relaying, where a source node transmits information through a relay node to a destination node. This communication is usually carried out in two time-slots. In the first time-slot,

the source sends its signals to the relay which is then processed using a relaying scheme. A modified signal is produced and sent to the destination in the second time-slot. Different transmission schemes can be applied to the signal received by the relay before retransmitting [7–9, 12]. In the Amplify-and-Forward (AF) scheme, the relay node amplifies the signal received from the source before retransmitting to the destination. In the Decode-and-Forward (DF) scheme, the relay decodes the signal received from the source before forwarding it to the destination. The destination has to know the characteristics of the relay channel for best detection. DF is usually used when noise at the relay is high and amplifying the signals will amplify the noise as well. However, the drawback of DF is the increase in processing power consumption and the design complexity of the relay. In the Compress-and-Forward (CF) scheme, the relays quantize and compress the information from the source and then send it to the destination. At the destination, the transmitted signal is decoded to obtain the original signal.

When two source nodes exchange information through one or multiple relay nodes, the network created is referred to as a two-way relay network. It is a cooperation method proposed to reduce the spectral efficiency loss that arises in half-duplex one-way relay systems.

1.2.4 Two-Way Relaying

The two-way communication channel was first introduced for point-to-point communication [13]. The use of relays for two-way communication is known as bidirectional relaying or Two-Way Relaying (TWR). In TWR, two source nodes exchange information through one or more relay nodes. The spectral efficiency loss in one-way relay networks is combatted.

In [14], the relaying techniques used for two-way relaying are described. The traditional two-way relaying technique uses four time-slots to complete the transmission and reception of signals between two end-sources. Each source sends its information to the relay one at a time using one time-slot each. The relay then sends the received processed information to both end-nodes using two time-slots. Since the exchange of information requires four time-slots, this technique is not efficient.

Another method discussed in [14] uses three time-slots. Both end-sources transmit their signals to the relay one at a time in the first two time-slots. During the third time-slot, the relay forwards the exclusive OR (XOR) of the decoded signal to both end-terminals. The end-terminals obtain the signal they want by executing an XOR on its transmitted symbol and its received signal. As an alternative, the received signal at relay could be demodulated and decoded to obtain the transmitted data. This data is then re-encoded and re-modulated before retransmitting to the end-nodes. This process is known as Digital Network Coding (DNC) or MAC Layer Network Coding (MLNC) [15,16].

A third method presented in [14] is the two time-slot scheme, where both sources send their information to the relay in the first time-slot simultaneously and then the relay retransmits the processed signal in the second time-slot. The use of two time-slots for bidirectional data exchange improves the spectral efficiency incurred in one-way relay networks. This process is known as Analog Network Coding (ANC) or Physical Layer Network Coding (PLNC). The PLNC scheme improves the throughput by taking one less time-slot as compared to the MLNC scheme and hence achieves larger rate regions. However, this rate improvement comes at the cost of coding and transceiver complexity since it requires strict synchronization requirements [17,18]. MLNC is only considered in this thesis since it is more practical in today's wireless networks. Figure 1.2 illustrates examples of traditional relaying, MLNC and PLNC

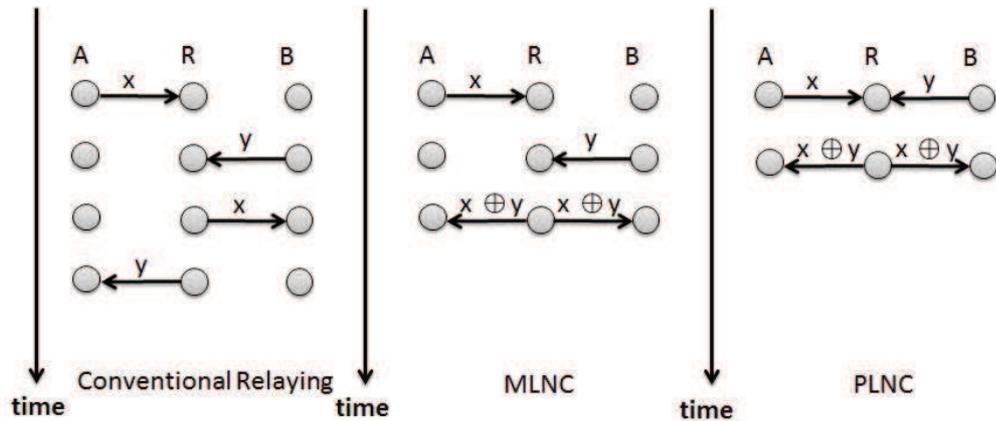


Figure 1.2: Conventional Relaying vs. Network Coding

transmission schemes in a bidirectional communication system.

In the context of cooperative diversity, network coding can also be employed at the relay station to combine packets from different sessions provided that the relay overhears and successfully decodes multiple transmissions and that these transmissions share a common destination [19].

1.3 Motivation

The growing increase in resource-demanding wireless applications requires an efficient management and allocation of the scarce radio resources. Mobile broadband traffic is typically asymmetric as it is dominated by downlink traffic. Therefore, most research on resource allocation in wireless networks studies uplink and downlink problems separately. However, with the emergence of mobile services, such as video conferencing, multi-player mobile gaming, collaborative downloading, and peer-to-peer file sharing, there is a trend towards symmetric traffic. Hence, it is important to investigate joint uplink-downlink resource allocation in next-generation wireless networks. The need for joint uplink-downlink scheduling becomes more important as the evolution

of next-generation wireless communication paved the way for the advent of two-way bidirectional communication.

In recent years, the concepts of cooperative diversity and network coding in two-way relaying have attracted significant attention. Network coding has been shown to achieve substantial throughput gains in wireless networks. However, it is unclear how channel realizations and traffic patterns influence the achievable rates and whether network coding is always better than conventional direct communications or cooperative diversity in terms of end-to-end throughput. Characterizing the achievable rate regions of such transmission schemes will disclose the performance difference between them. It will be observed that none of the supported transmission schemes is superior to the others, i.e., no transmission scheme rate region is a superset of the other. This motivates us to provide a hybrid transmission scheme that contains the benefits of all the transmission schemes (e.g., the combined rate region).

It is commonly assumed that the same channel resource is used for the different phases of the cooperative communication. However, due to independent channel fading over the different time-slots, the system performance may not be optimal. A unique feature of OFDM-based relaying is the utilization of frequency diversity by channel pairing. Channel pairing matches the incoming and outgoing channels at the relay based on channel dynamics. Therefore, it provides significant system performance improvement. However, in multi-user environments, network channels should be carefully allocated and paired. The use of multiple relays further complicates the problem because relay selection becomes tightly coupled with channel assignment. Thus, resource allocation in multi-user multi-relay multichannel networks remains a highly challenging problem. This motivates us to present low-complex channel pairing and resource allocation algorithms for multi-user multi-relay environments in two-way relay networks. Our resource allocation algorithms are modeled such that

they fit practical time-slotted wireless networks.

1.4 Objective

In this thesis, we consider the DF half-duplex two-way relay network and propose novel cross-layer resource allocation schemes for relay-assisted bidirectional multi-channel networks. We define the problem of joint channel assignment/pairing, relay selection and bidirectional transmission scheme selection as a combinatorial optimization problem and provide optimal and suboptimal solutions.

1.5 Contributions and Organization of the Thesis

This thesis investigates resource allocation in multichannel relay-based wireless networks. The thesis problem is defined and optimization frameworks are developed to solve the resource allocation problem.

In Chapter 3, we propose a cross-layer design for optimal joint resource allocation and power control in bidirectional relaying networks. The work assumes that each bidirectional link has a dedicated channel in the multi-user multichannel system in which channel-pairing is not considered. The bidirectional achievable rate regions for direct transmission, Cooperative Relaying (CoR) and Network Coding with CoR (NC/CoR) are characterized and it is shown analytically that NC/CoR does not always perform better than CoR and direct transmission. We observe that each transmission scheme can outperform the others depending on certain conditions in the channel and topology state of the system as well as the uplink and downlink offered traffic rates. This motivates us to propose a hybrid transmission scheme which dynamically allocates communication resources to achieve maximum rate regions. In

our hybrid scheme, the different transmission schemes are selected dynamically over different time-frames and optimal time allocation is proposed based on the channel state and the traffic backlog at the UE and eNB. The channel information is assumed to be known at eNB, which dynamically allocates network resources (subchannels). The simulation results show that the proposed hybrid scheme outperforms the conventional schemes in terms of delay and throughput. The hybrid scheme is then extended by adding power control at the relay station. The simulation results show that joint power control and resource allocation significantly enhances the throughput and delay performance.

In Chapter 4, an OFDMA-based bidirectional cellular network is considered. An enhanced three-time-slot per cycle Time-Division Duplexing (TDD) frame architecture is proposed, in which a dedicated Relay Node (RN) subframe is introduced. The first two subframes in each cycle are used for transmission by eNB only (downlink) and UE only (uplink). However, the RN subframe is used for both uplink (to eNB) and downlink (to UE). Using the enhanced TDD frame architecture, an optimization framework is proposed for the joint resource allocation, subchannel pairing, transmission scheme selection and relay selection problem. The proposed optimization framework is a combinatorial optimization problem maximizing the throughput stability region of the system and stabilizing the queues. First, we adopt a graph-based approach by establishing an equivalent Maximum Weighted Clique Problem (MWCP). A modified Bron-Kerbosch algorithm is proposed to provide an optimal solution. It is used as a benchmark to evaluate our suboptimal algorithms for small size problems. In addition, a sequential greedy heuristic algorithm is used to provide a fast suboptimal solution to the MWCP problem. Second, the joint combinatorial optimization problem is also mapped into a 3-dimensional assignment problem. An enhanced 3-dimensional Ant Colony Optimization (ACO) algorithm is proposed to

provide a suboptimal solution for the combinatorial assignment problem in polynomial time. The simulation results show that the proposed hybrid Cooperative Relaying (CoR) and Network Coding (NC) scheme outperforms the conventional non-hybrid schemes. Moreover, we observe that the proposed ACO-based algorithm provides better approximation results than the greedy heuristic algorithm in real environments.

The remainder of the thesis is structured as follows. In Chapter 2, we present the related work to our research. In Chapter 3, we introduce our optimal joint resource allocation and power control in bidirectional wireless relaying networks. This chapter assumes that each bidirectional link has a dedicated channel in the multichannel system. In Chapter 4, an optimization framework is proposed for OFDMA-based systems to jointly optimize subchannel allocation/pairing, transmission scheme selection and relay selection. Finally, we summarize the thesis and discuss future directions in Chapter 5.

1.6 List of Publications and Patent Filings

1. Ahmed Zainaldin, Hassan Halabian, and Ioannis Lambadaris, “Joint Resource Allocation and Relay Selection in LTE-Advanced Network using Hybrid Cooperative Relaying and Network Coding,” *Submitted to IEEE Transactions on Wireless Communications (TWC)*, Status: minor revision, Jan. 2015.
2. Ahmed Zainaldin, Hassan Halabian, and Ioannis Lambadaris, “Ant Colony Optimization for Joint Resource Allocation and Relay Selection in LTE-Advanced Networks,” *in Proc. IEEE GLOBECOM14*, Austin, TX, USA, Dec. 2014.
3. Hassan Halabian, Ahmed Zainaldin, and Ioannis Lambadaris, “Optimal Joint Resource Allocation and Power Control in Bidirectional Relaying Networks,” *IEEE Transactions on Vehicular Technology*, vol.63, pp.4520–4535, Nov. 2014.

4. Ahmed Zainaldin, Hassan Halabian, and Ioannis Lambadaris, “Hybrid Transmission Scheme and Relay Selection in LTE-Advanced network: Graph-based approach,” *in Proc. IEEE Next Generation Networks and Services (NGNS’14)*, Casablanca, Morocco, May 2014.
5. Ahmed Zainaldin, Hassan Halabian, and Ioannis Lambadaris, “Optimal Resource Allocation in LTE-Advanced Network Using Hybrid Cooperative Relaying and Network Coding,” *in Proc. IEEE International Conference on Communications (ICC’13)*, Budapest, Hungary, June 2013.
6. Ahmed Zainaldin, Hassan Halabian, and Ioannis Lambadaris, “Dynamic Resource Allocation using Hybrid Cooperative Relaying and Network Coding in LTE-Advanced Network,” *in Proc. IEEE Next Generation Networks and Services (NGNS’12)*, Algarve, Portugal, Dec. 2012.
7. Ahmed Zainaldin, “Network Coding in a Cellular Communications System.” U.S. Patent 20,130,229,968, issued September 5, 2013.
8. Ahmed Zainaldin, “Cooperative Relaying and Network Coding in a Cellular Communications System.” U.S. Patent 8,913,543, issued December 16, 2014.
9. Ahmed Zainaldin, “Selection of a Transmission Mode for Wireless Communications with a Mobile Device.” U.S. Patent 20,140,226,547, issued August 14, 2014.
10. Ahmed Zainaldin, Hassan Halabian, and Ioannis Lambadaris, “Downlink Resource Allocation in OFDM Networks.” U.S. Patent 14,266,257, filed April 30, 2014.

Chapter 2

Background and Related Work

In this chapter, we review relative background and recent research works in the field of bidirectional (Two-Way) wireless relay communication. The chapter reviews the relative research work on resource allocation for bidirectional multichannel relay-assisted networks. In addition, background on the optimization frameworks used in our work will be shown. An overview of the Ant Colony Optimization (ACO) Algorithm is presented to facilitate the understanding of research results to be presented in the following chapters.

2.1 Two-Way Multichannel Relaying

2.1.1 Opportunistic Scheduling in Two-Way Relaying

Cooperative communication has received significant attention recently as an emerging communication concept for wireless networks [7–9, 20, 21]. Due to the broadcast nature of wireless transmissions, cooperative communications enables neighboring network nodes to share resources and cooperate to send information to an intended node. Distributed transmissions from source and relay nodes provide spatial diversity and multiplexing gain for information detection at a destination node. Cooperative

Relaying (CoR) can significantly improve system performance and robustness of wireless networks especially in severe fading environments. In two-way CoR, information exchange between two users via a relay is accomplished in four time slots.

Network Coding (NC) has been employed in Two-Way Relay Channels (TWRCs) in order to improve the spectral efficiency of the relaying network [15, 16, 22–36]. In [25, 26, 30, 31, 33–35], the authors studied a two-time-slot TWRC protocol with an Amplify-and-Forward (AF) relay. The protocol is called Analog NC (ANC) or Physical Layer NC (PLNC), where the knowledge of the first phase allows the receiving nodes to perform interference cancellation before decoding. The outage probability, rate region and diversity-multiplexing tradeoff have been analyzed. However, this requires very sensitive timers, which make PLNC difficult to adopt in today’s networks. Another approach is based on the MAC Layer NC (MLNC), also called Time Division Broadcasting (TDBC) protocol, where the relay performs an XOR operation on the decoded bit streams [15, 22–24, 26–28, 33, 36]. MLNC is a more practical protocol since it operates on the decoded data and does not require synchronization between the two source nodes. In [31, 32], the closed-form expression of outage probability and achievable rate region of MLNC have been obtained for Amplify-and-Forward (AF) and Decode-and-Forward (DF) relaying. In these works, it is assumed that only the relay station decodes the packets and then sends a network coded packet to the destination, i.e., the direct channel between the two source nodes is not considered. It is easy to see that the PLNC and ANC protocols achieve the highest spectral efficiency since two-traffic flows are concurrently supported at each time-slot. However, the direct link between the two end-sources cannot be utilized even if such direct link physically exists. Hence, these protocols can only achieve diversity order one. On the other hand, the direct link can be utilized in MLNC since the two end-sources transmit in two different time-slots. Therefore, MLNC can achieve diversity order

two. In [33, 36], the outage probability and rate region have been obtained when the direct link is considered.

NC can usually achieve higher achievable rate regions than CoR due to the fewer number of time-slots required in the bidirectional transmission. However, the disadvantage of NC is that asymmetric traffic ratio conditions might cause significant decrease in the bidirectional throughput. In addition, transmission power allocations are different in NC as one transmission often needs to reach multiple destinations. In order for a packet to reach all the destinations, the transmission power of the relay node should be based on the worst link condition. This can increase the interference level and reduce network performance [31]. Thus, the advantage of using network coding can be reduced significantly in a fading channel. This happens because the worst link gain can be very small, which either requires very high transmission power in order to keep a reasonably high transmission rate, or results in very low transmission rate for the same transmission power. A large body of research on opportunistic scheduling has been proposed to tackle this problem [31, 34, 37–39]. Opportunistic schedulers take into account network information such as channel quality and traffic flow to dynamically select the best transmission scheme. The hybrid scheme achieves a larger rate region than those achieved independently.

In addition, Optimal Time Allocation (OTA) for one-way and two-way relay networks has been studied in the literature for CoR and NC [20, 21, 29–31, 39]. The analysis starts by characterizing the achievable rate regions and then finding the optimal resource fraction allocated to each communication channel. In [20, 21, 39], OTA is proposed for the CoR one-way relaying using two-time-slots to minimize the outage probability. In [29, 30], OTA is proposed for the two-phase PLNC protocol. The resource allocation is simpler for such problems due to the use of two-time-slots.

In [31,32], OTA is proposed for the MLNC scheme. In [31], OTA is proposed for a traditional four-time-slot TWRC, MLNC and PLNC. A hybrid scheme is then proposed to select the optimal transmission mode for each time-frame. However, the direct channel between the two end-sources was not considered, which adds complexity into the problem.

To the best of our knowledge, the optimal time allocation for the combined NC and CoR schemes (NC/CoR) has not been proposed in the literature. This thesis builds on the results in [26,31,40] by characterizing an explicit form of the rate region of the combined NC and CoR schemes (NC/CoR) and comparing it with that of CoR and direct transmission schemes. It is shown that each transmission scheme (direct transmission, CoR and NC/CoR) performs differently according to certain conditions. As such, a new hybrid transmission scheme is proposed. Chapter 3 proposes an optimal time and relay power allocation scheme for our hybrid transmission scheme. The optimal channel resource fractions for each time-slot are determined analytically at each time-frame in order to maximize the achievable hybrid rate region.

2.1.2 Cross-layer Resource Allocation and Stability Analysis in Two-Way Relaying

Cross-layer optimization is a topic of great interest in wireless networks. In two-way relaying networks, cross-layer approaches, in which physical layer decisions are made jointly with higher layers, have recently been attracting considerable attention [25, 27, 28, 31–34, 41]. One essential issue is to consider the stochastic nature of the traffic to be communicated and to determine the allocated power and achievable rates for nodes according to the queue and channel states for the supported transmission schemes. A fundamental performance measure of a communication network

is the stability of its queues. The works in [34, 41] characterized the queue stability region for infinite backlogs with the XOR and super-position coding based DF protocols, respectively. The optimal time division between the uplink and the downlink phases was investigated. In addition, a power allocation strategy in the downlink to optimize throughput with current queue length at each node is provided. The allocation problem requires that the central controller knows instantaneous channel state and queue length information. In [31], the stability of the hybrid transmission scheduling under Poisson arrivals is shown via queuing analysis. These works aim at maximizing the bidirectional sum rate. In [32], a cross-layer design is proposed for MLNC to minimize total energy usage in the bidirectional transmission while ensuring queue stability. The resource allocation problem is solved for both static and fading channels. In [35], the maximum throughput region for relay-assisted wireless broadcast with digital and analog NC is derived. For stochastic packet traffic arrivals, variations of the maximum differential backlog [42] policy is shown to be throughput optimal. In [33], the cross-layer optimization of two-way relaying under statistical delay-QoS constraints is studied. The work assumes both MLNC and PLNC transmission protocols. By integrating the theory of effective capacity, the optimization problem for weighted sum throughput maximization in physical layer and delay provisioning in data-link layer was modeled as a long-term weighted sum effective capacity maximization problem. In [27, 28], a suboptimal delay-constrained buffer-aided relaying protocol is proposed which can guarantee a bounded delay and a suboptimal distributed protocol, in which each node requires only the local instantaneous channel state information of the links to determine the best transmission scheme.

Chapters 3 and 4 propose a cross-layer design of the Physical Layer and the Data-Link Layer, which includes the Medium Access Sublayer. We are interested in an efficient power control and channel allocation strategy with respect to the higher

layer traffic. Accordingly, we are concerned with an adaptive resource allocation policy which takes the exogenous stochastic arrivals at the source nodes from the higher layer into account. To this end, we look at time-slotted queueing processes with ergodic arrival processes where the information theoretic bidirectional achievable rates serve as service rates. At the beginning of each time-frame, a centralized controller with queue and channel state information adjusts the service rates so that stability at all queues is achieved. The throughput optimal rate allocation policy presented in this work is an adaptation from the maximum differential backlog (back-pressure) algorithm introduced in [42]. This work concentrates on the characterization of the throughput optimal rate allocation policy, the bidirectional achievable rate region, and the resulting stability region. To the best of our knowledge, this is the first work on the throughput optimal stability region for our proposed hybrid transmission scheme.

2.2 Resource Allocation in Two-Way Relaying OFDMA Networks

Orthogonal Frequency Division Multiple Access (OFDMA) is an efficient technique used to mitigate the problems of frequency selective fading and inter-symbol interference. In OFDMA networks, a total bandwidth is divided into a number of subchannels or subcarriers and multiple users transmit their information simultaneously on the different subchannels. Therefore, OFDMA has been accepted as the most appropriate air interface for next generation wireless networks, such as LTE and LTE-Advanced, due to its inherent ability to combat frequency-selective multipath fading and higher spectral efficiency [43]. An overview of the Long Term Evolution (LTE) standard is presented in Appendix A.

There exists a great body of research on the design of Network Coding (NC) and

Cooperative Relaying (CoR) in wireless OFDMA networks [26, 44]. Resource allocation optimization has been extensively studied in the literature [45–48]. However, there are fewer studies which analyze the optimization of bidirectional OFDMA-based networks using cooperative relaying and network coding [37, 49]. The majority of previous research used the Lagrange dual decomposition method to solve the joint optimization problem by decomposing the subchannel allocation problem into per-subchannel subproblems that can be solved independently [49]. The work in [37] proposed an XOR based cooperative diversity scheme and presented an algorithm for joint optimization of relay selection, transmission mode selection, power allocation and subchannel allocation. In [50], the authors proposed a hierarchical protocol for one-way and two-way relaying in a two time-slot Time-Division Duplexing (TDD) mode. In this protocol, the transmission mode of each UE as well as its assisting relay node is pre-fixed, and the downlink and uplink transmission modes for each UE are the same. Then, joint power and subcarrier allocation is considered and solved by Lagrange dual decomposition method. The system studied in [51] operates in Frequency-Division Duplexing (FDD) mode with fixed sets of data subchannels and relay subchannels. None of these studies considered subchannel pairing which provides additional gains to the optimization framework.

2.2.1 Joint Channel Assignment and Pairing in Two-Way OFDMA Relaying

Most research on resource allocation in OFDMA-based relay networks assume that cooperative relaying communication is performed in the same channel [44, 52, 53]. This category of resource allocation schemes enforces that the two hops, i.e., the source-to-relay and relay-to-destination, to use the same channel. This strategy significantly

simplifies the optimization problems but does not fully exploit the channel dynamics. Due to independent channel fading on the same channel over the two hops, the system performance may not be optimal. An important property of OFDM-based relaying is that the frequency diversity can be exploited by channel pairing, which matches the incoming and outgoing subchannels at the relay based on channel dynamics and hence, improves system performance. In multi-user environments, subchannels should not only be carefully paired at the relay but also be assigned adaptively for different users. If used with multiple relays, it further complicates the problem because relay selection is tightly coupled with channel pairing. Thus, channel-pairing based resource allocation in multi-user multi-relay OFDMA networks remains a highly-challenging open problem.

Channel-pairing based resource allocation has been originally investigated for single-user single-relay one-way relaying [54–57]. In [54], it is proved that the ordered pairing is optimal for the AF protocol. However, such ordered pairing on per-channel based pairing is not sufficient for the DF protocol, where the information from one set of channels in the first hop can be decoded and re-encoded jointly and then transmitted over a different set of channels in the next hop. Joint optimization of system resources, such as channel pairing, power allocation and channel assignment, has been investigated in [56, 57], where efficient numerical algorithms were devised to solve the complex joint optimization problems.

Channel-pairing was then extended for two-way relaying communication [58, 59]. The joint channel assignment and pairing is more complex for bidirectional relaying since the choices of incoming and outgoing channels between the relay and the two end nodes are tied to each other. Authors in [58] investigated separated power allocation and subchannel pairing in bidirectional relaying using single relay, where the power allocation is first employed by water-filling and then subchannels are paired at the

relay by a heuristic method. Most of the literature work on two-way relaying resource allocation and pairing consider the two-phase PLNC protocol. The optimal channel assignment and pairing turns out to be a standard two-dimensional linear assignment problem which can be solved by the Hungarian method [60]. Under given power allocation, the pairing problem for PLNC two-way relaying is considered in [59, 61, 62], where a numerical optimization algorithm and low-complexity pairing strategies were proposed, respectively. For MLNC-based two-way relaying, the problem of channel pairing is even more complex. Since the transmission from each source node to the relay is performed in two separate time-slots, channel pairing is no longer just between incoming and outgoing channels, but is among the two incoming channels and the outgoing channel. For communication systems with large number of subchannels, designing efficient channel pairing strategies is thus important and challenging. There is limited research addressing channel pairing design in a three-phase MLNC two-way relay network. A joint load balancing and proportional fairness based resource allocation has been investigated in two-way relaying cooperative networks with both ANC and DNC protocols in [52]. The two-way communication is assumed to happen in one channel, i.e., no channel pairing is considered. In [63], the problem of subchannel pairing at the relay for three-phase MLNC-based two-way AF relaying is investigated. The work proposed a low complexity sorting-based algorithm based on SNR over each subchannel. In [64], the authors studied relay-assisted bidirectional DF OFDMA cellular networks, wherein the subchannel-pairing based joint optimization of bidirectional transmission mode selection, relay selection, and subchannel assignment were investigated by a graph approach. The work did not assume the direct link between the source nodes in the MLNC protocol.

Chapter 3 builds on the results in [26] by characterizing the rate region of the combined direct, CoR and NC/CoR hybrid transmission schemes in an explicit form. Each

bidirectional link has a dedicated channel in the multichannel system, i.e., channel-pairing is not considered. In Chapter 4, a joint optimization of transmission scheme selection, relay selection, and subchannel pairing and assignment in DF two-way relay OFDMA networks is studied. This joint problem is mapped first into a Maximum Weighted Clique Problem (MWCP) and then optimal and suboptimal solutions are proposed. In addition, the joint problem is mapped into a three-dimensional problem in which a metaheuristic algorithm based on Ant Colony Optimization (ACO) is employed to find a suboptimal solution in polynomial time.

2.2.2 Optimization vs. Heuristic Approaches

The optimal solution for the problem of resource allocation for a multi-user OFDMA system is generally difficult to solve. Exact solution methods, such as branch and bound, are limited to solving discrete combinatorial problems of small size. Optimization-based resource allocation approaches are computationally intensive, especially when the number of users and subchannels are large. Therefore, it is difficult to compute the optimal solution in real-time. Hence, for large-scale applications, more research efforts have been devoted to developing heuristic and metaheuristic suboptimal solution methods, including linear programming, genetic, simulated annealing and tabu search algorithms. The optimization algorithms used to provide suboptimal solutions to the resource allocation problem of OFDMA systems can be classified into deterministic algorithms [65, 66] and bio-inspired algorithms which are part of stochastic algorithms [67, 68]. Linear Programming (LP) [69] and Lagrangian Relaxation (LR) [37, 49] are deterministic algorithms which have been used extensively in the design of approximation algorithms for a variety of combinatorial problems. LR is typically used to relax a set of complex side constraints and accordingly compensate a penalty term in the objective function. LR method provides a fast solution

which is usually updated by a sub-gradient method [66], although it may converge slowly. Consequently, the quality of the final solution is not guaranteed. A common heuristic approach is to use randomized methods such as Genetic Algorithms (GA). Such stochastic algorithms have been developed by inspiration from nature and have demonstrated superior performance both in computational time and accuracy [70]. GA methods are very popular for resource allocation problems but yet to be fully exploited in the context of OFDMA networks [67, 68]. Ant Colony Optimization (ACO), proposed in [71] to find suboptimal solutions for NP-hard problems, is one of the most well-known GA techniques. ACO has also been widely used in OFDMA cellular networks to solve the resource allocation problem [64, 72].

An enhanced form of ACO is proposed in Chapter 4 to solve our three-dimensional resource allocation optimization problem. The resource allocation problem considered is a combinatorial optimization problem which is difficult to solve in polynomial time. Metaheuristic methods are proposed to provide a suboptimal solution to such problems. A conventional approach to solve mixed-integer combinatorial problems is to use a branch and bound algorithm which finds an optimal solution. However, such algorithms are computationally inefficient for OFDMA resource allocation since the algorithm is called at each time-frame. Bron-Kerbosch [73] is a well-known branch and bound algorithm. As this algorithm provides an optimal solution, we will use it as a benchmark to evaluate our suboptimal heuristic algorithms.

2.3 Ant Colony Optimization

Swarm Intelligence is a problem-solving approach inspired by the social behavior of insects and other animals. In particular, ants have inspired the development of a general purpose optimization technique known as Ant Colony Optimization (ACO).

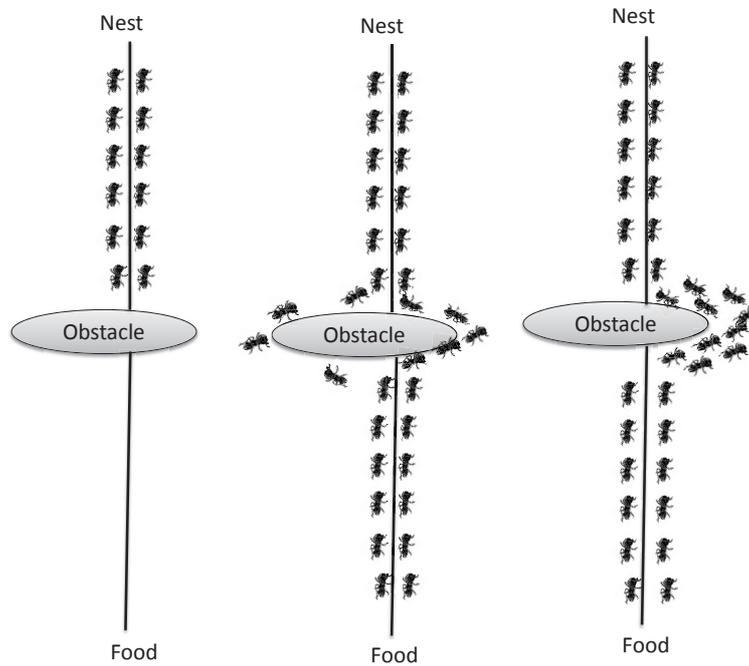


Figure 2.1: Ants Facing an Obstacle

ACO is an algorithm for finding optimal paths based on the behaviour of ants searching for food sources using the shortest path [74]. During their trips back, a chemical trail (pheromone) is left on the ground to guide the other ants towards the food source. The path is chosen according to the quantity of pheromones. Furthermore, pheromones have a decreasing action over time, and the quantity left by one ant depends on the amount of food found and the number of ants using this trail. As illustrated in Figure 2.1, when ants face an obstacle, there is an equal probability for each ant to choose the left or right path. As the right trail is shorter than the left one and thus requires less travel time, it will end up with higher level of pheromone. The more the ants take the right path, the higher the pheromone trail is.

This principle of communicating ants has been used as a framework for solving combinatorial optimization problems. Algorithm 1 presents the generic ACO algorithm. The first step consists mainly of the initialization of the pheromone trail. In

Algorithm 1: Generic ACO Algorithm

Step 1: Initialization

- Initialize the pheromone trail

Step 2: Iteration**for each Ant m do**

- Solution construction using pheromone trail
 - Update the pheromone trail
- Continue until stopping criterion
-

the iteration step, each ant constructs a complete solution to the problem according to a probabilistic state transition rule. The state transition rule depends mainly on the state of the pheromone. Once all ants generate a solution, a global pheromone updating rule is applied in two phases; an evaporation phase where a fraction of the pheromone evaporates, and an intensification (reinforcement) phase where each ant deposits an amount of pheromone. This process is iterated until a stopping criterion is met.

2.3.1 ACO Algorithms

Several ACO algorithms have been proposed in the literature. The first published algorithm is known as the Ant System (AS) [75]. The AS algorithm was first applied to the Traveling Salesman Problem (TSP). In the TSP solution, an ant is placed on each city as it traverses from its current city, visiting other cities only once and returning to its origin when a tour is completed. The ants deposit pheromone on the edges connecting each city during their journey. The pheromone concentration on the edges is constantly adjusted so that pheromones on unused edges eventually evaporate completely. The AS algorithm produced encouraging results, yet its performance was not competitive with state-of-the-art algorithms for the TSP problem.

Therefore, one important focus of research on ACO algorithms has been the introduction of algorithmic improvements to achieve better performances for the TSP and other combinatorial problems.

The ACO algorithm became a widely used method for solving the Quadratic Assignment Problem (QAP). In [76, 77], the ant algorithm was shown to provide an efficient suboptimal solution to the QAP problem. The MAX-MIN Ant System (MMAS) algorithm proposed in [76] allows only the best solution to add pheromone trail during the pheromone trail update. A bound is used for trail levels to avoid premature convergence of the search. In [77], Gambardella proposed a Hybrid Ant System-Quadratic Assignment Problem (HAS-QAP) to the QAP problem. The originality of this approach is that the pheromone trail was not used to construct solutions but to modify them in the local search. An enhanced ACO algorithm is proposed in Chapter 4 to solve our resource allocation three-phase assignment problem. The proposed algorithm is a modified version of the HAS-QAP algorithm.

Chapter 3

Optimal Joint Resource Allocation and Power Control in Bidirectional Relaying Networks

In this chapter, the end-to-end achievable rate regions for direct transmission, Cooperative Relaying (CoR) and Network Coding with CoR (NC/CoR) are characterized. A new hybrid transmission scheme is proposed. The hybrid scheme dynamically allocates communication resources to achieve the maximum achievable rate regions. The hybrid scheme is then extended by adding power control at the relay station (RS), also called relay node (RN).

3.1 System Description and Assumptions

In bidirectional relaying networks, the bidirectional channel between the end-point entities can be modeled by a Two-Way Relay Channel (TWRC). A TWRC consists of one UE, one RS and one eNB in which UE and eNB transmit their information to each other by the cooperation of RS. In this Chapter, we consider a TWRC model in which the RS is capable of performing MAC layer network coding and UE and

eNB are capable of performing cooperative diversity detection. More specifically, we assume that the system is capable of using Cooperative Relaying (CoR) and Network Coding/Cooperative Relaying (NC/CoR) which are introduced in detail below.

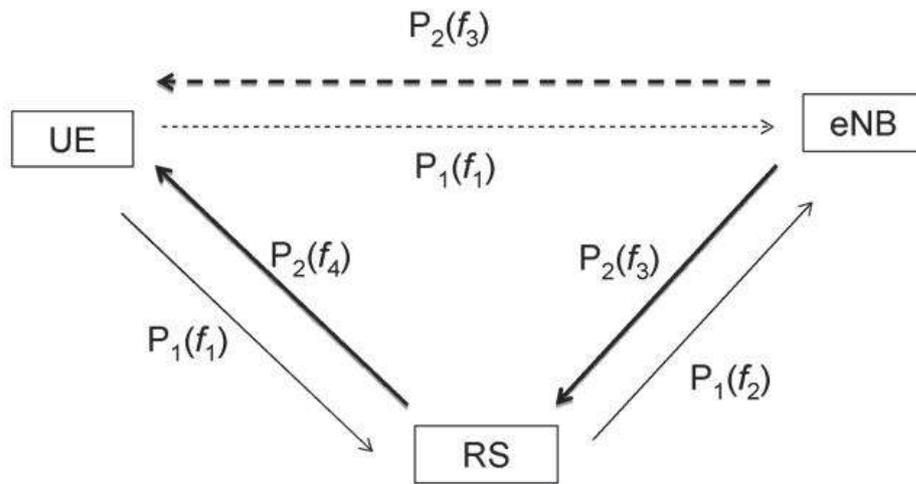
Cooperative Relaying (CoR): In CoR as shown in Figure 3.1(a), UE and eNB send packets P_1 and P_2 to each other respectively via RS using four subchannels (slots)¹. P_1 is hopped twice from UE to eNB (uplink) through the relay at subchannels f_1 and f_2 . Similarly, P_2 is hopped twice from eNB to UE (downlink) through the relay at subchannels f_3 and f_4 . P_1 and P_2 are also overheard by eNB and UE at f_1 and f_3 , respectively. The signals received in f_1 and f_2 (resp. f_3 and f_4) are used by eNB (resp. UE) to cooperatively detect the packet P_1 (resp. P_2) and hence improve the system efficiency and reliability.

Network Coding/Cooperative Relaying (NC/CoR): Although CoR increases the system coverage and robustness due to use of cooperative diversity, it suffers from low spectral efficiency². By employing network coding at RS, we may achieve higher spectral efficiency.

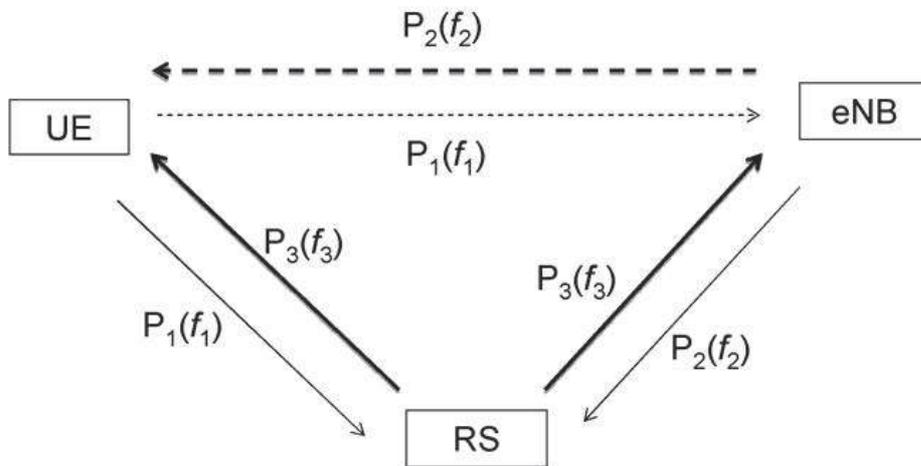
Figure 3.1(b) illustrates an example of a TWRC with the use of network coding. As shown in Figure 3.1(b), UE and eNB can send data to each other via a relay node using just 3 subchannels. P_1 is hopped from UE to RS at subchannel f_1 and P_2 is hopped from eNB to RS at subchannel f_2 . Then, packets P_1 and P_2 are network coded (bit-wise XOR) as $P_3 = P_1 \oplus P_2$ and broadcast as one packet to both UE and eNB. P_1 and P_2 are also overheard by eNB and UE at f_1 and f_2 , respectively. Therefore, eNB will have two versions of P_1 (one from overhearing and the second by XORing the network coded P_3 with its original packet P_1). Similarly, UE will have two versions of P_2 . Cooperative detection can then be applied on the two copies of

¹Time Division Duplexing (TDD) is used to separate the uplink and downlink transmissions.

² It requires four subchannels f_1 to f_4 in each bidirectional transmission time-frame.



(a) Cooperative Relaying (CoR) scheme



(b) Network Coding with Cooperative Relaying (NC/CoR) scheme

Figure 3.1: Two-Way Relaying Schemes

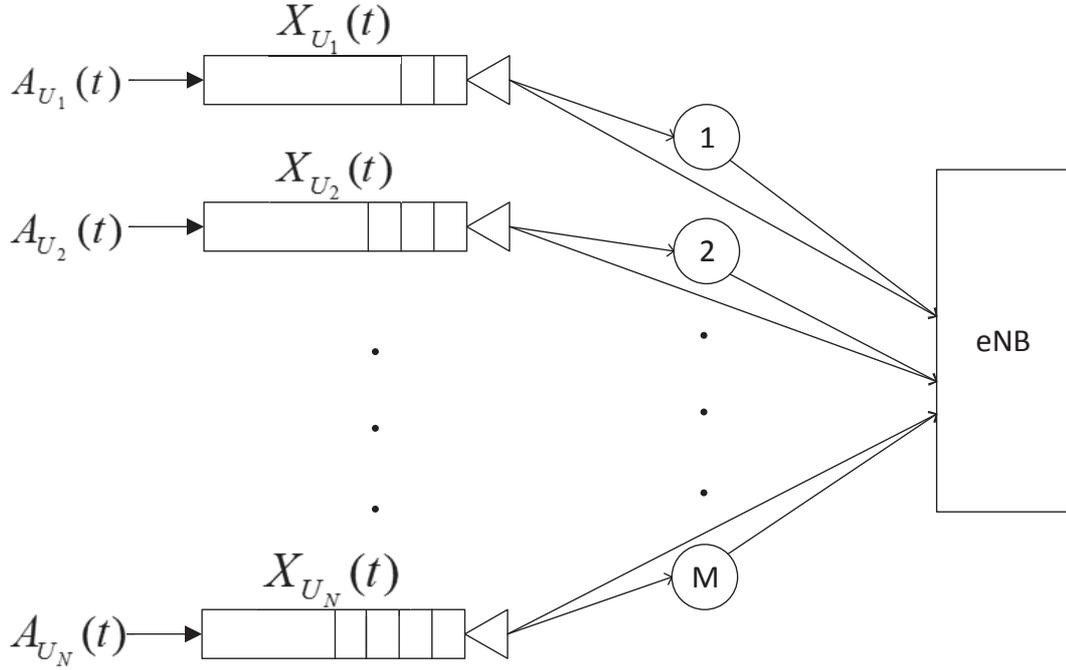


Figure 3.2: Uplink Model

each packet to detect it.

The length of subchannels f_i affects the performance of the system in terms of achievable throughput. In Section 3.2, we will derive the optimum values for the duration of subchannels f_i that lead into the maximum achievable rate region for CoR and NC/CoR schemes.

The bidirectional cooperative model considered in this chapter consists of the set of UEs $\mathcal{N} = \{1, 2, \dots, N\}$, the set of relay nodes $\mathcal{M} = \{1, 2, \dots, M\}$ and the eNB. Assume that there are N orthogonal channels in the system and each UE n is tuned to a separate channel C_n . The eNB node is capable of transmitting/receiving information on different channels, simultaneously. The system is assumed to be time slotted with equal length intervals. Therefore, all UEs can transmit/receive at each time slot

(time-frame) using their own frequency channels. Within each UE-eNB bidirectional communication, in each bidirectional scheme, the channel is divided into subchannels separated into time-fractions. For instance, the channels in direct transmission, CoR and NC/CoR schemes are divided into two, four and three subchannels, respectively separated in the time domain.

Figure 3.2 illustrates the uplink model used in this Chapter. Similar modelling is followed by the downlink transmission in which there are N queues at the eNB corresponding to the different active UEs in the network. Each UE n can transmit directly to eNB or it can use one of the relays to send its information cooperatively using CoR or NC/CoR.

3.1.1 Channel Model

In this thesis, we assume the eNodeB possesses complete knowledge of channel state information (CSI). In information theory, channel (Shannon) capacity is the upper-bound on the rate at which information can be reliably transmitted over of a communication channel. Shannon's Capacity Theorem states that error-free transmission is possible as long as the transmitter does not exceed the channel's capacity.

To briefly describe the notion of channel capacity, consider the additive white Gaussian noise (AWGN) channel is characterized by $y(t) = x(t) + n(t)$, where $x(t)$, $y(t)$ and $n(t)$ are the transmitted (modulated) signal, received signal and receiver noise, respectively. The channel capacity is evaluated by the maximum mutual information over all possible channel realizations and is given by Shannons well-known formula [78].

$$C = \max I(X; Y) = B \log_2(1 + \gamma)$$

where $I(X;Y)$ is the mutual information over the channel $x - y$, B is the channel bandwidth and γ is the average channel signal-to-noise ratio (SNR). Shannons coding theorem proves that there exists a code for x that achieves data rates arbitrarily close to the capacity C with arbitrarily small probability of bit errors for a sufficiently long coding block.

We now introduce the notation that will be frequently used in this chapter. We denote the link capacity from node x to node y at time-frame t by $C_{xy}(t)$ which is defined by the Shannon capacity formula, i.e., $C_{xy}(t) = B \log_2(1 + \gamma_{xy}(t))$, where $C_{xy}(t)$ is measured in bits/s, $\gamma_{xy}(t)$ is the average received SNR at node y at time-frame t and B is the bandwidth of the channel. Note that $\gamma_{xy}(t) = \frac{|h_{xy}(t)|^2 P_x}{\sigma_{xy} d_{xy}^\alpha}$, where P_x is the transmission power of node x , σ_{xy} is the noise power, $h_{xy}(t)$ is the channel gain of link $x \rightarrow y$ which is assumed to be i.i.d. $\mathcal{CN}(0, 1)$ and does not change within each time-frame³, d_{xy} is the distance of nodes x and y and α is the path-loss exponent. If the $x \rightarrow y$ channel is an $M \times N$ MIMO channel (with N transmitting and M receiving antennas), we use the MIMO channel capacity formula [79] $C_{xy}(t) = B \log_2 \left[\det(\mathbf{I}_M + \frac{P_x}{N \sigma_{xy} d_{xy}^\alpha} \mathbf{H} \mathbf{H}^*) \right]$ in which \mathbf{H} is the $M \times N$ channel matrix ($[\mathbf{H}]_{m,n} \sim \mathcal{CN}(0, 1)$) and \mathbf{I}_M denotes the identity matrix of size M . It is assumed that the relays operate in DF mode. Henceforth, we use subscripts U,R and E to denote UE, RS and eNB, respectively.

The channel state of the TWRC at time-frame t is denoted by vector $C(t) = (C_{UE}(t), C_{EU}(t), C_{UR}(t), C_{RU}(t), C_{ER}(t), C_{RE}(t))$ which contains the capacities of all the communication links in the TWRC. We will assume that the channel state of the system is a stationary process with stationary probabilities $P(c)$, $\forall c \in \mathcal{C}$ where \mathcal{C} is the system channel state space.

³This is a common assumption when modeling a wireless channel as a slow fading channel.

3.1.2 Queuing Model

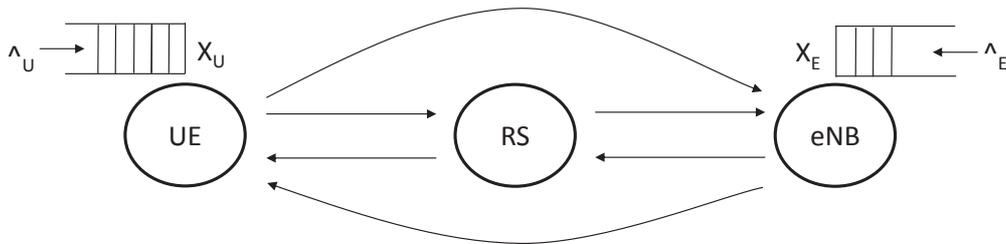


Figure 3.3: Two-Way Relay Queuing Model

We consider a two-way relay network with sources UE and eNB, and relay RS, as shown in Figure 3.3. We assume a slotted system, in which each UE and eNB buffer their incoming packets in queues $X_U(t)$ and $X_E(t)$ at time-frame t , respectively. We consider a queueing system with N queues at the eNB corresponding to the different active UEs in the network. On the other hand, one queue is available per UE for uplink transmission. Sources UE and eNB choose the service rates at time-frame t such that $(R_{UE}(t), R_{EU}(t)) \in R(t)$, where $R(t)$ is the bidirectional achievable rate region at time-frame t . We assume the two sources communicate with the relay over orthogonal interference-free channels, i.e. separate time-slots (TDD).

Let $\{A_U(t)\}_{t=1}^{\infty}$ and $\{A_E(t)\}_{t=1}^{\infty}$ denote the packet arrival processes to UE and eNB, respectively. More specifically, $A_U(t)$ and $A_E(t)$ are the random Poisson variables of the number of packets arrived exogenously to UE and eNB at time-frame t . We assume that $\{A_U(t)\}_{t=1}^{\infty}$ and $\{A_E(t)\}_{t=1}^{\infty}$ are stationary processes with rates Λ_U and Λ_E , respectively. UE and eNB keep the arriving packets in their queues (with infinite buffering). We use $X_U(t)$ and $X_E(t)$ to denote the queue lengths of UE and eNB at

the end of time-frame t . The queue lengths at sources UE and eNB evolves as

$$\begin{aligned} X_U(t) &= (X_U(t-1) - R_{UE}(t))^+ + A_U(t) \\ X_E(t) &= (X_E(t-1) - R_{EU}(t))^+ + A_E(t) \end{aligned}$$

where $(\cdot)^+$ outputs the term inside the brackets if it is non-negative and zero otherwise since the number of transmitted packets is limited by the number of packets available in the queue.

We assume that relay RS does not buffer the incoming packets and immediately forwards any received data over a channel that is orthogonal to the channels used by each source. The RS decodes and re-encodes (DF) the packets incoming from sources before the subsequent transmission in the next time-frame. Our work would still be applicable for hop-by-hop forwarding in which the RS is able to buffer packets for future retransmissions. It was shown in [41] that hop-by-hop forwarding and immediate forwarding result in identical stability regions for the two-way single relay network. To achieve this performance, immediate forwarding involves time-slot optimization and more complex scheduling of two hops jointly, whereas hop-by-hop scheduling necessitates buffers at the relays.

The channel state (SNR) of all the links are estimated by the receivers of the links and forwarded to eNB through separate control channels. The queue length at UE is also forwarded by UE to eNB through a separate control channel. At each time-frame t , a centralized scheduler (usually at the eNB) observes the queue lengths of UE and eNB as well as the system channel conditions and determines the following; First, the bidirectional transmission scheme (Direct⁴, CoR or NC/CoR) that should be used by UE and eNB at each time-frame; second, the optimal resource allocation that results

⁴In Direct transmission scheme, no relay is deployed in the system and UE and eNB communicate through a direct link.

in the largest achievable rate region for the selected transmission scheme. At the end of the time-frame and after serving the packets according to the scheduling policy, the new arrivals are added to the queues and the queue lengths are updated.

3.2 Achievable Rate Regions for Direct, CoR and NC/CoR Schemes

In this section, we characterize the achievable rate regions, in an explicit form, for the Direct, CoR and NC/CoR schemes.

3.2.1 Direct Transmission

In direct transmission, it is assumed that no relay is deployed in the system and each bidirectional transmission time-frame consists of two subchannels f_1 and f_2 whose lengths are λ_1 and λ_2 , respectively. In this case, λ_1 and λ_2 are the time sharing parameters that determine the fraction of each bidirectional time-frame in which UE or eNB are transmitting. For bidirectional direct transmission, it is known that the uplink and downlink achievable rate pairs at time-frame t are described by the constraints given in (3.1).

$$R_{\text{UE}}(t) \leq \lambda_1 C_{\text{UE}}(t) \quad R_{\text{EU}}(t) \leq \lambda_2 C_{\text{EU}}(t) \quad \lambda_1 + \lambda_2 = 1 \quad (3.1)$$

As shown in Figure 3.4(a), the rate region of direct transmission at time-frame t denoted by $C_{\text{D}}(t)$ is a triangle formed by the origin $(0, 0)$, $(C_{\text{UE}}(t), 0)$ and $(0, C_{\text{EU}}(t))$. Since the region is restricted by the linear constraint $\lambda_1 + \lambda_2 = 1$, by using the

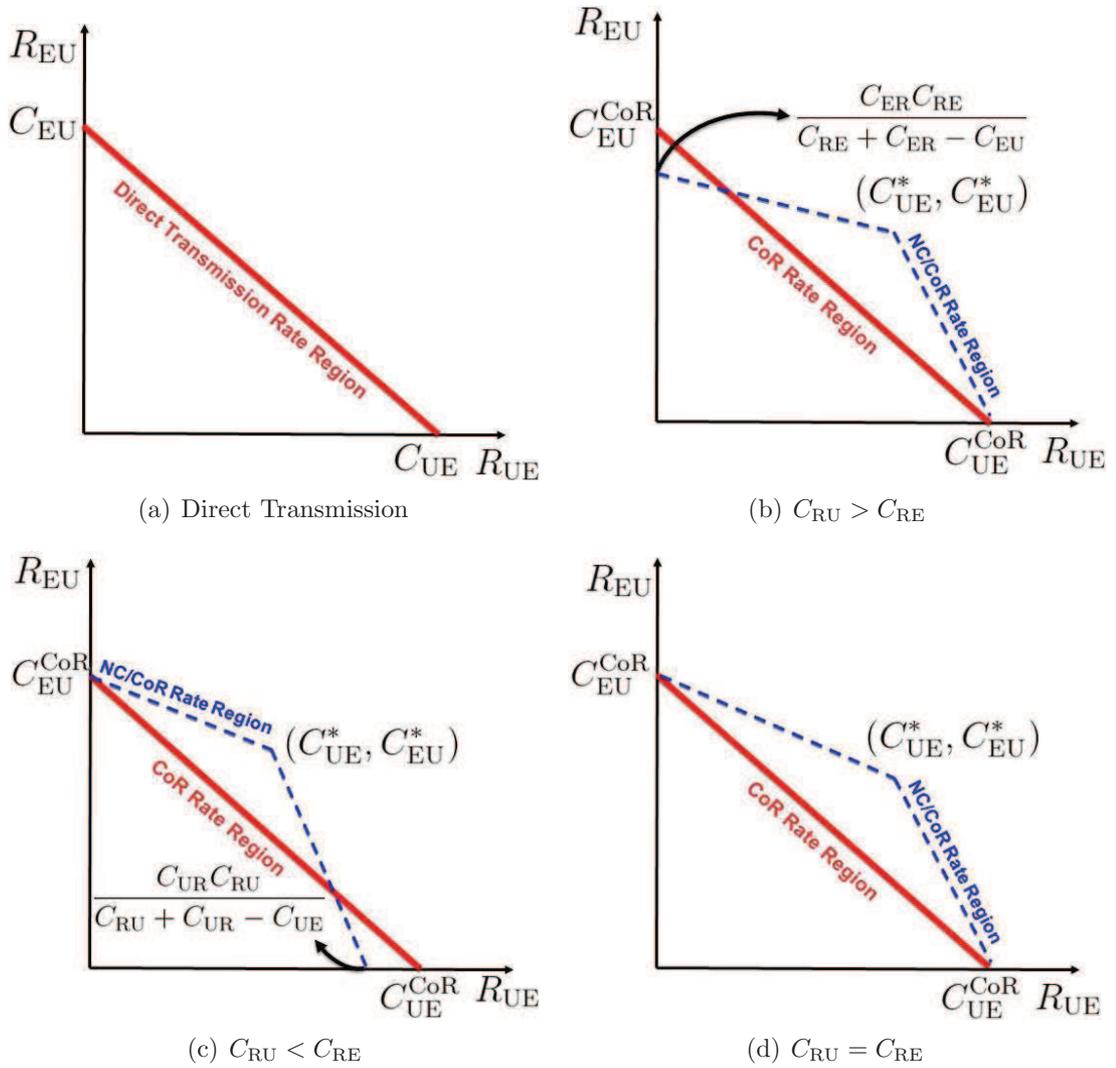


Figure 3.4: Achievable Rate Regions for Direct, CoR and NC/CoR Schemes in Example 1

inequalities in (3.1), we have

$$\frac{R_{\text{UE}}(t)}{C_{\text{UE}}(t)} + \frac{R_{\text{EU}}(t)}{C_{\text{EU}}(t)} \leq \lambda_1 + \lambda_2 = 1. \quad (3.2)$$

The achievable rate region in this case is the triangular convex region below the line connecting the two points $(C_{\text{UE}}(t), 0)$ and $(0, C_{\text{EU}}(t))$.

3.2.2 CoR Transmission

According to [26, 80], the uplink and downlink achievable rate pairs for CoR scheme (which is using relays in DF mode) are described by the inequalities in (3.3) based on parallel coding.

$$\begin{aligned} R_{\text{UE}}(t) &\leq \lambda_1 C_{\text{UR}}(t) & R_{\text{UE}}(t) &\leq \lambda_1 C_{\text{UE}}(t) + \lambda_2 C_{\text{RE}}(t) \\ R_{\text{EU}}(t) &\leq \lambda_3 C_{\text{ER}}(t) & R_{\text{EU}}(t) &\leq \lambda_3 C_{\text{EU}}(t) + \lambda_4 C_{\text{RU}}(t) \end{aligned} \quad (3.3)$$

In (3.3), $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ are constrained to satisfy $\sum_{i=1}^4 \lambda_i = 1$. In the following proposition, we will introduce a linear algebraic representation of the achievable rate region for CoR scheme.

Proposition 1. *The CoR achievable rate region at time-frame t denoted by $C_{\text{CoR}}(t)$ is described by all the positive rate pairs $(R_{\text{UE}}(t), R_{\text{EU}}(t)) \in \mathbb{R}_+^2$ that are constrained by (3.4).*

$$\frac{R_{\text{UE}}(t)}{C_{\text{UE}}^{\text{CoR}}(t)} + \frac{R_{\text{EU}}(t)}{C_{\text{EU}}^{\text{CoR}}(t)} \leq 1 \quad (3.4)$$

In (3.4), $C_{\text{UE}}^{\text{CoR}}(t)$ and $C_{\text{EU}}^{\text{CoR}}(t)$ are computed using (3.5) and (3.6).

$$C_{\text{UE}}^{\text{CoR}}(t) = \begin{cases} \min \{C_{\text{UR}}(t), C_{\text{UE}}(t)\} & \text{if } C_{\text{UE}}(t) \geq C_{\text{RE}}(t) \\ C_{\text{UR}}(t) \min \left\{ \frac{C_{\text{RE}}(t)}{C_{\text{UR}}(t) + C_{\text{RE}}(t) - C_{\text{UE}}(t)}, 1 \right\} & \text{if } C_{\text{UE}}(t) < C_{\text{RE}}(t) \end{cases} \quad (3.5)$$

$$C_{\text{EU}}^{\text{CoR}}(t) = \begin{cases} \min \{C_{\text{ER}}(t), C_{\text{EU}}(t)\} & \text{if } C_{\text{EU}}(t) \geq C_{\text{RU}}(t) \\ C_{\text{ER}}(t) \min \left\{ \frac{C_{\text{RU}}(t)}{C_{\text{ER}}(t) + C_{\text{RU}}(t) - C_{\text{EU}}(t)}, 1 \right\} & \text{if } C_{\text{EU}}(t) < C_{\text{RU}}(t) \end{cases} \quad (3.6)$$

Thus, the rate region is a triangle formed by the origin $(0, 0)$, $(C_{\text{UE}}^{\text{CoR}}(t), 0)$ and $(0, C_{\text{EU}}^{\text{CoR}}(t))$.

The proof of Proposition 1 is presented in Appendix B.

3.2.3 NC/CoR Transmission

According to [26], the uplink and downlink achievable rate pairs for NC/CoR transmission scheme are described by the inequalities in (3.7).

$$\begin{aligned} R_{\text{UE}}(t) &\leq \lambda_1 C_{\text{UR}}(t) & R_{\text{UE}}(t) &\leq \lambda_1 C_{\text{UE}}(t) + \lambda_3 C_{\min}(t) \\ R_{\text{EU}}(t) &\leq \lambda_2 C_{\text{ER}}(t) & R_{\text{EU}}(t) &\leq \lambda_2 C_{\text{EU}}(t) + \lambda_3 C_{\min}(t) \end{aligned} \quad (3.7)$$

In (3.7), we have $C_{\min}(t) = \min \{C_{\text{RU}}(t), C_{\text{RE}}(t)\}$ and $\lambda_1, \lambda_2, \lambda_3$ are constrained to satisfy $\sum_{i=1}^3 \lambda_i = 1$. In the following proposition, we will introduce a linear algebraic representation of the achievable rate region for NC/CoR scheme.

Proposition 2. *The NC/CoR achievable rate region at time-frame t denoted by $C_{\text{NC/CoR}}(t)$ is described in the following cases.*

Case 1: If $C_{\text{UE}}(t) \geq C_{\text{UR}}(t)$ or $C_{\text{EU}}(t) \geq C_{\text{ER}}(t)$, the rate region is described by

$$\frac{R_{\text{UE}}(t)}{C_{\text{UE}}^{\text{NC/CoR}}(t)} + \frac{R_{\text{EU}}(t)}{C_{\text{EU}}^{\text{NC/CoR}}(t)} \leq 1 \quad (3.8)$$

where $C_{\text{UE}}^{\text{NC/CoR}}(t)$ and $C_{\text{EU}}^{\text{NC/CoR}}(t)$ are defined in (3.9) and (3.10).

$$C_{\text{UE}}^{\text{NC/CoR}}(t) = \begin{cases} \min\{C_{\text{UR}}(t), C_{\text{UE}}(t)\} & \text{if } C_{\text{UE}}(t) \geq C_{\min}(t) \\ C_{\text{UR}}(t) \min\left\{\frac{C_{\min}(t)}{C_{\text{UR}}(t) + C_{\min}(t) - C_{\text{UE}}(t)}, 1\right\} & \text{if } C_{\text{UE}}(t) < C_{\min}(t) \end{cases} \quad (3.9)$$

$$C_{\text{EU}}^{\text{NC/CoR}}(t) = \begin{cases} \min\{C_{\text{ER}}(t), C_{\text{EU}}(t)\} & \text{if } C_{\text{EU}}(t) \geq C_{\min}(t) \\ C_{\text{ER}}(t) \min\left\{\frac{C_{\min}(t)}{C_{\text{ER}}(t) + C_{\min}(t) - C_{\text{EU}}(t)}, 1\right\} & \text{if } C_{\text{EU}}(t) < C_{\min}(t) \end{cases} \quad (3.10)$$

Case 2: $C_{\text{UE}}(t) < C_{\text{UR}}(t)$ and $C_{\text{EU}}(t) < C_{\text{ER}}(t)$, then the rate region is specified as follows:

- **Sub-case 2.1:** $C_{\min}(t) < C_{\text{UE}}(t)$ and $C_{\min}(t) < C_{\text{EU}}(t)$: The rate region is

$$\frac{R_{\text{UE}}(t)}{C_{\text{UE}}(t)} + \frac{R_{\text{EU}}(t)}{C_{\text{EU}}(t)} \leq 1. \quad (3.11)$$

- **Sub-case 2.2:** $C_{\min}(t) \geq C_{\text{UE}}(t)$ or $C_{\min}(t) \geq C_{\text{EU}}(t)$: The rate region is specified as the intersection of the regions defined in (3.12).

$$\frac{R_{\text{UE}}(t)}{A} + \frac{R_{\text{EU}}(t)}{B} \leq 1 \quad \frac{R_{\text{UE}}(t)}{C} + \frac{R_{\text{EU}}(t)}{D} \leq 1 \quad (3.12)$$

$$A = \begin{cases} \frac{C_{\text{UR}}(t)C_{\text{min}}(t)}{C_{\text{min}}(t) + C_{\text{UR}}(t) - C_{\text{UE}}(t)} & \text{if } C_{\text{min}}(t) \geq C_{\text{UE}}(t) \text{ and } C_{\text{min}}(t) < C_{\text{EU}}(t) \\ C_{\text{UR}}(t) & \text{if } C_{\text{min}}(t) < C_{\text{UE}}(t) \text{ and } C_{\text{min}}(t) \geq C_{\text{EU}}(t) \\ \frac{C_{\text{UR}}(t)C_{\text{min}}(t)}{C_{\text{min}}(t) + C_{\text{UR}}(t) - C_{\text{UE}}(t)} & \text{if } C_{\text{min}}(t) \geq C_{\text{UE}}(t) \text{ and } C_{\text{min}}(t) \geq C_{\text{EU}}(t) \end{cases} \quad (3.13)$$

$$B = \begin{cases} C_{\text{ER}}(t) & \text{if } C_{\text{min}}(t) \geq C_{\text{UE}}(t) \text{ and } C_{\text{min}}(t) < C_{\text{EU}}(t) \\ \frac{C_{\text{ER}}(t)C_{\text{min}}(t)}{C_{\text{min}}(t) + C_{\text{ER}}(t) - C_{\text{EU}}(t)} & \text{if } C_{\text{min}}(t) < C_{\text{UE}}(t) \text{ and } C_{\text{min}}(t) \geq C_{\text{EU}}(t) \\ C_{\text{ER}}(t) & \text{if } C_{\text{min}}(t) \geq C_{\text{UE}}(t) \text{ and } C_{\text{min}}(t) \geq C_{\text{EU}}(t) \end{cases} \quad (3.14)$$

$$C = \begin{cases} \frac{\frac{1}{C_{\text{min}}(t) + C_{\text{UR}}(t) - C_{\text{UE}}(t)} - \frac{C_{\text{UR}}(t) - C_{\text{UE}}(t)}{C_{\text{UR}}(t)C_{\text{min}}(t)}}{\frac{C_{\text{UR}}(t)C_{\text{min}}(t)}{C_{\text{min}}(t) + C_{\text{UR}}(t) - C_{\text{UE}}(t)} - \frac{C_{\text{UR}}(t) - C_{\text{UE}}(t)}{C_{\text{UR}}(t)C_{\text{EU}}(t)}} & \text{if } C_{\text{min}}(t) \geq C_{\text{UE}}(t) \text{ and } C_{\text{min}}(t) < C_{\text{EU}}(t) \\ C_{\text{UE}}(t) & \text{if } C_{\text{min}}(t) < C_{\text{UE}}(t) \text{ and } C_{\text{min}}(t) \geq C_{\text{EU}}(t) \\ C_{\text{UR}}(t) & \text{if } C_{\text{min}}(t) \geq C_{\text{UE}}(t) \text{ and } C_{\text{min}}(t) \geq C_{\text{EU}}(t) \end{cases} \quad (3.15)$$

$$D = \begin{cases} C_{\text{EU}}(t) & \text{if } C_{\text{min}}(t) \geq C_{\text{UE}}(t) \text{ and } C_{\text{min}}(t) < C_{\text{EU}}(t) \\ \frac{\frac{1}{C_{\text{min}}(t) + C_{\text{ER}}(t) - C_{\text{EU}}(t)} - \frac{C_{\text{ER}}(t) - C_{\text{EU}}(t)}{C_{\text{ER}}(t)C_{\text{min}}(t)}}{\frac{C_{\text{ER}}(t)C_{\text{min}}(t)}{C_{\text{min}}(t) + C_{\text{ER}}(t) - C_{\text{EU}}(t)} - \frac{C_{\text{ER}}(t) - C_{\text{EU}}(t)}{C_{\text{ER}}(t)C_{\text{UE}}(t)}} & \text{if } C_{\text{min}}(t) < C_{\text{UE}}(t) \text{ and } C_{\text{min}}(t) \geq C_{\text{EU}}(t) \\ \frac{C_{\text{ER}}(t)C_{\text{min}}(t)}{C_{\text{min}}(t) + C_{\text{ER}}(t) - C_{\text{EU}}(t)} & \text{if } C_{\text{min}}(t) \geq C_{\text{UE}}(t) \text{ and } C_{\text{min}}(t) \geq C_{\text{EU}}(t) \end{cases} \quad (3.16)$$

In (3.12), A , B , C and D are computed using (3.13)–(3.16).

The proof of Proposition 2 is presented in Appendix C.

So far, we have characterized the convex rate region corresponding to each transmission scheme at any given time-frame. The long-run achievable rate region of each scheme is obtained by taking the mean of the rate region with respect to the channel state distribution as follows.

$$C_{\text{D}} = \sum_{c \in \mathcal{C}} P(c)C_{\text{D}}(c) \quad C_{\text{CoR}} = \sum_{c \in \mathcal{C}} P(c)C_{\text{CoR}}(c) \quad C_{\text{NC/CoR}} = \sum_{c \in \mathcal{C}} P(c)C_{\text{NC/CoR}}(c) \quad (3.17)$$

$C_D(c)$, $C_{\text{CoR}}(c)$ and $C_{\text{NC/CoR}}(c)$ are the achievable rate regions corresponding to the Direct, CoR and NC/CoR transmission schemes when the system channel is in state $c \in \mathcal{C}$.

3.3 Achievable Rate Region for PLNC/CoR

Physical Layer Network Coding (PLNC)/CoR employs another type of network coding designed to be applied in relay-assisted wireless access networks [26, 31, 81–83]. Unlike the MAC layer network coding (in NC/CoR) where packets are combined before channel coding, in PLNC/CoR the operation of network coding is performed on the channel coded packets. This makes PLNC/CoR achieve the broadcast capacities of individual channels from the relay to the UE and eNB while in NC/CoR the minimum of these capacities is achievable for broadcasting from the relay node to the UE and eNB. Unlike Physical Layer Network Coding (PLNC) [81] which does not use CoR, PLNC/CoR is incorporating the notions of CoR and PLNC together. Since the wireless nodes are assumed to be half-duplex, PLNC/CoR must use three subchannels (unlike PLNC where two subchannels are used). Although PLNC/CoR has lower spectral efficiency than PLNC, it has larger diversity gain and thus lower outage probability. It is shown in [26, 31] that PLNC/CoR achieves a larger rate region with respect to NC/CoR. However, this rate improvement comes at the cost of *coding and transceiver complexity*.

According to [26, 40], the achievable rate region for PLNC/CoR is characterized by the following inequalities.

$$R_{\text{UE}} \leq \lambda_1 C_{\text{UR}} \qquad R_{\text{UE}} \leq \lambda_1 C_{\text{UE}} + \lambda_3 C_{\text{RE}} \qquad (3.18)$$

$$R_{\text{EU}} \leq \lambda_2 C_{\text{ER}} \qquad R_{\text{EU}} \leq \lambda_2 C_{\text{EU}} + \lambda_3 C_{\text{RU}} \qquad (3.19)$$

Proposition 3. *The PLNC/CoR achievable rate region at time-frame t denoted by $C_{\text{PLNC/CoR}}(t)$ is described in the following cases.*

Case 1: *If $C_{\text{UE}}(t) \geq C_{\text{UR}}(t)$ or $C_{\text{EU}}(t) \geq C_{\text{ER}}(t)$, the rate region is described by*

$$\frac{R_{\text{UE}}(t)}{C_{\text{UE}}^{\text{PLNC/CoR}}(t)} + \frac{R_{\text{EU}}(t)}{C_{\text{EU}}^{\text{PLNC/CoR}}(t)} \leq 1 \qquad (3.20)$$

where $C_{\text{UE}}^{\text{PLNC/CoR}}(t)$ and $C_{\text{EU}}^{\text{PLNC/CoR}}(t)$ are defined in (3.21) and (3.22).

$$C_{\text{UE}}^{\text{PLNC/CoR}}(t) = \begin{cases} \min\{C_{\text{UR}}(t), C_{\text{UE}}(t)\} & \text{if } C_{\text{UE}}(t) \geq C_{\text{RE}}(t) \\ C_{\text{UR}}(t) \min\left\{\frac{C_{\text{RE}}(t)}{C_{\text{UR}}(t) + C_{\text{RE}}(t) - C_{\text{UE}}(t)}, 1\right\} & \text{if } C_{\text{UE}}(t) < C_{\text{RE}}(t) \end{cases} \qquad (3.21)$$

$$C_{\text{EU}}^{\text{PLNC/CoR}}(t) = \begin{cases} \min\{C_{\text{ER}}(t), C_{\text{EU}}(t)\} & \text{if } C_{\text{EU}}(t) \geq C_{\text{RU}}(t) \\ C_{\text{ER}}(t) \min\left\{\frac{C_{\text{RU}}(t)}{C_{\text{ER}}(t) + C_{\text{RU}}(t) - C_{\text{EU}}(t)}, 1\right\} & \text{if } C_{\text{EU}}(t) < C_{\text{RU}}(t) \end{cases} \qquad (3.22)$$

Case 2: *If $C_{\text{UE}}(t) < C_{\text{UR}}(t)$ and $C_{\text{EU}}(t) < C_{\text{ER}}(t)$, then the rate region is specified as follows:*

- **Sub-case 2.1:** *If $C_{\text{RE}}(t) < C_{\text{UE}}(t)$ and $C_{\text{RU}}(t) < C_{\text{EU}}(t)$: Then, the rate region*

is

$$\frac{R_{\text{UE}}(t)}{C_{\text{UE}}(t)} + \frac{R_{\text{EU}}(t)}{C_{\text{EU}}(t)} \leq 1. \qquad (3.23)$$

$$A' = \begin{cases} \frac{C_{UR}(t)C_{RE}(t)}{C_{RE}(t)+C_{UR}(t)-C_{UE}(t)} & \text{if } C_{RE}(t) \geq C_{UE}(t) \text{ and } C_{RU}(t) < C_{EU}(t) \\ C_{UR}(t) & \text{if } C_{RE}(t) < C_{UE}(t) \text{ and } C_{RU}(t) \geq C_{EU}(t) \\ \frac{C_{UR}(t)C_{RE}(t)}{C_{RE}(t)+C_{UR}(t)-C_{UE}(t)} & \text{if } C_{RE}(t) \geq C_{UE}(t) \text{ and } C_{RU}(t) \geq C_{EU}(t) \end{cases} \quad (3.25)$$

$$B' = \begin{cases} C_{ER}(t) & \text{if } C_{RE}(t) \geq C_{UE}(t) \text{ and } C_{RU}(t) < C_{EU}(t) \\ \frac{C_{ER}(t)C_{RU}(t)}{C_{RU}(t)+C_{ER}(t)-C_{EU}(t)} & \text{if } C_{RE}(t) < C_{UE}(t) \text{ and } C_{RU}(t) \geq C_{EU}(t) \\ C_{ER}(t) & \text{if } C_{RE}(t) \geq C_{UE}(t) \text{ and } C_{RU}(t) \geq C_{EU}(t) \end{cases} \quad (3.26)$$

$$C' = \begin{cases} \frac{1}{\frac{C_{RE}(t)+C_{UR}(t)-C_{UE}(t)}{C_{UR}(t)C_{RE}(t)} - \frac{C_{RU}(t)(C_{UR}(t)-C_{UE}(t))}{C_{UR}(t)C_{EU}(t)C_{RE}(t)}} & \text{if } C_{RE}(t) \geq C_{UE}(t) \text{ and } C_{RU}(t) < C_{EU}(t) \\ C_{UE}(t) & \text{if } C_{RE}(t) < C_{UE}(t) \text{ and } C_{RU}(t) \geq C_{EU}(t) \\ C_{UR}(t) & \text{if } C_{RE}(t) \geq C_{UE}(t) \text{ and } C_{RU}(t) \geq C_{EU}(t) \end{cases} \quad (3.27)$$

$$D' = \begin{cases} C_{EU}(t) & \text{if } C_{RE}(t) \geq C_{UE}(t) \text{ and } C_{RU}(t) < C_{EU}(t) \\ \frac{1}{\frac{C_{RU}(t)+C_{ER}(t)-C_{EU}(t)}{C_{ER}(t)C_{RU}(t)} - \frac{C_{RE}(t)(C_{ER}(t)-C_{EU}(t))}{C_{ER}(t)C_{UE}(t)C_{RU}(t)}} & \text{if } C_{RE}(t) < C_{UE}(t) \text{ and } C_{RU}(t) \geq C_{EU}(t) \\ \frac{C_{ER}(t)C_{RU}(t)}{C_{RU}(t)+C_{ER}(t)-C_{EU}(t)} & \text{if } C_{RE}(t) \geq C_{UE}(t) \text{ and } C_{RU}(t) \geq C_{EU}(t) \end{cases} \quad (3.28)$$

-
- **Sub-case 2.2:** If $C_{RE}(t) \geq C_{UE}(t)$ or $C_{RU}(t) \geq C_{EU}(t)$: Then, the rate region is specified as the intersection of the regions described by the following equations.

$$\frac{R_{UE}(t)}{A'} + \frac{R_{EU}(t)}{B'} \leq 1 \quad \frac{R_{UE}(t)}{C'} + \frac{R_{EU}(t)}{D'} \leq 1 \quad (3.24)$$

In (3.24), A' , B' , C' and D' are defined by (3.25)–(3.28).

3.4 Hybrid Transmission Scheme

According to the results obtained Section 3.2, we can observe that NC/CoR may not always achieve higher throughput than CoR or direct transmission. In fact,

depending on the channel condition and the uplink and downlink traffic ratio, a policy can perform better than the others as shown in the following examples.

Example 1. Consider the case that we have $C_{UE}(t) < \min\{C_{\min}(t), C_{UR}(t)\}$ and $C_{EU}(t) < \min\{C_{\min}(t), C_{ER}(t)\}$. The achievable rate region of CoR is specified by

$$\frac{\frac{R_{UE}(t)}{C_{UR}(t)C_{RE}(t)}}{C_{RE}(t) + C_{UR}(t) - C_{UE}(t)}} + \frac{\frac{R_{EU}(t)}{C_{ER}(t)C_{RU}(t)}}{C_{ER}(t) + C_{RU}(t) - C_{EU}(t)}} \leq 1. \quad (3.29)$$

For the achievable rate region of NC/CoR we have the following cases:

- If $C_{RU}(t) > C_{RE}(t)$: The rate region is the intersection of the regions described by the following equations.

$$\begin{aligned} \frac{\frac{R_{UE}(t)}{C_{UR}(t)C_{RE}(t)}}{C_{RE}(t) + C_{UR}(t) - C_{UE}(t)}} + \frac{R_{EU}(t)}{C_{ER}(t)} &\leq 1 \\ \frac{R_{UE}(t)}{C_{UR}(t)} + \frac{\frac{R_{EU}(t)}{C_{ER}(t)C_{RE}(t)}}{C_{RE}(t) + C_{ER}(t) - C_{EU}(t)}} &\leq 1 \end{aligned} \quad (3.30)$$

- If $C_{RU}(t) < C_{RE}(t)$: The rate region is the intersection of the regions described by the following equations.

$$\begin{aligned} \frac{\frac{R_{UE}(t)}{C_{UR}(t)C_{RU}(t)}}{C_{RU}(t) + C_{UR}(t) - C_{UE}(t)}} + \frac{R_{EU}(t)}{C_{ER}(t)} &\leq 1 \\ \frac{R_{UE}(t)}{C_{UR}(t)} + \frac{\frac{R_{EU}(t)}{C_{ER}(t)C_{RU}(t)}}{C_{RU}(t) + C_{ER}(t) - C_{EU}(t)}} &\leq 1 \end{aligned} \quad (3.31)$$

- If $C_{RU}(t) = C_{RE}(t)$: The rate region is the intersection of the regions described by

the following equations.

$$\begin{aligned} \frac{R_{\text{UE}}(t)}{\frac{C_{\text{UR}}(t)C_{\text{RE}}(t)}{C_{\text{RE}}(t) + C_{\text{UR}}(t) - C_{\text{UE}}(t)}} + \frac{R_{\text{EU}}(t)}{C_{\text{ER}}(t)} &\leq 1 \\ \frac{R_{\text{UE}}(t)}{C_{\text{UR}}(t)} + \frac{R_{\text{EU}}(t)}{\frac{C_{\text{ER}}(t)C_{\text{RU}}(t)}{C_{\text{RU}}(t) + C_{\text{ER}}(t) - C_{\text{EU}}(t)}} &\leq 1 \end{aligned} \quad (3.32)$$

Figure 3.4 compares the rate regions of CoR and NC/CoR for the different cases considered above. It can be observed from the figures that the rate region of NC/CoR is not always a superset of the rate region of CoR. Especially in cases (b) and (c) of Figure 3.4, depending on the ratio of the uplink and downlink traffic rates, CoR scheme can outperform NC/CoR. The assumptions on the channel conditions in this example are not unusual in realistic wireless networks. For example, if the UE is close to the edge of the cell, UE and eNB will be far from each other, and hence the channels over links UE→eNB and eNB→UE are weak. We also observe that the rate region of NC/CoR in case (d) is a superset of the rate region of CoR scheme. In the last case, since the channels from the relay station to UE and eNB are symmetric, i.e., $C_{\text{min}} = C_{\text{RU}} = C_{\text{RE}}$, when re-encoding the network coded message at the relay station and broadcasting it to UE and eNB, no channel capacity is wasted. However, in the asymmetric case where $C_{\text{RU}} \neq C_{\text{RE}}$, since channel coding and broadcasting is performed by rate C_{min} , channel capacity is wasted either in uplink (if $C_{\text{min}} = C_{\text{RU}}$) or downlink (if $C_{\text{min}} = C_{\text{RE}}$). Therefore, NC/CoR has the best performance when the channel capacities from the relay station to UE and eNB are equal.

Example 2. We assume that UE and eNB are very close to each other. In this case, we can readily see that direct transmission outperforms CoR and NC/CoR. In this case, the rate regions of CoR and NC/CoR are specified by $\frac{R_{\text{UE}}(t)}{C_{\text{UR}}(t)} + \frac{R_{\text{EU}}(t)}{C_{\text{ER}}(t)} \leq 1$ while

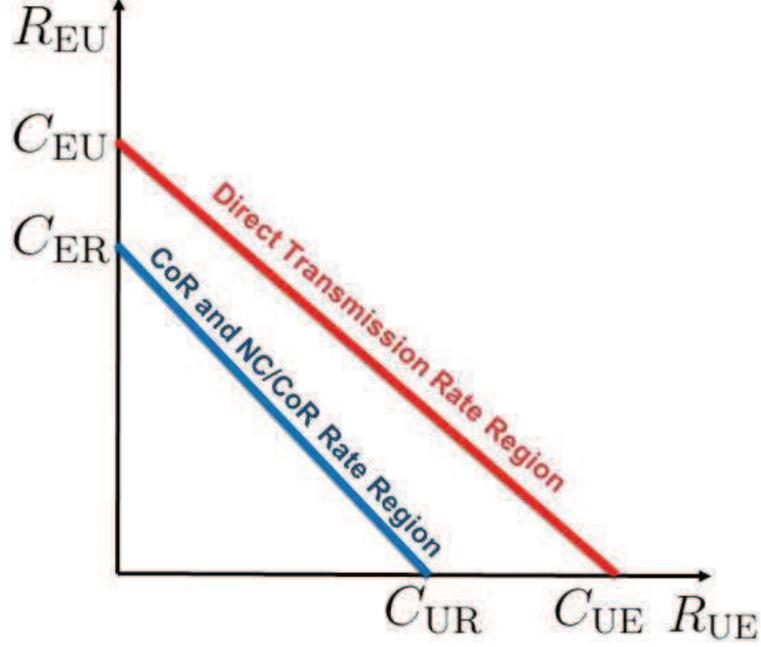


Figure 3.5: Achievable Rate Regions for Direct, CoR and NC/CoR Schemes in Example 2

the capacity of direct transmission is given by $\frac{R_{UE}(t)}{C_{UE}(t)} + \frac{R_{EU}(t)}{C_{EU}(t)} \leq 1$ which is a superset of the rate region of CoR and NC/CoR schemes. These regions are compared in Figure 3.5. The dominance of the rate region of direct transmission with respect to CoR and NC/CoR can be explained by considering the following points: First, in CoR and NC/CoR, the end-to-end achievable rates are restricted to the capacity of relay links $C_{UR}(t)$ and $C_{ER}(t)$ while in direct transmission no relay is involved in bidirectional transmission. In this example, the direct links are more powerful than the relay links. Second, in direct transmission only two subchannels are used while in CoR and NC/CoR four and three subchannels are used, i.e., in this example, direct transmission is dominant due to its higher spectral efficiency.

Our discussion above motivates us to look for a dynamic resource allocation scheme that can take advantages of all the individual transmission schemes. Such a hybrid scheme determines the best transmission scheme at each time-frame based on the

channel state of the system as well as the queue length information.

The hybrid scheme addresses the following scheduling problems: First, it determines the best transmission scheme to be used at each time-frame; second, it optimizes the resource allocation and the corresponding achievable rates at that time-frame.

When the system channel is in state c , the rate region $\mathcal{R}(c) = \text{Conv}\{C_D(c) \cup C_{\text{CoR}}(c) \cup C_{\text{NC/CoR}}(c)\}$ is also achievable [84] by using a randomized policy among Direct, CoR and NC/CoR schemes. This is due to the fact that each rate pair $r = (R_{\text{UE}}, R_{\text{EU}}) \in \mathcal{R}(c)$ can be written as a convex combination of a set of points in $C_D(c) \cup C_{\text{CoR}}(c) \cup C_{\text{NC/CoR}}(c)$, i.e., for any $r \in \mathcal{R}(c)$, there exist $r_1, r_2, \dots, r_N \in C_D(c) \cup C_{\text{CoR}}(c) \cup C_{\text{NC/CoR}}(c)$ and $0 \leq p_1, p_2, \dots, p_N \leq 1$ with $\sum_{n=1}^N p_n = 1$ such that $r = \sum_{n=1}^N p_n r_n$. Hence, each rate pair in $\mathcal{R}(c)$ is achievable if we select $r_n, n = 1, 2, \dots, N$ (by applying the appropriate transmission scheme, i.e., Direct, CoR or NC/CoR with the corresponding time sharing parameters) with probability p_n when the system is in state c . Therefore, by applying a randomized strategy the following rate region is achievable.

$$\mathcal{R} = \sum_{c \in \mathcal{C}} P(c) \mathcal{R}(c) \quad (3.33)$$

To apply the randomized strategy, we need to know the rate region of each of the transmission schemes at each state c and also identify r_1, r_2, \dots, r_N and p_1, p_2, \dots, p_N . In the following, we will introduce a queue-length based hybrid scheme which is based on the well-known back-pressure approach [42] to achieve the combined rate region \mathcal{R} . In this approach, we just need to know the channel state of the system at each transmission time-frame as well as the queue lengths. The back-pressure algorithm operates in slotted time in which at each time-slot it seeks to route data in directions that maximize the differential backlog between neighboring nodes. The decisions of

the current time-slot will affect the results in the subsequent time-slots. The main attractive features of the backpressure algorithm are that it leads to maximum network throughput, it is provably robust to time-varying network conditions, and it can be implemented without knowing traffic arrival rates or channel state probabilities [42]. We first present the scheme for a system in which the traffic arrivals and departures at each time-frame are treated as continuous flows of information. We call such a system “continuous-flow” system.

3.4.1 Hybrid Scheme for the Continuous-flow System

Consider a TWRC system in which the traffic arrivals and departures at each transmission time-frame are *continuous* variables. For such a system, the variables $X_U(t)$ and $X_E(t)$ are called the *system unfinished work* (instead of queue lengths) at the end of time-frame t . The hybrid scheme for the continuous-flow system is defined as follows.

- For the Direct, CoR and NC/CoR schemes, the hybrid approach solves the following linear optimization problems at each time-frame.

$$R_D = \max_{\substack{R_{UE}(t) \leq \lambda_1 C_{UE}(t) \\ R_{EU}(t) \leq \lambda_2 C_{EU}(t) \\ \lambda_1 + \lambda_2 \leq 1 \\ \lambda_1, \lambda_2 \geq 0}} \{X_U(t-1)R_{UE}(t) + X_E(t-1)R_{EU}(t)\} \quad (3.34)$$

$$R_{CoR} = \max_{\substack{R_{UE}(t) \leq \lambda_1 C_{UR}(t) \\ R_{UE}(t) \leq \lambda_1 C_{UE}(t) + \lambda_2 C_{RE}(t) \\ R_{EU}(t) \leq \lambda_3 C_{ER}(t) \\ R_{EU}(t) \leq \lambda_3 C_{EU}(t) + \lambda_4 C_{RU}(t) \\ \sum_{n=1}^4 \lambda_n \leq 1 \\ \lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0}} \{X_U(t-1)R_{UE}(t) + X_E(t-1)R_{EU}(t)\} \quad (3.35)$$

$$R_{\text{NC/CoR}} = \max_{\substack{R_{\text{UE}}(t) \leq \lambda_1 C_{\text{UR}}(t) \\ R_{\text{UE}}(t) \leq \lambda_1 C_{\text{UE}}(t) + \lambda_3 C_{\text{min}}(t) \\ R_{\text{EU}}(t) \leq \lambda_2 C_{\text{ER}}(t) \\ R_{\text{EU}}(t) \leq \lambda_2 C_{\text{EU}}(t) + \lambda_3 C_{\text{min}}(t) \\ \sum_{n=1}^3 \lambda_n \leq 1 \\ \lambda_1, \lambda_2, \lambda_3 \geq 0}} \{X_{\text{U}}(t-1)R_{\text{UE}}(t) + X_{\text{E}}(t-1)R_{\text{EU}}(t)\} \quad (3.36)$$

- The transmission scheme corresponding to $\max\{R_{\text{D}}, R_{\text{CoR}}, R_{\text{NC/CoR}}\}$ is selected as the transmission scheme for that time-frame.
- The time sharing parameters λ_n corresponding to the maximum values of R_{D} , R_{CoR} and $R_{\text{NC/CoR}}$ are selected as the resource allocation parameters for that time-frame.

It is worth mentioning that by having a linear algebraic characterization of the achievable rate regions for Direct, CoR and NC/CoR schemes (which we derived in Section 3.2), we can readily obtain a closed-form solution for the above problems as follows.

$$R_{\text{D}} = \max\{X_{\text{U}}(t-1)C_{\text{UE}}(t), X_{\text{E}}(t-1)C_{\text{EU}}(t)\} \quad (3.37)$$

$$R_{\text{CoR}} = \max\{X_{\text{U}}(t-1)C_{\text{UE}}^{\text{CoR}}(t), X_{\text{E}}(t-1)C_{\text{EU}}^{\text{CoR}}(t)\} \quad (3.38)$$

$C_{\text{UE}}^{\text{CoR}}(t)$ and $C_{\text{EU}}^{\text{CoR}}(t)$ were defined in (3.5) and (3.6). For NC/CoR, we have the following cases:

- If $C_{\text{UE}}(t) \geq C_{\text{UR}}(t)$ or $C_{\text{EU}}(t) \geq C_{\text{ER}}(t)$,

$$R_{\text{NC/CoR}} = \max\{X_{\text{U}}(t-1)C_{\text{UE}}^{\text{NC/CoR}}(t), X_{\text{E}}(t-1)C_{\text{EU}}^{\text{NC/CoR}}(t)\} \quad (3.39)$$

where $C_{\text{UE}}^{\text{NC/CoR}}(t)$ and $C_{\text{EU}}^{\text{NC/CoR}}(t)$ are defined in (3.9) and (3.10).

- If $C_{\text{UE}}(t) < C_{\text{UR}}(t)$ and $C_{\text{EU}}(t) < C_{\text{ER}}(t)$, then we have the following sub-cases:

- $C_{\min}(t) < C_{\text{UE}}(t)$ and $C_{\min}(t) < C_{\text{EU}}(t)$:

$$R_{\text{NC/CoR}} = \max\{X_{\text{U}}(t-1)C_{\text{UE}}(t), X_{\text{E}}(t-1)C_{\text{EU}}(t)\} \quad (3.40)$$

- $C_{\min}(t) \geq C_{\text{UE}}(t)$ or $C_{\min}(t) \geq C_{\text{EU}}(t)$:

$$R_{\text{NC/CoR}} = \max\{X_{\text{U}}(t-1) \min\{A, C\}, X_{\text{E}}(t-1) \min\{B, D\}, \\ X_{\text{U}}(t-1)C_{\text{UE}}^* + X_{\text{E}}(t-1)C_{\text{EU}}^*\} \quad (3.41)$$

where $(C_{\text{UE}}^*, C_{\text{EU}}^*) = \left(\frac{1-\frac{B}{D}}{\frac{1}{C}-\frac{B}{AD}}, \frac{1-\frac{A}{C}}{\frac{1}{D}-\frac{A}{BC}}\right)$, and A, B, C, D are defined in (3.13)–(3.16).

A sketch of the proof for deriving Equations (3.37)–(3.41) is provided.

Sketch of the proof for Equations (3.37)–(3.41): In the optimization problems in (3.34), (3.35) and (3.36), the objective is a linear function of $R_{\text{UE}}(t)$ and $R_{\text{EU}}(t)$ and the constraints are polytope regions (triangles in Direct, CoR and NC/CoR schemes and quadrilateral in NC/CoR) derived in Section 3.2. According to a known result in linear programming, the optimal point is achieved at the extreme points in the feasible polytope region. In our problems, the extreme points of the polytope regions are triangle and quadrilateral vertices. Thus, the optimal solutions are derived by replacing these vertices into the objective functions and taking the maximum over all the vertices which results into Equations (3.37)–(3.41).

Stability

In the following proposition, we will state that the hybrid approach is rate stable for all the rate pairs in \mathcal{R} . Hence, the hybrid approach can stabilize the system backlog (the system backlog remains bounded as time progresses) when for the arrival rate pairs we have $(\Lambda_U, \Lambda_E) \in \mathcal{R}$. \mathcal{R} is called the network stability region (also called network layer capacity region in literature). In other words, the network stability region is the closure of the set of all bidirectional arrival rate vectors for which there exists a policy that can stabilize the network. Note that the network stability region is independent of the resource allocation policy and in fact is a specific characteristic of any network. In this work, we apply the stability definition used in [84].

Definition 1. *A queue is called strongly stable if it has bounded time averaged expected backlog, i.e.*

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} E[X(\tau)] < \infty$$

A resource allocation policy is called throughput optimal if it can stabilize the system for all the arrival rate vectors strictly inside the network stability region.

Proposition 4. *The hybrid transmission scheme can stabilize the continuous-flow system for all the traffic rates inside the combined rate region \mathcal{R} .*

The proof is based on applying the Lyapunov stability technique [84] and is presented in Appendix D. Note that $C_D, C_{CoR}, C_{NC/CoR} \subset \mathcal{R}$ and thus the hybrid scheme supports more arrival rates stably.

3.4.2 Hybrid Scheme with Discrete Packet Arrivals

In this case, the queues are fed by packets of size L instead of a continuous stream of information data and the arrival variables $A_U(t)$ and $A_E(t)$ are positive integer numbers. Moreover, data can be served from each queue in the form of discrete packets. Thus, the maximum number of successful packet transmissions over time-frame t from UE to eNB would be $W_{UE}(t) = \left\lfloor \frac{R_{UE}(t)}{L} \right\rfloor$ and the achievable rate would be $L \left\lfloor \frac{R_{UE}(t)}{L} \right\rfloor$. Similarly we can define $W_{EU}(t)$ for transmissions from eNB to UE. Since $LW_{UE}(t) \leq R_{UE}(t)$, for the discrete packet system, the achievable rate region when the system is in state c is a subset of the rate regions derived in (3.2), Propositions 1 and 2. According to the back-pressure principle [42, 84], the throughput-optimal hybrid scheme for the discrete packet system would be similar to the continuous-flow system except we determine R_D , R_{CoR} and $R_{NC/CoR}$ by solving the following *non-linear* optimization problems.

$$R_D^{(p)} = \max_{\substack{W_{UE}(t) \leq \left\lfloor \frac{\lambda_1 C_{UE}(t)}{L} \right\rfloor \\ W_{EU}(t) \leq \left\lfloor \frac{\lambda_2 C_{EU}(t)}{L} \right\rfloor \\ W_{UE}(t), W_{EU}(t) \in \{0, 1, 2, \dots\} \\ \lambda_1 + \lambda_2 \leq 1 \\ \lambda_1, \lambda_2 \geq 0}} \{X_U(t-1)W_{UE}(t) + X_E(t-1)W_{EU}(t)\} \quad (3.42)$$

$$R_{CoR}^{(p)} = \max_{\substack{W_{UE}(t) \leq \left\lfloor \frac{\lambda_1 C_{UR}(t)}{L} \right\rfloor \\ W_{UE}(t) \leq \left\lfloor \frac{\lambda_1 C_{UE}(t) + \lambda_2 C_{RE}(t)}{L} \right\rfloor \\ W_{EU}(t) \leq \left\lfloor \frac{\lambda_3 C_{ER}(t)}{L} \right\rfloor \\ W_{EU}(t) \leq \left\lfloor \frac{\lambda_3 C_{EU}(t) + \lambda_4 C_{RU}(t)}{L} \right\rfloor \\ W_{UE}(t), W_{EU}(t) \in \{0, 1, 2, \dots\} \\ \sum_{n=1}^4 \lambda_n \leq 1 \\ \lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0}} \{X_U(t-1)W_{UE}(t) + X_E(t-1)W_{EU}(t)\} \quad (3.43)$$

$$R_{\text{NC/CoR}}^{(\text{p})} = \max_{\substack{R_{\text{UE}}(t) \leq L \left\lfloor \frac{\lambda_1 C_{\text{UR}}(t)}{L} \right\rfloor \\ R_{\text{UE}}(t) \leq L \left\lfloor \frac{\lambda_1 C_{\text{UE}}(t) + \lambda_3 C_{\text{min}}(t)}{L} \right\rfloor \\ R_{\text{EU}}(t) \leq L \left\lfloor \frac{\lambda_2 C_{\text{ER}}(t)}{L} \right\rfloor \\ R_{\text{EU}}(t) \leq L \left\lfloor \frac{\lambda_2 C_{\text{EU}}(t) + \lambda_3 C_{\text{min}}(t)}{L} \right\rfloor \\ W_{\text{UE}}(t), W_{\text{EU}}(t) \in \{0, 1, 2, \dots\} \\ \sum_{n=1}^3 \lambda_n \leq 1 \\ \lambda_1, \lambda_2, \lambda_3 \geq 0}} \{X_{\text{U}}(t-1)R_{\text{UE}}(t) + X_{\text{E}}(t-1)R_{\text{EU}}(t)\} \quad (3.44)$$

Similar to the continuous-flow model, we select the transmission scheme which results in the maximum of $R_{\text{D}}^{(\text{p})}$, $R_{\text{CoR}}^{(\text{p})}$, $R_{\text{NC/CoR}}^{(\text{p})}$ for packet forwarding in the current time-frame and its corresponding λ values as the optimal resource allocation parameters.

The above optimization problems are Mixed-Integer-Linear Programs (MILP) in contrast to the continuous-flow model where we are dealing with linear programs. Therefore, we have to apply non-linear optimization algorithms to solve these problems which may impose real-time computational complexity to the system. However, if $W_{\text{UE}}(t)$ and $W_{\text{EU}}(t)$ are large quantities⁵ (which is usually the case), we may relax the MILP problems in (3.42), (3.43) and (3.44) and use the relaxed linear optimization problems introduced in (3.34), (3.35) and (3.36). In this case, the maximum number of successful packet transmissions for uplink and downlink are calculated by $\left\lfloor \frac{R_{\text{UE}}(t)}{L} \right\rfloor$ and $\left\lfloor \frac{R_{\text{EU}}(t)}{L} \right\rfloor$, respectively and therefore the achievable uplink and downlink rates are given by $L \left\lfloor \frac{R_{\text{UE}}(t)}{L} \right\rfloor$ and $L \left\lfloor \frac{R_{\text{EU}}(t)}{L} \right\rfloor$. Thus, for the discrete packet system, instead of solving the MILP problems, we solve the linear approximation of these problems given in (3.34), (3.35) and (3.36) and then we encode and transmit in uplink and downlink according to the rates $L \left\lfloor \frac{R_{\text{UE}}(t)}{L} \right\rfloor$ and $L \left\lfloor \frac{R_{\text{EU}}(t)}{L} \right\rfloor$. The *relative* performance loss due to such a relaxation is minor if packet length is small relative to the achievable rates $R_{\text{UE}}(t)$ and $R_{\text{EU}}(t)$. This is due to the fact that in both the discrete-packet system and continuous-flow system, the solutions of the back-pressure

⁵This is equivalent to the fact that $R_{\text{UE}}(t)$ and $R_{\text{EU}}(t)$ are much larger than L .

optimization problems are on the boundary of the capacity regions. When $\frac{R_{\text{UE}}(t)}{L}$ and $\frac{R_{\text{eNB}}(t)}{L}$ are large quantities, the boundaries of the capacity region in the discrete-packet and continuous-flow systems will be relatively very close. By applying the linear relaxation, the complexity of the scheduling and resource allocation phase is negligible since we already have the closed-form solutions for the linear problems. Hence, the real-time implementation of the hybrid scheme in practical relay-based mobile networks does not impose additional computational overhead on the system. Moreover, as we will see in the simulation results, the hybrid scheme substantially outperforms the conventional schemes (Direct, CoR and NC/CoR) in terms of throughput and average queueing delay.

3.4.3 Hybrid Scheme with Power Control

In Section 3.1, we mentioned that the average transmission power of UE, eNB and RS in each time-frame are constrained to fixed values P_{U} , P_{E} and P_{R} . In the Direct and NC/CoR schemes during a time-frame, each node only transmits in one subchannel while in CoR, the RS transmits in both subchannels f_2 and f_4 . Thus, in Direct and NC/CoR schemes UE, eNB and RS transmit in the corresponding subchannels with the maximum power, i.e., P_{U} , P_{E} and P_{R} , respectively. However, in CoR the relay station can adjust its power consumption in subchannels f_2 and f_4 such that it can satisfy the power constraint while improving the achievable rate rather than distributing the RS power equally. This provides more degrees of freedom since we can determine the jointly optimal values of the time sharing parameters $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ as well as the allocated relay power in subchannels f_2 and f_4 .

Assume that in the CoR scheme, the relay transmission power at subchannels f_2 and f_4 of time-frame t are denoted by $P_{\text{R}}^{(2)}(t)$ and $P_{\text{R}}^{(4)}(t)$, respectively. Then, these

variables must satisfy the following constraint.

$$\frac{\lambda_2 P_R^{(2)}(t) + \lambda_4 P_R^{(4)}(t)}{\lambda_2 + \lambda_4} \leq P_R \quad (3.45)$$

Therefore, the hybrid scheme with power control is similar to what was introduced in (3.35) except R_{CoR} is determined by solving the following non-linear problem.

$$R_{\text{CoR}} = \max_{\substack{R_{\text{UE}}(t) \leq \lambda_1 C_{\text{UR}}(t) \\ R_{\text{UE}}(t) \leq \lambda_1 C_{\text{UE}}(t) + \lambda_2 C_{\text{RE}}(t, P_R^{(2)}(t)) \\ R_{\text{EU}}(t) \leq \lambda_3 C_{\text{ER}}(t) \\ R_{\text{EU}}(t) \leq \lambda_3 C_{\text{EU}}(t) + \lambda_4 C_{\text{RU}}(t, P_R^{(4)}(t)) \\ \sum_{n=1}^4 \lambda_n \leq 1 \\ \lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0 \\ \frac{\lambda_2 P_R^{(2)}(t) + \lambda_4 P_R^{(4)}(t)}{\lambda_2 + \lambda_4} \leq P_R}} \{X_{\text{U}}(t-1)R_{\text{UE}}(t) + X_{\text{E}}(t-1)R_{\text{EU}}(t)\} \quad (3.46)$$

where $C_{\text{RE}}(t, P_R^{(2)}(t)) = B \log(1 + \frac{|h_{\text{RE}}(t)|^2 P_R^{(2)}(t)}{\sigma_{\text{RE}} d_{\text{RE}}^\alpha})$ and $C_{\text{RU}}(t, P_R^{(4)}(t)) = B \log(1 + \frac{|h_{\text{RU}}(t)|^2 P_R^{(4)}(t)}{\sigma_{\text{RU}} d_{\text{RU}}^\alpha})$. In our simulation results, we will observe that power control improves the performance of the system considerably in terms of achievable throughput and average queuing delay.

3.4.4 Complexity Analysis

In terms of implementation complexity, the four transmission schemes can be ranked from low to high as direct, CoR, NC/CoR, PLNC/CoR. PLNC/CoR requires stringent synchronization and joint decoding and could suffer from decoding error propagation due to channel estimation and quantization errors. Explicit characterizations of the achievable rate regions have been derived for all transmission schemes and solved using linear programming (LP). This makes the implementation practical for real networks. However, the hybrid scheme with power control problem defined in 3.4.3 is formulated as a Mixed Integer Nonlinear Programming (MINLP) problem, which is NP-hard [85].

Due to high evaluation costs, such problems are difficult to solve. Hence, centralized heuristic algorithms are used to generate approximate solutions. KNITRO solver offers a nonlinear branch-and-bound method for solving MINLP problems [86]. The solver is used to provide a near-optimal solution to our MINLP power control problem. Branch-and-bound seemed to be a robust method, however the time required to solve increases exponentially as the number of variables increases [87, 88]. Therefore, computational complexity becomes extremely prohibitive in dense networks.

3.5 Simulations and Discussion

A time slotted bidirectional relaying system is simulated in which UEs are randomly distributed in a $10\text{km} \times 10\text{km}$ plane. The UEs' positions are randomly selected across the cell for each time-frame. The UEs and the eNB are fed by exogenous Poisson packet arrival processes with rates Λ_U and Λ_E , respectively. Simulation experiments have been performed under various scenarios in order to assess the performance of the proposed hybrid transmission scheme and to quantify the performance improvement in terms of stable end-to-end throughput and queue backlog (i.e., average queueing delay) over the direct, CoR and NC/CoR transmission schemes. We also performed the simulations for different number of RSs which are placed at 50% of the distance of eNB and the cell edge and determined the best number of relays to deploy in the system. The relay selection algorithm used in our simulations is based on selecting the best relay station that has the shortest distance to UE. Furthermore, we considered the Physical Layer Network Coding/Cooperative Relaying (PLNC/CoR) scheme in our simulations and compared its performance with the other transmission schemes.

The average system throughput (capacity) and queue backlog are evaluated for each transmission scheme under SISO and MIMO configurations. We generated the

Table 3.1: Simulation Parameters for System Model

<i>Parameter</i>	<i>Setting</i>
System Bandwidth	10MHz
Cell Coverage	10km x 10km
Available Subchannels (N)	600
Number of time-frames	100,000
time-frame Duration	10 msec
Packet Length	1000 bits
Traffic Model	Poisson arrivals ($\Lambda_E = 10 \times \Lambda_U$)
Path-loss Exponent α	4
Channel Model	Rayleigh fading
Transmission Powers of eNB, RS and UE	46, 30 and 23 dBm
Noise PSD	-174 dBm/Hz
Number of UEs (N)	100 – 600
Number of RSs (M)	1, 2, 4, 6, 8, 10
Number of MIMO Antennas	eNB(4x4) RS(4x4) UE(2x4)

simulation codes in C++ and for solving the non-linear programs we used KNITRO solver [86]. Each data point collected is averaged over 100,000 time-frames and averaged over 10 simulation trials using a different random seed (UE position). The detailed simulation parameters are shown in Table 3.1.

3.5.1 Average System Queue Backlog (Delay) and Capacity Evaluation

Average Total Delay and the End-to-end Stable Throughput Region

We compute the average queue backlog performance which is indicative of the end-to-end delay performance of the above mentioned transmission schemes under a random

network topology using one relay station. Figure 3.6 shows the average queue backlog performance with different combinations of Λ_E and Λ_U for Direct, CoR, NC/CoR and the hybrid approach (without power control). In this 3D plot, the x and y axes are Λ_E and Λ_U and the z axis represents the total queue backlog in the system. It is implied from the figure that the hybrid transmission scheme provides a higher stable throughput region than the other transmission schemes and also results in lower backlog in the system. From Little's law [89] it follows that the hybrid scheme outperforms the other transmission schemes in terms of average total delay. In the rest of the simulations, we assume that $\Lambda_E = 10 \times \Lambda_U$ and we plot the total backlog vs. Λ_U in 2D plots.

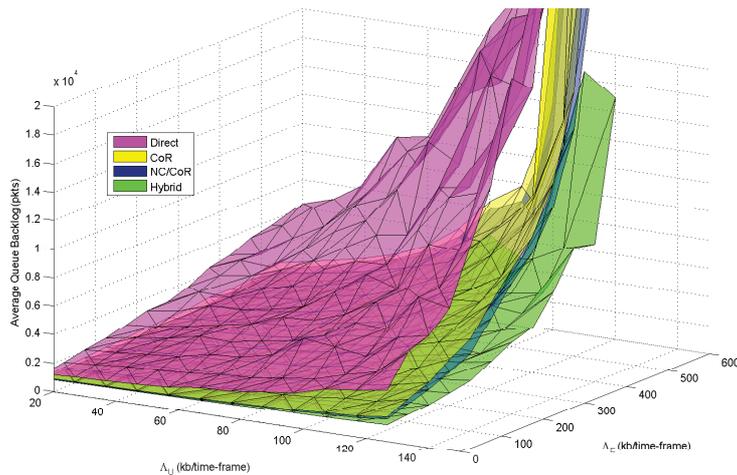


Figure 3.6: Delay and Throughput Performance of Direct, CoR, NC/CoR and Hybrid schemes (SISO)

Figures 3.7(a) and 3.7(b) show the performance plots for the average system queue backlog of different transmission schemes for SISO and MIMO configurations, respectively. It can be observed that as we increase the traffic rate, the average queue length also increases. However, for a certain rate the queue lengths will become excessively

large (i.e., in theory infinite) such that adding more traffic makes the system unstable and imposes significant delays. This rate marks the system end-to-end throughput region. As expected, it is shown that the hybrid scheme outperforms Direct, CoR and NC/CoR in terms of delay and end-to-end throughput.

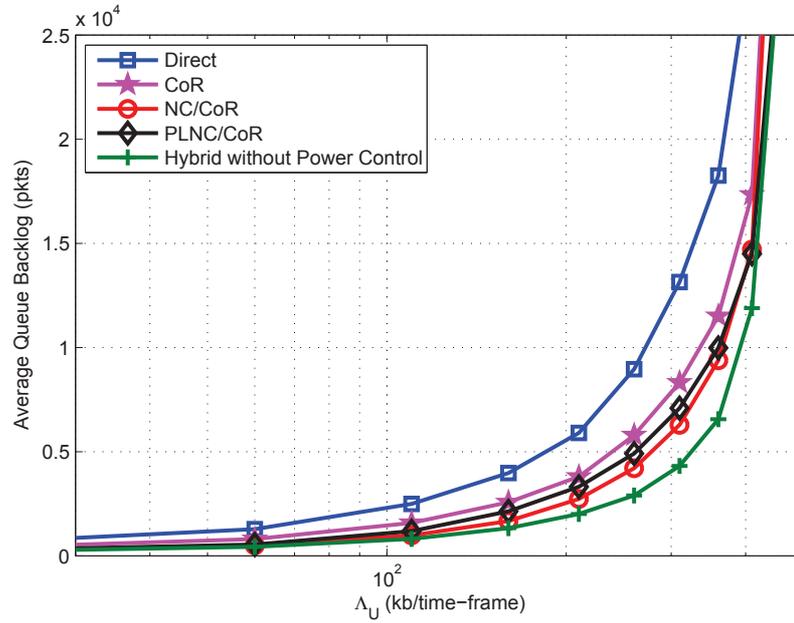
It is also shown in Figure 3.7(a) that the hybrid scheme under a SISO configuration provides 80 kb/time-frame and 20 kb/time-frame improvements over direct and NC/CoR transmission schemes.

Similar performance results for MIMO simulations are shown in Figure 3.7(b). However, the improvements are even higher for MIMO in which the hybrid scheme provides 600 kb/time-frame and 200 kb/time-frame improvement over the direct and NC/CoR transmission schemes which map to approximately 80% and 20% improvements, respectively.

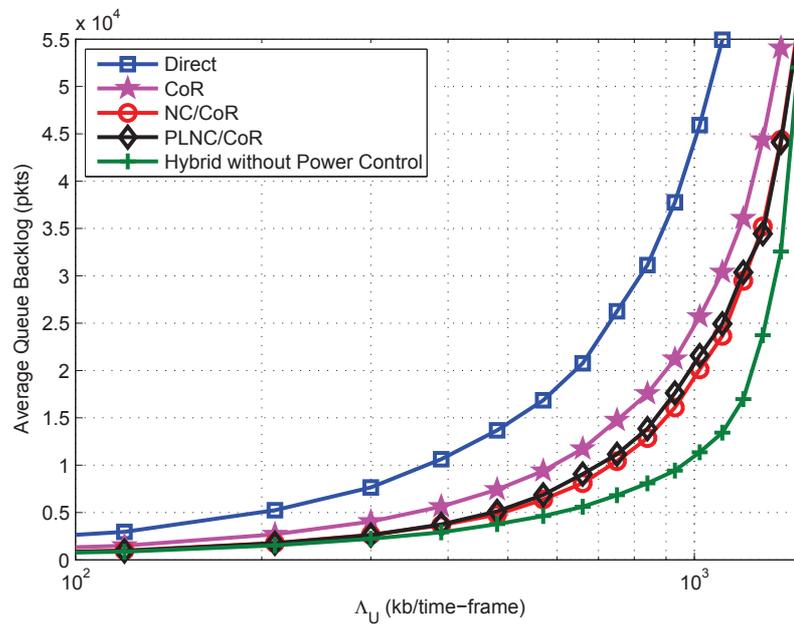
Impact of Number of Relays on Average System Queue Backlog

The average queue backlog performance is computed for the hybrid transmission scheme by increasing the number of UEs and RSs. The number of RSs is assumed to be 2, 4, 6, 8 and 10, the eNB is located at the middle of the cell and the UEs are randomly placed across the network. The simulations are performed for different number of RSs which are placed at 50% of the distance of eNB from the cell edge.

Figure 3.8(a) shows the average system queue length of the hybrid transmission scheme without power control for different number of RSs and for SISO configuration. It is shown that as we increase the number of RSs, the average stable throughput also increases. However, increasing the number of RSs above 8 relays does not provide significant impact on the average system capacity. This is because most of the UEs across the cell will be covered by the relay nodes and hence the maximum reachable

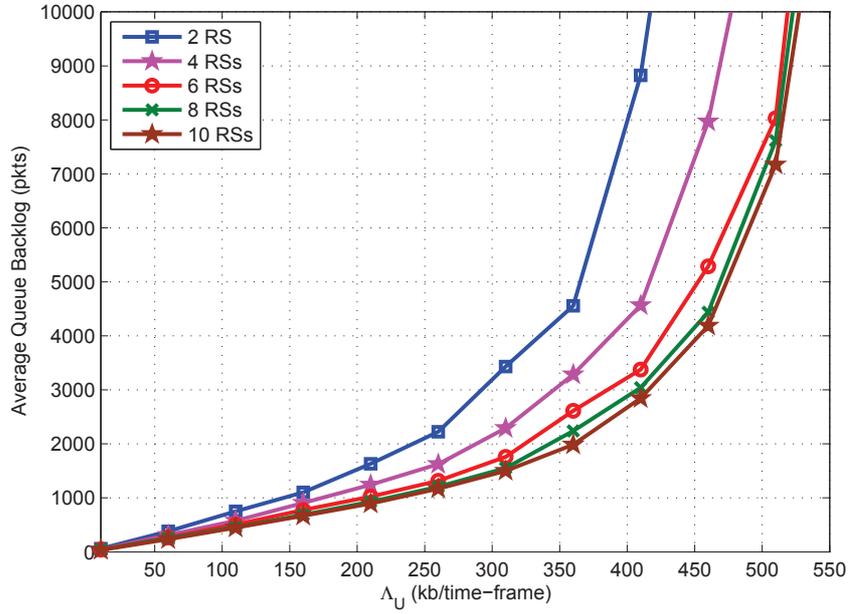


(a) SISO

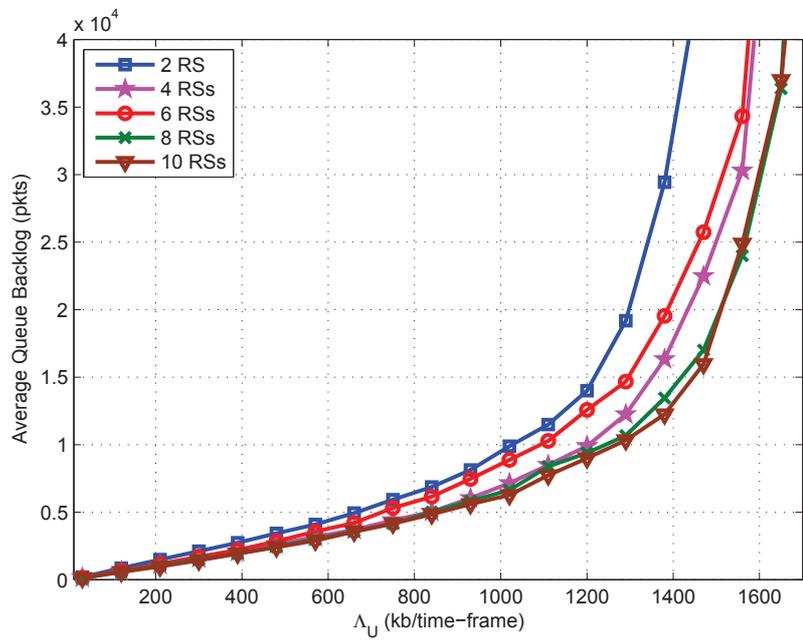


(b) MIMO

Figure 3.7: Capacity and Delay Performance of Different Transmission Schemes under SISO and MIMO Configurations



(a) SISO



(b) MIMO

Figure 3.8: Average Queue Backlog vs. Number of RSs

performance and improvements will eventually saturate. Therefore, adding more RSs would increase the operator cost without providing any gain. Figure 3.8(b) shows similar results under MIMO configuration.

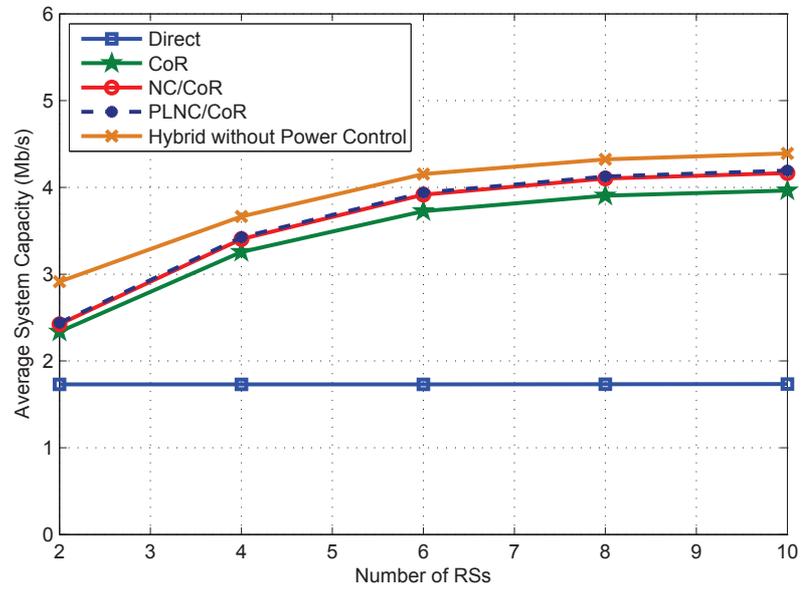
Average System Capacity (SISO and MIMO)

Figures 3.9(a) and 3.9(b) show the average system capacity (combined downlink and uplink) for various transmission schemes under SISO and MIMO configurations, respectively.

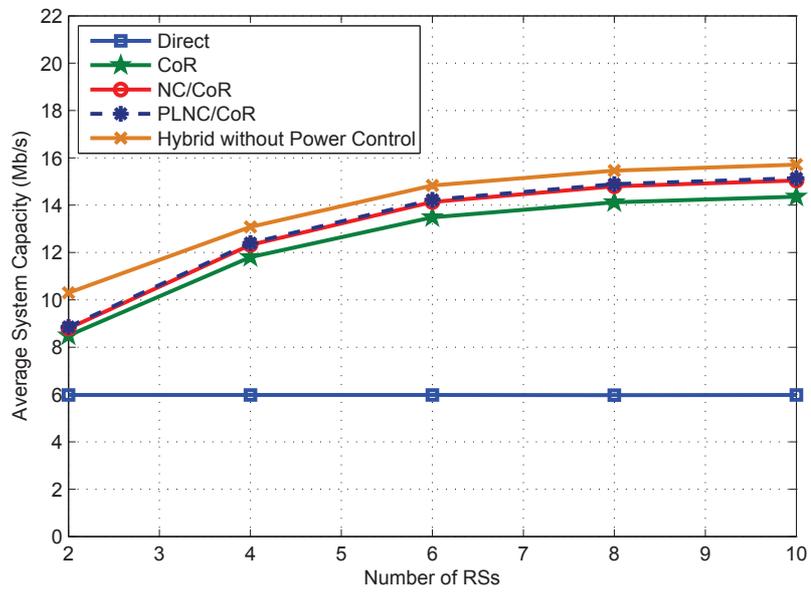
It is shown that the hybrid scheme outperforms the other schemes in terms of average system capacity for different number of relays as expected and proved in Section 3.4. The simulations are performed by increasing the number of RSs which are placed at 50% of the distance of eNB from the cell edge. It is also concluded that by increasing the number of RSs, the system capacity increases. However, increasing the number of RSs above 8 relays does not provide significant enhancement on the average system capacity.

3.5.2 Comparison with Physical Layer Network Coding

As it is shown in Figures 3.7(a) and 3.7(b), PLNC/CoR is performing very close to the NC/CoR from delay and capacity perspective. However, it is observed that in low traffic load NC/CoR outperforms PLNC/CoR while as the traffic load increases PLNC/CoR will dominate NC/CoR. This can be explained based on the following two points: First, PLNC/CoR in general provides higher transmission rates for the queues as compared to NC/CoR. Second, in the back-pressure algorithms provided in Section 3.4, the objective is to maximize the product of queue backlog and the achievable rate at each time-frame and the resources are assigned to maximize such



(a) SISO



(b) MIMO

Figure 3.9: Average System Capacity vs. Number of RSs

an objective. In PLNC/CoR with low traffic rates, since the queue backlogs are small, between queue backlog and achievable rate, the latter will play the dominant role in maximizing the back-pressure objective. This may result in offering larger rates to the queue with lower backlog. However, since there are not enough packets to be served, a big portion of the service may be wasted. In NC/CoR however the achievable rates are smaller than PLNC/CoR and therefore both queue backlog and achievable rate affect the resource allocation. Thus, the probability of wasting service is smaller. Hence, although the service offered in NC/CoR is slightly smaller than PLNC/CoR, it is being used more efficiently. In high traffic load, the queue backlog in both UE and eNB is large. Since there are sufficient amount of packets in both queues to be served, the probability of wasting service tends to zero in both PLNC/CoR and NC/CoR. Since PLNC/CoR is offering larger service rates to the queues, the queue backlog in PLNC/CoR will be smaller than NC/CoR.

We have also evaluated the average capacity performance of the PLNC/CoR and compared it with that of other schemes in Figures 3.9(a) and 3.9(b). As it is expected, PLNC/CoR is performing slightly better than NC/CoR. However, the hybrid approach still outperforms both schemes. As we can observe in these figures, NC/CoR and PLNC/CoR perform very closely in terms of average capacity/throughput. This is explained by considering the following points: First, comparing the derived rate region for NC/CoR and PLNC/CoR in Propositions 2 and 3 reveals that the rate region of PLNC/CoR is similar to the rate region of NC/CoR in most of the channel state realizations. There are cases where PLNC/CoR achieves a larger rate region. However, in those cases the difference of the rate regions is not significant and depending on the ratio of uplink and downlink rates can be even zero. Second, in our graphs we plot the *average* throughput (taken over possible channel states) of PLNC/CoR and NC/CoR calculated through simulations. Therefore, we expect to see slight improvement for

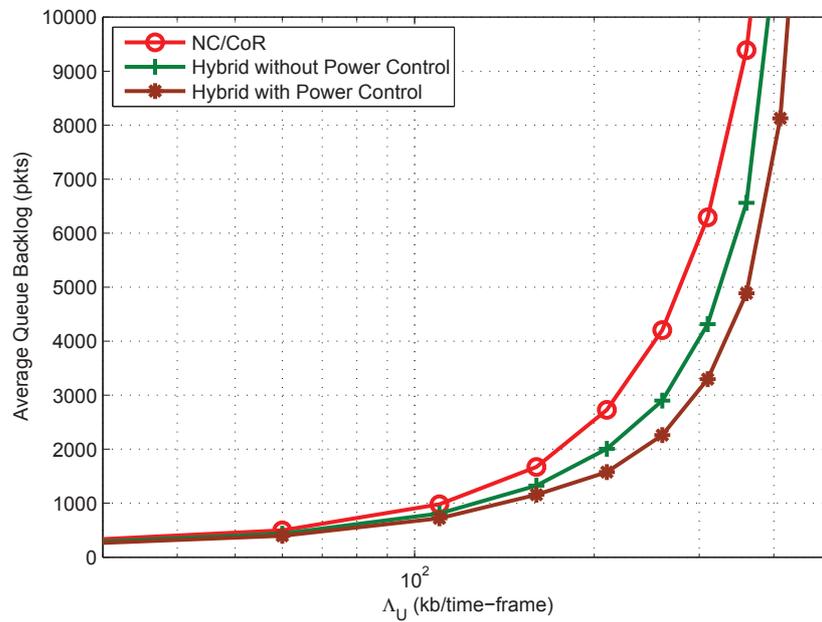
the average capacity performance of PLNC/CoR with respect to NC/CoR.

3.5.3 Performance Evaluation of the Hybrid Scheme with Power Control

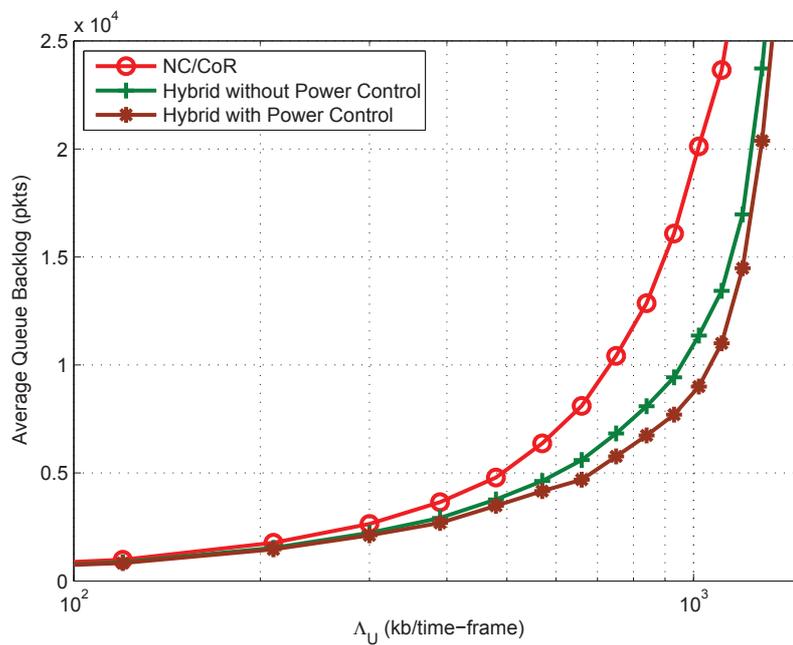
In this section, we will evaluate the performance of the hybrid approach with power control described in Section 3.4.3 and compare it with the hybrid approach without power control and NC/CoR. As it is observed in Figures 3.10(a) and 3.10(b), application of power control in the hybrid scheme improves the delay performance of the system significantly. Moreover, we evaluate the effect of power control in the capacity of the hybrid scheme in Figures 3.11(a) and 3.11(b). As we observe in the figures, power control significantly improves the total capacity of the system. However this performance improvement comes at the cost of increasing the computational and real-time implementation complexity of the system as the network controller must solve a separate non-linear problem for each user at each time-frame. Overall, we conclude that while the hybrid approach without power control is a promising technique to be applied in bidirectional relaying networks (by improving the throughput and delay performance of the system), it does not impose real-time complexity into the system for its implementation. Furthermore, at the expense of the complexity of power control, the performance of the system can be further enhanced.

3.6 Chapter Summary

In this chapter, a cross-layer design for optimal joint resource allocation and power control in bidirectional relaying networks is presented. The work assumes that each bidirectional link has a dedicated channel in the multichannel system. Direct, CoR,

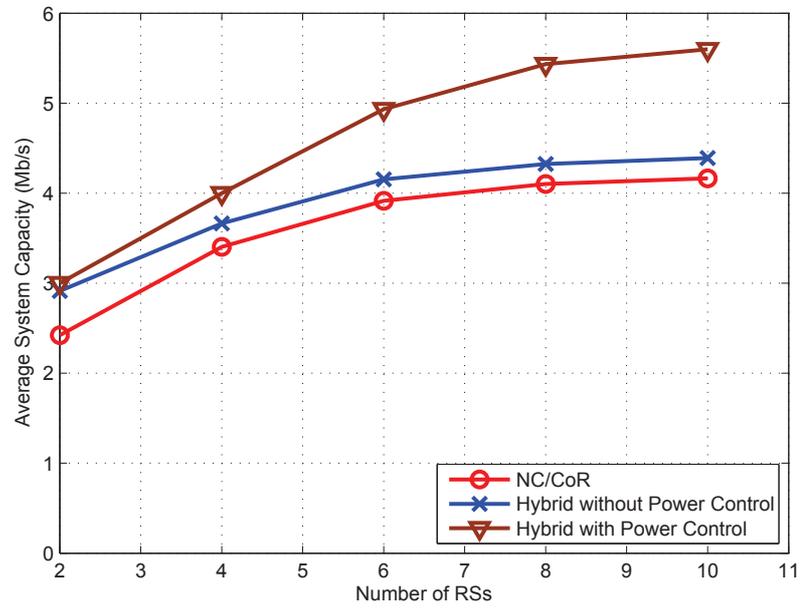


(a) SISO

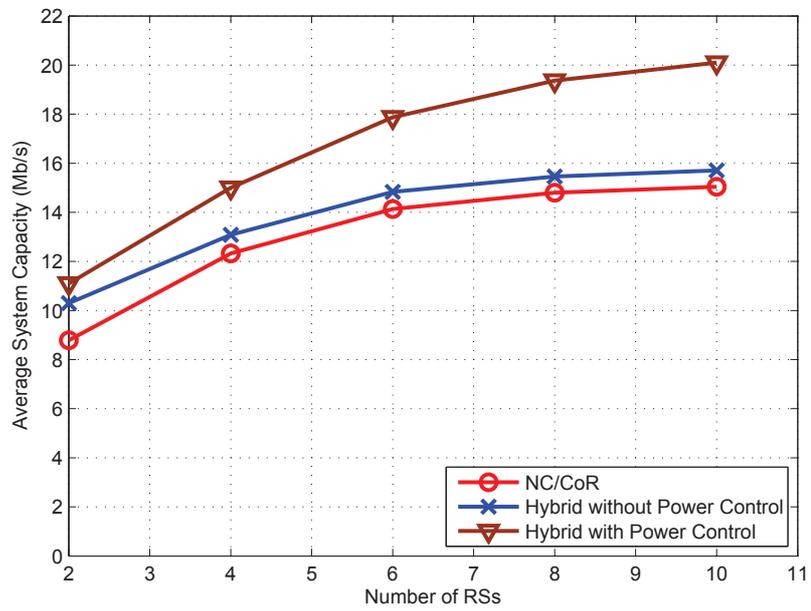


(b) MIMO

Figure 3.10: Network Capacity and Delay Performance of the Hybrid Scheme with Power Control



(a) SISO



(b) MIMO

Figure 3.11: Average System Capacity vs. Number of RSs - Hybrid Scheme with Power Control

and NC/CoR TWRC transmission schemes are considered and their rate regions are explicitly characterized. It is shown that each of the supported schemes can outperform the others depending on the channel states, system backlog and downlink and uplink traffic. Therefore, a hybrid scheme which dynamically determines the optimal transmission scheme and optimal time allocation at each time-frame is proposed. In addition, power control in the system is considered and optimal joint power control and resource allocation scheme is proposed. The new scheme outperforms the previous transmission schemes in terms of delay and throughput. However, this improvement comes at the cost of adding real-time computational complexity into the system. It is also shown that the average system capacity increases by deploying more relay nodes because more user equipment (UEs) benefit from the resource savings using the NC/CoR scheme. However, deploying more than 6 relays in the system does not have considerable impact on the system performance.

Chapter 4

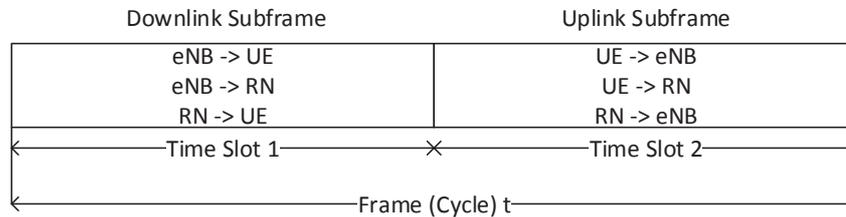
Joint Resource Allocation and Relay Selection in OFDMA Networks

In this chapter, a relay-assisted OFDMA wireless cellular network is considered. An enhanced three-time-slot per cycle Time-Division Duplexing (TDD) frame architecture is proposed. An optimization framework is proposed for the joint resource allocation, subchannel pairing, transmission scheme selection and relay selection problem. The joint optimization problem is a combinatorial NP-complete problem. To make it more tractable, we adopt two approaches to find optimal and suboptimal solutions to our joint optimization problem. First, we adopt a graph-based approach by establishing an equivalent Maximum Weighted Clique Problem (MWCP). A modified Bron-Kerbosch algorithm is proposed to provide an optimal solution. The Bron-Kerbosch algorithm is a well-known branch and bound algorithm which finds an exact solution for the MWCP Problem [73]. It is used as a benchmark to evaluate our suboptimal algorithms for small size problems. In addition, a sequential greedy heuristic algorithm is used to provide a fast suboptimal solution to the MWCP problem. Second, the joint combinatorial optimization problem is also mapped into a Three-Dimensional Assignment Problem (3DAP). An enhanced 3-dimensional Ant

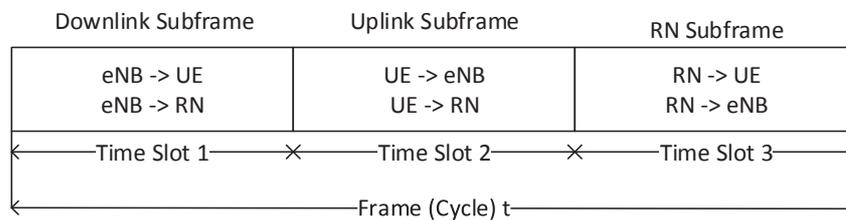
Colony Optimization (ACO) algorithm [90] is proposed to provide a suboptimal solution in polynomial time. From simulation results and analysis, we can conclude that the proposed ACO-based algorithm provides better approximation results than the greedy heuristic algorithm in real environments.

4.1 System Description and Assumptions

In this work, a bidirectional network is modeled by a Two-Way Relay Channel (TWRC) where the end-nodes (UE and eNB) communicate with each other either directly or with cooperation of a relay node (RN). The supported transmission schemes are direct, Cooperative Relaying (CoR) and Network Coding/Cooperative Relaying (NC/CoR) which have been discussed in Section 3.1. In CoR, four network resources are used to send a pair of packets P_1 and P_2 between the UE and eNB. In uplink (UL), the UE transmits P_1 to the RN and is overheard by the eNB in the first network resource (subchannel). The relay then forwards the packet to the eNB in the second subchannel. In CoR, it is required that both the relay and eNB decode P_1 without error [8]. Similarly, two network resources are required to transmit P_2 in downlink (DL). In NC/CoR, three network resources are used to send P_1 and P_2 between the UE and eNB. In UL, the UE transmits P_1 to RN using one subchannel. In DL, the eNB transmits P_2 to RN using a second subchannel. Then, RN performs network coding (bit-wise XOR) on P_1 and P_2 as $P_3 = P_1 \oplus P_2$ and broadcasts it as one packet to both UE and eNB using a third subchannel. Each receiver can then obtain its intended packet by XORing the received packet from RN with its own packet. Similar to CoR, P_1 (resp. P_2) is overheard by the eNB (resp. UE) in the UL (resp. DL) where each end-node will have two versions of its intended packet. Cooperative diversity detection is then be applied on the two versions of the received packet [26].



(a) Legacy Frame Architecture



(b) Proposed Frame Architecture

Figure 4.1: OFDMA Frame Architectures

4.1.1 Proposed Frame Architecture

We consider a single cell OFDMA wireless Time-Division Duplexing (TDD) network in which each transmission frame (cycle) is divided into a downlink (DL) subframe and an uplink (UL) subframe. The UL is separated from DL by the allocation of different time slots in the same frequency band. The proposed frame architecture can be adapted to next-generation networks such as LTE and LTE-Advanced [4]. Figure 4.1(a) illustrates a traditional legacy TDD frame architecture with relay support, in which a bidirectional cycle is divided into DL and UL subframes. The legacy frame architecture is not suitable to support the Network Coding (NC) transmission scheme as can be seen in Figure 4.1(a). More specifically, the relay node (RN) performs network coding (bit-wise XOR) on packets received from the UE and eNB and broadcasts it as one packet to both UE and eNB. This requires the use of both UL and

DL resources (subchannels) in one transmission which is not possible in Figure 4.1(a). Hence, to support relay-assisted cooperative transmission and network coding, we introduce a new RN subframe to support CoR and/or NC/CoR transmissions. The resulting frame architecture consists of a three-time-slot per cycle frame architecture as shown in Figure 4.1(b). Each subframe consists of \mathcal{N} orthogonal network resources (subchannels) in which each UE can be tuned to one or more subchannels. The eNB node is capable of transmitting/receiving information on different channels during a given time slot. Moreover, to avoid wasting unused subchannels in the RN subframe during direct transmission, we assume that the subchannels in the RN subframe are split equally between UL and DL.

4.1.2 Channel Model

We denote the link capacity from node x to node y ¹ at cycle t by $C_{xy}(t)$, which is given by $C_{xy}(t) = F(B, h_{xy}(t), P_x, \sigma_{xy}, d_{xy}, \alpha)$ where $C_{xy}(t)$ is measured in bits/s. $F(\cdot)$ represents a general function to express the channel capacity. The channel capacity is expressed by the Shannon capacity formula ($\log_2(1 + \gamma_{xy}(t))$), where $\gamma_{xy}(t)$ is the signal-to-noise-ratio (SNR) between nodes x and y . The bandwidth of the employed subchannel is B , P_x is the transmission power of node x , and σ_{xy} is noise power. $h_{xy}(t)$ is the channel gain of link $x \rightarrow y$ which is modelled by an independent and identically distributed complex Gaussian process that remains constant within each cycle t (slow fading channels). The distance of nodes x and y is given by d_{xy} and α is the path-loss exponent. If the $x \rightarrow y$ channel is an $M \times N$ MIMO channel (with N transmitting and M receiving antennas), $h_{xy}(t)$ will be a complex matrix. It is assumed that the relays operate in Decode-and-Forward (DF) mode. Henceforth, we use subscripts U, R and E , in our formulas to denote UE, RN and eNB, respectively. The

¹A node can be a UE, an eNB or an RN

channel state of the TWRC at cycle t is denoted by vector $C(t) = (C_{UE}(t), C_{EU}(t), C_{UR}(t), C_{RU}(t), C_{ER}(t), C_{RE}(t))$, which contains the capacities of all the communication links in the TWRC.

4.1.3 Queuing Model

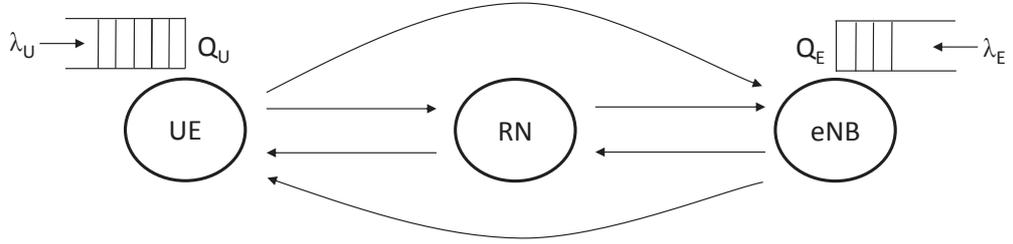


Figure 4.2: Two-Way Relay Queuing Model

We consider a two-way relay network with sources UE and eNB, and relay RN, as shown in Figure 4.2. Each UE and eNB buffer their incoming packets in queues with lengths $Q_U(t)$ and $Q_E(t)$ at cycle t , respectively. Each active UE in the system can use one or more subchannels. The eNB contains M queues, one for each active UE, for downlink transmissions. For uplink transmissions, one queue is available per UE. Sources UE and eNB choose the service rates $(R_{UE}(t), R_{EU}(t))$ at cycle t , where $R_{UE}(t)$ and $R_{EU}(t)$ are the end-to-end achievable rates in UL and DL, respectively. The service rates are chosen as the maximum of the achievable rates of the supported transmission schemes.

The packet arrival processes to UE and eNB are denoted by $\{A_U(t)\}_{t=1}^{\infty}$ and $\{A_E(t)\}_{t=1}^{\infty}$, respectively. $A_U(t)$ and $A_E(t)$ are the random Poisson processes representing the number of packets arriving exogenously to UE and eNB at cycle t . These processes are assumed to be stationary with rates Λ_U and Λ_E , respectively.

The packets are assumed to have fixed length L . The UE and eNB nodes keep the incoming packets in separate buffers. We assume that these buffers are not limited in size, i.e., we do not have a packet drop in the system. Let $Q_U(t)$ and $Q_E(t)$ denote the queue lengths of UE and eNB at the end of cycle t , respectively. At each cycle t , a centralized scheduler (usually in the eNB) observes the queue state of the system as well as the channel condition and solves the joint resource allocation, subchannel pairing, transmission scheme selection and relay selection optimization problem. At the end of each cycle, and after serving the packets according to the scheduling policy, the new arrivals are added to the queues, and the queue lengths are updated.

We assume the uplink and downlink traffic rates are the same (i.e., symmetric case). Therefore, the number of UL and DL subframes is assumed to be equal. However, the frame architecture is still applicable for asymmetric traffic demands in which the new three-time-slots can be formed for DL and UL subframes with bidirectional traffic. Such symmetric and asymmetric frame configurations exist in LTE networks.

4.2 Optimization Framework

In this section, we present the joint resource allocation, subchannel pairing, transmission scheme selection and relay selection optimization framework.

4.2.1 Achievable Bidirectional Rate Pairs

In this section, we will derive and study the achievable DL and UL rate pairs of direct, CoR and NC/CoR transmission schemes for Decode-and-Forward (DF) relaying. Tuples (i, j, k) correspond to the subchannel indexes of the DL, UL and RN subframes, respectively, where $i, j, k \in \mathcal{N}$ and \mathcal{N} is the number of subchannels per subframe.

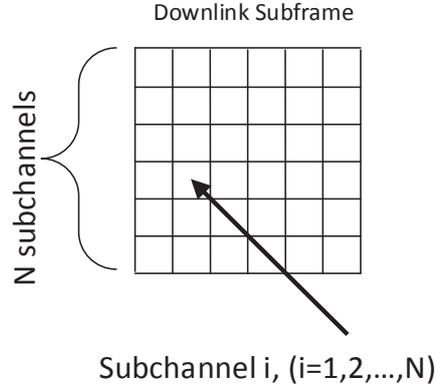


Figure 4.3: Downlink Subframe with \mathcal{N} Subchannels (one time-slot)

Resources are allocated in DL, UL and RN subframes in units called subchannels. Figure 4.3 illustrates a sample DL subframe with \mathcal{N} subchannels where $i \in \mathcal{N}$ is a subchannel in the DL subframe. Similar subframes are used for UL (index j) and RN (index k).

Direct Transmission

In direct transmission, the relay node (RN) is not used in the bidirectional transmission. However, we assume that the subchannel in the unused RN subframe is split equally between UL and DL. In this way, the subchannel in the RN subframe is utilized. Figure 4.4 illustrates the bidirectional communication for the direct transmission scheme with respect to the proposed frame architecture shown in Figure 4.1(b). Note that the subchannel pairing in the bidirectional direct transmission scheme involves three subchannels: subchannel i in the DL subframe from the eNB to UE bidirectional pair, subchannel j in the UL subframe from the UE to eNB bidirectional pair, and subchannel k in the RN subframe in which the RN subframe is split

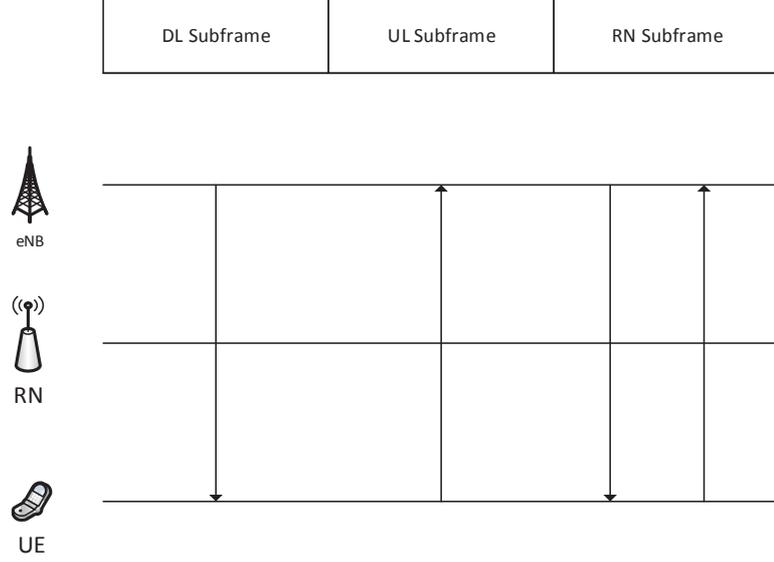


Figure 4.4: Direct Transmission Scheme

equally between UL and DL for the same bidirectional pair. The uplink and downlink achievable rate pair for the direct transmission scheme over the DL, UL and RN subframes for the subchannel tuples (i, j, k) is denoted by $(R_{EU}^{i,k}(t), R_{UE}^{j,k}(t))$, in which

$$R_{EU}^{i,k}(t) = \frac{1}{3} \left(C_{EU}^i(t) + \frac{1}{2} C_{EU}^k(t) \right) \quad , \quad R_{UE}^{j,k}(t) = \frac{1}{3} \left(C_{UE}^j(t) + \frac{1}{2} C_{UE}^k(t) \right) \quad (4.1)$$

where the factor $\frac{1}{3}$ is a result of the use of the proposed three-time-slot frame architecture. The achievable downlink rate is denoted by $R_{EU}^{i,k}(t)$, where subchannel i in the DL subframe is used for eNB to UE transmission and half of subchannel k in the RN subframe is used for the same eNB to UE transmission. Similarly, the achievable uplink rate is denoted by $R_{UE}^{j,k}(t)$, where subchannel j in the UL subframe is used for UE to eNB transmission and half of subchannel k in the RN subframe is used for the same UE to eNB transmission.

CoR Transmission

The cooperative relaying (CoR) transmission scheme can be explained using a 3-node network (source, relay and destination). In CoR, the relay processes the signal it receives from the source. Under a repetition-coded scheme, the relay transmits the signal in the RN subframe. Since we assume Decode-and-Forward (DF) relaying, it entails the possibility of forwarding erroneously-detected symbols to the destination. Therefore, it is required that both the relay and destination decode all information without error. This leads to the expression of maximum average mutual information as the minimum of the two rates: one at which the relay can reliably decode the source message, and the other at which the destination can reliably decode the source message given repeated transmissions from the source and relay. It can be observed that the performance of CoR is limited by the worst link between the (source-destination) and (source-relay-destination). The cooperative rate requires that both the relay and destination fully decode the received data [8]. According to [8], the uplink and downlink achievable rate pair for the CoR scheme over the DL, UL and RN subframes for the subchannel tuples (i, j, k) is denoted by $(R_{EU}^{i,k}(t), R_{UE}^{j,k}(t))$, in which

$$\begin{aligned} R_{EU}^{i,k}(t) &= \frac{1}{3} \min (C_{ER}^i(t), DF^{i,k}(\gamma_{EU}^i(t), \gamma_{eff}^1(t))), \\ R_{UE}^{j,k}(t) &= \frac{1}{3} \min (C_{UR}^j(t), DF^{j,k}(\gamma_{UE}^j(t), \gamma_{eff}^2(t))) \end{aligned} \quad (4.2)$$

where $DF(\gamma_1(t), \gamma_2(t))$ is the capacity of the cooperative transmission using the DF scheme where $\gamma_1(t)$ and $\gamma_2(t)$ are the SNR of source-destination and relay-destination links, respectively. Since in our frame architecture, the subchannels in the RN subframe are split for UL and DL transmissions, we need to use effective SNR for the relay links reflected in (4.2) as $\gamma_{eff}^1(t) = \sqrt{1 + \gamma_{RU}^k(t)} - 1$ and $\gamma_{eff}^2(t) = \sqrt{1 + \gamma_{RE}^k(t)} - 1$ [91, 92]. The $DF(\gamma_1(t), \gamma_2(t))$ for the case of Shannon capacity formula is given by a

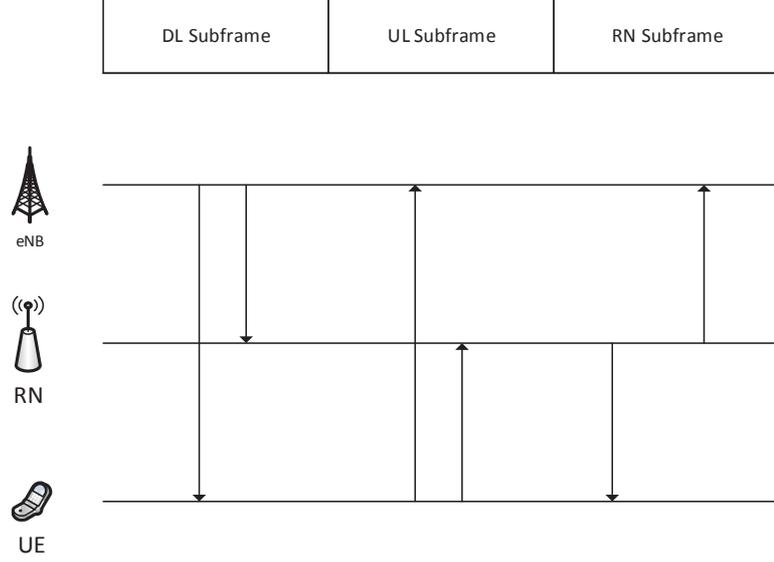


Figure 4.5: CoR Transmission Scheme

\log function as $\log(1 + \gamma_1(t) + \gamma_2(t))$. We only use three subchannels for the bidirectional communication to complete the four transmissions. The achievable downlink rate is denoted by $R_{EU}^{i,k}(t)$, where subchannel i in the DL subframe is used for eNB to UE/RN transmission and subchannel k in the RN subframe is used for the RN to UE transmission. Similarly, the achievable uplink rate is denoted by $R_{UE}^{j,k}(t)$, where subchannel j in the UL subframe is used for UE to eNB/RN transmission and subchannel k in the RN subframe is used for the RN to eNB transmission.

Figure 4.5 illustrates the bidirectional communication for the CoR scheme with respect to the proposed frame architecture. Note that the subchannel pairing in CoR involves three subchannels: the two incoming subchannels (i, j) to the relay from the UE and eNB bidirectional pair, and a third common outgoing subchannel k to the same bidirectional pair. Half of subchannel k is used for the RN to UE transmission and the second half for the RN to eNB transmission.

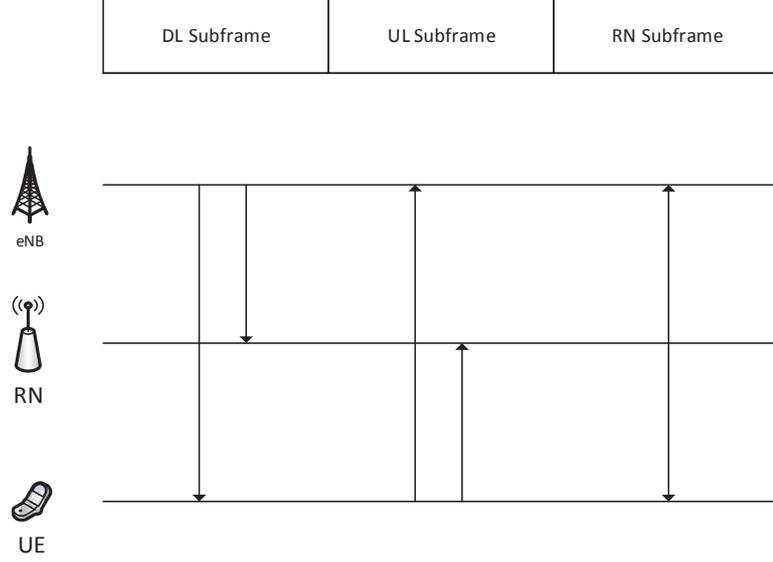


Figure 4.6: NC/CoR Transmission Scheme

NC/CoR Transmission

In NC/CoR, the RN combines the received signals (bitwise XOR) from UE and eNB and broadcasts it to both of them in the RN subframe. Figure 4.6 illustrates the bidirectional communication for the NC/CoR scheme with respect to the proposed frame architecture. Note that for the three-phase NC/CoR two-way relaying scheme, the subchannel pairing involves three subchannels: the two incoming subchannels (i, j) to the relay from the UE and eNB, respectively, and the common outgoing subchannel k to the same UE-eNB bidirectional pair. According to [26], the uplink and downlink achievable rate pair for the NC/CoR scheme over the DL, UL and RN subframes for the subchannel tuples (i, j, k) is denoted by $(R_{EU}^{i,k}(t), R_{UE}^{j,k}(t))$, in which

$$\begin{aligned} R_{EU}^{i,k}(t) &= \frac{1}{3} \min (C_{ER}^i(t), C_{EU}^i(t) + C_{\min}^k(t)), \\ R_{UE}^{j,k}(t) &= \frac{1}{3} \min (C_{UR}^j(t), C_{UE}^j(t) + C_{\min}^k(t)), \end{aligned} \quad (4.3)$$

where $C_{\min}^k(t) = \min(C_{RU}^k(t), C_{RE}^k(t))$. $C_{\min}^k(t)$ represents the maximum rate at which both eNB and UE can reliably decode the signals from RN during the RN subframe. The achievable downlink rate is denoted by $R_{EU}^{i,k}(t)$, where subchannel i in the DL subframe is used for eNB to UE/RN transmission. Similarly, the achievable uplink rate is denoted by $R_{UE}^{j,k}(t)$, where subchannel j in the UL subframe is used for UE to eNB/RN transmission. Subchannel k in the RN subframe is used for both DL and UL where the RN performs network coding and then broadcasts an XORed packet to both the UE and eNB.

4.2.2 Optimization Problem Formulation

We consider a wireless communication system which contains one eNB with a number of fixed associated relay nodes and UE nodes. Let $\mathcal{L} = \{1, 2, \dots, L\}$ be the set of RNs, $\mathcal{M} = \{1, 2, \dots, M\}$ the set of UEs and $\mathcal{N} = \{1, 2, \dots, N\}$ the set of subchannels. The decision binary variable $\rho_{m,direct}^{i,j,k}(t)$ is equal to 1 if subchannel tuple (i, j, k) in the DL, UL and RN subframes is assigned to UE m using direct transmission mode or 0 otherwise. The decision binary variable $\rho_{m,l,CoR}^{i,j,k}(t)$ is equal to 1 if subchannel tuple (i, j, k) in the DL, UL and RN subframes is assigned to UE m and RN l using CoR or 0 otherwise. The decision binary variable $\rho_{m,l,NC/CoR}^{i,j,k}(t)$ is equal to 1 if subchannel tuple (i, j, k) in the DL, UL and RN subframes is assigned to UE m and RN l using NC/CoR or 0 otherwise.

We assume that each subchannel can only operate in one of the three transmission modes. Our objective is to maximize the total product of backlog and achievable rate by allocating subchannel tuples optimally, selecting the best RNs and transmission schemes for each bidirectional communication per-cycle. The objective function is based on the back-pressure principle which provides throughput-optimality and queue

$$\begin{aligned}
\max \quad & \sum_{m \in \mathcal{M}, i, j, k \in \mathcal{N}} (R_{EU, m, \text{direct}}^{i, k}(t) Q_{E_m}(t-1) + R_{UE, m, \text{direct}}^{j, k}(t) Q_{U_m}(t-1)) \rho_{m, \text{direct}}^{i, j, k} \\
& + \sum_{m \in \mathcal{M}, l \in \mathcal{L}, i, j, k \in \mathcal{N}} (R_{EU, m, l, \text{CoR}}^{i, k}(t) Q_{E_m}(t-1) + R_{UE, m, l, \text{CoR}}^{j, k}(t) Q_{U_m}(t-1)) \rho_{m, l, \text{CoR}}^{i, j, k} \\
& + \sum_{m \in \mathcal{M}, l \in \mathcal{L}, i, j, k \in \mathcal{N}} (R_{EU, m, l, \text{NC/CoR}}^{i, k}(t) Q_{E_m}(t-1) + R_{UE, m, l, \text{NC/CoR}}^{j, k}(t) Q_{U_m}(t-1)) \rho_{m, l, \text{NC/CoR}}^{i, j, k} \\
\text{s.t.} \quad & \sum_{m \in \mathcal{M}, j, k \in \mathcal{N}} \rho_{m, \text{direct}}^{i, j, k} + \sum_{m \in \mathcal{M}, l \in \mathcal{L}, j, k \in \mathcal{N}} (\rho_{m, l, \text{CoR}}^{i, j, k} + \rho_{m, l, \text{NC/CoR}}^{i, j, k}) \leq 1, \forall i \in \mathcal{N} \\
& \sum_{m \in \mathcal{M}, i, k \in \mathcal{N}} \rho_{m, \text{direct}}^{i, j, k} + \sum_{m \in \mathcal{M}, l \in \mathcal{L}, i, k \in \mathcal{N}} (\rho_{m, l, \text{CoR}}^{i, j, k} + \rho_{m, l, \text{NC/CoR}}^{i, j, k}) \leq 1, \forall j \in \mathcal{N} \\
& \sum_{m \in \mathcal{M}, i, j \in \mathcal{N}} \rho_{m, \text{direct}}^{i, j, k} + \sum_{m \in \mathcal{M}, l \in \mathcal{L}, i, j \in \mathcal{N}} (\rho_{m, l, \text{CoR}}^{i, j, k} + \rho_{m, l, \text{NC/CoR}}^{i, j, k}) \leq 1, \forall k \in \mathcal{N}
\end{aligned} \tag{4.4}$$

$$\tag{4.5}$$

$$\tag{4.6}$$

$$\tag{4.7}$$

stability [42]. Therefore, the proposed hybrid transmission scheme can stabilize the system for all the traffic rates inside the network stability region, which is proved based on applying the Lyapunov stability technique [84] and is presented in Appendix D. The objective function is formulated as the maximization Problem (4.4) subject to the constraints (4.5)–(4.7). $Q_{E_m}(t)$ and $Q_{U_m}(t)$ are the queue lengths associated to UE m in the eNB and UE at the end of cycle t , respectively.

Problem (4.4) is a characterization of the maximum weighted rate (total product of backlog and achievable downlink/uplink rates) of each UE-eNB bidirectional communication over the direct, CoR and NC/CoR transmission schemes. The decision binary variables in constraints (4.5)–(4.7) indicate that each subchannel tuple (i, j, k) can only be assigned to one UE m in direct transmission and one RN l in CoR and NC/CoR per-cycle. Moreover, the bidirectional communication session on each subchannel tuple (i, j, k) can only operate in one of the three transmission schemes. For

$$\max \sum_{i,j,k \in \mathcal{N}} W_1(i, j, k) \rho_{m^*, \text{direct}}^{i,j,k} + \sum_{i,j,k \in \mathcal{N}} W_2(i, j, k) \rho_{m^*, l^*, s^*}^{i,j,k} \quad (4.8)$$

$$W_1(i, j, k) = \max_{m \in \mathcal{M}} (R_{EU, m, \text{direct}}^{i,k}(t) Q_{E_m}(t-1) + R_{UE, m, \text{direct}}^{j,k}(t) Q_{U_m}(t-1)) \quad (4.9)$$

$$W_2(i, j, k) = \max_{m \in \mathcal{M}, l \in \mathcal{L}} ((R_{EU, m, l, \text{CoR}}^{i,k}(t) Q_{E_m}(t-1) + R_{UE, m, l, \text{CoR}}^{j,k}(t) Q_{U_m}(t-1)), \quad (4.10)$$

$$\text{s.t.} \quad \sum_{j,k \in \mathcal{N}} \rho_{m^*, \text{direct}}^{i,j,k} + \sum_{j,k \in \mathcal{N}} \rho_{m^*, l^*, s^*}^{i,j,k} \leq 1, \forall i \in \mathcal{N} \quad (4.11)$$

$$\sum_{i,k \in \mathcal{N}} \rho_{m^*, \text{direct}}^{i,j,k} + \sum_{i,k \in \mathcal{N}} \rho_{m^*, l^*, s^*}^{i,j,k} \leq 1, \forall j \in \mathcal{N} \quad (4.12)$$

$$\sum_{i,j \in \mathcal{N}} \rho_{m^*, \text{direct}}^{i,j,k} + \sum_{i,j \in \mathcal{N}} \rho_{m^*, l^*, s^*}^{i,j,k} \leq 1, \forall k \in \mathcal{N} \quad (4.13)$$

example, based on constraints (4.5)–(4.7), the tuple assignments (1, 1, 0) and (1, 0, 0) are not feasible since subchannel i in the DL subframe and subchannel k in the RN subframe are the same. The tuple assignments (1, 0, 0) and (0, 1, 1) are feasible as the subchannel tuple (i, j, k) for both assignments are disjoint. In addition, each tuple assignment can only operate in one of the direct, CoR and NC/CoR transmission schemes.

4.2.3 Modified Optimization Problem

Problem (4.4) is a combinatorial optimization problem whose computational complexity increases drastically with the number of binary variables [93]. By observing the first summation term of Problem (4.4), it is easy to find that there is at most one non-zero element for a given subchannel tuple (i, j, k) due to the constraints (4.5)–(4.7). This implies that among the M UEs, at most one UE m can occupy the subchannel

tuple (i, j, k) for direct transmission in the DL, UL and RN subframes. By observing the second and third summation terms of Problem (4.4), one also finds at most one non-zero element for a given subchannel tuple (i, j, k) for CoR and NC/CoR, respectively. This implies that the subchannel tuple can only be assigned to one relay node l to help one UE m for CoR or NC/CoR.

Based on the above observations, we define $W_1(i, j, k)$ in (4.9) as the weight function for the subchannel tuple (i, j, k) in direct transmission and $W_2(i, j, k)$ in (4.10) as the weight function for the subchannel tuple (i, j, k) in CoR and NC/CoR. Then, for maximizing the system total objective function, Problem (4.4) can be represented as the maximization Problem (4.8) subject to the new constraints (4.11)–(4.13). The term m^* in the first summation term represents the UE that maximizes the term in (4.9). The terms m^* , l^* and s^* in the second summation represent the UE index, RN index and transmission scheme (CoR or NC/CoR) that maximizes the term in (4.10), respectively. The new problem boils down to the subchannel assignment problem in the DL, UL and RN subframes for direct, CoR and NC/CoR transmissions.

4.3 Graph-Based Optimization

Conventional approaches to solve mixed-integer combinatorial problems, such as exhaustive search, are computationally inefficient since our optimization problem needs to be called at each cycle. We propose to solve our problem defined in (4.8) by transforming it into a Maximum Weighted Clique Problem (MWCP). A clique is a fully connected subgraph of a graph and a maximum clique is the clique with the largest number of vertices in a given graph. In other words, a clique is a subset S of vertices in a graph G such that each pair of vertices in S is connected by an edge. The maximum clique problem is the problem of finding, in a given graph, the clique with

the largest number of vertices. A Maximum Weighted Clique Problem (MWCP) is a clique of a weighted graph with maximum total vertex weights.

Finding all cliques using exhaustive search is expensive as the number of cliques can grow exponentially with every node (vertex) added. The Bron-Kerbosch algorithm [73] is an algorithm used to solve the maximal clique problem. However, the Bron-Kerbosch algorithm for clique search is only investigated for graphs without weights. In this section, a modified Bron-Kerbosch algorithm is used to solve our combinatorial MWCP problem in a weighted graph.

4.3.1 Maximum Weighted Clique Problem (MWCP)

Let $G = (V, E)$ be an undirected graph with vertex set V and edge set $E \subset V \times V$. A clique C of G is a subset of V such that every two vertices are pairwise adjacent, i.e., $\forall u, v \in C$ we have $u, v \in E$. A clique is maximal if it is not contained in any other clique. A clique is maximum if its cardinality is the largest among all the cliques of the graph. The Maximum Clique Problem (MCP) is the problem of finding the maximal cliques for a given graph G .

A generalization of the MCP problem is the Maximum Weighted Clique Problem (MWCP). A weighted graph G is defined by $G = (V, E, W)$, where W is the vertex weighting function. Each vertex v is associated with a positive weight ω_v . For a clique C of G , the weight is defined by $W(C) = \sum_{v \in C} \omega_v$. Then, the MWCP problem is to determine a clique C^* of maximum weight, i.e., $\forall C \in \mathcal{C}, W(C^*) \geq W(C)$ where \mathcal{C} is the set of all possible cliques of the graph. Figure 4.7 shows an example of a sample weighted graph used in this work.

Algorithm 2: Basic Bron-Kerbosch Algorithm

```

BronKerbosch( $R, P, X$ )
1 if  $P = \phi$  AND  $X = \phi$  then
2   | report  $R$  as a maximal clique
3 for each vertex  $v$  in  $P$  do
4   | BronKerbosch( $R \cup v, P \cap N(v), X \cap N(v)$ )
5   |  $P \leftarrow P \setminus \{v\}$ 
6   |  $X \leftarrow X \cup \{v\}$ 

```

4.3.2 Basic Bron-Kerbosch Algorithm

Algorithm 2 outlines the standard Bron-Kerbosch procedure for enumerating the maximal unweighted cliques of a given graph G using a recursive backtracking algorithm. The algorithm operates on three disjoint subsets R , P , and X of vertices from V . The set R contains the vertices belonging to the current clique. Set P keeps all prospective vertices, each of which is connected to all vertices of R . Vertices from P are used for expanding the current clique R . Finally, the set X contains all vertices that can no longer be used for completion of R , because all maximal cliques containing these vertices have already been reported. The Bron-Kerbosch algorithm is initially called with $R = X = \phi$ and $P = V$, i.e., $\text{BronKerbosch}(\phi, V, \phi)$. $N(v)$ indicates the neighbors of the vertex v .

The algorithm operates as follows: Pick a vertex v from P to expand. Add v to R and remove its non-neighbors from P and X . Then pick another vertex from the new P set and repeat the process. Continue until P is empty. Once P is empty, if X is empty then report the content of R as a new maximal clique. If it is not then R contains a subset of an already found clique. Now backtrack to the last vertex picked and restore P , R and X as they were before the choice, remove the vertex from P and add it to X . In this algorithm, set X will control the avoidance of reporting duplicate cliques.

Algorithm 3: Modified Bron-Kerbosch with pivoting (MWCP)

```

BronKerbosch( $R, P, X$ )
1 if  $P = \phi$  AND  $X = \phi$  then
  |   if  $W(R^*) \leq W(R)$  then
  |   |   report  $R$  as the maximum weighted clique, i.e.,  $R^* = R$ 
3 for each vertex  $v$  in  $P \setminus N(u)$  do
4   |   BronKerbosch( $R \cup v, P \cap N(v), X \cap N(v)$ )
5   |    $P \leftarrow P \setminus \{v\}$ 
6   |    $X \leftarrow X \cup \{v\}$ 

```

4.3.3 Modified Bron-Kerbosch Algorithm for MWCP with Pivoting

The basic Bron-Kerbosch procedure described in Algorithm 2 is inefficient in the case of graphs with many non-maximal cliques since it makes a recursive call for every clique (maximal or not). Bron and Kerbosch also introduced a second version of the standard algorithm involving a pivot vertex u chosen from P . Any maximal clique of X either includes the pivot vertex u or one of the vertices not adjacent to u , i.e., $v \in P \setminus N(u)$. Therefore, only the pivot vertex u and vertices from P not adjacent to u need to be considered as expansions of the current clique R in each recursive call of the Bron-Kerbosch algorithm. Vertex v from P adjacent to u can be skipped, because any clique containing v must also contain u . Such a clique will be discovered in a subsequent recursive call once u has been added to R . Algorithm 3 presents the Bron-Kerbosch procedure with pivoting and an enhancement for enumerating all maximal weight cliques of G . To find the maximum weighted clique then we need to compare the weight of any recently found clique with the global maximum weight. Initially the global maximum clique, denoted by R^* , is an empty set with weight zero. As the algorithm progresses, the maximum weighted clique and its weight are updated.

4.3.4 MWCP Graph Model

It is shown that Problem (4.8) is equivalent to a Maximum Weighted Clique Problem (MWCP). We introduce an arbitrary undirected and weighted graph $G = (V, E, W)$ where V is the set of vertices, E is the set of edges and W is the weighting function.

Proposition 5. *The hybrid optimization problem defined in (4.8) is equivalent to a Maximum Weighted Clique Problem (MWCP).*

Proof. Let $G = (V, E, W)$ be a graph in which each vertex represents a subchannel tuple (i, j, k) with subchannels i for DL subframe, j for UL subframe and k for RN subframe. Two vertices intersect if they share at least one common index element at a given subframe, and are otherwise said to be disjoint. Two vertices will have an edge if they are disjoint. For example $(1, 1, 0)$ and $(1, 0, 0)$ intersect but $(1, 0, 0)$ and $(0, 1, 1)$ are disjoint. Each vertex has a corresponding weight, denoting the maximum product of backlog and achievable rate. Specifically, the weighting functions defined in (4.9) and (4.10) are found over all possible assignments of this vertex's subchannels to all UEs and RNs, and then the UE and/or RN with the maximum weight function are selected since the subchannel tuple (i, j, k) can only be assigned to one UE in direct transmission and one RN in CoR and NC/CoR, i.e., $W_v(i, j, k) = \max\{W_1(i, j, k), W_2(i, j, k)\}$. Our optimization problem is to find the optimal strategy to choose the transmission scheme and assign RN and subchannel tuples to each UE in order to maximize the aggregated total product of backlog and achievable bidirectional rate. Equivalently, we can also interpret it as to find the optimal strategy to select disjoint subchannel combinations and assign RNs and UEs to them to maximize our objective function. Therefore, jointly optimizing the subchannel assignment, transmission scheme selection and RN selection for our maximization problem is to find a subset C of pairwise adjacent vertices in the graph G having

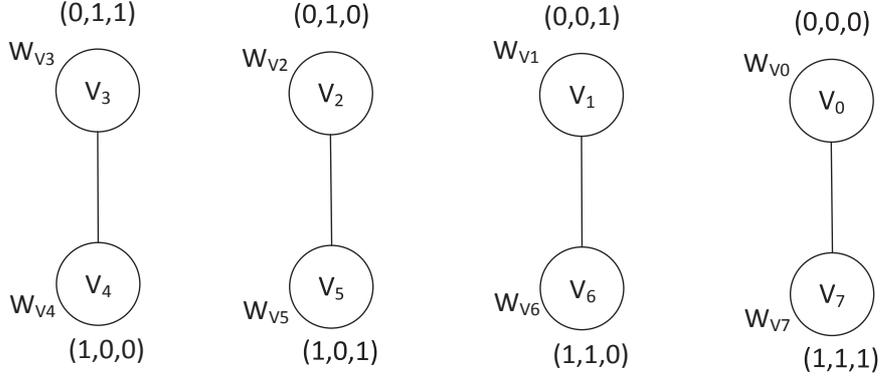


Figure 4.7: MWCP graph with $N = 2$ where $i \in N$, $j \in N$ and $k \in N$

the largest total weight, i.e., the so called Maximum Weighted Clique. An example of a complete graph for our problem for set of subchannels with $N = 2$ where $i, j, k \in \{0, 1\}$ is shown in Figure 4.7. \square

4.3.5 Complexity

The total number of distinct vertices in our three-time-slot TDD frame architecture is N^3 , i.e., $|V| = N^3$ where $|V|$ is the cardinality of the set V . The complexity of the weighting function (4.9), corresponding to the direct transmission scheme, is $O(M)$ while the complexity of the weighting function (4.10) is $O(2ML)$ where the factor 2 corresponds to the CoR and NC/CoR transmission schemes. Therefore, the total complexity of the defined weighting functions is $O(MN^3 + 2MLN^3)$. It is shown in [94] that the worst-case time complexity for generating all maximal cliques of an undirected graph in the Bron-Kerbosch algorithm is $O(3^{\frac{|V|}{3}})$, where $|V|$ is the number of vertices in the graph. Therefore, the total complexity of our algorithm for solving Problem (4.8) is $O(MN^3 + 2MLN^3 + 3^{\frac{N^3}{3}})$.

4.3.6 Sequential Greedy Allocation for MWCP problem

Exact approaches for the Maximum Clique Problem (MCP) (e.g., branch and bound) are algorithms which guarantee finding an optimal solution. However, they require a runtime that often grows exceeding large with respect to graph size. As shown in Subsection 4.3.5, the complexity grows exponentially with the size of the graph. Using heuristic techniques is a popular alternative which is less bleak than attempting to solve the MCP exactly. Such methods can be practical since they are able to provide suboptimal solutions of acceptable quality with reasonable computing efforts. A greedy algorithm is an algorithm that follows a technique of making a locally-optimal decision at each stage with the expectation of finding a global optimum solution for an optimization problem [95]. We implemented a greedy algorithm which builds a maximal clique by the repeated addition of a vertex into a partial clique, starting from an empty clique. The algorithm constructs feasible solutions from scratch by making, in each step, the most favorable choice for a decision variable. Each choice depends on the decisions made in the previous steps, but the effects that this choice will have in the following steps are unknown. Such choices can thus be viewed as local decision rules that generally lead to suboptimal solutions since future decisions may have a large impact on the resulting quality of the solution. The sequential greedy heuristics are very fast by nature, but their quality is generally unsatisfactory for large graph sizes. The run-time of our algorithm is limited to a pre-defined duration.

4.4 ACO-Based Optimization

Many heuristic methods currently used in combinatorial optimization are inspired by natural behaviors. Examples of such methods are genetic algorithms, simulated

annealing and neural networks. The Ant Colony Optimization (ACO) belongs to the class of biologically inspired heuristics. The basic idea is to imitate the cooperative behavior of ant colonies in order to provide an efficient suboptimal solution to combinatorial optimization problems. The ACO is a general purpose heuristic (meta-heuristic) algorithm that has been proposed in [90] and achieved widespread success in solving different optimization problems [77, 96].

4.4.1 Three-Dimensional Assignment Problem ACO (3DAP-ACO) Algorithm

In this section, the 3DAP-ACO algorithm is used to provide a suboptimal solution for our optimization problem defined in (4.8). The 3DAP-ACO algorithm is a modified version of the Hybrid Ant System-Quadratic Assignment Problem (HAS-QAP) [77]. HAS-QAP is an application of the ant system to the QAP problem, which is one of the fundamental combinatorial optimization problems. Our modified problem defined in (4.8) is a three-dimensional assignment problem based on the subchannel tuple (i, j, k) assignment. The new proposed algorithm is an enhanced algorithm based on HAS-QAP to provide a suboptimal solution for our 3DAP optimization problem. 3DAP can be stated as a binary integer programming problem as follows:

$$\begin{aligned}
\min \quad & \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} c_{ijk} x_{ijk} \\
\text{s.t.} \quad & \sum_{j \in J} \sum_{k \in K} x_{ijk} = 1, \forall i \in I \\
& \sum_{i \in I} \sum_{k \in K} x_{ijk} = 1, \forall j \in J \\
& \sum_{i \in I} \sum_{j \in J} x_{ijk} = 1, \forall k \in K \\
& x_{ijk} \in \{0, 1\}, \forall (i, j, k) \in I \times J \times K
\end{aligned} \tag{4.14}$$

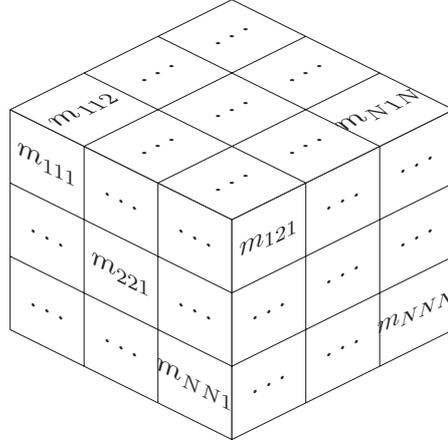


Figure 4.8: 3DAP Cost Matrix

where I , J and K are disjoint sets with $|I| = |J| = |K| = N$.

An example of the 3DAP is the assignment of N managers to N locations to N projects, where N is the number of elements and where the managers, locations, and projects represent the three dimensions. A cost exists for each assignment, and the goal is to minimize the total cost after all managers have been assigned a single location and project. Therefore, a three-dimensional matrix ($N \times N \times N$) can be formed with each element c_{ijk} which corresponds to the cost of manager i going to location j doing project k .

Figure 4.8 shows the three dimensional cost matrix used in 3DAP. Our subchannel assignment problem defined in (4.8) is a three-dimensional (subchannel tuple (i, j, k)) problem. The problem can be revisited as a 3DAP problem by defining

$$c_{ijk} = \max\{W_1(i, j, k), W_2(i, j, k)\}, \quad (4.15)$$

where $W_1(i, j, k)$ and $W_2(i, j, k)$ are given by (4.9) and (4.10), respectively and each element c_{ijk} corresponds to the maximum product of the achievable rate and backlog for each of the transmission schemes. Hence, our problem is equivalent to revenue

maximization 3DAP problem.

HAS-QAP algorithm is selected as the basis to the proposed 3DAP-ACO algorithm. Normally in standard ACO algorithms, the pheromone trails are used to generate completely new solutions. However in HAS-QAP, pheromone trails are used to modify the current solutions by keeping a global best solution to shorten the finding process. The algorithm also performs an additional local search phase to help improving the convergence speed. The idea of intensification and diversification were devised to strengthen or weaken the pheromone trail, respectively.

4.4.2 Enhanced Hybrid 3DAP-ACO Algorithm

Similar to HAS-QAP, in 3DAP-ACO each ant a of the m ants is associated with an integer permutation (solution) π^a . Each solution π^a is an $N \times 3$ matrix that determines the subchannel assignments in each subframe. It is also called permutation as it is being updated during the ACO algorithm. Modifications based on the pheromone trails are then applied to each permutation. The solutions found up until the stage of the process are then optimized using a 2-exchange local search method. Each pheromone trail is represented in the algorithm by a pheromone matrix τ . The pheromone matrix can be seen as a shared memory holding the desirability of the assignments which have produced the best solutions. The pheromone levels are updated globally after each iteration based on the best solution found. The major difference between 3DAP-ACO and HAS-QAP is that 3DAP-ACO requires two $N \times N$ pheromone matrices rather than one, where N is the number of elements per dimension (number of subchannels per subframe). The first pheromone matrix $\tau^{(1)}$ represents the desirability (strength) of the first and second dimension assignments (subchannels of the first and second subframes) and the second pheromone matrix $\tau^{(2)}$ represents the desirability of the second and third dimension assignments (subchannels of the second and third

subframes).

The pheromone levels are also used for exploration and exploitation. Exploration is a stochastic process in which a solution to the problem is made probabilistically. Exploitation utilizes the pheromone levels to choose a path that maximizes the sum of pheromone levels between available edge assignments. After building a new solution, our algorithm updates the pheromone trails. First, the pheromone trails are decreased globally to simulate the evaporation of pheromone. Then, the pheromone trails corresponding to the best solution obtained up until the stage of the process are intensified, taking into consideration the strength of the solution. This makes the search more aggressive and requires less time to reach good solutions. Intensification and diversification are used to decide where each ant starts its solution. Intensification is used to explore the neighbourhood of good solutions. When intensification is active in an iteration, the ant returns to the previous solution if the previous solution is better. In the case where the new solution is better, the ant simply continues working on its current solution. Such a process imposes a risk of an early convergence of the algorithm. Therefore, a diversification mechanism implements a partial restart of the algorithm when the solutions seem to not be improving any more. It consists of a reinitialization of both pheromone trails matrices. For a detailed description of the HAS-QAP algorithm, the reader is encouraged to consult [77].

The complete architecture of the modified 3DAP-ACO heuristic system can be found in Algorithm 4. A brief explanation of the parameters and the algorithm is provided as follows:

- m indicates the number of ants in the system, where each ant is shown by the parameter a .
- $\tau^{(1)}$ indicates the first $(N \times N)$ pheromone matrix which represents the strength

Algorithm 4: 3DAP-ACO Algorithm

Problem Initialization:Generate m random initial permutations π^1, \dots, π^m associated to m antsImprove π^1, \dots, π^m with 2-exchange local search procedureLet π^* be the best solutionInitialize the pheromone trail matrices $\tau^{(1)}$ and $\tau^{(2)}$

Activate intensification

Main Loop:**for** $iter = 1$ to I^{max} **do** **for** $a = 1$ to m **do** /* pheromone swap procedure */ **for** $swaps = 1$ to R **do**

Select a random element index as 1st swap position

 Generate a random number r **if** $r \leq q$ **then**

/* exploit pheromone trails */

Based on maximizing pheromone sum do:

Choose an element from the 2nd column to swap

Choose an element from the 3rd column to swap

else

Do the default actions /* explore */

Based on a probabilistic method do:

Choose an element from the 2nd column to swap

Choose an element from the 3rd column to swap

 Perform swaps on π^a within 2nd and 3rd column and get $\hat{\pi}^a$ Apply 2-exchange local search to $\hat{\pi}^a$ to get $\tilde{\pi}^a$ **for** $a = 1$ to m **do** **if** intensification is active **then** $\pi^a(iter + 1) \leftarrow$ best permutation between $\hat{\pi}^a(iter)$ and $\tilde{\pi}^a(iter)$ **else** $\pi^a(iter + 1) \leftarrow \tilde{\pi}^a(iter)$ **if** $\forall a \pi^a(iter + 1) = \pi^a(iter)$ **then**

/* No more improvement to the solution */

deactivate intensification

if $\exists a f(\pi^a(iter)) > f(\pi^*)$ **then** Update π^* , the best solution found so far

activate intensification

pheromone trail updatingUpdate pheromone trail matrices $\tau^{(1)}$ and $\tau^{(2)}$ by evaporation and reinforcement**diversification****if** S iterations performed without improving π^* **then** Perform a diversification

of the first and second dimension assignments (subchannels of the DL and UL subframes).

- $\tau^{(2)}$ indicates the second ($N \times N$) pheromone matrix which represents the strength of the second and third dimension assignments (subchannels of the UL and RN subframes).
- π indicates a 3D assignment permutation solution which is represented by an $N \times 3$ matrix. The elements of the matrix represent the subchannel number assigned. The 3 columns represent the DL, UL and RN subframes, respectively.
- π^a indicates the assignment solution of ant a .
- $\hat{\pi}^a$ indicates the permutation solution of ant a after performing R swaps during the swap procedure.
- $\tilde{\pi}^a$ indicates the permutation solution of ant a after performing the local search procedure.
- π^* indicates the best assignment solution found up until the stage of the process.
- $f(\pi^*)$ indicates the objective function of the best solution permutation.

Initialization Phase

The initial solution, which is assigned for each ant a , is randomly generated and goes through a fast 2-exchange local search procedure as will be explained later. Initially, all the elements of the pheromone matrices $\tau^{(1)}$ and $\tau^{(2)}$ are set to the same initial scalar τ_0 and the best initial solution to the 3DAP is stored in π^* . Similar to [77], τ_0 is set to $\tau_0 = \frac{1}{Q \times f(\pi^*)}$, where Q is a fixed parameter.

Pheromone Swap Procedure

The pheromone swap procedure is performed by each ant on its own permutation. Parameter R controls the number of 2-exchange swaps. The two element swap is performed to check if the solution improves the current solution. The swap procedure continues until all R swaps are considered. The two elements chosen to be swapped in each pheromone matrix are chosen according to the following rule:

An index r (between 1 and N) is chosen randomly and used to perform the swap with the first and second dimensions. Depending on the value of r , the indices s_1 and s_2 ($s_1, s_2 \neq r$) can be chosen by employing one of two different policies:

- Exploiting the pheromone trail with probability q . Indices s_1 and s_2 are chosen such that $\tau_{r\pi_{s_1,2}^a}^{(1)} + \tau_{s_1\pi_{r,2}^a}^{(1)}$ is maximized for $\tau^{(1)}$ and $\tau_{r\pi_{s_2,3}^a}^{(2)} + \tau_{s_2\pi_{r,3}^a}^{(2)}$ is maximized for $\tau^{(2)}$.
- Exploring the pheromone trail with probability $1 - q$. Indices s_1 and s_2 are chosen with probability $\frac{\tau_{r\pi_{s_1,2}^a}^{(1)} + \tau_{s_1\pi_{r,2}^a}^{(1)}}{\sum_{u \neq r} (\tau_{r\pi_{u,2}^a}^{(1)} + \tau_{u\pi_{r,2}^a}^{(1)})}$ and $\frac{\tau_{r\pi_{s_1,3}^a}^{(2)} + \tau_{s_1\pi_{r,3}^a}^{(2)}}{\sum_{u \neq r} (\tau_{r\pi_{u,3}^a}^{(2)} + \tau_{u\pi_{r,3}^a}^{(2)})}$ for $\tau^{(1)}$ and $\tau^{(2)}$, respectively.

After the selection of the indices r , s_1 and s_2 , the elements of the current solution π^a (i.e., $\pi_{r,2}^a$ and $\pi_{s_1,2}^a$ in the second column, $\pi_{r,3}^a$ and $\pi_{s_2,3}^a$ in the third column) are swapped if improved to produce a new permutation solution $\widehat{\pi}^a$.

2-Exchange Local Search Procedure

The local search procedure consists of applying a complete neighbourhood examination. An exhaustive 2-exchange local search is also performed during solution initialization phase. This procedure systematically examines exchanging two element positions (subchannel indices). If an improving swap is found, the solution is updated and the local search continues until all possible swaps are considered. The local search

procedure does not necessarily reach a local optimum, but is fast and may produce different solutions when starting with the same initial solution. Comparison of the different local search methods is out of the scope of this work. The R pheromone swaps result in all m ants to have a new modified solution $\hat{\pi}^a$ in addition to the initial solution π^a . After the local search procedure, the modified solution $\hat{\pi}^a$ is changed to $\tilde{\pi}^a$.

Intensification Phase

The intensification mechanism explores the neighborhood of the best solution up until the stage of the process. Intensification is active as long as at least one ant is capable of improving the current global best solution. The intensification is deactivated after iteration $iter$ if the modified solutions $\hat{\pi}^a(iter)$ and $\tilde{\pi}^a(iter)$ do not improve the current solution $\pi^a(iter)$ for all m ants. If a better solution is found, $\pi^a(iter + 1)$ is set to $\tilde{\pi}^a(iter)$ and intensification remains active.

Pheromone Trail Updating

The two pheromone matrices are updated according to the global best solution π^* leading to a faster convergence of the algorithm. Before doing so, the usual pheromone evaporation has to be realized by choosing parameter α_1 and using the equations defined in (4.16).

$$\begin{aligned}\tau^{(1)}(iter + 1) &= (1 - \alpha_1)\tau^{(1)}(iter) \\ \tau^{(2)}(iter + 1) &= (1 - \alpha_1)\tau^{(2)}(iter)\end{aligned}\tag{4.16}$$

where α_1 is a parameter that controls the evaporation of the pheromone trail. Then, the pheromone trails of the best solution π^* are reinforced using the equations defined

in (4.17).

$$\begin{aligned}\tau^{(1)}(iter + 1) &= \tau^{(1)}(iter + 1) + (\Delta\tau)_{N \times N} \\ \tau^{(2)}(iter + 1) &= \tau^{(2)}(iter + 1) + (\Delta\tau)_{N \times N}\end{aligned}\tag{4.17}$$

where $(\Delta\tau)_{N \times N}$ is an $N \times N$ matrix whose elements are identically equal to $\Delta\tau$ which is the amount of deposited pheromone and is given by $\frac{\alpha_2}{f(\pi^*)}$. Parameter α_2 controls the amount of new pheromone deposits (reinforcement).

Diversification Phase

Diversification helps on generating new starting solutions for the ants (only one of the m ants keeps the global best solution). The pheromone matrices are reinitialized for the other $(m - 1)$ ants by using a new random permutation. It is activated if no new best solution is generated during the last S iterations, where S is a fixed parameter.

Similarly, our modified ACO algorithm can be applied to any multi-dimensional assignment problem. As it is observed, the ACO algorithm is controlled through a set of parameters. The factors along with descriptions of each parameter and default values are listed in Table 4.1 and are chosen according to the default values in [77].

4.4.3 Complexity

The complexity of the HAS-QAP algorithm is defined in [77]. The most time consuming part of the algorithm is the local search procedure with complexity $O(N^3)$ and it is repeated $I^{max}m$ times. Therefore, the total complexity of HAS-QAP is $O(I^{max}mN^3)$, where m is the number of permutations (ants). Since our proposed 3DAP-ACO algorithm uses two pheromone matrices, the total complexity is

Table 4.1: ACO System Parameters

<i>Parameter</i>	<i>Value</i>
α_1 (Pheromone evaporation rate)	0.1
α_2 (Pheromone strengthening factor)	0.1
m (Number of ants)	10
Q (Influences initial pheromone levels)	100
R (Number of pheromone swaps)	3
S (No improvement threshold)	5
q (Probability of doing an exploit or explore type swap)	0.9
I^{max} (Number of algorithm iterations)	100

$O(2I^{max}mN^3)$. Similar to Section 4.3, the total complexity of the defined weighting functions is $O(MN^2 + 2MLN^3)$. Therefore, the total complexity of Problem (4.8) using Algorithm 4 is $O(MN^3 + 2MLN^3 + 2I^{max}mN^3)$.

4.4.4 Fast-3DAP-ACO

As shown in the Subsection 4.4.3, the proposed 3DAP-ACO algorithm provides a suboptimal solution for our three-dimensional resource assignment problem in polynomial time. However, the most consuming parts of the algorithm are the local search and the number of algorithm iterations I^{max} followed by all ants. Since a proper suboptimal solution can be achieved with a lower number of iterations, we developed another version of the proposed ACO algorithm called Fast-3DAP-ACO. Fast-3DAP-ACO limits the number of iterations whenever the time required to provide a suboptimal solution for our 3DAP-ACO problem reaches a threshold. The threshold is set to the LTE frame duration in order to check the algorithm validity in real wireless networks.

4.5 Simulations and Discussion

We developed our own C++ simulator to implement the Bron-Kerbosch, Greedy and 3DAP-ACO algorithms. The TDD frame architecture is simulated for an OFDMA system in which UEs are randomly distributed and moving over a 10Km x 10Km plane. RNs are placed at 50% of the eNB transmission range. UE and eNB arrival processes (Λ_U and Λ_E , respectively) are symmetric and modeled as Poisson processes. Simulation experiments have been conducted under various scenarios in order to assess the performance of the proposed hybrid scheme, i.e., transmission scheme selection, relay selection and subchannel resource assignment in terms of stable end-to-end throughput and queue backlog (average queueing delay) for SISO and MIMO configurations. Each data point collected is averaged over 10,000 cycles (three-time-slots) and averaged over 10 simulation trials using a different random seed (UE position). The detailed simulation parameters are shown in Table 4.2.

The simulations performed reflect an LTE-Advanced network configuration. In the LTE radio interface, two consecutive time-slots create a subframe where each time-slot spans 0.5 msec. Resources are allocated to UEs in units of resource blocks (RBs) or Resource Block Groups (RBGs). An RB is a set of 12 consecutive subcarriers (180 kHz) over the duration of one slot. The number of RBs included in an RBG is determined by the system bandwidth. Moreover, in order to maintain the same subframe durations as the legacy TDD frame architecture, the subframe lengths of the new three-slot cycle is the same as the length of the combined UL and DL subframes, i.e., the total number of resources are divided equally among the three slots. For example, for 1.4 MHz LTE carriers, 6 usable RBs (PRBs) per subframe are available and the number of RBs per RBG group is 1. Therefore, the 12 RBGs can be divided equally such that each subframe in our three-time-slot cycle contains 4

Table 4.2: Simulation Parameters in the System

<i>Parameter</i>	<i>Setting</i>
System Bandwidth	1.4MHz, 10MHz
Available Subchannels (N) per Slot	4, 11
Cycle Duration	2 subframes
Packet Length	1 Kb
Traffic Model	Poisson arrivals ($\Lambda_E = \Lambda_U$)
Path-loss Exponent α	3.5
Channel Model	Rayleigh fading
Tx Powers of eNB, RN, UE	46, 30 and 23 dBm
Noise PSD	-174 dBm/Hz
Number of UEs	10, 15
Number of RNs	1, 2, 3, 4, 5, 6, 7, 8
Number of MIMO Antennas	eNB(4x4) RN(4x4) UE(4x2)

RBGs. Similarly, for 10 MHz carriers (a more common configuration), each subframe consists of 50 RBs with a maximum RBG size of 3 RBs. Therefore, the 100 total RBs can be divided such that each subframe contains 11 RBGs. In this chapter, the simulations are performed for a 1.4MHz bandwidth carrier ($N = 4$) with 10 UEs and 10MHz bandwidth ($N = 11$) with 15 UEs.

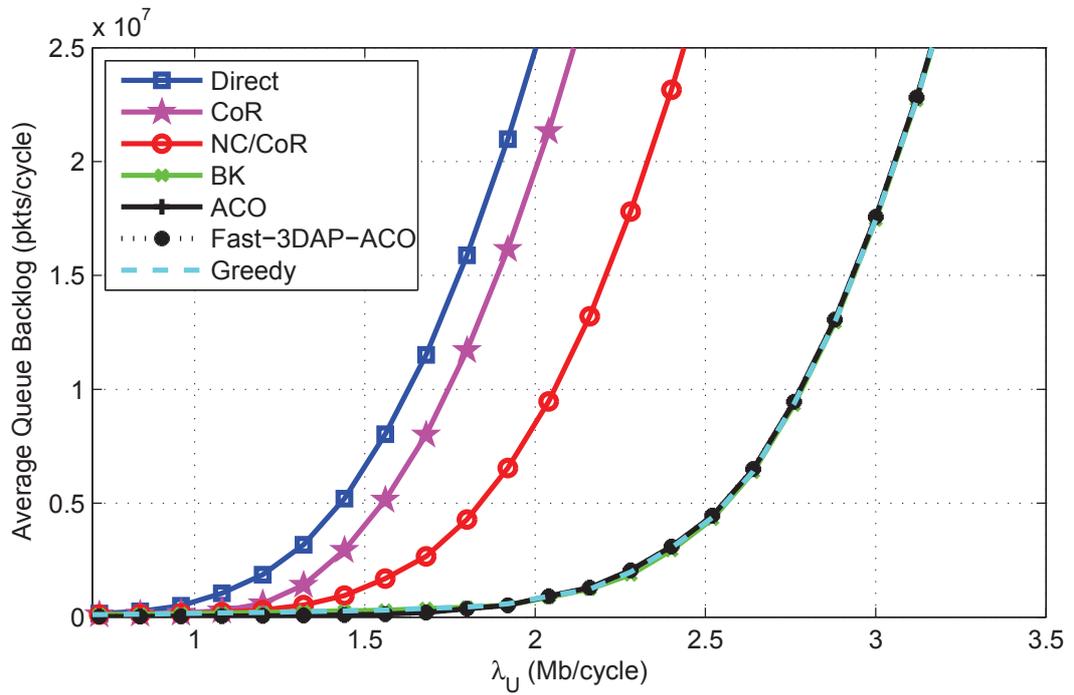
We compare our results with our Greedy Allocation algorithm in Subsection 4.3.6 and Fast-3DAP-ACO algorithm in Subsection 4.4.4. In this work, we do not consider power allocation for simplicity. It is known that power allocation can bring significant improvement in relay networks when the source and relay nodes are subject to a total power constraint [97]. However, as demonstrated in [97, 98], the gain obtained by power adaptation is very limited in OFDM-based relay networks if each transmitting node is subject to an individual peak power constraint as the case in most cellular networks. In our considered model, all the eNB, UEs and RNs are subject to their

own individual peak power constraints and, therefore, the transmit power is assumed to be fixed and uniformly distributed among all subchannels for each of them.

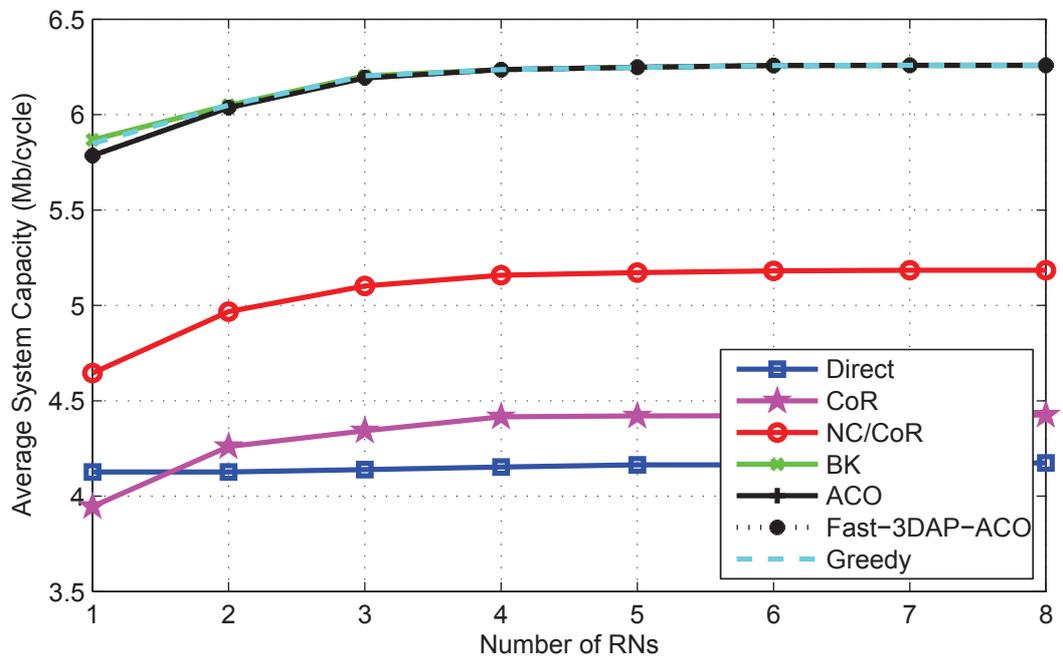
4.5.1 Average System Queue Backlog Evaluation

We first compute the average queue backlog performance, which is analogous to the end-to-end delay performance of the supported transmission schemes (direct, CoR, NC/CoR and hybrid) under a random network topology using 4 relay nodes (RNs), 10 UEs and $N = 4$ using the optimal Bron-Kerbosch MWCP algorithm. Figures 4.9(a) and 4.10(a) show the performance plots which represent the average system queue length of the different transmission schemes for SISO and MIMO configurations, respectively. It can be shown that as we increase the traffic rate for the UE, the average queue length also increases. However, at some traffic rates, the average queue length tends to infinity. Therefore, adding more traffic would make the system unstable and impose significant delay. This traffic rate marks the network stability region. As expected, it is shown that our proposed hybrid transmission scheme outperforms direct, CoR and NC/CoR.

It is also shown in Figure 4.9(a) that the hybrid scheme under a SISO configuration provides 2 Mb/cycle (46%) and 1 Mb/cycle (19%) improvements over direct and NC/CoR schemes. Similar performance results for MIMO simulations are shown in Figure 4.10(a). However, the improvements are even higher for MIMO in which the hybrid scheme provides 9 Mb/cycle (80%) and 4 Mb/cycle (25%) improvement over the direct and NC/CoR schemes. This would provide a great interest for LTE-Advanced networks.

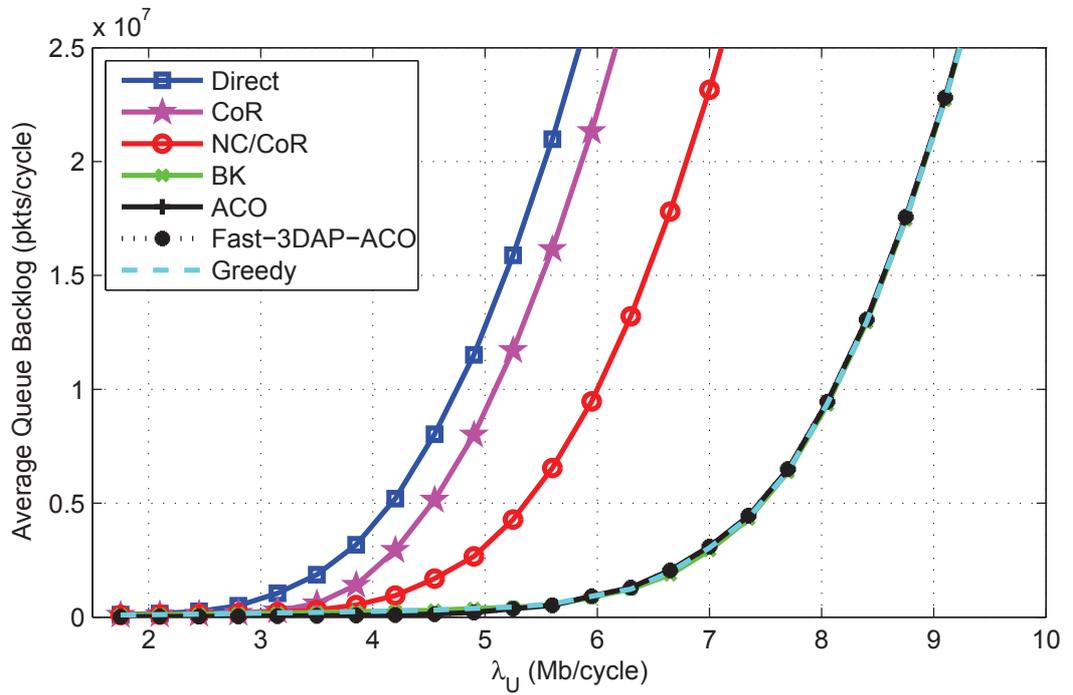


(a) Queue backlog (delay) performance of different transmission schemes (4 RN)

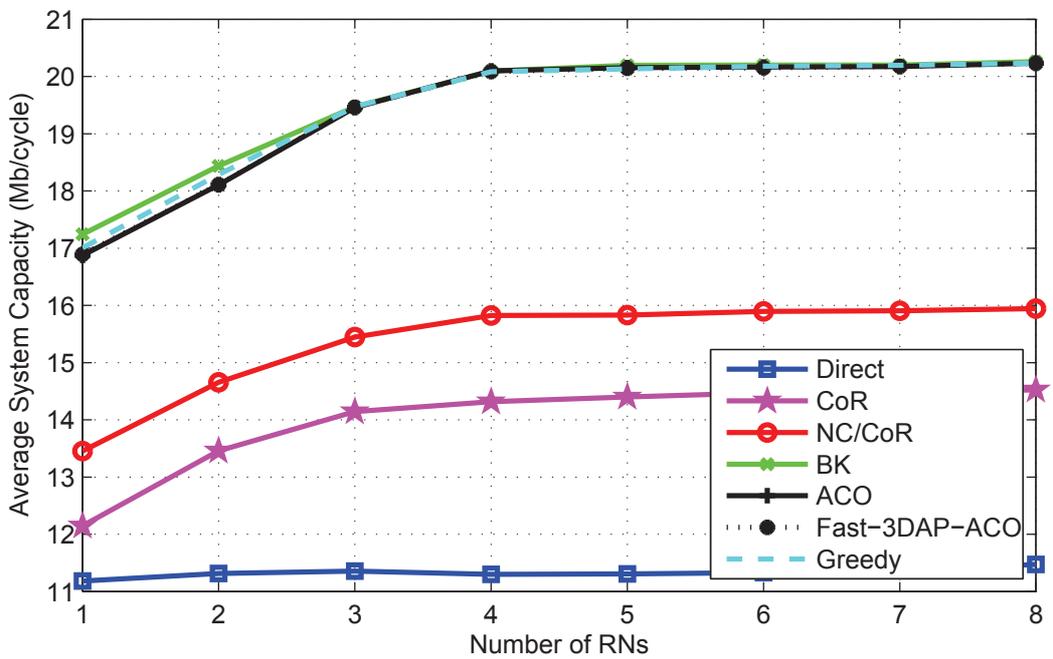


(b) Average system capacity vs. number of RNs

Figure 4.9: Network capacity and delay performance of different transmission schemes (SISO)



(a) Queue backlog (delay) performance of different transmission schemes



(b) Average system capacity vs. number of RNs

Figure 4.10: Network capacity and delay performance of different transmission schemes (MIMO)

4.5.2 Impact of Number of RNs on Average System Capacity

The average system capacity performance is computed for our proposed algorithm by increasing the number of RNs. The number of RNs is assumed to be 1-8 and the UEs are randomly placed across the network. Figure 4.9(b) shows the average system capacity of the proposed optimization algorithm for different numbers of RNs for SISO configuration. The capacity shown is the average achievable bidirectional rate for each bidirectional transmission scheme. Packets are added to the queues at the beginning of each cycle. It is observed that increasing the number of RNs makes the average stable throughput increase. However, increasing the number of RNs above 6 relays does not have significant impact on the average system capacity. This is because most of the UEs across the cell will be covered by the relay nodes and thus, the maximum performance that can be reached will eventually saturate. Therefore, adding more RNs would increase the operator cost without providing any gain. Figure 4.10(b) shows similar results for MIMO configuration.

4.5.3 Optimization Algorithms Comparison

We compare the optimal solution (Bron-Kerbosch) with 3DAP-ACO, Fast-3DAP-ACO and Greedy suboptimal solutions to our resource allocation problem. It can be seen in Figure 4.9 that the suboptimal 3DAP-ACO, Fast-3DAP-ACO and Greedy algorithms for our hybrid assignment problem provide approximate results to the Bron-Kerbosch MWCP algorithm for $N = 4$. The run-time of our Greedy algorithm is limited to the LTE frame duration (10 msec). It provides results closest to the optimal solution since it is dealing with a small graph (low number of vertices) to solve the MWCP problem. In addition, both ACO and Fast-3DAP-ACO algorithms

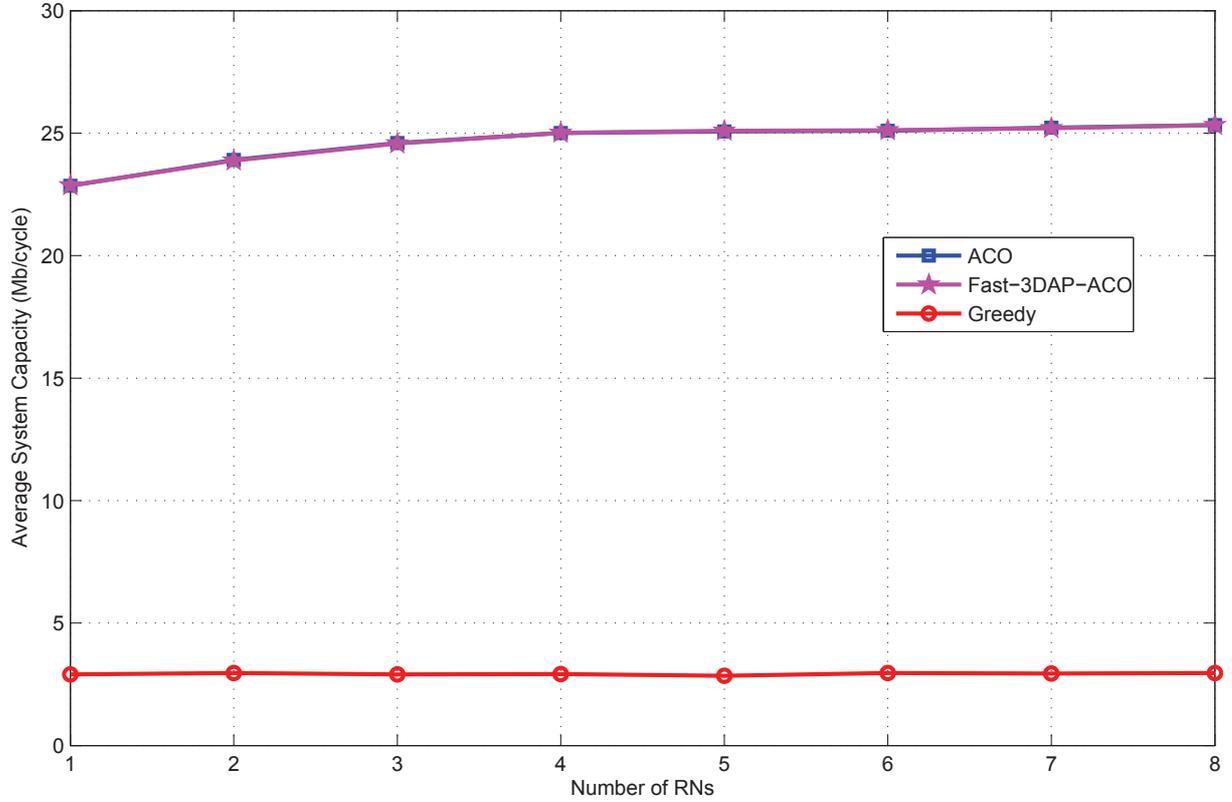


Figure 4.11: ACO vs. Greedy for $N = 11$ for SISO

produce results close to the optimal result with a slight degradation in Fast-3DAP-ACO because the number of iterations is limited to the LTE frame duration. The same simulations are performed for 10MHz LTE carrier bandwidth with 15 UEs and $N = 11$ as shown in Figure 4.11. Both 3DAP-ACO and Fast-3DAP-ACO have similar performances while the Greedy algorithm produces unsatisfactory results because the graph size is large (11^3). Therefore, the proposed ACO based algorithms produce the best results in more general network scenarios while the Greedy algorithm can still be used when the number of subchannels is relatively small. The Bron-Kerbosch was not feasible due to the exponential increase in complexity and cannot be used to solve

such problems with large graph sizes.

4.5.4 Optimal vs Suboptimal Algorithm Simulation Time

Figure 4.12 shows the simulation times in milliseconds for the hybrid transmission scheme defined in Problem (4.8) using Bron-Kerbosch, Greedy, 3DAP-ACO and Fast-3DAP-ACO algorithms for 4 RNs. Bron-Kerbosch, Greedy and Fast-3DAP-ACO algorithms produce good results when $N = 4$ with simulation times 6, 3 and 10 msec, respectively. The 3DAP-ACO takes more time (63 msec) because of the local search algorithm and the iterations needed by the ants. The Greedy allocation algorithm provided the best results when the graph is small. The case is different with $N = 11$. The Bron-Kerbosch algorithm is not feasible and unable to solve our resource allocation problem. This Bron-Kerbosch algorithm is required to go through all the vertices of the undirected graph. It can also be shown that as the number of subchannels increase, the assignment problem takes significantly longer to be solved using the Bron-Kerbosch algorithm. This is because to get the maximal weighted clique, the calculation needs to happen for each UE and RN for all vertices. This results in a significant increase in the simulation time because the complexity of the algorithm is exponentially dependent on the number of vertices as shown in Section 4.3. 3DAP-ACO is able to provide a good suboptimal solution to the assignment problem but with increased delays (600 msec). Therefore, applying Fast-3DAP-ACO and Greedy algorithms with limiting the complexity to the frame duration (10 msec) is important. However, as mentioned in Subsection 4.5.3, the Greedy allocation is not a feasible algorithm to be deployed when the number of subchannels is higher. It is important to note that the simulations are performed in a personal laptop while in real environments, eNBs have high processing and scheduling power. More simulations should be performed in real environments to test such optimization algorithms. We hope this

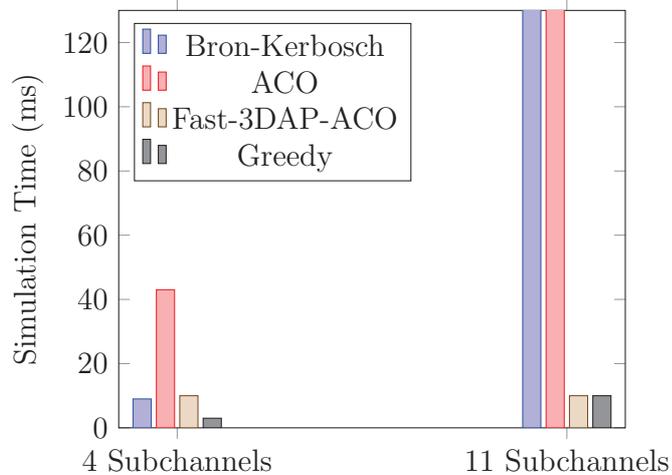


Figure 4.12: Average Simulation Time per Cycle

work will pave the way towards real-environment validations.

4.6 Chapter Summary

In this chapter, the impact of CoR and NC on the performance of relay-assisted OFDMA network is studied. An enhanced three-time-slot per cycle is proposed to support direct, CoR, and NC/CoR bidirectional transmission schemes. A combinatorial optimization problem is formulated to jointly optimize subchannel allocation/pairing, transmission scheme selection and relay selection with the objective of maximizing the system total product of backlog and achievable rate. An equivalent problem is formulated using Maximum Weighted Clique Problem (MWCP) in graph theory. A modified optimal Bron-Kerbosch algorithm is employed to solve the MWCP problem. In addition, suboptimal solutions are proposed based on an enhanced 3-dimensional Ant Colony Optimization (3DAP-ACO) algorithm and sequential greedy heuristic to solve the subchannel assignment problem. Simulation results confirm that the proposed algorithm provides the least average system queue backlog which maps

to the improved throughput and end-to-end delay performance. Moreover, the enhanced 3DAP-ACO suboptimal algorithm significantly reduces the complexity and simulation time compared to the optimal solution.

Chapter 5

Conclusion and Future Work

5.1 Conclusion and Discussion

In this thesis, we studied bidirectional relaying and proposed cross-layer resource allocation designs in multichannel relay-enhanced cellular networks. With the deployment of relay nodes, feasibility and optimal resource allocation become a more complicated problem to solve. The impact of Cooperative Relaying (CoR) and Network Coding (NC) on the performance of Two-Way Relay Channel (TWRC) networks was investigated. The scarce radio resources and limited energy supply of mobile nodes in wireless networks make efficient resource and power allocation strategies necessary. It is widely known that performance and efficiency of wireless networks can be substantially enhanced if we allow interaction between layers, which leads to the cross-layer design concept.

In Chapter 3, we proposed a cross-layer design for optimal joint resource allocation and power control in bidirectional relaying networks. The work assumed that each bidirectional link has a dedicated channel in the multichannel system, i.e. channel-pairing is not considered. Direct, CoR, and NC/CoR TWRC transmission schemes were considered and their rate regions were characterized explicitly. It was shown

that each of the aforementioned schemes can outperform the others depending on the channel states, system backlog and downlink and uplink traffic. Therefore, a hybrid solution that dynamically determines the optimal transmission scheme and optimal time allocation at each time-frame is proposed. We further considered power control in the system and proposed an optimal joint power control and resource allocation scheme. The new scheme outperformed the previous transmission schemes in terms of delay and throughput. However, this improvement comes at the cost of adding real-time computational complexity into the system.

In Chapter 4, the impact of CoR and NC on the performance of relay-assisted OFDMA networks was studied. An enhanced three-time-slot per cycle was proposed to support direct, CoR, and NC/CoR bidirectional transmission schemes. A combinatorial optimization problem was formulated to jointly optimize subchannel allocation/pairing, transmission scheme selection and relay selection. An equivalent problem was formulated using Maximum Weighted Clique Problem (MWCP) in graph theory. A modified optimal Bron-Kerbosch algorithm was employed to solve the MWCP problem. In addition, suboptimal solutions were proposed based on an enhanced 3-dimensional Ant Colony Optimization (3DAP-ACO) algorithm and sequential greedy heuristic to solve the subchannel assignment problem. Simulation results confirmed that our proposed algorithm provided the least average system queue backlog which maps to the improved throughput and end-to-end delay performance. Moreover, the enhanced 3DAP-ACO suboptimal algorithm significantly reduces the complexity and simulation time compared to the optimal solution with minimal effect on our simulation results for the most common deployments. The proposed optimization algorithms are practical models that can be applied to any TDD cellular networks such as LTE. The simulations showed promising results, and we hope this work paves the way towards more optimized resource allocation schemes

in such networks.

5.2 Future Work

In the present work, cooperative relaying and network coding are studied such that the relay nodes are assumed to be dedicated, i.e., the relaying nodes are fixed and do not have information to transmit. A possible future work is to address the problem of user pairing (i.e., partner selection) in a multi-user environment, where users employ network coding to transmit to a common destination (e.g., a base station in a cellular environment). In the absence of dedicated relay nodes, users mutually pair among themselves to realize network coding [99].

One important constraint in LTE scheduling is that all network resources for a given user need to use the same Modulation and Coding Scheme (MCS) in any given Transmission Time Interval (TTI) or subframe for the purpose of signaling cost reduction and device simplicity [100]. The effect of this restriction on LTE scheduling has not been adequately evaluated for two-way relay channels. Efficient power allocation schemes should be adapted to the resource allocation problem by taking into account the MCS constraint.

The work in this thesis assumes fixed relays such that they are located between the serving eNodeB and the cell edge. In addition, cooperation is limited to one relay which results in the saturation on the capacity with the increase in the number of relays and our model is limited to a single cell. Although the proposed algorithms show promising results, they have not been analyzed in a multi-cell network architecture that includes inter-cell interference nor has the work considered architectures that allow mobiles to connect to multiple relays simultaneously. Such an approach provides more diversity gain and will not limit the number of relays in the system.

This work is based on centralized resource allocation algorithms in which a control unit collects all the channel and queue state information of every user in the system and allocates the available resources to maximize the system capacity. Therefore, centralized scheduling can be ineffective due to the stringent time required to exchange the scheduling information and the large feedback required by the users to send all the information. A possible future direction would be to consider distributed resource allocation algorithms in order to reduce the complexity of centralized algorithms [101].

List of References

- [1] A. Ghosh, R. Ratasuk, B. Mondal, N. Mangalvedhe, and T. Thomas, “Lte-advanced: next-generation wireless broadband technology,” *IEEE Wireless Communications*, vol. 17, pp. 10–22, June 2010.
- [2] T.-Y. Lee, “Lte-advanced: Future of mobile broadband.” Presentation, TATA Consultancy Services, Oct. 2013.
- [3] H. Holma and A. Toskala, *LTE for UMTS - OFDMA and SC-FDMA Based Radio Access*. Chichester, UK: John Wiley & Sons Ltd, 2009.
- [4] “3GPP LTE.” <http://www.3gpp.org/lte/>. [Online; accessed 1-Dec-2015].
- [5] S. Sesia, I. Toufik, and M. Baker, “Lte, the umts long term evolution: From theory to practice,” *John Wiley and Sons*, 2009.
- [6] T. Rappaport, *Wireless Communications: Principles and Practice*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2nd ed., 2001.
- [7] A. Sendonaris, E. Erkip, and B. Aazhang, “User cooperation diversity-part i: System description,” *IEEE Transactions on Communications*, vol. 51, pp. 1927–1938, Nov. 2003.
- [8] J. Laneman, D. Tse, and G. W. Wornell, “Cooperative diversity in wireless networks: Efficient protocols and outage behavior,” *IEEE Transactions on Information Theory*, vol. 50, pp. 3062–3080, Dec. 2004.
- [9] K. Yamamoto, H. Maruyama, T. Shimizu, H. Murata, and S. Yoshida, “Spectral efficiency of fundamental cooperative relaying in interference-limited environments,” *IEICE transactions on communications*, vol. 91, pp. 2674–2682, Aug. 2008.

- [10] H. Alves, R. D. Souza, G. Fraidenraich, and M. E. Pellenz, "Throughput performance of parallel and repetition coding in incremental decode-and-forward relaying," *Wireless Networks*, vol. 18, pp. 881–892, Nov. 2012.
- [11] D. Benevides da Costa and S. Aissa, "Performance of cooperative diversity networks: Analysis of amplify-and-forward relaying under equal-gain and maximal-ratio combining," in *IEEE International Conference on Communications*, (Dresden, Germany), pp. 1–5, June 2009.
- [12] T. Cover and A. Gamal, "Capacity theorems for the relay channel," *IEEE Transactions on Information Theory*, vol. 25, pp. 572–584, Sep. 1979.
- [13] C. E. Shannon, "Two-way communication channels," *4th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 611 – 644, 1961.
- [14] V. Havary-Nassab, S. ShahbazPanahi, and A. Grami, "Optimal distributed beamforming for two-way relay networks," *IEEE Transactions on Signal Processing*, vol. 58, pp. 1238–1250, Mar. 2010.
- [15] R. Ahlswede, N. Cai, S.-Y. Li, and R. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, pp. 1204–1216, July 2000.
- [16] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Medard, and J. Crowcroft, "Xors in the air: Practical wireless network coding," *IEEE Transactions on Networking*, vol. 16, pp. 497–510, June 2008.
- [17] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: analog network coding," in *Proc. of ACM SIGCOMM Computer Communication Review*, (Kyoto, Japan), Aug. 2007.
- [18] Z. Li, X.-G. Xia, and B. Li, "Achieving full diversity and fast ml decoding via simple analog network coding for asynchronous two-way relay networks," *IEEE Transactions on Communications*, vol. 57, pp. 3672–3681, Dec. 2009.
- [19] Y. Chen, S. Kishore, and J. Li, "Wireless diversity through network coding," in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC 2006)*, (Las Vegas, Nevada, USA), Apr. 2006.
- [20] Y. Liang and V. Veeravalli, "Gaussian orthogonal relay channels: Optimal resource allocation and capacity," *IEEE Transactions on Information Theory*, vol. 51, pp. 3284–3289, Sep. 2005.

- [21] Z. Mo, W. Su, S. Batalama, and J. Matyjias, “Cooperative communication protocol designs based on optimum power and time allocation,” *IEEE Transactions on Wireless Communications*, vol. 13, pp. 4283–4296, Aug. 2014.
- [22] P. Larsson, N. Johansson, and K.-E. Sunell, “Coded bi-directional relaying,” in *Proc. of IEEE 63rd Vehicular Technology Conference (VTC Spring)*, (Melbourne, Australia), May 2006.
- [23] H. Hamdoun, P. Loskot, T. O’Farrell, and J. He, “Practical network coding for two way relay channels in lte networks,” in *Proc. of IEEE 73rd Vehicular Technology Conference (VTC Spring)*, (Budapest, Hungary), May 2011.
- [24] H. Hamdoun, P. Loskot, and T. O’Farrell, “Scalability of two-way relay channel network coding in a relay enhanced lte cell,” in *Proc. of 11th International Symposium on Communications and Information Technologies (ISCIT)*, (Hengzhou, China), Oct. 2011.
- [25] T. J. Oechtering and H. Boche, “Optimal resource allocation for a bidirectional regenerative half-duplex relaying,” in *Proc. IEEE International Symposium on Information Theory and its Applications (ISITA '06)*, (Seoul, Korea), pp. 528–533, Nov. 2006.
- [26] S. J. Kim, N. Devroye, P. Mitran, and V. Tarokh, “Achievable rate regions and performance comparison of half duplex bi-directional relaying protocols,” *IEEE Transactions on Information Theory*, vol. 57, pp. 6405–6418, Oct. 2011.
- [27] V. Jamali, N. Zlatanov, A. Ikhlef, and R. Schober, “Achievable rate region of the bidirectional buffer-aided relay channel with block fading,” *IEEE Transactions on Information Theory*, vol. 60, pp. 7090–7111, Nov. 2014.
- [28] V. Jamali, N. Zlatanov, and R. Schober, “Bidirectional buffer-aided relay networks with fixed rate transmissionpart i: Delay-unconstrained case,” *IEEE Transactions on Wireless Communications*, vol. 14, pp. 1323–1338, Mar. 2015.
- [29] D.-S. Cho, Y. Shim, and H. Park, “Optimal time allocation for two-way relay channel using physical-layer network coding,” *IET Communications*, vol. 8, pp. 2469–2475, Sep. 2014.
- [30] C. Hausl, O. Iscan, and F. Rossetto, “Optimal time and rate allocation for a network-coded bidirectional two-hop communication,” in *Proc. of European Wireless Conference*, (Lucca, Italy), Apr. 2010.

- [31] C.-H. Liu and F. Xue, "Network coding for two-way relaying: Rate region, sum rate and opportunistic scheduling," in *Proc. of IEEE International Conference on Communications*, (Beijing, China), May 2008.
- [32] Z. Chen, T. J. Lim, and M. Motani, "Digital network coding aided two-way relaying: Energy minimization and queue analysis," *IEEE Transactions on Wireless Communications*, vol. 12, pp. 1947–1957, Apr. 2013.
- [33] C. Lin, Y. Liu, and M. Tao, "Cross-layer optimization of two-way relaying for statistical qos guarantees," *IEEE Journal on Selected Areas in Communications*, vol. 31, pp. 1583–1596, Aug. 2013.
- [34] T. Oechtering and H. Boche, "Stability region of an optimized bidirectional regenerative half-duplex relaying protocol," *IEEE Transactions on Communications*, vol. 56, pp. 1519–1529, Sep. 2008.
- [35] Y. Sagduyu, R. Berry, and D. Guo, "Throughput and stability for relay-assisted wireless broadcast with network coding," *IEEE Journal on Selected Areas in Communications*, vol. 31, pp. 1506–1516, Aug. 2013.
- [36] P. Liu and I.-M. Kim, "Performance analysis of bidirectional communication protocols based on decode-and-forward relaying," *IEEE Transactions on Communications*, vol. 58, pp. 2683–2696, Sep. 2010.
- [37] H. Xu and B. Li, "Xor-assisted cooperative diversity in ofdma wireless networks: Optimization framework and approximation algorithms," in *Proc. of IEEE INFOCOM*, (Rio de Janeiro, Brazil), Apr. 2009.
- [38] Y. Xu, X. Xia, K. Xu, and D. Zhang, "On the hybrid relaying protocol for time division broadcasting," *Transactions on Emerging Telecommunications Technologies*, vol. 26, pp. 893–904, Nov. 2013.
- [39] M. Upadhyay and D. Kothari, "Optimal resource allocation techniques for cooperative af and df wireless networks," in *Proc. of International Conference on Signal Processing and Integrated Networks (SPIN)*, pp. 382–387, Feb. 2014.
- [40] L. Shan, H. Murata, S. Aissa, and S. Yoshida, "Hybrid network coding and cooperative relaying schemes for bi-directional communication systems," in *Proc. of Vehicular Technology Conference (VTC Spring)*, (Budapest, Hungary), May 2011.

- [41] E. N. Ciftcioglu, A. Yener, and R. Berry, "Stability regions of two-way relaying with network coding," in *Proc. of IEEE WICON*, (Hawaii, USA), Nov. 2008.
- [42] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Transactions on Automatic Control*, vol. 37, pp. 1936–1949, Dec. 1992.
- [43] K. Etemad, "Overview of mobile wimax technology and evolution," *IEEE Communications Magazine*, vol. 46, pp. 31–40, 2008.
- [44] A. Zainaldin, H. Halabian, and I. Lambadaris, "Optimal resource allocation in lte-advanced network using hybrid cooperative relaying and network coding," in *Proc. of IEEE International Conference on Communications (ICC)*, (Budapest, Hungary), June 2013.
- [45] G. Song and Y. Li, "Cross-layer optimization for ofdm wireless networks-part i: theoretical framework," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 614–624, Mar. 2005.
- [46] G. Song and Y. Li, "Cross-layer optimization for ofdm wireless networks-part ii: algorithm development," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 625–634, Mar. 2005.
- [47] P. Tarasak and S. Sun, "Resource allocation for downlink two-user ofdma systems with wireless network coding," in *Proc. of IEEE 12th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, (San Francisco, CA, USA), June 2011.
- [48] D. Zhang, Y. Wang, and J. Lu, "Qos aware relay selection and subcarrier allocation in cooperative ofdma systems," *IEEE Communications Letters*, vol. 14, pp. 294–296, 2010.
- [49] C. K. Ho, R. Zhang, and Y.-C. Liang, "Two-way relaying over ofdm: Optimized tone permutation and power allocation," in *Proc. of IEEE International Conference on Communications (ICC '08)*, (Beijing, China), May 2008.
- [50] K. Jitvanichphaibool, R. Zhang, and Y.-C. Liang, "Optimal resource allocation for two-way relay-assisted ofdma," *IEEE Transactions on Vehicular Technology*, vol. 58, pp. 3311–3321, Feb. 2009.

- [51] M. Tao, Y.-C. Liang, and F. Zhang, "Resource allocation for delay differentiated traffic in multiuser ofdm systems," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 2190–2201, 2008.
- [52] M. Abrar, X. Gui, and A. Punchihewa, "Two-way relaying cooperative wireless networks: Resource allocation and performance analysis," *International Journal of Electronics and Telecommunications*, vol. 59, pp. 229–235, Dec. 2013.
- [53] T. Wang and L. Vandendorpe, "Sum rate maximized resource allocation in multiple df relays aided ofdm transmission," *IEEE Journal on Selected Areas in Communications*, vol. 29, pp. 1559–1571, Sep. 2011.
- [54] A. Hottinen and T. Heikkinen, "Optimal subchannel assignment in a two-hop ofdm relay," in *Proc. of Signal Processing Advances in Wireless Communications*, (Helsinki, Finland), June 2007.
- [55] C.-N. Hsu, H.-J. Su, and P.-H. Lin, "Joint subcarrier pairing and power allocation for ofdm transmission with decode-and-forward relaying," *IEEE Transactions on Signal Processing*, vol. 59, pp. 399–414, Jan. 2011.
- [56] M. Abrar, X. Gui, and A. Punchihewa, "Low complexity joint sub-carrier pairing, allocation and relay selection in cooperative wireless networks," *International Journal of Communication Networks and Information Security (IJCNIS)*, vol. 6, pp. 182–188, Dec. 2014.
- [57] B. Han, M. Peng, Z. Zhao, and W. Wang, "A multidimensional resource-allocation optimization algorithm for the network-coding-based multiple-access relay channels in ofdm systems," *IEEE Transactions on Vehicular Technology*, vol. 62, pp. 4069–4078, Oct. 2013.
- [58] W. Dang, M. Tao, H. Mu, and J. Huang, "Subcarrier-pair based resource allocation for cooperative multi-relay ofdm systems," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 1640–1649, May 2010.
- [59] H. Zhang, Y. Liu, and M. Tao, "Resource allocation with subcarrier pairing in ofdma two-way relay networks," *IEEE Wireless Communications Letters*, vol. 1, pp. 61–64, Apr. 2012.
- [60] H. W. Kuhn, "The hungarian method for the assignment problem," *Naval research logistics quarterly*, vol. 2, pp. 83–97, Mar. 1955.

- [61] F. Zuo and M. Dong, "Low-complexity snr-based channel pairing in multichannel mabc-based two-way relaying," in *Proc. of IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 355–359, June 2013.
- [62] M. Chang, M. Dong, and S. Shahbazpanahi, "Joint pairing and power allocation optimization in multichannel mabc-based two-way relaying," in *Proc. of IEEE 15th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, (Toronto, ON, Canada), June 2014.
- [63] M. Chang and M. Dong, "Snr-based channel pairing design in multichannel tdbc-based two-way relaying," in *Proc. of Asilomar Conference on Signals, Systems and Computers*, (Pacific Grove, CA, USA), Nov. 2013.
- [64] Y. Liu, M. Tao, B. Li, and H. Shen, "Optimization framework and graph-based approach for relay-assisted bidirectional ofdma cellular networks," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 3490–3500, Oct. 2010.
- [65] W. Yu, "A dual decomposition approach to the sum power gaussian vector multiple access channel sum capacity problem," in *Proc. of Conference on Information Science and Systems (CISS)*, (Princeton, NJ, USA), Mar. 2003.
- [66] S. Boyd, L. Xiao, and A. Mutapcic, "Subgradient methods." University Lecture, Stanford University, Autumn Quarter 2003-2004.
- [67] X. Gao and M. Naraghi-Pour, "Computationally efficient resource allocation for multiuser ofdm systems," in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC 2006)*, (Las Vegas, NV, USA), Apr. 2006.
- [68] Y. Yu and W. Zhou, "Resource allocation for ofdma system based on genetic algorithm," in *Proc. of International Workshop on Cross Layer Design (IWCLD '07)*, (Jinan, China), Sep. 2007.
- [69] I. Kim, I.-S. Park, and Y. H. Lee, "Use of linear programming for dynamic sub-carrier and bit allocation in multiuser ofdm," *IEEE Transactions on Vehicular Technology*, vol. 55, pp. 1195–1207, July 2006.
- [70] A. Sinha and T. Kirubarajan, "A randomized heuristic approach for multi-dimensional association in target tracking," in *Proc. SPIE, Signal and Data Processing of Small Targets*, (Orlando, FL, USA), Apr. 2004.

- [71] M. Dorigo, M. Birattari, and T. Stutzle, “Ant colony optimization – artificial ants as a computational intelligence technique,” *IEEE Computational Intelligence Magazine*, vol. 1, pp. 28–39, Nov. 2006.
- [72] H. Ahmadi, Y. H. Chew, and C. C. Chai, “Multicell multiuser ofdma dynamic resource allocation using ant colony optimization,” in *Proc. of IEEE 73rd Vehicular Technology Conference (VTC Spring 2011)*, (Budapest, Hungary), May 2011.
- [73] C. Bron and J. Kerbosch, “Algorithm 457: finding all cliques of an undirected graph,” *Communications of the ACM*, vol. 16, pp. 575–577, Sep. 1973.
- [74] M. Dorigo, M. Birattari, and T. Stutzle, “Ant colony optimization,” *IEEE Computational Intelligence Magazine*, vol. 1, pp. 28–39, Nov. 2006.
- [75] M. Dorigo, V. Maniezzo, and A. Colorni, “Ant system: optimization by a colony of cooperating agents,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 26, pp. 29–41, Feb. 1996.
- [76] T. Stutzle and H. Hoos, “Max-min ant system and local search for the traveling salesman problem,” in *Proc. of IEEE International Conference on Evolutionary Computation*, (Indianapolis, IN, USA), Apr. 1997.
- [77] L. M. Gambardella, E. Taillard, and M. Dorigo, “Ant colonies for the quadratic assignment problem,” *Journal of the Operational Research Society*, vol. 50, pp. 167–176, Feb. 1999.
- [78] C. SHANNON, *A Mathematical Theory of Communication*. 1948.
- [79] E. Telatar, “Capacity of multi-antenna gaussian channels,” *Europ. Transactions on Telecommunications*, vol. 10, pp. 585–596, Nov. 1999.
- [80] A. Host-Madsen and J. Zhang, “Capacity bounds and power allocation for the wireless relay channel,” *IEEE Transactions on Information Theory*, vol. 51, pp. 2020–2040, June 2005.
- [81] S. Zhang, S. chang Liew, and P. P. Lam, “Physical-layer network coding,” in *Proc. of ACM Mobicom’06*, (Los Angeles, CA, USA), Sep. 2006.
- [82] V. Muralidharan and B. Rajan, “Physical layer network coding for the multiple access relay channel,” in *Proc. of IEEE International Conference on Communications (ICC)*, (Budapest, Hungary), June 2013.

- [83] S. C. Liew, S. Zhang, and L. Lu, “Physical-layer network coding: Tutorial, survey, and beyond,” *Physical Communication*, vol. 6, pp. 4–42, May 2013.
- [84] L. Georgiadis, M. J. Neely, and L. Tassiulas, *Resource Allocation and Cross Layer Control in Wireless Networks*. Now Publisher, 2006.
- [85] P. Bonami, M. Kilinc, and J. Linderoth, “Algorithms and software for convex mixed integer nonlinear programs.” Technical Report 1664, Computer Sciences Department, University of Wisconsin-Madison, Nov. 2009.
- [86] R. Byrd, J. Nocedal, and R. Waltz, “Knitro: An integrated package for nonlinear optimization,” *Large-Scale Nonlinear Optimization*, vol. 83, pp. 35–59, Feb. 2006.
- [87] I. Quesada and I. E. Grossmann, “An lp/nlp based branch and bound algorithm for convex minlp optimization problems,” *Computers & chemical engineering*, vol. 16, pp. 937–947, 1992.
- [88] B. Rustem, I. Akrotirianakis, and I. Maros, “An outer approximation based branch and cut algorithm for convex 0-1 minlp problems,” *Optimization Methods and Software*, vol. 16, pp. 21–47, June 2001.
- [89] A. Leon-Garcia, *probability, Statistics, and Random Processes For Electrical Engineering*. P. Prentice Hall, 2008.
- [90] M. Dorigo, V. Maniezzo, and A. Colorni, “Ant system: optimization by a colony of cooperating agents,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 26, pp. 29–41, 1996.
- [91] K.-K. Wong and E. Elsheikh, “Optimized cooperative diversity for a three-node decode-and-forward relay channels,” in *Proc. of 2nd International Symposium on Wireless Pervasive Computing*, (San Juan, Puerto Rico), Feb. 2007.
- [92] E. M. Elsheikh and K.-K. Wong, “Optimizing time and power allocation for cooperation diversity in a decode-and-forward three-node relay channel,” *Journal of Communications*, vol. 3, pp. 43–52, Apr. 2008.
- [93] C. H. Papadimitriou and K. Steiglitz, *Combinatorial optimization: algorithms and complexity*. Courier Corporation, 1998.

- [94] E. Tomita, A. Tanaka, and H. Takahashi, “The worst-case time complexity for generating all maximal cliques and computational experiments,” *Theoretical Computer Science*, vol. 363, pp. 28–42, Oct. 2006.
- [95] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms, Third Edition*. The MIT Press, 3rd ed., 2009.
- [96] M. Dorigo, “Optimization, learning and natural algorithms,” *Ph. D. Thesis, Politecnico di Milano, Italy*, 1992.
- [97] I. Hammerstrom and A. Wittneben, “On the optimal power allocation for non-regenerative ofdm relay links,” in *Proc. of IEEE International Conference on Communications*, (Istanbul, Turkey), June 2006.
- [98] W. Dang, M. Tao, H. Mu, and J. Huang, “Subcarrier-pair based resource allocation for cooperative multi-relay ofdm systems,” *IEEE Transactions on Wireless Communications*, vol. 9, pp. 1640–1649, May 2010.
- [99] J. Manssour, “Practical user pairing and selective network coding for multi-user bidirectional relaying,” in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC)*, (Shanghai, China), April 2012.
- [100] S. Sesia, I. Toufik, and M. Baker, *The UMTS Long Term Evolution: From Theory to Practice*. Wiley Online Library, 2009.
- [101] M. Moretti, A. Abrardo, and M. Belleschi, “On the convergence and optimality of reweighted message passing for channel assignment problems,” *IEEE Signal Processing Letters*, vol. 21, pp. 1428–1432, Nov. 2014.
- [102] “Feasibility study for further advancements for eutra (lte-advanced).” 3GPP TR 36.913 version 10.0.0, Technical Report, Apr. 2011.
- [103] S. P. Meyn and R. L. Tweedie, *Markov chains and stochastic stability*. Springer Science & Business Media, 2012.
- [104] M. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems*. Morgan & Claypool Publishers, 2010.
- [105] E. Leonardi, M. Melia, F. Neri, and M. A. Marson, “Bounds on average delays and queue size averages and variances in input-queued cell-based switches,” in *Proc. of IEEE INFOCOM*, (Anchorage, Alaska, USA), Apr. 2001.

Appendix A

Long Term Evolution (LTE) Background

Long Term Evolution (LTE) is a cellular standard developed by the 3rd Generation Partnership Project (3GPP) [4]. Historically, the 3GPP standard body also produced the standard for Global System for Mobile Communication (GSM), General Packet Radio Service (GPRS), Enhanced GPRS (EGPRS) EDGE, Wideband Code Division Multiple Access (WCDMA) and High Speed Packet Access (HSPA) technology. The newest LTE-Advanced standard is the next generation technology and is known as 4th Generation (4G) cellular technology that meets the ITU criteria for IMT Advanced 4G requirements. Figure A.1 shows the evolution of radio access technologies [2].

Compared to its predecessors, LTE represents a radical new step forward for the wireless industry, targeting order-of-magnitude increases in data rates by means of wider bandwidths and improved spectral efficiency. In addition, LTE aims to provide a highly efficient, low-latency, packet-optimized radio access technology offering enhanced spectrum flexibility. One big difference between LTE and other 3GPP technologies is that LTE uses Orthogonal Frequency Division Multiplex (OFDM) on the physical layer.

In this work, we consider 4G technology with multi-user OFDMA. OFDMA is the modulation and multiple access method currently being used in Long Term Evolution

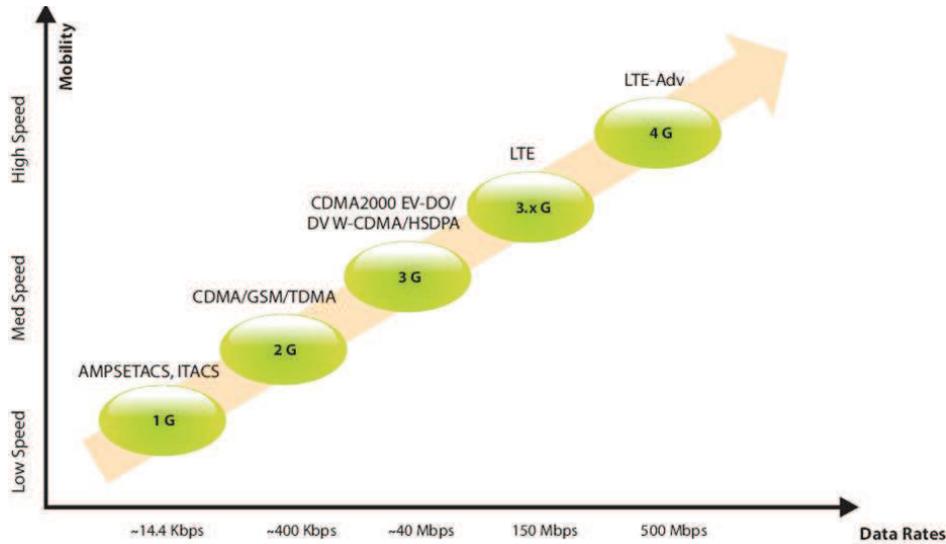


Figure A.1: Evolution of Radio Access Technologies [2]

(LTE) and LTE-Advanced. OFDMA is an extension of OFDM which divides a broadband channel into many narrowband subcarriers. Each subcarrier carries a modulated signal like a Quadrature Amplitude Modulation (QAM) signal or Quadrature Phase Shift Keying (QPSK) signal. The subcarriers are combined in a computationally efficient way and fed in to an Inverse Fast Fourier transform (IFFT) in the transmitter. OFDMA allows multiple users to transmit simultaneously on the different subcarriers per OFDM symbol to provide better rate and capacity.

A.1 LTE Network Architecture

The LTE network consists of the E-UTRA Network (E-UTRAN) and the Evolved Packet Core (EPC).

E-UTRA Network (E-UTRAN)

The E-UTRAN is comprised of the Evolved NodeB (eNodeB) base stations and the UE devices (smartphones, tablets, air cards, etc). The eNodeB provides the air

interface to the UEs, modulation, channel coding, and error handling. The E-UTRAN is responsible for all radio-related functions, which can be summarized briefly as:

- **Radio Resource Management (RRM):** This covers all functions related to the radio, such as radio bearer control, radio admission control, radio mobility control, as well as scheduling and dynamic allocation of resources to UEs in both uplink and downlink.
- **Header Compression:** This helps ensure efficient use of the radio interface by compressing the IP packet headers that could otherwise represent significant overhead, especially for small packets such as VoIP.
- **Security:** All data sent over the radio interface is encrypted.
- **Connectivity to the EPC:** This consists of the signaling toward Mobility Management Entity (MME) and the bearer path toward the Serving Gateway (S-GW).

The eNodeB and UE communicate over OFDMA in the downlink and Single Carrier Frequency Division Multiple Access (SC-FDMA) in the uplink, providing more robust and higher-bandwidth communication links.

Evolved Packet Core (EPC)

The EPC is the gateway from the E-UTRAN to Packet Data Networks (PDNs) and is also responsible for management of the UE devices and the connections between the UEs and PDNs. The EPC is responsible for the overall control of the UE and establishment of the bearers. The main logical nodes of the EPC are:

- **Mobility Management Entity (MME):** MME is responsible for establishing and managing the connections between the UE and the network. As such, the MME caches the UE subscriber data/profile and maintains the current location of the

UE. It is also responsible for security, and facilitating key delivery used for encrypting connections.

- **Serving Gateway (S-GW):** S-GW is noted as the mobility anchor for the UEs, as they move between eNodeBs during service or to/from legacy radio access technologies, such as UMTS (WCDMA). All IP packets between the UE and the PDNs are proxied through the S-GW.
- **PDN Gateway (P-GW):** P-GW is tasked with assigning IP addresses to the UEs as well as enforcing QoS on the individual data flows between the PDNs and UEs. The QoS rate shaping is performed by a Policy and Charging Enforcement Function (PCEF), which is part of the Policy and Charging Control (PCC) role. The PCC provides flow-based charging and QoS rate shaping per subscriber profile, which is becoming a key request for service providers in the face of the alarming growth rate of data traffic. It also reports events, such as location change, QoS enforcement, and IP address changes.

A.2 Relay Technologies in LTE-Advanced Cellular Networks

The Third Generation Partnership Project (3GPP) Long Term Evolution (LTE) is one of the most promising standards for next generation wireless systems. LTE is a candidate for the International Mobile Telecommunication (IMT) of International Telecommunications Union (ITU). The LTE standard was developed with the goal of providing a high-data rate, low-latency, and packet-optimized radio-access technology supporting flexible bandwidth deployments. However, as the achieved rates were still under the ITU requirements for the 4th Generation (4G) of mobile services [102],

new features were added to develop the so called LTE-Advanced (LTE-A) standard. Advanced relaying is one of the most prominent features in LTE-A systems, targeting at improving the cell-edge user performance and fairness among users. Relays Nodes (RNs) are designed to be low cost and lower power nodes which can be easily and quickly deployed when needed. In its release 10, LTE-A has standardized the use of relays [1]. The relay transmission can be achieved by forwarding information from a local base station (eNB) to a mobile station (UE) using a relay node (RN). Therefore, an RN can effectively extend the signal and service coverage of an eNB and enhance the overall throughput performance of a wireless communication system.

Relays Classifications

A variety of different classifications have been used to categorize relay nodes in LTE-A standard. In one category, relays may be distinguished based on the functionality [1] as follows:

- **Repeater:** This type of relays is the simplest in terms of implementation and functionality. The RN simply receives the signals from the eNodeB, amplifies it and then forwards it to the UE.
- **Decoder/Encoder:** This relay is able to decode the received signals and re-code the transmit signals in order to achieve higher RN-UE link quality. The advantage of achieving higher link quality comes at the expense of higher cost and complexity of the RN and also adds delay to the communication link.
- **Base Station:** This type of relay has the functionality of a base station like mobility management, session set-up, and handover. Such functionality adds more complexity to the implementation of the RN and the delay budget is further increased.

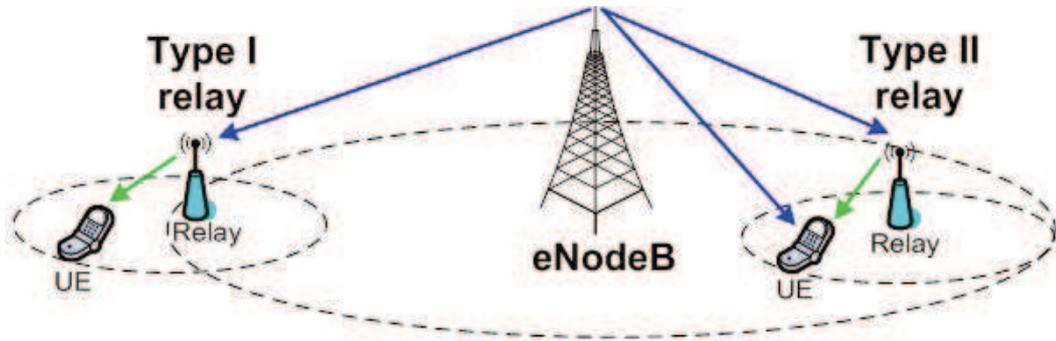


Figure A.2: LTE-Advanced Relay Architecture

A different classification is used in 3GPP standardization where two types of relays have been defined in 3GPP LTE-A standard [1], Type I (non-transparent) and Type II (transparent):

- **Type I:** This relay type can help a remote UE unit, which is located far away from a base station, to access the eNodeB. Type I relays transmit the common Reference Signal (RS) and control information to the eNodeB, and its main objective is to extend signal and service coverage, as shown in Figure A.2. Type I relays can mainly make some contributions to the overall system capacity by enabling communication services and data transmissions for remote UEs that are outside the eNodeBs coverage.
- **Type II:** This relay type can help a local UE unit, which is located within or outside the coverage of an eNodeB and has a direct communication link with the eNodeB, to improve its service quality and link capacity, as shown in Figure A.2. The main objective of a Type II RN is to increase the overall system capacity by achieving diversity and transmission gains for local UEs.

This thesis only considers Type II Decode-and-Forward (DF) relays. Our work uses the direct communication link to perform cooperative diversity and network coding.

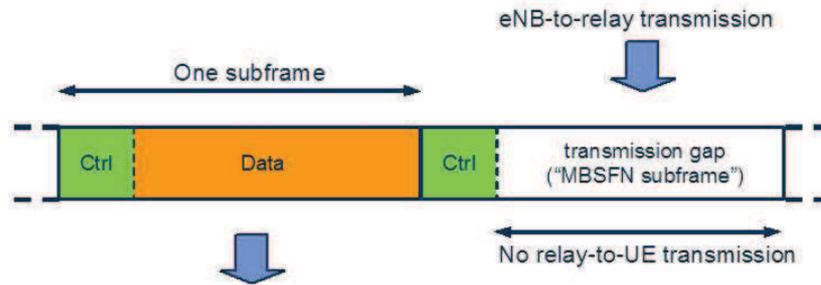


Figure A.3: MBSFN Subframe with Relay Support

Relay Support in LTE-Advanced Networks

According to the way that the wireless backhaul link is allocated, the RNs can be classified as:

- **In-band:** when the backhaul link is allocated on carriers at the same frequency Band as the access link.
- **Out-band:** when the backhaul link and the access link are allocated on carries at different frequency Bands.

For in-band relays, as both access and backhaul links use carriers on the same frequency Band, further constraints have to be imposed to avoid self-interference on the RN antennas. In LTE-A, the links are time multiplexed with the help of Multicast Broadcast Single Frequency Network (MBSFN) subframes. These are special subframes initially introduced for multimedia broadcasting during which the UEs will only receive if they are engaged in the broadcast service. In an enhanced relay scenario, MBSFN subframes are used to inform the subordinate UEs of an RN that no transmission is expected in the downlink during that subframe. In this case, UE would automatically turn off the receiver to save battery. In the uplink case, the RN may choose not to schedule any data from UEs for the subframes during which it

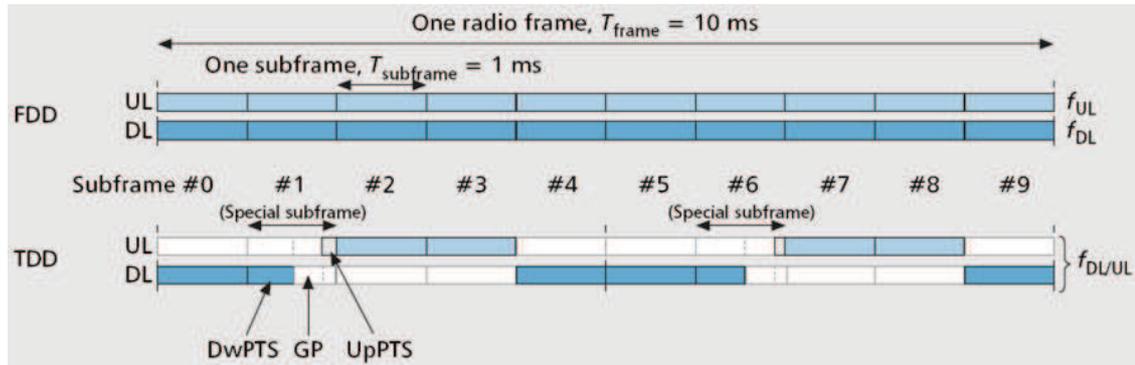


Figure A.4: FDD vs. TDD LTE Frame Structure [3]

will transmit to the eNodeB. Figure A.3 shows an MBSFN subframe configuration to support relays in LTE-A.

A.3 Frequency-Division Duplexing (FDD) versus Time-Division Duplexing (TDD)

LTE offers two implementation styles for the radio air interface: Frequency-Division Duplexing (FDD) and Time-Division Duplexing (TDD). In FDD, paired spectrum deployments allow uplink and downlink transmissions to occur at the same time while being separated in the frequency domain. This configuration allows full-duplex operations. For example, E-UTRA operating band 17 between 704-746 MHz is split into an uplink band from 704-716 MHz and a downlink band from 734-746 MHz. In contrast, TDD deployments will operate in half-duplex mode, sharing frequency but separating in the time domain. This allows a subset of the frequency spectrum to be allocated for either uplink or downlink operations (but not both) at a given time slot. Thus, there is no need for splitting the frequency band. Therefore, LTE TDD provides unpaired spectrum deployments.

The use of FDD versus TDD will dictate the structure and behavior of the radio frame. In either case, the radio frame is a 10ms duration superframe consisting of a set of subframes. In the FDD deployment, the radio frame structure will consist of ten equally-sized subframes of 1ms each. Each subframe will consist of two 0.5ms slots. As this deployment operates in full-duplex mode, there are 10 subframes in the uplink frequencies and 10 subframes in the downlink frequencies during the 10ms duration. The half-duplex TDD deployment consists of two 5ms half-frames. Each half-frame contains four subframes (of two 0.5ms slots each) and three special fields, called the Downlink Pilot Timeslot (DwPTS), Guard Period (GP) and Uplink Pilot Timeslot (UpPTS). In this deployment, downlink-to-uplink switch-points of 5ms and 10ms are supported. Figure A.4 illustrates the FDD and TDD LTE frame structure [3].

A.4 Spectrum Allocation and Resource Blocks

The MAC sublayer must allocate specific resources for the active connections, in both downlink and uplink directions, to allow transmissions to occur. The MAC sublayer allocates these resources in a unit called a Resource Block (RB), also called a Physical RB (PRB). An RB is defined as 12 consecutive OFDM frequency subcarriers known as a Resource Element (RE), each 15 kHz in size. The resulting RB is 180 kHz in size. In the time domain, RBs are defined as a 0.5ms time slot and are allocated in pairs. A 0.5ms time slot corresponds to 7 OFDM symbols using normal cyclic prefix or 6 OFDM symbols for extended cyclic prefix in both downlink and uplink directions. The resulting resource allocation map is a two-dimensional structure in the time and frequency domains which maps RBs to time slots as depicted in Figure A.5. A Resource Block Group (RBG) is a group of RBs where the size of the RBG depends

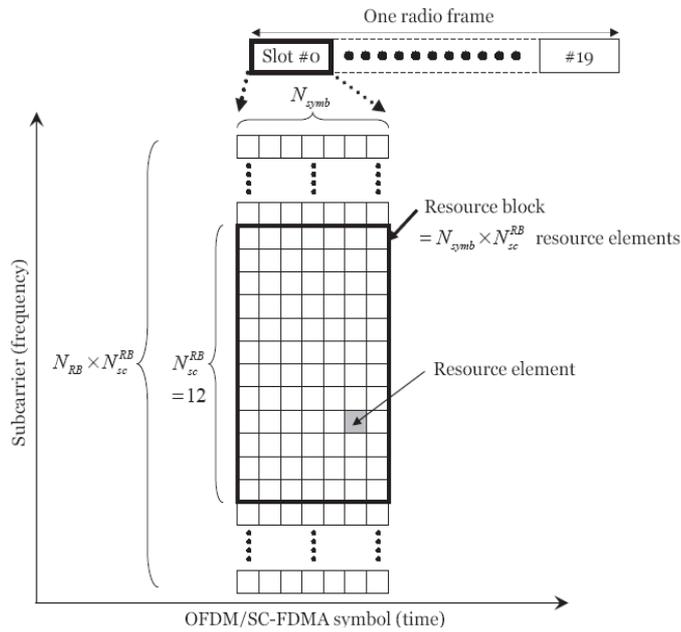


Figure A.5: LTE Resource Grid [3]

Table A.1: Transmission Bandwidth Configuration in LTE Channel

Channel Bandwidth (MHz)	1.4	3	5	10	15	20
Transmission Bandwidth Configuration (N_{RB})	6	15	25	50	75	100

on the system bandwidth in the cell. Resources are allocated in units of RBGs.

The size of transmission bandwidth varies based on its configuration as shown on Table A.1 [3]. The current 3GPP specifications allow channel bandwidths from 1.4MHz to 20MHz. Each channel bandwidth consists of resource blocks in the time and frequency domains. For example, when the transmission bandwidth configuration has 50 RBs, the channel bandwidth is 10MHz. All UEs and eNodeBs must support all of the defined channel bandwidths in the LTE specifications. This allows flexibility in network planning and design with assurance that any vendor equipment will comply with the chosen channel bandwidths.

Appendix B

CoR Rate Region

Proof of Proposition 1

For brevity, in this proof we remove the time index from our formulations. According to (3.3), the achievable rate region for Cooperative Relaying (CoR) is characterized by the following inequalities.

$$R_{\text{UE}} \leq \lambda_1 C_{\text{UR}} \quad R_{\text{UE}} \leq \lambda_1 C_{\text{UE}} + \lambda_2 C_{\text{RE}} \quad (\text{B.1})$$

$$R_{\text{EU}} \leq \lambda_3 C_{\text{ER}} \quad R_{\text{EU}} \leq \lambda_3 C_{\text{EU}} + \lambda_4 C_{\text{RU}} \quad (\text{B.2})$$

We define $\lambda_1^* = \lambda_1 + \lambda_2$, $\lambda_2^* = \lambda_3 + \lambda_4$ and bound R_{UE} and R_{EU} using (B.1) and (B.2) as follows.

$$R_{\text{UE}} \leq \min\{\lambda_1 C_{\text{UR}}, \lambda_1(C_{\text{UE}} - C_{\text{RE}}) + \lambda_1^* C_{\text{RE}}\} \quad (\text{B.3})$$

$$R_{\text{EU}} \leq \min\{\lambda_3 C_{\text{ER}}, \lambda_3(C_{\text{EU}} - C_{\text{RU}}) + \lambda_2^* C_{\text{RU}}\} \quad (\text{B.4})$$

To maximize the rate region of CoR, we need to maximize the right hand side of (B.3) and (B.4).

First, we will elaborate on the inequality in (B.3). We consider the following cases:

Case 1: If $C_{\text{UE}} \geq C_{\text{RE}}$, then

$$R_{\text{UE}} \leq \max_{\substack{\lambda_1 \\ 0 \leq \lambda_1 \leq \lambda_1^*}} \min\{\lambda_1 C_{\text{UR}}, \lambda_1(C_{\text{UE}} - C_{\text{RE}}) + \lambda_1^* C_{\text{RE}}\} = \lambda_1^* \min\{C_{\text{UR}}, C_{\text{UE}}\}. \quad (\text{B.5})$$

In this case, since $C_{\text{UE}} - C_{\text{RE}} \geq 0$ by choosing $\lambda_1 = \lambda_1^*$ both terms $\lambda_1 C_{\text{UR}}$, $\lambda_1(C_{\text{UE}} - C_{\text{RE}}) + \lambda_1^* C_{\text{RE}}$ are maximized and therefore, their minimum will be maximized as well.

Case 2: If $C_{\text{UE}} < C_{\text{RE}}$, then we consider the following sub-cases:

Sub-case 2.1: If $C_{\text{UR}} < C_{\text{UE}}$, then

$$\min\{\lambda_1 C_{\text{UR}}, \lambda_1(C_{\text{UE}} - C_{\text{RE}}) + \lambda_1^* C_{\text{RE}}\} = \min\{\lambda_1 C_{\text{UR}}, \lambda_1 C_{\text{UE}} + C_{\text{RE}}(\lambda_1^* - \lambda_1)\} = \lambda_1 C_{\text{UR}}.$$

This is concluded by considering the facts that $C_{\text{RE}}(\lambda_1^* - \lambda_1) \geq 0$ and $C_{\text{UR}} < C_{\text{UE}}$.

Hence,

$$\max_{\substack{\lambda_1 \\ 0 \leq \lambda_1 \leq \lambda_1^*}} \min\{\lambda_1 C_{\text{UR}}, \lambda_1(C_{\text{UE}} - C_{\text{RE}}) + \lambda_1^* C_{\text{RE}}\} = \max_{\substack{\lambda_1 \\ 0 \leq \lambda_1 \leq \lambda_1^*}} \lambda_1 C_{\text{UR}} = \lambda_1^* C_{\text{UR}}. \quad (\text{B.6})$$

Sub-case 2.2: If $C_{\text{UR}} \geq C_{\text{UE}}$, increasing λ_1 increases $\lambda_1 C_{\text{UR}}$ but decreases $\lambda_1(C_{\text{UE}} - C_{\text{RE}}) + \lambda_1^* C_{\text{RE}}$. Hence, their minimum is maximized

when $\lambda_1 C_{\text{UR}} = \lambda_1(C_{\text{UE}} - C_{\text{RE}}) + \lambda_1^* C_{\text{RE}}$. We can derive λ_1 based on λ_1^* and therefore, we have the following equality.

$$\max_{\substack{\lambda_1 \\ 0 \leq \lambda_1 \leq \lambda_1^*}} \min\{\lambda_1 C_{\text{UR}}, \lambda_1(C_{\text{UE}} - C_{\text{RE}}) + \lambda_1^* C_{\text{RE}}\} = \lambda_1^* \frac{C_{\text{UR}} C_{\text{RE}}}{C_{\text{UR}} + C_{\text{RE}} - C_{\text{UE}}} \quad (\text{B.7})$$

Using (B.6) and (B.7) in Case 2, R_{UE} can be bounded as

$$R_{UE} \leq \lambda_1^* C_{UR} \min \left\{ \frac{C_{RE}}{C_{UR} + C_{RE} - C_{UE}}, 1 \right\}. \quad (\text{B.8})$$

using a similar argument as what was used for inequality (B.4), we can have similar arguments for bounding R_{EU} . In particular, we have the following cases:

Case 3: If $C_{EU} \geq C_{RU}$, then

$$R_{EU} \leq \max_{\substack{\lambda_3 \\ 0 \leq \lambda_3 \leq \lambda_2^*}} \min \{ \lambda_3 C_{ER}, \lambda_3 (C_{EU} - C_{RU}) + \lambda_2^* C_{RU} \} = \lambda_2^* \min \{ C_{ER}, C_{EU} \}. \quad (\text{B.9})$$

In this case, since $C_{EU} - C_{RU} \geq 0$ by choosing $\lambda_3 = \lambda_2^*$ both terms $\lambda_3 C_{ER}$, $\lambda_3 (C_{EU} - C_{RU}) + \lambda_2^* C_{RU}$ are maximized and therefore, their minimum will be maximized as well.

Case 4: If $C_{EU} < C_{RU}$, then we consider the following sub-cases:

Sub-case 4.1: If $C_{ER} < C_{EU}$, then

$$\min \{ \lambda_3 C_{ER}, \lambda_3 (C_{EU} - C_{RU}) + \lambda_2^* C_{RU} \} = \min \{ \lambda_3 C_{ER}, \lambda_3 C_{EU} + C_{RU} (\lambda_2^* - \lambda_3) \} = \lambda_3 C_{ER}.$$

This is concluded by considering the facts that $C_{RU} (\lambda_2^* - \lambda_3) \geq 0$ and $C_{ER} < C_{EU}$.

Hence,

$$\max_{\substack{\lambda_3 \\ 0 \leq \lambda_3 \leq \lambda_2^*}} \min \{ \lambda_3 C_{ER}, \lambda_3 (C_{EU} - C_{RU}) + \lambda_2^* C_{RU} \} = \max_{\substack{\lambda_3 \\ 0 \leq \lambda_3 \leq \lambda_2^*}} \lambda_3 C_{ER} = \lambda_2^* C_{ER}. \quad (\text{B.10})$$

Sub-case 4.2: If $C_{ER} \geq C_{EU}$, increasing λ_3 increases $\lambda_3 C_{ER}$ but decreases $\lambda_3 (C_{EU} - C_{RU}) + \lambda_2^* C_{RU}$. Hence, their minimum is maximized

when $\lambda_3 C_{ER} = \lambda_3 (C_{EU} - C_{RU}) + \lambda_2^* C_{RU}$. We can derive λ_3 based on λ_2^* and therefore, we have the following equality.

$$\max_{\substack{\lambda_3 \\ 0 \leq \lambda_3 \leq \lambda_2^*}} \min\{\lambda_3 C_{ER}, \lambda_3(C_{EU} - C_{RU}) + \lambda_2^* C_{RU}\} = \lambda_2^* \frac{C_{ER} C_{RU}}{C_{ER} + C_{RU} - C_{EU}} \quad (\text{B.11})$$

Using (B.10) and (B.11) in Case 4, R_{EU} can be bounded as

$$R_{EU} \leq \lambda_2^* C_{ER} \min\left\{\frac{C_{RU}}{C_{ER} + C_{RU} - C_{EU}}, 1\right\}. \quad (\text{B.12})$$

Since $\lambda_1^* + \lambda_2^* = 1$, by considering Cases 1, 2, 3 and 4 above, the rate region of CoR is characterized by (B.13) in which $C_{UE}^{\text{CoR}}(t)$ and $C_{EU}^{\text{CoR}}(t)$ are defined in (3.5) and (3.6).

$$\frac{R_{UE}}{C_{UE}^{\text{CoR}}(t)} + \frac{R_{EU}}{C_{EU}^{\text{CoR}}(t)} \leq 1 \quad (\text{B.13})$$

where $C_{UE}^{\text{CoR}}(t)$ and $C_{EU}^{\text{CoR}}(t)$ are defined as follows:

$$\bullet C_{UE}^{\text{CoR}}(t) = \begin{cases} \min\{C_{UR}, C_{UE}\} & \text{if } C_{UE} \geq C_{RE} \\ C_{UR} \min\left\{\frac{C_{RE}}{C_{UR} + C_{RE} - C_{UE}}, 1\right\} & \text{if } C_{UE} < C_{RE} \end{cases}$$

$$\bullet C_{EU}^{\text{CoR}}(t) = \begin{cases} \min\{C_{ER}, C_{EU}\} & \text{if } C_{EU} \geq C_{RU} \\ C_{ER} \min\left\{\frac{C_{RU}}{C_{ER} + C_{RU} - C_{EU}}, 1\right\} & \text{if } C_{EU} < C_{RU} \end{cases}$$

Appendix C

NC/CoR Rate Region

Proof of Proposition 2

For brevity, in this proof we remove the time index from our formulations. According to (3.7), the achievable rate region for the Network Coding/Cooperative Relaying (NC/CoR) is characterized by the following inequalities.

$$R_{\text{UE}} \leq \min\{\lambda_1 C_{\text{UR}}, \lambda_1 C_{\text{UE}} + \lambda_3 C_{\text{min}}\} \quad (\text{C.1})$$

$$R_{\text{EU}} \leq \min\{\lambda_2 C_{\text{ER}}, \lambda_2 C_{\text{EU}} + \lambda_3 C_{\text{min}}\} \quad (\text{C.2})$$

In (C.1) and (C.2), $C_{\text{min}} = \min\{C_{\text{RE}}, C_{\text{RU}}\}$. We consider the following cases:

Case 1: $C_{\text{UE}} \geq C_{\text{UR}}$ and $C_{\text{EU}} \geq C_{\text{ER}}$: Then, we have the following equalities.

$$\min\{\lambda_1 C_{\text{UR}}, \lambda_1 C_{\text{UE}} + \lambda_3 C_{\text{min}}\} = \lambda_1 C_{\text{UR}} \quad \min\{\lambda_2 C_{\text{ER}}, \lambda_2 C_{\text{EU}} + \lambda_3 C_{\text{min}}\} = \lambda_2 C_{\text{ER}}$$

These equalities follow since $C_{\text{min}} \geq 0$ and $\lambda_1, \lambda_2, \lambda_3 \geq 0$. To maximize the rate region, we need to have $\lambda_3 = 0$ and $\lambda_1 + \lambda_2 = 1$. Hence, the rate region is given by

$$\frac{R_{\text{UE}}}{C_{\text{UR}}} + \frac{R_{\text{EU}}}{C_{\text{ER}}} \leq 1. \quad (\text{C.3})$$

Case 2: $C_{\text{UE}} \geq C_{\text{UR}}$ and $C_{\text{EU}} < C_{\text{ER}}$: In this case, since $C_{\text{UE}} \geq C_{\text{UR}}$, we always have

$$R_{\text{UE}} \leq \min\{\lambda_1 C_{\text{UR}}, \lambda_1 C_{\text{UE}} + \lambda_3 C_{\text{min}}\} = \lambda_1 C_{\text{UR}}.$$

For the second inequality, i.e., (C.2), we consider the following sub-cases:

Sub-case 2.1: If $C_{\text{min}} < C_{\text{EU}}$, then

$$\max_{\substack{\lambda_3 \\ \lambda_1, \lambda_2, \lambda_3 \geq 0 \\ \lambda_2 + \lambda_3 = 1 - \lambda_1}} \min\{\lambda_2 C_{\text{ER}}, \lambda_2 C_{\text{EU}} + \lambda_3 C_{\text{min}}\} = \lambda_2 C_{\text{EU}}. \quad (\text{C.4})$$

We prove this as follows: Assume that $\lambda_2 + \lambda_3 = \lambda^*$. Thus, we have $\min\{\lambda_2 C_{\text{ER}}, \lambda_2 C_{\text{EU}} + \lambda_3 C_{\text{min}}\} = \min\{\lambda_2 C_{\text{ER}}, \lambda_2(C_{\text{EU}} - C_{\text{min}}) + \lambda^* C_{\text{min}}\}$. Since $C_{\text{EU}} - C_{\text{min}} > 0$, we observe that $\min\{\lambda_2 C_{\text{ER}}, \lambda_2(C_{\text{EU}} - C_{\text{min}}) + \lambda^* C_{\text{min}}\}$ is maximized if we let λ_2 be maximized, i.e., $\lambda_2 = \lambda^*$ (or $\lambda_3 = 0$) and therefore the result follows. Thus, the rate region is specified by

$$\frac{R_{\text{UE}}}{C_{\text{UR}}} + \frac{R_{\text{EU}}}{C_{\text{EU}}} \leq 1. \quad (\text{C.5})$$

Sub-case 2.2: If $C_{\text{min}} \geq C_{\text{EU}}$, to maximize the right hand side of (C.2), we must have $\lambda_2 C_{\text{ER}} = \lambda_2 C_{\text{EU}} + \lambda_3 C_{\text{min}}$. Thus, $\lambda_3 = \lambda_2 \frac{C_{\text{ER}} - C_{\text{EU}}}{C_{\text{min}}}$. Since $\lambda_1 + \lambda_2 + \lambda_3 = 1$, we obtain one equality as a constrain on the time sharing parameters $\lambda_1, \lambda_2, \lambda_3$ as $\lambda_2 \frac{C_{\text{min}} + C_{\text{ER}} - C_{\text{EU}}}{C_{\text{min}}} + \lambda_1 = 1$. Therefore, the rate region is specified by all the positive points $(R_{\text{UE}}, R_{\text{EU}})$ that satisfy $R_{\text{UE}} \leq \lambda_1 C_{\text{UR}}$ and $R_{\text{EU}} \leq \lambda_2 C_{\text{ER}}$ while $\lambda_2 \frac{C_{\text{min}} + C_{\text{ER}} - C_{\text{EU}}}{C_{\text{min}}} + \lambda_1 = 1$. We can verify that the region is described by the following inequality.

$$\frac{R_{\text{UE}}}{C_{\text{UR}}} + \frac{R_{\text{EU}}}{\frac{C_{\text{min}} C_{\text{ER}}}{C_{\text{min}} + C_{\text{ER}} - C_{\text{EU}}}} \leq 1 \quad (\text{C.6})$$

Case 3: $C_{UE} < C_{UR}$ and $C_{EU} \geq C_{ER}$: For this case, we use similar arguments as what we did for the previous case. Since $C_{EU} \geq C_{ER}$, we can conclude that

$$R_{EU} \leq \min\{\lambda_2 C_{ER}, \lambda_2 C_{EU} + \lambda_3 C_{\min}\} = \lambda_2 C_{ER}.$$

For the other inequality, i.e., (C.1), we have the following sub-cases:

Sub-case 3.1: If $C_{\min} < C_{UE}$, then

$$\max_{\substack{\lambda_3 \\ \lambda_1, \lambda_2, \lambda_3 \geq 0 \\ \lambda_1 + \lambda_3 = 1 - \lambda_2}} \min\{\lambda_1 C_{UR}, \lambda_1 C_{UE} + \lambda_3 C_{\min}\} = \lambda_1 C_{UE}. \quad (\text{C.7})$$

The proof is similar to the proof in Sub-case 2.1. Thus, the rate region is specified by

$$\frac{R_{UE}}{C_{UE}} + \frac{R_{EU}}{C_{ER}} \leq 1. \quad (\text{C.8})$$

Sub-case 3.2: If $C_{\min} \geq C_{UE}$, to maximize the right hand side of (C.1), we must have $\lambda_1 C_{UR} = \lambda_1 C_{UE} + \lambda_3 C_{\min}$ and thus $\lambda_3 = \lambda_1 \frac{C_{UR} - C_{UE}}{C_{\min}}$. Using a similar argument used in Sub-case 3.2 (we do not repeat it here), we can verify that the rate region is specified by all the positive points (R_{UE}, R_{EU}) that satisfy $R_{UE} \leq \lambda_1 C_{UR}$ and $R_{EU} \leq \lambda_2 C_{ER}$ while $\lambda_1 \frac{C_{\min} + C_{UR} - C_{UE}}{C_{\min}} + \lambda_2 = 1$. Thus, the region is specified by the following inequality.

$$\frac{R_{UE}}{\frac{C_{\min} C_{UR}}{C_{\min} + C_{UR} - C_{UE}}} + \frac{R_{EU}}{C_{ER}} \leq 1 \quad (\text{C.9})$$

Case 4: $C_{UE} < C_{UR}$ and $C_{EU} < C_{ER}$: In this case, we consider the following four cases:

Sub-case 4.1: $C_{\min} < C_{UE}$ and $C_{\min} < C_{EU}$: Then, using the arguments for (C.4) in Sub-case 2.1 and (C.7) in Sub-case 3.1, the region is specified by $\frac{R_{UE}}{C_{UE}} + \frac{R_{EU}}{C_{EU}} \leq 1$.

Sub-case 4.2: $C_{\min} \geq C_{UE}$ and $C_{\min} < C_{EU}$: In this case, to maximize the right hand side in (C.1), we must have $\lambda_1 C_{UR} = \lambda_1 C_{UE} + \lambda_3 C_{\min}$ and hence $\lambda_3 = \lambda_1 \frac{C_{UR} - C_{UE}}{C_{\min}}$. Since $\lambda_1 + \lambda_2 + \lambda_3 = 1$, we obtain $\lambda_1 \frac{C_{\min} + C_{UR} - C_{UE}}{C_{\min}} + \lambda_2 = 1$.

For the inequality in (C.2), R_{EU} must be less than or equal to both terms $\lambda_2 C_{ER}$ and $\lambda_2 C_{EU} + \lambda_3 C_{\min}$. If we put $R_{EU} \leq \lambda_2 C_{ER}$, the rate region is specified by $R_{UE} \leq \lambda_1 C_{UR}$ and $R_{EU} \leq \lambda_2 C_{ER}$ while $\lambda_1 \frac{C_{\min} + C_{UR} - C_{UE}}{C_{\min}} + \lambda_2 = 1$. If $R_{EU} \leq \lambda_2 C_{EU} + \lambda_3 C_{\min} = \lambda_2 C_{EU} + \lambda_1 (C_{UR} - C_{UE})$, the rate region is specified by $R_{UE} \leq \lambda_1 C_{UR}$ and $R_{EU} \leq \lambda_2 C_{EU} + \lambda_1 (C_{UR} - C_{UE})$ while $\lambda_1 \frac{C_{\min} + C_{UR} - C_{UE}}{C_{\min}} + \lambda_2 = 1$. Consequently, the rate region is given by the intersection of the following regions.

$$\frac{\frac{R_{UE}}{C_{UR} C_{\min}}}{\frac{C_{\min} + C_{UR} - C_{UE}}{C_{\min}}} + \frac{R_{EU}}{C_{ER}} \leq 1 \quad (\text{C.10})$$

$$\frac{\frac{R_{UE}}{1}}{\frac{C_{\min} + C_{UR} - C_{UE}}{C_{UR} C_{\min}} - \frac{C_{UR} - C_{UE}}{C_{UR} C_{EU}}} + \frac{R_{EU}}{C_{EU}} \leq 1 \quad (\text{C.11})$$

Sub-case 4.3: $C_{\min} < C_{UE}$ and $C_{\min} \geq C_{EU}$: The argument for this sub-case is similar to the previous one. If we only swap the subscript U and E, we will have the same condition as in Sub-case 4.3. We do not repeat the argument and just present the final result. In this case, the rate region is given by the intersection of

the following regions.

$$\frac{R_{\text{UE}}}{C_{\text{UR}}} + \frac{R_{\text{EU}}}{\frac{C_{\text{ER}}C_{\text{min}}}{C_{\text{min}} + C_{\text{ER}} - C_{\text{EU}}}} \leq 1 \quad (\text{C.12})$$

$$\frac{R_{\text{UE}}}{C_{\text{UE}}} + \frac{R_{\text{EU}}}{\frac{1}{\frac{C_{\text{min}} + C_{\text{ER}} - C_{\text{EU}}}{C_{\text{ER}}C_{\text{min}}} - \frac{C_{\text{ER}} - C_{\text{EU}}}{C_{\text{ER}}C_{\text{UE}}}}} \leq 1 \quad (\text{C.13})$$

Sub-case 4.4: $C_{\text{min}} \geq C_{\text{UE}}$ and $C_{\text{min}} \geq C_{\text{EU}}$. To maximize the right hand side of (C.1) and (C.2), we must have $\lambda_1 C_{\text{UR}} = \lambda_1 C_{\text{UE}} + \lambda_3 C_{\text{min}}$ and $\lambda_2 C_{\text{ER}} = \lambda_2 C_{\text{EU}} + \lambda_3 C_{\text{min}}$. Since $\lambda_1 + \lambda_2 + \lambda_3 = 1$, thus $\lambda_1 \frac{C_{\text{min}} + C_{\text{UR}} - C_{\text{UE}}}{C_{\text{min}}} + \lambda_2 = 1$ and $\lambda_2 \frac{C_{\text{min}} + C_{\text{ER}} - C_{\text{EU}}}{C_{\text{min}}} + \lambda_1 = 1$. Hence, the rate region is specified by $R_{\text{UE}} \leq \lambda_1 C_{\text{UR}}$ and $R_{\text{EU}} \leq \lambda_2 C_{\text{ER}}$ while $\lambda_1 \frac{C_{\text{min}} + C_{\text{UR}} - C_{\text{UE}}}{C_{\text{min}}} + \lambda_2 = 1$ and $\lambda_2 \frac{C_{\text{min}} + C_{\text{ER}} - C_{\text{EU}}}{C_{\text{min}}} + \lambda_1 = 1$. Thus, the rate region is given by the intersection of the following inequalities.

$$\frac{R_{\text{UE}}}{\frac{C_{\text{UR}}C_{\text{min}}}{C_{\text{min}} + C_{\text{UR}} - C_{\text{UE}}}} + \frac{R_{\text{EU}}}{C_{\text{ER}}} \leq 1 \quad (\text{C.14})$$

$$\frac{R_{\text{EU}}}{\frac{C_{\text{ER}}C_{\text{min}}}{C_{\text{min}} + C_{\text{ER}} - C_{\text{EU}}}} + \frac{R_{\text{UE}}}{C_{\text{UR}}} \leq 1 \quad (\text{C.15})$$

The cases considered above cover all the possible cases for the channel capacity of all the communication links in the system. The rate region derived above can be summarized in Cases 1 and 2 of Proposition 2 given by (3.8), (3.11) and (3.12).

Appendix D

Lyapunov Stability Analysis

Proof of Proposition 4

Definition 2. A queue is called strongly stable if it has bounded time averaged expected backlog [84], i.e.

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} E[X(\tau)] < \infty$$

Definition 2 can be extended to the bidirectional queueing network as follows.

Definition 3. A queueing system is called to be strongly stable if all the queues in the system are strongly stable. Specifically, for a bidirectional network consisting of N uplink and downlink queues, the system is strongly stable if the time averaged expected aggregated backlog in the network is bounded, i.e.

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \sum_{n=1}^N \{E[X_{U_n}(\tau)] + E[X_{E_n}(\tau)]\} < \infty$$

Consider the queueing network in this work evolves in discrete time with normalized time slots t in $\{0, 1, 2, \dots\}$. There are N uplink and downlink queues in the

bidirectional network. The proof is shown for one bidirectional UE-eNB pair. However, the same conclusions can be drawn for all N UE-eNB pairs. Let us define the vector of queue backlogs in uplink and downlink at the beginning of time-frame t by: $X(t) = (X_U(t), X_E(t))$.

For each slot t , define $L(X(t))$ as the quadratic Lyapunov function (sum of the squares of the current queue backlogs)

$$L(X(t)) = X_U(t)^2 + X_E(t)^2,$$

Define the Lyapunov drift as the change in this function from one slot to the next

$$\Delta(X(t)) = L(X(t+1)) - L(X(t)),$$

Suppose the queue backlogs at the UE and eNB change over time as follows:

$$X_U(t+1) = (X_U(t) - R_{UE}(t))^+ + A_U(t)$$

$$X_E(t+1) = (X_E(t) - R_{EU}(t))^+ + A_E(t)$$

The above equations can be used to compute a bound on the Lyapunov drift for slot t :

$$X(t+1)^2 = \max[(X(t) - R(t))^+ + A(t)]^2 < (X(t) - R(t) + A(t))^2$$

This leads to the following Equation:

$$\Delta(X(t)) \leq B(t) + X_U(t)(A_U(t) - R_{UE}(t)) + X_E(t)(A_E(t) - R_{EU}(t)) \quad (\text{D.1})$$

where $B(t) = B_1(t) + B_2(t)$ and $B_1(t)$ and $B_2(t)$ are defined as:

$$\begin{aligned} B_1(t) &= \frac{1}{2}[A_U(t)^2 + R_{UE}(t)^2 - 2A_U(t)R_{UE}(t)] \\ B_2(t) &= \frac{1}{2}[A_E(t)^2 + R_{EU}(t)^2 - 2A_E(t)R_{EU}(t)] \end{aligned}$$

Suppose the second moments of arrivals and service in the uplink and downlink queues are bounded, so that there is a finite constant $B > 0$ such that for all t and all possible queue pairs $X(t)$, the following property holds: $E[B(t)|X(t)] \leq B$. Taking conditional expectations of Equation D.1 leads to the following bound on the conditional expected Lyapunov drift:

$$\begin{aligned} E[\Delta(X(t))|X(t)] &\leq B + X_U(t)E[A_U(t) - R_{UE}(t)|X(t)] \\ &\quad + X_E(t)E[A_E(t) - R_{EU}(t)|X(t)] \end{aligned} \quad (\text{D.2})$$

By Foster-Lyapunov criterion [103], Equation D.2 reduces to the drift condition as [104, 105]: Suppose there are constants $B \geq 0$, $\epsilon > 0$ such that for all t and all possible vectors $X(t)$ the conditional Lyapunov drift satisfies:

$$E[\Delta(X(t))|X(t)] \leq B - \epsilon(X_U(t) + X_E(t))$$

Then for all slots $t > 0$ the time average queue size in the network satisfies:

$$\frac{1}{t} \sum_{\tau=0}^{t-1} \{E[X_U(\tau)] + E[X_E(\tau)]\} \leq \frac{B}{\epsilon} + \frac{E[L(X(0))]}{\epsilon t}$$

This proves Definition 3 by following the same steps for all N bidirectional UE-eNB pairs.