ENHANCING THE PERFORMANCE OF MULTICAST SYSTEMS WITH LAYERED TRANSMISSIONS

by

Rashed Alsakarnah

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
April 2019

To all beloved ones who have been supporting me through my journey

Table of Contents

List of Tables

List of Figures

Abstract

Multicast services are an indispensable part of present and future wireless networks due to their ability to transmit the same data to a group of users within the same radio resources. To improve the efficiency of multicast data distribution, this dissertation investigates layered transmissions in cooperative networks with multi-antenna terminals. Layered transmissions are particularly promising for multicast services because they provide flexible use of the broadcast spectrum through the transmission of several data streams over the same radio resources. The focus of this dissertation is to avoid using a conservative approach, which foresees a system data rate bounded by the user with the worst channel conditions. With these motivations, this dissertation presents three main contributions to mitigate interference effects in multicast networks with relays and to improve bandwidth efficiency and fairness among multicast users and groups.

Firstly, to improve throughput and data reliability for cell-edge users, layered transmissions and successive relaying are considered. This is accomplished by utilizing superposition coding (SC) at the base station (BS) and successive interference cancellation (SIC) at both relays and cell-edge users. In the developed decode-and-forward (DF) transmission schemes, transmissions from the BS and the relays are aligned in the time, power, and spatial domains to improve the system throughput and reliability.

Secondly, a transmission scheduling system for downlink cellular networks which incorporates cooperative relaying and a virtual MIMO method is presented for a multi-group multicast multiple-input single-output (MISO) system. Cooperative relaying is used to recover independent messages for desired users in a multicast group (MG) by taking advantage of independent signal replicas and channel reuse at different locations.

Finally, a multicast transmission system based on pattern division multiple access (PDMA) is developed to allocate the available radio resources fairly among users and groups. A joint design for both transmitter and receiver is considered. At the receiver, multiple groups are detected by using a SIC detection method.

This dissertation advances the theory of multicast data distribution using layered division multiplexing within the framework of cooperative communications system design. The results presented in this work are applicable to wireless networks ranging from infrastructure-based to ad-hoc networks with different antenna configurations at the terminals.

List of Abbreviations Used

The following abbreviations and acronyms are used in this dissertation.

AF	amplify-and-forward
AWGN	additive white Gaussian noise
BER	bit error rate
BPSK	binary phase shift keying
BS	base station
CMS	conventional multicast scheme
CSI	channel state information
DF	decode-and-forward
FDMA	frequency division multiple access
i.i.d	independent, identically distributed
IRI	inter-relay interference
MAI	multiple access interference
MDC	multiple description coding
MIMO	multiple-input multiple-output
MISO	multiple-input single-output
MS	mobile station
MU	multiuser
MG	multicast group
OFDM	orthogonal frequency-division multiplexing
pdf	probability density function
PDMA	pattern division multiple access
QAM	quadrature amplitude modulation
QPSK	quadrature phase shift keying
RE	radio element
RF	radio frequency

RV	random variable
SIC	successive interference canceler
SIMO	single input multiple-output
SISO	single input single-output
SIR	signal to interference ratio
SNIR	signal-to-interference-to-noise ratio
SNR	signal-to-noise ratio
SC	superposition coding
SU	single user
TDMA	time division multiple access
TS	time slot
ZF	zero-forcing

Acknowledgements

I wish to express my sincere gratitude and appreciation to Professor Jacek Ilow for his guidance and support throughout the duration of my study at Dalhousie University. I would like also to thank him for always being available for assistance and advice over the period of my studies and for the personal experiences that he has shared with me. I appreciate all of the time, ideas, and funding he contributed to my Ph.D studies.

I wish to thank Dr. Dmitry Trukhachev and Dr. William Phillips for serving on my PhD. Supervisory Committee. My thanks extended to Dr. Cheng Li for accepting to be the external examiner of my PhD defence.

I wish to extend my thanks to Daniela Valente for always being there to help and encourage me, and Dr. Ilow's research group, especially Fadhel Alhumaidi, Javad Hoseyni, and Aasem Alyahya.

I would like to thank my family back in Jordan, with special thanks to my brother Omar for the support he has shown since my arrival in Canada. Special thanks is due to my father and mother for supporting me emotionally and spiritually my whole life.

Finally, I would like to express my deep appreciation and thanks to Alzaytouneh University for the generous financial support during my studies in Canada.

Chapter 1

Introduction

With increasing demands to support higher data rates in mobile applications, there is a universal need for the wireless communication industry to improve the bandwidth efficiency of new communication systems [1]. Of particular interest are some applications which require transmission the same data to selected groups of users that naturally lend themselves towards multicasting. For instance, local news, geographic information updates such as traffic reports, weather forecast and location-based adverts. Such wireless multicasting may as well appear in wireless sensor networks where a gateway sensor node needs to simultaneously send the same control information to a group of sensor nodes [2], [3]. Multimedia entertainments in all forms, which currently account for one-third of mobile Internet market , are one of the main applications that can be deployed using multicast technology [4], [5]. Over the last few years, wireless multicasting has become an important technology in this area as it allows the transmission of the same information to multiple users at the same time.

To address the requirements of new multicast services where there are more and more users requiring improved quality of service (QoS), enhancements to multicast transmission schemes have to be introduced [6], [7]. In order to efficiently manage radio resources in these systems, layered transmissions are emerging as a promising technique to deliver more diversified multicast services using broadcasting-type infrastructure including relays. LDM is a non-orthogonal multiplexing (NOM) technique, which has been included in the ATSC 3.0 standard [8]. Splitting the main data stream into sub-streams with different importance is characteristic in multimedia applications such as video and audio transmissions. In this dissertation, the focus is on designing advanced signal processing techniques to aid the operation of wireless multicasting systems with significantly improved capacities.

Wireless multicasting is based on serving a group of users simultaneously by using the same radio resources, such as time, frequency, power, and spatial streams. The multicast groups are usually created by the BS, and users are usually added to each group according to their request or interest. Therefore, wireless multicasting enables interested users to form groups so that they can share allocated resources in order to receive the same data of interest.

In conventional multicast systems, the BS imposes the requirement to select a single data rate that allows each user in the multicast group (MG) to decode the same transmitted data based on the instantaneous achievable rates of the minimum users at the cell edge. As a result, the user with the worst channel conditions limits the multicast group throughput. Moreover, because of variation in channels conditions among users within the same multicast group, conventional radio resource management (RRM) favors users with good channel conditions over cell-edge users (users with bad channel conditions). One of the most successful and efficient solutions for addressing this and other issues in multicast wireless is deploying layered transmissions [9], [10]. Conventionally, layered transmission is a technique to make flexible use of the broadcast spectrum by transmitting several data streams on the same radio resources. Splitting the data stream into sub-streams of different importance is characteristic in multimedia applications such as video and audio transmissions. When combining hierarchical transmission of video sequences with hierarchical or layer modulations, different levels of resilience can be obtained when decoding video at the terminal with different signal to noise ratios (SNRs). This hierarchal transmission has already been used in point-to-point video broadcasting systems such as Digital Video Broadcasting (DVB) using asymmetric (hierarchical) modulations. In this dissertation, we refer to this approach of representing the original bit stream with signaling technique corresponding to bit sub-streams as a bit level layered transmission. When the different streams are separated in power domain, this can be considered as non orthogonal multiple access (NOMA), or in general as Layered-Division Multiplexing (LDM). An important motivation of this research is the prospect of applying layered transmission protocols to multicast services and tailoring them to existing solutions to improve the performance of cell-edge users while providing users having good channel conditions with an acceptable QoS.

There are several approaches when deploying NOMA to improve the utilization of radio spectrum and power resources when users share the same radio resources while transmitting and receiving [11], [12]. Compared to orthogonal multiple access (OMA), NOMA relaxes the orthogonality constrain for decoding purposes in the radio resource allocation. Therefore, the user/group scheduling problem constrained by the limited time and bandwidth resources in conventional multicast systems is not a binary selection any longer, but the joint optimization of power, frequency, code signature, and receiver design. Therefore, fairness among users/groups in terms of throughput and QoS can be improved by exploiting the flexibility offered by NOMA as pursued in this disertation [13].

In wireless networks, cooperative communications with relays have long been used because of the ability to extend the range of the network [14]–[16]. Relays allows densification of mobile networks without incurring additional fiber deployment cost because they do not need to be physically connected to the BS. The deployment of cooperative communications with NOMA can further enhance the performance of wireless multicast systems in terms of reliability and capacity as pursued in this dissertation. When all relays operate in half-duplex (HD) mode, traditionally, the full transmission cycle consumes more time than a direct communications process, because the relays require additional time slot to send signals to the destinations [17]. In this area, a promising solution to overcome the loss in spectral efficiency (the pre-log factor is $\frac{1}{2}$ in capacity calculations) consists of alternate relaying (AR), which reuses source transmission time slots (TSs) for relay transmissions. Because of the broadcast nature of the wireless medium, when multiple signals are transmitted at the same time and frequency (from the BS and relays) to improve bandwidth utilization, multiple access interference (MAI) is inevitably introduced. When alternate relaying is deployed, the impact of inter-relay interference (IRI) at the two relays has to be considered. In general, uncoordinated interference reduces wireless network throughput. In order to address this problem, this dissertation seeks to design signal processing strategies and wireless network protocols that manage interference in HD decode-and-forward (DF) alternate relaying (AR) systems to achieve the advantageous network performance with improved capacity of cell-edge users. In particular, this dissertation introduces system level improvements to combat the IRI in AR multicast systems by exploiting power and/or spatial domains of NOMA or super-position coding (SC) transmissions between different nodes in the wireless system.

The use of multiple transmit and/or multiple receive antennas, where different data streams are transmitted using the same radio resources, opens new possibilities to improve spectrum utilization [18]–[22]. In addition to time and frequency aspects, a new spatial dimension is being explored to find more efficient approaches to deal with long-standing issues in the design of communication systems, such as bandwidth and power limitations. Our effort for improving throughput and reliability in multicast transmissions when deploying multiple-input multiple-output (MIMO) techniques does not follow the basic single user MIMO configurations. In our cooperative MIMO transmissions (and signal decoding to remove the IRI), multiple antenna nodes share their antennas to achieve the effect of a virtual antenna array, and as a result, the overall network capacity is improved significantly. As such, this dissertation follows the trend in radio communications system designs based on multiuser (MU) MIMO and more advanced relaying strategies. However, rather than maximizing the diversity or multiplexing gains of unicast/broadcast transmissions, we target to optimize the effectiveness of multicast transmissions.

In multicast MIMO networks which are still under development, spatial degrees of freedom offered by multiple antennas at the BS and relays are exploited to enhance the system capacity, e.g., by scheduling users within a MG to share the spatial channel simultaneously. This dissertation works with time scheduling, frequency reuse strategies and transceiver methodologies in virtual MU-MIMO setup to obtain performance gains and beneficial trade-offs. In this dissertation, we refer to this approach of representing the original bit stream with signaling technique corresponding to bit sub-streams as spatial level layered transmissions.

In some parts of this work, motivated by the design of relay-assisted downlink transmissions to support increased data rates in unicast communications, a multicast transmission scheme is developed which creates a "reusable" virtual antenna array for terminals within the same multicast group. Specifically, in the first stage of the communication process, the BS transmits to different multicast groups in their designated time slots (TSs). In the next stage, within the same multicast group, users and their assigned relays form an independent network. Then, in multiple TSs, different single antenna relays redistribute (within one multicast group) different multiple-input single-output (MISO) versions of the original BS data. In this system, an acceptable level of MAI is controlled by using a specialized frequency reuse plan and power control and there is no need for the IRI cancellation. In this area of our work, we effectively combine the cooperation gains with the ability to effectively multiplex multicast sessions with the purpose of maximizing the aggregate data rate in the multicast system.

In wireless multicast network with multiple multicast groups (MGs), as a result of high variations in channel conditions among users within the same MG and QoS requirements for each MGs, ensuring fairness among users/groups (e.g. comparable data rates for users with various channel conditions) poses a serious challenge in the design of multicast transmission systems. To allocate the radio resources fairly among users and MGs, a promising solution that improves the spectral efficiency while providing an acceptable level of both inter- and intra-group fairnesses is the pattern division multiple access (PDMA) which is another form of NOMA. PDMA is based on mapping the information for different users to radio resources such as power, code, time, and spatial streams [23]. It utilizes SIC based detection at the receiver and SIC amenable pattern design at the transmitter. In this dissertation, we developed a scheme based on PDMA principles to assign a unique pattern or signature to each MG sharing the same radio resources to optimize the fairness among users/ groups. Fairness indices are developed in order to maintain an acceptable level of fairness among users and maximizing fairness among MGs.

1.1 Dissertation Objectives, Contributions, and Organization

1.1.1 Objectives

This dissertation strives to improve the performance of multicast systems with layered transmissions. More specifically, the goals are: firstly, to improve the capacity of cell-edge users via HD relaying with a layered transmission process; secondly, to ensure efficient cancellation of the IRI introduced at the relays due to the use of alternate relaying; and thirdly, to maximize inter-group fairness while maintaining an acceptable level of intra-group fairness in a multi-group multicast network.

The first objective of improving the throughput of cell-edge users is addressed by using AR relaying with SC coding at the BS, while adopting various strategies for IRI cancellation. The IRI cancellation is first achieved by utilizing SIC decoding schemes that consider different power allocations for the transmitted layers. Specifically, based on the channel gains of the transmitter-receiver link, the transmissions from the BS and the relays are aligned in the time and power domains to optimize the system throughput and cancel the IRI. The second strategy for IRI cancellation in MISO networks uses both virtual MIMO and SIC techniques, where the receiving relay decouples two spatial streams from the BS, and the second relay representing the data of interest uses the virtual MIMO channel created from the BS and the transmitting relay. In the third approach, the operation of the second system is generalized to work in a full MIMO configuration. In this dissertation, due to the system and algorithm complexity, it was not feasible to use only analytical methods to evaluate the performance of the algorithms developed. Therefore, a combination of analytical and simulation results has been used. All proposed schemes and algorithms have been validated and compared via a large number of computer simulations run in the MATLAB computing environment. For this type of research, this is an acceptable research methodology for solving networking problems and developing new concepts and designs.

The second objective of this work is to increase the throughput of a MG in a single-cell cellular system, with a BS located at the center of the cell and equipped with multiple antennas. Within the same MG, this BS serves multiple single-antenna users, which are distributed randomly in the cell according to the Poisson point process. Here it is assumed, based on the adopted scheme, that any idle mobile station (MS) can act as a cooperative relay node. Without the aid of relays, when the channel state information (CSI) is not available at the BS, only one spatial stream can be transmitted simultaneously in the downlink. However, the key feature of cooperative transmission is to encourage multiple users within the same MG to share their antennas cooperatively. In this way, a virtual antenna array can be formed and, as a result, the overall network capacity is improved.

The third objective is to use a PDMA-based transmission system in a multi-group multicast network to maximize the inter-group fairness while maintaining an acceptable level of intra-group fairness. Here, because of the flexibility of PDMA in allocating radio resources, a PDMA-based transmission scheme is used as an efficient tool to allocate the available radio resources fairly among users and MGs.

1.1.2 Contributions

Results of the research described in this thesis have been published in the form of conference papers in [24], [25], [26], and [27]. In addition, one journal article has been submitted and another is in preparation. The details of these publications are outlined below. **Refereed Conference Proceeding Publications**

[C-1] **R. Alsakarnah** and J. Ilow, "*Superposition coding in alternate DF relaying systems with inter-relay interference cancellation*," in Proc. off Wireless and Mobile Computing, Networking and Communications (WiMob), Rome, 2017 IEEE 13th International Conference, pp. 104-109.

[C-2] **R. Alsakarnah** and J. Ilow, "*Superposition Coding in Alternate DF Relaying Systems with Virtual MIMO IRI Cancellation*," in Proc. off 2018 IEEE 29th Annual International Symposium on Personal, Bologna, Indoor and Mobile Radio Communications (PIMRC), 2018.

[C-3] **R. Alsakarnah**, Fadhel Alhumaidi, and J. Ilow, "*Increasing data rates in relay-assisted wireless multicast networks with single antenna receivers,*," in Proc. off 2018 Wireless Days (WD), Dubai, pp. 202-206.

[C-4] **R. Alsakarnah** and J. Ilow, "*Fairness-Aware Pattern-Based Multiple Access for Multi-Group Multicast Systems,*," in Proc. off 2018 11th International Symposium on Communication Systems, Budapest, Networks & Digital Signal Processing (CSNDSP), pp. 1-5.

Journal Papers (in preparation)

[IPJ-1] **R. Alsakarnah** and J. Ilow, "*Inter-relay Interference Cancellation in alternate DF relaying in MIMO Wireless Networks with SIC based detection,*," to be submitted to Computer Communications.

[IPJ-2] **R. Alsakarnah** and J. Ilow, "*Multicast Subgroup Formation using Pattern-Based Multiple Access,*," to be submitted to the IEEE Communications Letters.

The research in each of the papers cited above was initiated and carried out by the principal author of the papers, who is also the author of this dissertation.

The research contributions of this thesis can be classified into three areas, which correspond to the three main chapters of the dissertation. The chapters and the specific papers that relate to them are described below.

Chapter 2: **Superposition Coding in Alternate DF Relaying Systems**

In this chapter, various transmission schemes are proposed to improve the throughput of cell-edge users by using AR relaying with SC coding at the BS, while adopting various strategies for IRI cancellation. Three iterative methods for efficient IRI cancellation in AR systems in a multicast transmission system were developed, based on iterative SIC and zero-forcing (ZF). In the first approach, a SIC technique was implemented that considers different power allocations for the transmitted layers, to deal with IRI by considering it as a different layer (signal of interest). In the second approach, the IRI is addressed by using an additional spatial stream, where the relays and the destination are equipped with two antennas. In the third approach, a generalized method employing multiple antenna configurations at the relays and the destination is introduced. In comparison with existing techniques used in downlink transmissions, the proposed schemes performed better in terms of system capacity.

Chapter 2 is based on the submitted journal paper: [IPJ-1] and two conference publications: [C-1], [C-2].

Chapter 3: **Increasing Data Rates in Relay-Assisted Wireless Multicast Networks with a Single-Antenna Receiver**

Novel transmission scheduling for downlink MU-MISO multicast networks is proposed to increase the data rate of the network. In this chapter, cooperative relaying is deployed to recover independent messages transmitted by the BS at the MSs. Independent messages are recovered by the desired users within a MG by taking advantage of independent signal replicas and channel reuse at different locations, with reduced co-channel interference.

The material in Chapter 3 is based on the conference publication: [C-3]

Chapter 4: **Pattern-based Multiple Access for Multigroup Multicast Systems**

PDMA-based multicast transmission is proposed to allocate the available radio resources fairly in a multigroup environment. With a unique pattern or signature assigned to each MG in the system, the MGs can share the same radio resources. In the work developed, various indices are introduced for both intra- and inter-group fairness in order to maximize inter-group fairness among

MGs. The PDMA-based scheme performs much better in terms of fairness while providing a performance comparable to that of existing techniques in terms of aggregate data rate.

Chapter 4 is based on the conference publication: [C-4]

1.1.3 Organization

This dissertation is organized into five chapters. The first chapter outlines the objectives, contributions and thesis organization, and provides a review of the general concepts and models used throughout the dissertation, while the last chapter presents the conclusions and offers suggestions for future work that could be carried out to expand upon the results.

Chapter 2 discusses layered transmission between the source and the destination, aided by two HD relays to improve the capacity of cell-edge users. In Section 2.1, the first scheme is introduced. Here a single-antenna source continuously broadcasts superposition coded base and enhancement layers, while relays, with one antenna, retransmit the enhancement layer in turn. In Section 2.2, layered transmission between the source and the destination, aided by two HD relays, is developed in a SIMO network, where the BS is equipped with one antenna and both relays and destination are equipped with two antennas. In the proposed system, the single-antenna source continuously broadcasts SC base and enhancement layers, while relays, with one active antenna, retransmit the enhancement layer in turn. In Section 2.3, a general model of the system is established by working with multiple antennas at all nodes in the network. The BS is equipped with multiple antennas to communicate with the destination node, which represents the user with worst channel conditions within a MG. Here two relay nodes equipped with multiple antennas are employed.

Chapter 3 discusses downlink multicast transmission in wireless cellular networks from a cooperative point of view. In Section 3.1, the network topology and signal flow are introduced. This section also presents the relay retransmission strategy used in the system model. In Section 2.2, the principle behind the proposed data recovery algorithms is presented. In Section 3.3, various strategies for relay selection and management are developed. Section 3.4 presents the results in terms of capacity versus the signal-to-noise ratio (SNR).

In Chapter 4, a pattern-based multicast transmission scheme for a multigroup environment is introduced. In Section 4.1, an overview of PDMA and the system model are presented. This section explains the principles of assigning the patterns for each MG, as well as the principles of the transmitting and receiving processes. Fairness considerations are discussed in Section 4.2. Section 4.3 presents and discusses the simulation results, and Section 4.4 provides the conclusions for the chapter.

The following sections of Chapter 1 provide more detailed explanations of the building blocks concerned with this research, and review the corresponding literature in relation to: i) layered transmissions, more specifically NOMA, SC and SIC; ii) wireless network classifications; iii) principles underlying cooperative communications; iv) relaying systems; v) AR and capacity considerations; and vi) wireless channel models.

1.2 Layered Transmission

Layered transmission is a technique to make flexible use of the broadcast spectrum by transmitting several data streams on the same radio resources. Layered Division Multiplexing (LDM) has been proposed by broadcasters to enable mobile TV on top of conventional terrestrial digital TV services [28]. LDM is a non-orthogonal multiplexing (NOM) technique, which has been included in the ATSC 3.0 standard [8]. Splitting the main data stream into sub-streams with different importance is characteristic in multimedia applications such as video and audio transmissions. Splitting the main data stream into sub-stream can be based on different levels of signaling.

This work primarily adopts the layered transmission technique in a multicast environment.

1.2.1 NOMA

NOMA is fundamentally different from orthogonal MA scheme, where it allows multiple users to simultaneously use the entire bandwidth in the same time domain, frequency domain, and code domain. However, the multiple users are separated in power domain via SC technique at the transmitter and SIC techniques at the receivers. In principle, NOMA based techniques show a higher efficiency, especially when the throughput rate among different users is asymmetric [29]. In the following, a brief note about SC and SIC is presented, because these two basic techniques play important roles in understanding NOMA.

1.2.2 Superposition Coding

SC is a form of non-orthogonal multiplexing with hierarchical spectrum reuse in which the transmitter broadcasts to multiple users on the same radio resources. It was first proposed in [30] as a technique to communicate simultaneously to multiple users by a single transmission. In this scheme, several messages are linearly superimposed before being transmitted as one combined signal. SC is essentially a multiplexing scheme of data for a single user. It is similar to NOMA where one layer is viewed as one user [31].

SC has emerged as a promising technology to achieve high capacity in broadcast and multicast networks [32],[33]. Figure 1.1 shows a schematic diagram of SC for two users each using QPSK constellation. From the figure, it can be seen that the BS allocated higher transmitted power for s_1 and lower power for s_2. The superimposed signal of both s_1 and s_2 is given by:

$$x(n) = \sqrt{P_T} \cdot \Big(\alpha_1 s_1(n) + \alpha_2 s_2(n)\Big) \tag{1.1}$$

where $\alpha_{1,2}^2$ are the fraction of power assigned to each user, provided that (i) $0 < \alpha_i^2 \leq 1$ $(i \in \{0, 1\})$ (ii) $\alpha_1^2 + \alpha_2^2 = 1$, and P_T is the transmit power at the BS.

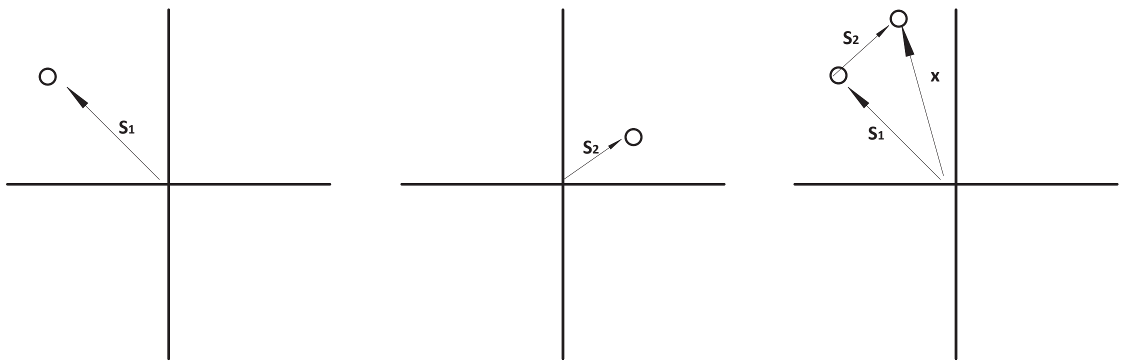

Figure 1.1: Superposition encoding

Figure 1.2 shows an example of the superimposed signal of two QPSK constellation creating a symmetric 16-QAM hierarchical constellation where the base layer has higher protection against the noise than the enhancement layer. In Chapters 2 and 4, SC is used by the BS to superimpose the transmitted layers to the relays and users within a MG.

1.2.3 Successive Interference Cancellation

The basic idea of SIC is that user's signals are decoded successively starting with the one that has the maximum received power and ending with the one that has the minimum allocated power. We assume here that s_1 and s_2 are transmitted to user 1 and user 2, respectively. The specific steps to decode the superimposed message at both user 1 and user 2 can be expressed as follows:

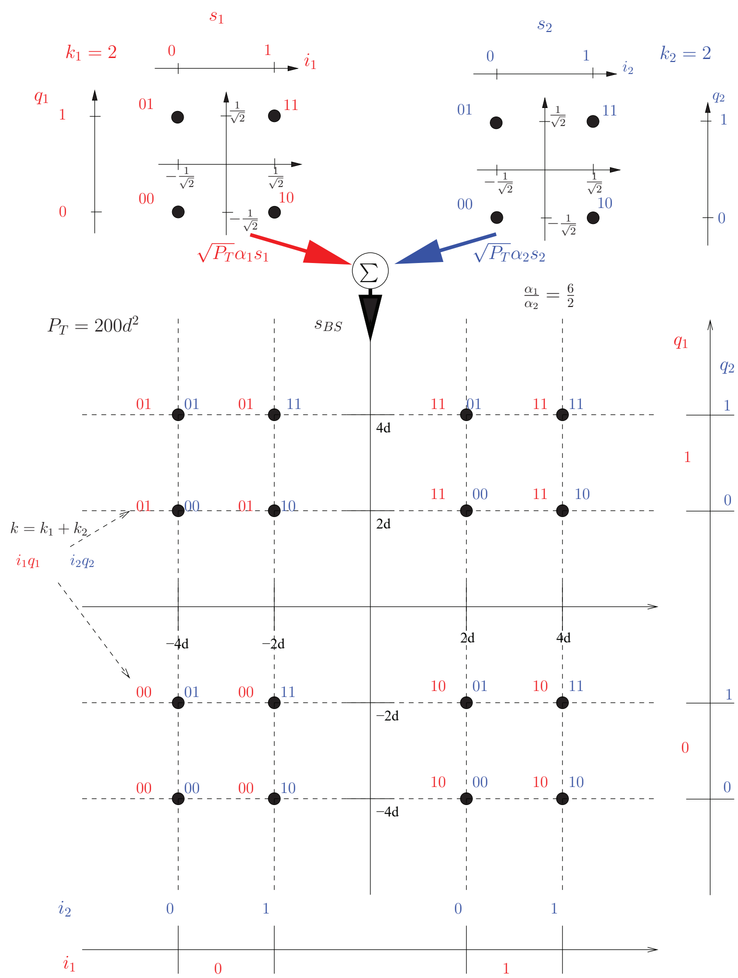

Figure 1.2: Layered coding with a 16-QAM constellation as a superposition of two QPSK constellations.

- User 1 decodes the message s_1 by treating s_2 as interference/noise.
- user 2 performs the following steps to recover the message from the superimposed signal:
  - user 2 first decodes s_1 by using single-user decoder.
  - after successfully recovering s_1, user 2 then subtract its effect from $x(n)$ leading to a new modified $x^{'}(n){=}x(n) - \sqrt{P_T}h_2\alpha_1 s_1(n)$, where h_2 is the

channel gain between the BS and user 2.

– user 2 then decodes s_2 from $x^{'}(n)$ by applying another single-user decoder.

Figure 1.3 shows the receiver used at user 2.

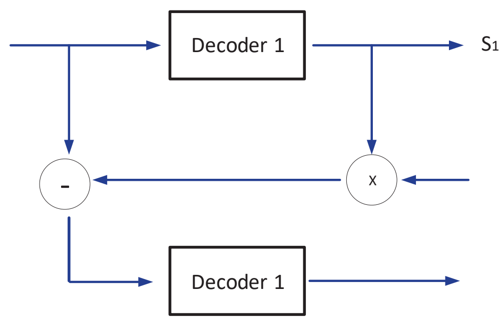

Figure 1.3: Successive interference cancellation at receiver

1.3 Wireless Network Classification

Many types of wireless communication networks exist with different topologies and applications. Each network has a different capacity based on the channel conditions, the number of nodes in the networks, the number of antennas at these nodes, and the availability of CSI at each node. In this section, various types of networks that play an essential role in the development of scheduling and signal processing in this dissertation are described.

1.3.1 One-to-One Networks

One-to-one networks consist of a transmitter and destination. In such a network the available channel is used exclusively by these two transceivers, so there will be no interference from external nodes [34].

In relay networks, a one-to-one system can use one or more intermediate relay nodes to aid in communications between the two nodes. Of particular interest is the $1 \times 2 \times 1$ network, where two intermediate relay nodes are being used to aid in the transmission process. The two relays take turns forwarding the signals from the transmitter to the destination, in a process referred to as alternate relaying [35]. In order for this task to be performed successfully, an interference cancellation scheme should be deployed to mitigate the effect the IRI.

1.3.2 Broadcast Networks

In a broadcast networks, the transmitter sends signals to all possible receivers [36]. The broadcasting nature of the wireless channel makes it naturally suitable for multicasting application, since every transmission by a node can be received by all nodes that lie within its communication range.

1.3.3 Multicast Networks

Multicast is a data delivery scheme in which a single source sends the same information to multiple receivers simultaneously. It is similar to broadcasting but it has an added bonus of receiver discretion, where the data is received by specific interested users or hosts. While multicast can be found on both application and physical layers, the focus of this work will be on multicast in physical layer; specifically, in a wireless channel. Multicasting is emerging as an enabling technology for wireless communications leading to support several group of users with flexible QoS requirements such as data rate. Although multicast is expected to be an enabler to push the limits of next generation communication systems; it is, however, one of the most challenging issues currently being addressed. Several open research issues still exist for multicasting mainly related to channel diversity among users within a MG and the QoS requirements of the users and MGs.

In wireless multicast networks, the BS sends signals to a group of interested users of the same information. Users can choose whether to participate in a multicast group or not. However, if users decide to join the MG, they have then to forward their CSI to the BS. According to the CSI information for all users within the MG, the BS selects the transmission parameters to choose the most appropriate transmission parameters for multicast content delivery, such as the modulation and coding scheme.

One the main challenges for supporting multicast services in wireless systems is related the efficiency of the radio resource management in allocating the available radio resources to each MG. Specifically, the selecting of transmission parameters according to the channel conditions experienced by users within the MG.

Many algorithms have been proposed for radio resource management in multicast systems to maximize the throughput of a MG while maintaining fairness among users/groups. Fairness among users within a MG is referred to as intra-group fairness which comes due the interaction and coexistence of multiple users within a MG, whereas fairness among MGs is usually referred to as inter-group fairness which results from the competitive coexistence in multiple MGs where MGs compete for system resources. The proposed algorithms can be divided into two general categories: single-rate multicast transmission and multiple-rate multicast transmission.

Single Rate Multicast Transmission

Three main algorithms have been widely used for single rate multicast transmissions. The first algorithm, predefined fixed default rate, is the one in which predefined fixed default rate is set [37]. This algorithm adopts a conservative approach, which defines a system data rate bounded by the worst channel conditions in the coverage area regardless whether there is a user experiencing those channel conditions or not. This approach can guarantee reliable transmission to a MG, and is also simple to implement; however, it fails to maximize the spectral efficiency by putting severe restriction on data rate, especially on those users who have good channel conditions.

The second approach is the least channel gain user rate (LCG) or Conventional Multicast Scheme (CMS). In this approach, the data rate is set adaptively based on the worst channel conditions experienced by the user within a MG [38]. LCG can be considered a special case of the pre-defined fixed rate algorithm. However, LCG outperforms the pre-defined fixed algorithm in terms of data rate as it sets the data rate adaptively based on the weakest link in a MG. both predefined fixed default rate and CMS suffer of low user fairness and weak spectral efficiency, respectively.

The third approach is average group throughput (AVG), in which the BS transmits to each MG based on their long-term average throughout [39].

In the first approach, pre-defined fixed default rate, fixed rate is always defined equal to the instantaneous achievable rate of worst channel conditions in the coverage area. Therefore, using pre-defined fixed default rate algorithm results in max-min fairness since the users with bad channel conditions are given resources priority to realize its maximum achievable rate (max fairness), while users experienced good channel conditions are forced to lower their data rate (min fairness). LCG is a special case of pre-defined fixed rate. Hence, it suffers from max-min fairness. In the AVG throughput algorithm, capacity maximization based on SNR averaging provides higher capacity than the LCG scheme. AVG approach has the potential to guarantee reliable multicast to half the users in a MG. Therefore, users with bad channel conditions might experience outages. Although single-rate schemes present advantages in terms of simple implementation and low complexity, they also suffer from poor spectrum efficiency and low fairness levels [40].

Multiple-rate Multicast Transmission

Multi-rate multiple transmission arises to give a solution for the sub-optimality which exists in single-rate multicast transmission considering the heterogeneous channel characteristics of wireless networks.

Several approaches have been developed for multi-rate multicast transmission, which enable users within a MG to receive multimedia traffic based on their channel conditions. The developed approaches can be divided into two main categories: multicast subgroup formation and the information decomposition technique. In the former, group splitting, the MG is divided and classified into smaller sub-groups [41], [42], [43]. In other words, this technique divides a cell into two or more regions; each region will have a different data stream according to the QoS definitions. The second technique uses stream splitting, which involves The Stream Splitting is based on splitting high-data rate multimedia streams into multiple substreams of lower data rate [44]. A base substream, receivable by all users, is transmitted in order to accomplish full coverage of the multicast group. Afterwards, users with good channel condition receive additional enhancements streams to improve information quality. Users' received information quality improves with the number of received sub-streams.

1.4 Relays and Cooperative Networks

Relays arise as an enabler to address many limitations that exist in wireless communications, e.g., by extending the coverage of a network [16],[45],[46] or by improving the quality of the service [47],[48].

Researchers have investigated various aspects and topologies of relay networks. Deploying relays has opened up many avenues for improving wireless systems by exploring relay types, relaying strategies, and the impact of relaying on the performance of the system. In this section, we present some aspects of relaying networks that are related to the work in this dissertation.

1.4.1 Cooperative Networks

Conventionally, a relay is a dedicated wireless network node which is an intermediate node between the communicating terminals. In cooperative networks, inactive nodes or terminals usually act as cooperative relay nodes [49].

In cooperative networks, relaying nodes are usually users that already exist in the network, which means that no additional hardware needs to be introduced to the network [49], [50]. Moreover, there could be more than one node that can act as a relay near the destination, allowing some freedom in the choice of relaying nodes to improve the signal quality [50]. However, since mobile users can act as relay nodes, the position of the relay nodes, therefore, is not fixed due to the user's mobility [51], [52]. This may result in some challenges in estimating CSI for the cooperative nodes at the transmitter, and relying on a certain node to act as a relay for a relatively long period of time.

In the cooperative relaying used in Chapter 2, it is assumed that the global CSI availability is required, and all CSI for all relays and destinations are available at the BS. Moreover, the CSI of the links between the relays and the destination are assumed to be available the BS as well. Learning about the relevant CSI could be accomplished by sending the training sequences and examining the effects at different points in a network along the data paths and the crosstalk paths. This may impose a limitation on the proposed schemes in fast changing channel environments. Since the BS and RSs have more processing power than the MSs, learning the relevant CSI at the BS and the RSs is a realistic task. The design of training vector sequences to learn the CSI at the receivers was not in the scope of this dissertation.

1.4.2 Dedicated Relay Nodes

Another approach to implement relays in networks is to use dedicated nodes that are responsible only for relaying signals from the source to the destination. In this approach, the location of the relays is usually determined based on factors that serve to optimize the performance of the network [53], [54].

Using dedicated nodes in wireless networks requires additional investment in both network hardware and resource element such as an additional TS. With a source located in the center of a cell, a single relay will cover only part of the cell, which is usually areas with bad reception .

The availability of CSI of different links at BS or relay node plays an important role in this dissertation because the power allocation depends mainly on the CSI of BS-R, BS-D, and R-D links. One way for the BS to obtain the CSI for the relays links is by exploiting channel reciprocity. This is defined in IEEE 802.16e (for nonrelay system) as a channel sounding mechanism [55]. The BS measures the up-link channel response from up-link pilot signals transmitted by the destination, and translates the measured up-link channel response to an estimated down-link channel response under the assumption of TDD reciprocity. In [56], a relay-destination link CSI estimation technique is proposed to efficiently obtain CSIs of all links in a relay network with minimized information exchange among nodes. The proposed method exploits the fact that the channel conditions between BS and dedicated relay tends to be more stable, and the frequent CSI update is not necessary for this link.

In the future generation of cellular systems with the shrinking size and BSs connected to the wired backhoul/backbone, the use of the dedicated relays may have a limited application as the coordinated multipoint transmissions. ComP may be better solutions to support edge users, however, in systems like terrestrial DVB or DAB, dedicated relays offer promising solution to extending the network range.

In Chapters 2 and 3, dedicated relays are used to improve the capacity of cell-edge users and MGs in half-duplex relay networks.

1.4.3 Relaying Strategies

Relaying strategies are usually classified based on the way the signals are processed at the relays. The most common relaying strategies are amplify-and-forward (AF) and decode-and-forward (DF), which dominates the relaying schemes [57].

The AF scheme deals with signals at the analog level. Relays receive the signals and simply amplify and retransmit them without decoding [57], [58]. If the CSI is available, the signals may be beamformed so that signals will be steered in a certain direction to improve the performance of the network [16]. This scheme is preferable for systems where the relays are not capable of decoding the signals or where decoding data at the bit level requires complex processing. Moreover, as is the case with any receiver, signals received at the relays are usually accompanied by additive white Gaussian noise (AWGN). Hence, amplification of the signals leads to the amplification of noise [59].

The DF scheme handles signals at the digital level. Relays decode the received signals by removing all of the effects from the receiving side, remodulate and retransmit the received signal to the next nodes [57]–[59]. Although the noise removal is an advantage of the DF scheme, this requires full data decoding.

Compared to the AF scheme, the complexity of the DF scheme is significantly higher due to its full processing capability. The DF scheme also requires a sophisticated media access control layer, which is unnecessary in the AF protocol.

This work adopts the AF scheme in Chapter 3, and DF scheme in Chapter 2 in downlink multicast cellular system.

1.5 Successive Relaying

To overcome the spectral efficiency loss in half-duplex wireless relay networks, researches have taken two paths. One path is based on utilizing two-way communications in which a transmitter acts also as a receiver and the data flow, therefore, is in two directions. However, network coding technique is required to reduce the transmission time [16], [45]. This approach is not suitable for networks with broadcast or multicast transmissions where data are to be transmitted in a single direction.

The second approach addresses one-way transmission. In an effort to improve diversity gain, channel capacity, and reduce relaying time, a transmitter sends signals to all users, then the relays which overhear these signals retransmit the desired signals at the same time [60]. This approach improves the prelog factor for the channel capacity from $\frac{1}{2}$ to $\frac{K}{K+1}$, where K is the number of users in the network.

To alleviate the complexity of full-duplex (FD) relays, the work described in [60] paved the way to mimic FD relaying via two HD relays which uses successive relaying techniques that can also overcome prelog factor limitations. Successive relaying, or alternate relaying, as shown in figure 1.4, allows the transmitter to transmit continuously, while the two HD relays take turns listening to the transmitter. Using the same radio resources as utilized by the transmitter, a relay which has completed a listening session retransmits their received signals to the destination successively. In other words, in each time slot, one of the relays receives new data from the transmitter while the other relay forwards the previous data to the destination. Then, the role is swapped in the next time slot. Listening and transmitting are performed by the relays successively and continuously.

The specific steps for each time are described as follows:

– In the 1^{st} time slot, R_1 listens to $s(1)$ from the BS, R_2 remains silent, and the destination listens to $s(1)$ from the BS.

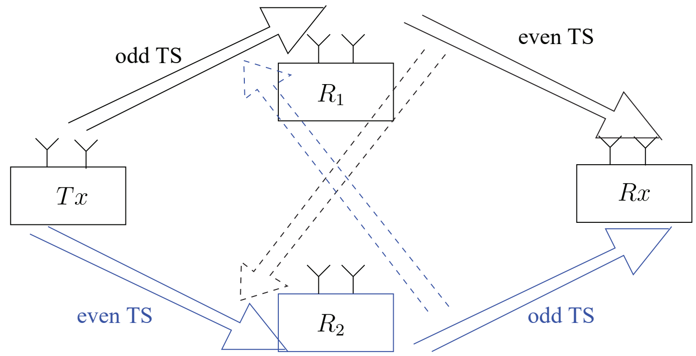

Figure 1.4: Two-path alternate relaying in a $1 \times 2 \times 1$ network

- In the 2^{nd} time slot, R_2 listens to $s(2)$ from the BS while being interfered by $s(1)$ from R_1, and destination receives simultaneously both $s(2)$ from the BS and $s(1)$ from the R_1.

- In the 3^{rd} time slot, the BS transmit $s(3)$, R_1 receives simultaneously both $s(3)$ from the BS and $s(2)$ from R_2, and the destination receives both $s(3)$ from the BS and $s(2)$ from R_2.

- In the t^{th} time slot (even TS), the BS transmits $s(t)$, R_2 receives both $s(t)$ from the BS and $s(t-1)$ from R_1, and the destination receives both $s(t)$ from the BS and $s(t-1)$ from R_1.

- In the $t+1^{th}$ time slot (odd TS), the BS transmits $s(t+1)$, R_1 simultaneously receives both $s(t+1)$ from the BS and $s(t)$ from R_2.

- Finally, in the last time slot $(T+1^{th})$, R_2 forwards $s(T)$ to the destination so that the destination only receives from R_1.

It can be seen that T symbols are transmitted from the BS to the destination in $T+1$ time slots. Therefore, the prelog factor will be improved to $\frac{T}{T+1}$.

The improvement in the prelog factor, however, comes at a cost. During the listening session and except for the first and last time slots, the transmitted signal from one relay can disturb the received signal by the other relay simultaneously as shown in figure 1.4. Therefore, IRI may degrade the whole performance of the system if not managed carefully.

IRI can be managed well in successive relaying either through (i) precoding by the transmitter, (ii) signal processing at the relays themselves, (iii) using designated DoF for IRI treatment, (iiii) post-processing at the destination, or any combination of (i), (ii), (iii) and (iiii).

In successive relaying, IRI cancellation is a major challenge. For AF relaying, IRI cancelling may be more challenging because interference and noise between relay nodes are inevitably amplified and forwarded to the destination, resulting in a reduced signal-to-interference-plus-noise ratio (SINR). One-to-one AF relaying networks treat interference at the bit level by canceling interference successively at the receiver [61]. This method requires extensive memory at the receiver, because it depends on previously received signals and canceling of the signals one by one in a decision feedback type of processing, and it removes only deterministic interference signals, leaving the desired signals with accumulated noise. Another self-interference cancellation method designed to be performed at one of the two relay nodes was proposed in [62] which treated the AF two-path relaying as an equivalent MIMO space time code.

In MIMO networks, in one-to-one alternate relay networks, IRI cancellation can be performed at the relays, using a zero-forcing (ZF) approach [63]. However, this method requires a large number of antennas at the relays, which may not be practical. AR can be performed by using interference alignment [64]. This approach is suitable only for networks where each node is equipped with an even number of antennas as this method recovers only part of the lost capacity. In a hybrid model demodulation-and-forward approach for SC relaying is based on the channel conditions switching between direct transmission and differential modulation.

In DF relaying, the IRI is decoded and subtracted from the received signal by simply treating it as Gaussian noise at the relays based on the interference level [65]. In [66], the IRI cancellation is managed by applying dirty-paper-coding technique at the source transmitter in a digital relaying scheme. In [67] , the prior knowledge of IRI at the relays was exploited to cancel the network interference. In [68], beam-forming based techniques were introduced to cancel the IRI in two-path successive relaying.

1.5.1 Capacity Challenges

Although relays have the ability to improve the performance of the system in terms of reliability and range extension, there are also challenges such as the capacity decrease. Communicating via conventional HD relays requires approximately twice as many radio resources compared to direct communication between the source and the destination. In conventional relay networks, a time slot is utilized by the source to transmit signals to the relays, and another time slot is used by the relays to retransmit the signals to the destination. This drop in system capacity between the source and the destination is referred to in the literature as the prelog factor [61].

The capacity is the maximum data rate that wireless channel can support. The capacity of a direct single-input single-output Rayleigh channel C can be expressed as:

$$C = \mathbf{E}\left(\log_2(1+\frac{|h|^2 P}{\sigma^2})\right) \qquad \text{[bps/Hz]} \tag{1.2}$$

where h is the channel between the source and the destination, P is the signal transmit power, σ^2 is the variance of the AWGN, and $\mathbf{E}(\cdot)$ is the expectation operator [18]. When a HD relay is used to aid in transmitting the signal to the destination, the capacity is scaled by $\frac{1}{2}$ due to the additional radio resources used in the process. This is known as the prelog factor. Note that the exact expression of the average capacity in relays depends on the relaying strategies used and the type of network.

In order to overcome deficiency in capacity, several solutions have been proposed for specific types of network; however, this remains an open problem for other networks. In a two-way communications networks, the prelog factor is improved to $\frac{2}{3}$ in DF relaying schemes. This is accomplished by adopting network coding so that the relay reduces its transmissions for two independent signals from two time periods to one time period [16]. The relay then waits for two terminals to send their signals consecutively. It then encodes the signals so that each terminal can remove its own data and can decode the desired data.

Another approach to eliminate the prelog factor is to allow the two terminals to transmit to the relay at the same time.Then the relay retransmits the received signal to the two terminals. The terminals know their signals and have the channel information, hence they can decode the desired signals [45].

These solutions do not work for broadcast and multicast networks, when the signals is transmitted in one direction. In this dissertation, this challenge is addressed in DF SISO and MIMO networks by using AR as discussed in Chapter 2.

In MIMO networks, ZF technique was adopted to decode the transmitted data. ZF detection is based on completely eliminates interference from other symbol layers when detecting a particular symbol layer, which results in suboptimal performance due to noise enhancement. However, in the developed work, we study the upper bound on the capacity of this network and there was no practical implementation of the setup to study the noise enhancement effect.

Chapter 3 goes a step further by not only removing the effects of the prelog factor but also improving the capacity of single antenna users within a MG beyond their one-to-one limitations.

1.5.2 Pattern Division Multiple Access

Pattern division multiple access (PDMA) is emerging as a promising candidate for non-orthogonal multiple access scheme for the the fifth generation (5G). The PDMA scheme is based on mapping of the transmitted data to radio resources that can consist of frequency, time, and spatial resources. Mapping the transmitted data is based on a pattern for each user. Therefore, the pattern is introduced to differentiate signals of users sharing the same RE. Figure 1.5 shows an example of a PDMA pattern matrix in a system with 3 users and 4 REs. It can be seen that each user has (i) a designated unique code and (ii) RE pattern. The PDMA patterns defines the rule of mapping data to REs. From figure 1.5, its can be noticed that $RE1$ and $RE2$ is shared by all three users, while $RE3$ and $RE4$ are shared by only two users. The flexibility of PDMA in the resource allocation makes it suitable for multicast transmission as it relaxes the constraint of orthogonal radio resource allocation; the group scheduling is no longer a binary selection, but the optimization of joint power, pattern, code signature, and receiver design.

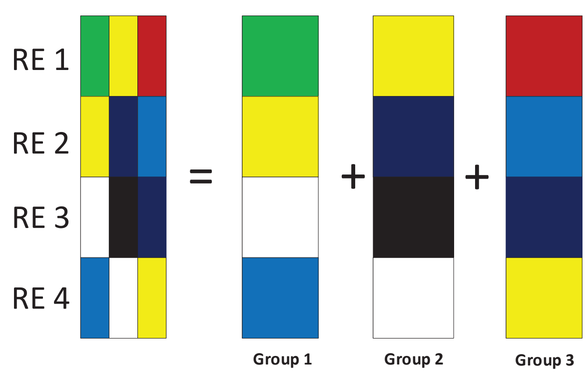

Figure 1.5: An example of a PDMA pattern for M=3 groups using L=4 REs.

PDMA is based on SIC amenable multiple access (SAMA) [69], which utilizes the SIC amenable design at the BS and SIC detection based at the receiver. The authors in [70] give an elaboration of PDMA including the concept, mechanism, key technologies, and evaluation. A cooperative PDMA in uplink network is proposed in [71] in which PDMA is designed with HD decode-and-forward relays.

In Chapter 4, PDMA-based scheme is adopted for multi-group multicast transmission to maximize the inter-group fairness while maintaining an acceptable level of intra-group fairness by exploiting the flexibility offered by the PDMA scheme.

1.6 Modeling Wireless Communication Channels

This dissertation contributes to the theoretical development of the field of signal processing algorithms for wireless multicast network. The results obtained are verified by using simulation models. This is an initial stage of design, which precedes practical implementation. In this dissertation, we first tested the simulations by using known systems so that the same results were obtained. Then, we moved to implement the proposed work and compared it with well-established works in the same field to compare the performance. This methodology is widely adopted and accepted in the field of communication system design. Sound modeling of wireless channels plays an important role in studying and analyzing wireless communication systems. Channel models play an essentials role in the conceptual stage of communication system design, where they are used as a way to model the various mediums that approximate the real world environment in order to predict and evaluate performance of wireless communications systems under realistic conditions. This section reviews some of the wireless communication channel models used in this dissertation.

1.6.1 Additive White Gaussian Noise

AWGN is a basic and generally accepted model for thermal noise in wireless communication channel. This noise is universally present in the RF end of wireless receivers resulting from a large number of small random variable (RVs). According to central limit theorem, the summation of large number of random variables (RVs) leads to a Gaussian distribution, which has a probability density function (pdf):

$$pdf(n) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp^{\frac{-(n-\mu)^2}{2\sigma_0^2}} \tag{1.3}$$

with a zero mean ($\mu = 0$) and a noise variance (σ_0^2) that represents the power spectral density ($\frac{N_0}{2}$ [W/Hz]).

This dissertation deals with received signals after down-converting and matched filtering, and noise is represented as a RV rather than as a stochastic process.

1.6.2 Rayleigh Fading Channels

When a signal is transmitted in a radio channel, it arrives at the receiver via a large number of paths due to refractions, reflections, and scattering as they pass various types of objects. Thus, the receiver combines multiple copies of the original signal at the output of the matched filtering, but every copy has its own attenuation and time delay. It is assumed that delays are smaller than the symbol duration, so that there will be no need to deal with time dispersive channels. This phenomenon is recognized as fading, where the copies of the original signal from all multipaths when combined are represented as the multiplicative effect of the received signal.

The most common situation in wireless communication is when there is no line of sight between the transmitter and the receiver. And the most common type of fading is slow fading, where the multiplicative factor is given by random variable denoted as h. Here h is a complex variable with an independent, and identically distributed (i.i.d.) Gaussian RV. Therefore, the gain (envelope) of the fading channel is: $|h| = \sqrt{\Re(h)^2 + \Im(h)^2}$, where $\Re(h)$ and $\Im(h)$ denote the real and imaginary values of h, respectively, and h is a complex normal RV: $h \sim \mathcal{CN}(0, \sigma_h^2)$. Therefore $|h|$ follows a Rayleigh distribution. The probability density function of the above mentioned gain

($|h|$) is given by :

$$pdf(|h|) = \frac{2|h|}{2\sigma_h^2} \exp^{\frac{|h|^2}{2\sigma_h^2}} \quad (1.4)$$

where σ_h^2 is a power scaling parameter. This applies only when there is no line-of-sight (LOS) path; equation 1.4 represents the most adverse type of fading.

1.6.3 Large-scale Attenuation

As is the case with any transmission medium, the signal power level decays with distance as the signals propagate through a channel. This phenomenon is known as signal attenuation or deterministic path loss. Various mathematical models are proposed for different radio environments in order to capture this phenomenon. This work adopts a generalized formula which links the power decay to the distance traveled $d > 1$ as follows:

$$p_r(d) = \frac{p_t}{d^\alpha} \quad (1.5)$$

where $p_r(d)$ and p_t represent the power of the received and transmitted signals, respectively. The α parameter corresponds to the propagation condition; this value normally ranges from 2 in free-space conditions to 6 in a dense urban area.

In this dissertation, the main goal of using multicast system is for the distribution of multimedia, DVB, and terrestrial television services.

1.7 Summary

This chapter has presented an overview of the research work to be presented in this dissertation. It additionally outlined the objectives, contributions, and organization of the dissertation. A review of the general concepts and elements used throughout the dissertation was then provided. This included a review of both wireless multicast systems and an overview of non orthogonal multiple access schemes.

Chapter 2

Superposition Coding in Alternate DF Relaying Systems

In the next generation of wireless systems, it is crucial to have an efficient multicast system that is capable of delivering the same service to all users within the same MG over the same radio resources. However, as a result of the wide variations in wireless channel conditions among users, and the high user mobility, delivering same quality of services to all users poses a serious challenge in the design of multicast transmission systems.

Cell-edge users are usually expected to receive the lowest data rate among the MG users because of their bad channel conditions due to their distance from the BS, leading to a fixed rate which is always equal to the instantaneous achievable rates of the minimum users at the cell edge. In the case of cell-edge users, some intermediate nodes (relays) might be the solution to extend the coverage and improve the link reliability. Relays have been used to meet the requirement of higher capacity, improve diversity and link reliability. Relays can receive and forward data from the BS to destination, which consumes more time than a direct transmission process when all relays operate in half duplex fashion. To overcome the loss of spectral efficiency in conventional relaying, alternate relaying (AR) has been used to mimic full duplex relays by reusing source transmission time slots for relay transmission as well [35]. Because of the simultaneous transmissions from both the relay and BS, the receiving relay can hear both transmitted signals from the BS and other relay leading to what is called IRI. IRI signals may limit the system performance if not mitigated efficiently.

This chapter analyzes alternate-relaying downlink transmission to improve the cell-edge users data rate and link reliability while dealing with IRI cancellation in several approaches. The IRI cancellation approaches developed in this chapter are based on layered transmission. This involves separating the signals transmitted by the BS and relays in either power or spatial domain while still using the interference signal caused by the simultaneous transmission from the BS and relays at the cell-edge users in such a way to improve data rate. Superposition coding (SC) is used in this chapter to superimpose several coded sequences (layers) to be transmitted as one signal to the MG users.

The availability of CSI of different links at BS or relay node plays an essential role in this dissertation because the power allocation depends mainly on the CSI of BS-R, BS-D, and R-D links. The IRI cancellation schemes developed in the chapter relies on the knowledge of channels gain matrices related to the inter-relay channels and channels between the BS, relays, and destination. learning about the relevant CSI could be accomplished by sending the training sequences and examining the effects at different points in a network along the data paths and the crosstalk paths. This may impose a limitation on the proposed schemes in fast changing channel environments. Since the BS and nodes have more processing power than the destination, learning the relevant CSI at the BS and each node is a realistic task. The design of training vector sequences to learn the CSI at each node is not in the scope of this dissertation.

2.1 Superposition Coding in Alternate DF Relaying Systems in SISO Networks

In this section, we introduce our first scheme, which is a combination of layered transmission and AR system to efficiently cancel the IRI using SIC principles based on power profiles of different transmitted layers, and improve the capacity of cell-edge users. We first start by presenting the system model in the next subsection. SIC detection and IRI cancellation principles are introduced in subsection 2.1.2. Information theoretic and reference schemes are considered in subsections 2.1.3 and 2.1.4, respectively. In subsection 2.1.5, the numerical results are discussed. The summary is presented at the end of this section.

2.1.1 System Model

We consider a downlink two-hop relay system in alternate two-path relaying network consisting of BS, two half-duplex relay nodes(R1;R2), and one destination as illustrated in Figure 2.1 where the BS continuously transmits information to the destination node with the aid of the relays. The BS, relays, and destination are all equipped with a single antenna.

All wireless links exhibit independent and identically distributed (i.i.d.) block Rayleigh fading and are affected by AWGN. This means that the fading coefficients (channel gains) remain constant during one TS, but change independently from one slot to another according to a complex Gaussian distribution with zero mean and variance σ_h^2; the variance captures path-loss. The channel gains between BS and relays are given by h_1 and between BS and D is h_2. The coefficient h_3 represents the channel gain between the relays. To simplify the presentation we assume the reciprocity of channels between R_1 and R_2. The channel gain between the relays and

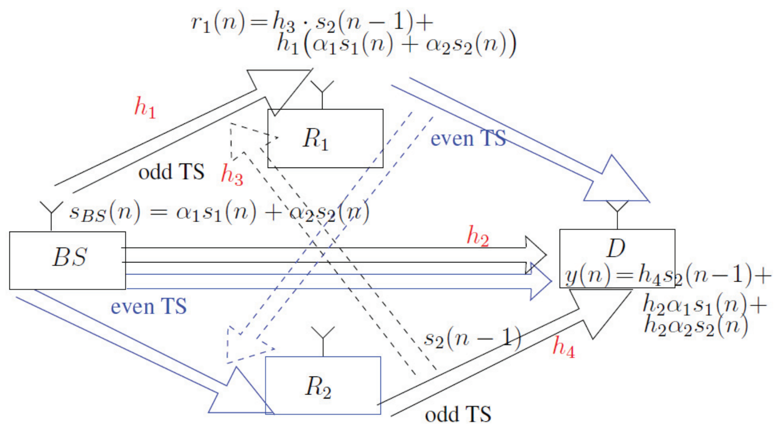

Figure 2.1: Two-path alternate relaying on a downlink with received asymmetric constellations.

D is h_4. All the channel gains are shown in figure 2.1 using red notation. The AWGN power levels at both relays and D are initially normalized to one.

Unlike conventional relaying, in alternate relaying the BS transmits continuously, while relays R1 and R2 takes turns in receiving from the BS and transmitting to the destination using the same spectrum that is utilized by the BS. We assume that relay nodes work in the practical half-duplex mode. Therefore, (i) the communication between the destination and BS is completed in two cycles, and (ii) the reception by R1, R2, and the destination is affected by signals from both the relay and BS which results in IRI. In our work we assume that the reception times at the relays and destinations for different signals from the BS and relays are the same time which implies the node/network synchronization. However, sometimes it is not sufficient to be synchronized when sending the data because the requirement is for synchronous reception and the latter is affected by the variability in the propagation delays of the signals. In that case, users are required to coordinate with each through network protocols and other methods with regard to their locations, which are not studied in this thesis. The timing problems can be solved with the deployment of OFDM where the symbol duration is extended in each subcarrier so that the delay over different nodes paths can be neglected. In this thesis, we assume single carrier transmissions, however the model adopted and schemes developed are applicable in OFDM systems if one considers multiple single subcarrier modulations with different channel gains along different frequencies.

The implementation of the SC by BS is done by splitting a source message (k bits) into two layers, the base and enhancement layers, with $k1$ and $k2$ bits in each layer where $k1 + k2 = k$, i.e., the quadrature amplitude modulation (QAM) symbols from M_1-QAM and M_2-QAM constellations where $M_1 = 2^{k_1}$ and $M_2 = 2^{k_2}$.

In the proposed scheme, SC is deployed by the BS to transmit the information to the destination. In SC, the source message is separated into two layers: base and enhancement layers. A simple representation of the superimposed signal is described by the following equation:

$$s_{BS}(n) = \sqrt{P_T} \cdot \Big(\alpha_1 s_1(n) + \alpha_2 s_2(n)\Big) \tag{2.1}$$

where $s_1(n)$ is base layer, $s_2(n)$ is the enhancement layer, these two layers are represented by two modulated signals s_1 and s_2 each with the unit energy, i.e., $\mathbf{E}\{|s_1|^2\} = \mathbf{E}\{|s_2|^2\} = 1$, where $\mathbf{E}(\cdot)$ is the expectation operator, and $\alpha_{1,2}^2$ are the fraction of power assigned to each layer, provided that (i) $0 < \alpha_i^2 \leq 1$ ($i \in \{0, 1\}$) (ii) $\alpha_1^2 + \alpha_2^2 = 1$, and P_T is the total transmit power at the BS. We assume further the following powers $P_1 = \alpha_1^2 \cdot P_T$ and $P_2 = \alpha_2^2 \cdot P_T$ are allocated to the base and the enhancement layers where $P_1 + P_2 = P_T$.

From the relay point of view, the main characteristic of the SC technique is the implementation of SIC for accessing the overlaid layers. At the BS side, the layers have been previously arranged according to the assigned fraction of total power as in Equation 2.1. Therefore, at the relays, the layers are processed following the same order. First, the strongest signal (s_1) is decoded, subtracted completely from the received signal, then s_2 is decoded with less interference.

The achievable data rate of s_1 at the relays denoted by R_1, is $R_1 \leq C = \log_2(1 + \frac{|h_1|^2 P_1}{|h_1|^2 P_2 + 1})$, while the achievable rate of s_2 after canceling the interference caused by s_1 denoted by R_2, is $R_2 \leq C = \log_2(1 + |h_1|^2 P_2)$, i.e., the relays decode both layers where decoding of the first layer is limited by the interference caused by the power allocated to the second layer. The total achievable data rate at the relays denoted by R_R, is $R_R = R_1 + R_2 \leq C = \log_2(1 + \frac{|h_1|^2 P_1}{|h_1|^2 P_2 + 1}) + \log_2(1 + |h_1|^2 P_2)$, which can written as $R_R \leq C = \log_2(1 + |h_1|^2 P_T)$.

The achievable data rate at node D from the direct transmission from the BS , denoted by R_2, is $R_2 \leq C = \log_2(1 + \frac{|h_1|^2 P_1}{|h_1|^2 P_2 + 1})$. The information from s_2 cannot be recovered by node D because of the channel conditions between the BS and node D, and the limited power at the BS.

In the proposed scheme, node D can recover both base and enhancement layers with the aid of relays, hence, improving the achievable data rate by canceling the effect of s_2 on the direct link transmission. Because of the channel conditions between the BS and relays, the relays are able to decode both s_1 and s_2, while node D can only decode s_1 in presence of the signal of s_2 which is considered noise. Therefore, there is a need for the relay to forward the partial source's information (the enhancement layer) to node D, allowing the recovery of the base layer with a new data rate without the effect of interference from the enhancement layer. The rationale behind adopting the alternate relaying is to overcome the loss of spectral efficiency when using conventional relaying scheme.

In the proposed scheme, R_1 and R_2 only send the delayed enhancement layers to the destination. The operation of SC alternate relaying scheme is summarized in Table. 2.1.

Table 2.1: Summary of transmission and decoding phases in the SC alternate relaying scheme.

	Transmission Phase
BS	In every TS n, BS sends $\sqrt{P_T} \cdot \Big(\alpha_1 s_1(n) + \alpha_2 s_2(n)\Big)$ with (i) P_T adjusted to meet the targeted SNR for decoding R_1 at D (without the noise of s_2) (ii) α_1 and α_2 adjusted to decode s_1 and s_2 at R_1/R_2.
$R_1/\ R_2$	R_1 transmits in even and R_2 in odd TSs In the designated TS n, using the maximum power P_T, the relay sends $s_2(n-1)$ decoded in the previous TS.
	Receiving Phase
$R_1/\ R_2$	In their respective TSs, e.g., n, using SIC, $R_1/\ R_2$ decode in order $s_2(n-1)$, $s_1(n)$ and $s_2(n)$
D	In every TS n, using SIC, D decodes $s_2(n-1)$ and use it to cancel the effect of stored $s_2(n-1)$ then decode $s_1(n-1)$, and store $s_1(n)+s_2(n)$ for the next time slot

2.1.2 SIC Detection with IRI Cancellation

In this section, we investigate the process of IRI cancellation and decoding the overlaid layers at both the relays and destination. From figure 2.1, transmissions/receptions in odd TSs are represented with black arrows, and the transmissions/receptions in even TSs are represented with blue arrows with the conventional interpretation for the beginning and the tip of the arrow representing transmission and reception. In this section, the focus is on the signal flow in the odd (black) TSs due to the system

symmetry, the signal analysis is the same in even (blue) TSs.
It can be seen from Figure 2.1 that the signal received at R_1 is composite of signals transmitted by the BS and R_2 which can be expressed as:

$$r_1(n)=h_1\cdot(\sqrt{P_T}\cdot\big(\alpha_1 s_1(n)+\alpha_2 s_2(n)\big)+h_3\sqrt{P_T}s_2(n-1)+n_{R_1}(n) \tag{2.2}$$

where n_{R_1} represents the AWGN introduced at R_1.

From the perspective of R_1, to transmit in the next (n+1) TS (even) the enhancement layer given by $s_2(n)$, R_1 has to decode $s_2(n-1)$, $s_1(n)$, and $s_2(n)$. From Equation 2.2, It can be noticed that R_1 has to be able to decode $s_2(n-1)$ and $s_1(n)$ first and suppress their effect from $r_1(n)$ to have the ability to successfully decode $s_2(n)$ by using SIC, i.e., the current base and enhancement layers and the delayed enhancement layer data. In the first stage of SIC, the relay decodes $s_2(n-1)$ which is usually received with the highest power because of the close proximity of relays ($s_2(n-1)$ is transmitted by R_2). In the second stage of SIC, the relay decodes $s_1(n)$ from $r_{l1}(n)$ after removing the effect of $s_2(n-1)$ as followings:

$$r_{l1}(n)=h_1\cdot(\sqrt{P_T}\cdot\big(\alpha_1 s_1(n)+\alpha_2 s_2(n)\big)+n_{R_1}(n) \tag{2.3}$$

After successfully decoding $s_1(n)$, the relays finally decode $s_2(n)$ from $R_{l2}(n)$ as follows:

$$r_{l2}(n)=h_1\cdot\sqrt{P_T}\cdot\alpha_1 s_1(n)+n_{R_1}(n) \tag{2.4}$$

The term $h_3\sqrt{P_T}s_2(n-1)$ in equation 2.2 represents the IRI. Therefore, the IRI cancellation process is based on considering the IRI signal as a signal of interest which can be effectively removed and canceled by using SIC strategy at the relays.

The signal received at the destination through over-the-air summation can be expressed as:

$$y(n)=h_2\cdot\big(\sqrt{P_T}\cdot\big(\alpha_1 s_1(n)+\alpha_2 s_2(n)\big)+h_4\sqrt{P_T}s_2(n-1)+n_D(n) \tag{2.5}$$

where $n_D(n)$ represents the AWGN at the destination, h_4 is the channel gain between the relay(s) and the destination, and h_2 is the channel gain between the BS and the destination. The signal $s_2(n-1)$ is received with highest power because of good channel conditions between the relay(s) and the destination, hence, it is recovered first. Table 2.2 shows the superimposed signal layers at the destination where the row represents time (TSs), and column represents the power of the received layer.

Table 2.2: Superimposed signals at the destination.

TS n	High power $s_2(n-1)$	Mid Power $s_1(n)$	Low power $s_2(n)$	Low power $n_D(n)$
1	$s_2(0)$	$s_1(1)$	$s_2(1)$	$n_D(1)$
2	$s_2(1)$	$s_1(2)$	$s_2(2)$	$n_D(2)$
3	$s_2(2)$	$s_1(3)$	$s_2(3)$	$n_D(3)$

At the destination, the first SIC process is to recover $s_2(n-1)$ and suppress its impact from $y(n)$ to become:

$$y_1(n) = h_2 \cdot \big(\sqrt{P_T} \cdot \big(\alpha_1 s_1(n) + \alpha_2 s_2(n)\big) + n_D(n) \tag{2.6}$$

Because of i) the bad channel conditions between the BS and relay(s), ii) the interference from $s_2(n)$, and iii) the power allocation, $s_1(n)$ cannot be decoded. However, under closer examination, by delaying decoding $s_1(n)$ by one TS to $n+1$, we can use $s_2((n+1)-1) = s_2(n)$ (decoded by SIC) to remove the impact of $s_2(n)$ from $y_1(n)$ so that $s_1(n)$ can be decoded with the presence of the noise only. The decoding of $s_1(n)$ would be affected only by the AWGN $n_D(n)$ and would be delayed by one TS. After all, decoding of s_2 is also delayed by one TS.

2.1.3 Information-theoretic Considerations

As presented so far, the decoding of the base and enhancement layers at node D is limited by the power allocation at the BS between $s_1(n)$ and $s_2(n)$. By working with power allocation parameters, α_1 and α_2, for the base and enhancement layers, we aim to make sure that the enhancement layer signal ($s_2(n)$) does not cause excessively high interference when decoding $s_1(n)$ at the relays, while at the same time , we aim to enable the relays to transmit $s_2(n)$ in such away that it can be decoded by node D with the interference from the BS.

For example, in the simulation section, $\alpha_1^2 = 0.95$ and $\alpha_2^2 = 0.05$, resulting in a $12.8\,dB$ SINR when decoding $s_1(n)$ because of the interference caused by $s_2(n)$. Because of the channel diversity in the received signals, and by considering the delayed decoding, there are alternative ways to recover $s_1(n)$,both at D and relays, by removing the effects of $s_2(n)$ as the interference.

The goal of this work is to select the coefficients α_1^2 and α_2^2 to improve the data rate at the destination while successfully canceling the IRI. In what follows, we provide detailed discussions on achieved data rates at both relays and destination.

Let $P_1 = \alpha_1^2 \cdot P_T$, $P_2 = \alpha_1^2 \cdot P_T$, and the instantaneous channel conditions given by the channel gains h_1, h_2, h_3 and h_4, the achievable data rate by the relays can be expressed as follows:

$$C_{Relay} = \log_2\left(1+\frac{|h_3|^2 P_T}{|h_1|^2 P_2 + |h_1|^2 P_1 + N_0}\right) + \log_2\left(1+\frac{|h_1|^2 P2}{|h_1|^2 P_1 + N_0}\right) + \log_2\left(1+\frac{|h_1|^2 P1}{N_0}\right) \quad (2.7)$$

The first term of this equation $\log_2\left(1+\frac{|h_3|^2 P_T}{|h_1|^2 P_2 + |h_1|^2 P_1 + N_0}\right)$ represents the data rate from R_1 to R_2, and $\log_2\left(1+\frac{|h_1|^2 P1}{N_0}\right)$ represents the data rate from the BS to the relays accounting for $s_2(n)$. However, it is worth mentioning that h_3 captures the channel conditions between the relays, which is usually higher than h_1 because of the short distance between the relays and the used deterministic path loss. The enhancement layer data rate is chosen so that relays are able to decode it (as IRI) in the presence of the signal transmitted by the BS as noise according to:

$$C_{D,s_2}=\min\Big(\log_2(1+\frac{|h_1|^2P_1}{N_0}),\log_2(1+\frac{|h_3|^2P_T}{|h_1|^2P_2+|h_1|^2P_1+N_0})\Big) \tag{2.8}$$

Since the transmitted signal from the second relay is not intended for the first relay, even though the relay decodes the three level sc signal, the effective achievable instantaneous data rate at the relay is:

$$\log_2\left(1+\frac{|h_1|^2P_2}{|h_1|^2P_1+N_0}\right)+\log_2\left(1+\frac{|h_1|^2P_1}{N_0}\right) \tag{2.9}$$

which can be also written as:

$$\log_2\left(1+\frac{|h_1|^2P_1+|h_1|^2P2}{N_0}\right) \tag{2.10}$$

The achievable data at the destination is given by:

$$C_D=\log_2\left(1+\frac{|h_4|^2P_T}{|h_2|^2P_2+|h_2|^2P_1+N_0}\right)+\log_2\left(1+\frac{|h_2|^2P_1}{N_0}\right) \tag{2.11}$$

The two terms in the equation represent the capacities of the corresponding links of (C_D^R and C_D^{BS}).

C_D can be simplified:

$$C_D=\log_2\left(1+\frac{|h_4|^2P_T+|h_2|^2P_1}{|h_2|^2P_2+N_0}\right) \tag{2.12}$$

From equation 2.7 and equation 2.10, it can be seen that both the destination and the relays have the same data rate. It is worth mentioning that the C_D^{BS} is limited by SINR at the relays from the BS that is allocated for s_1.

2.1.4 Reference Schemes

Here we consider three alternative schemes to the SC-alternate relaying scheme.

The first scheme is the direct transmission between the BS and destination ,without any intermediate relaying, at a data rate $C_{dir} = C_{BS-D}$, where C_{BS-D} is the achievable data rate for the BS-D link.

The second scheme is the traditional relaying scheme in which the relay alternates between reception and transmission in successive time slots. The capacity of this scheme is limited by the worst channel conditions between $BS - R_i$, and $R_i - D$ links. The data transmission in this scheme requires two time slots which leads to a two factor loss in capacity. Therefore, the achieved data rate by traditional relaying is easily found to be:

$$C_{tr} = \frac{1}{2}\min(C_{BS-R_i}, C_{R_i-D}) \tag{2.13}$$

The relay selection of this scheme is based on selecting the relay which has better channel conditions with BS and D.

The third scheme is the adaptive relaying in which the relays are equipped with buffers to adapt its transmission and reception based on the quality of the involved links. The capacity of the adaptive relaying scheme can be expressed as:

$$C_{ad} = \frac{C_{BS-R_i} C_{R_i-D}}{C_{BS-R_i} + C_{R_i-D}} \tag{2.14}$$

In this scheme, the BS decides to select the relay that maximizes the capacity based on the quality of both $BS - R$, and $R - D$ links.

In the proposed scheme, unlike all alternative schemes, the capacity is obtained by summing the capacities associated with the delivery of the two layers (s_1 and s_2) transmitted by the BS and relay(s). Specifically, considering successful cancellation of the IRI at the relays by using SIC, the maximum achieved capacity is obtained by adding s_1 capacity with maximum of quality of $BS - D$ link threshold (from the direct transmission) plus the minimum capacity of s_2 received at R and capacity from R at D. The capacity of our proposed scheme is obtained by $C_D^T = C_D^R + C_D^{BS}$, where C_D^R is the capacity from the signal transmitted from R, and C_D^{BS} is the capacity from the signal transmitted from BS.

2.1.5 Numerical Results

In this section, we evaluate the performance of the proposed scheme and compare it with existing alternative relaying schemes. The performance of all schemes is evaluated using Mote-Carlo simulation based on averaging SNRs and link capacities over 10^6 independent channel realizations for a given positions of the relays and node D. In our simulation, both relays are located at the same distance from the BS r_{BS,R_i} = 100 m with an angle of $-18/18$ as shown in Figure 2.2. The simulated end users are positioned randomly within a 10×10 m^2 square where the center of the square is at distance $\hat{r}_{BS,D}$. We assume that the destination is located at an obstructed area and the relays are in better locations. It was assumed that the channels between the BS and relays exhibits Rayleigh fading as it is one of the worst channels to deal with in wireless systems.

The ratio $\frac{r_{BS,R}}{\hat{r}_{BS,D}}$ will be used as a parameter to analyze the performance of the scheme developed. In our evaluation scenario, we consider two different values for SNR at node D from the direct transmission from the BS; the values for SNR are used, SNR_1 = 10 dB and SNR_2 = 15 dB. The reason behind using those values of SNRs is that with those SNR values, destination are not able to receive the data reliably, hence, some improvement are required.

We assume that the power of AWGN at D and relays are the same and that relays

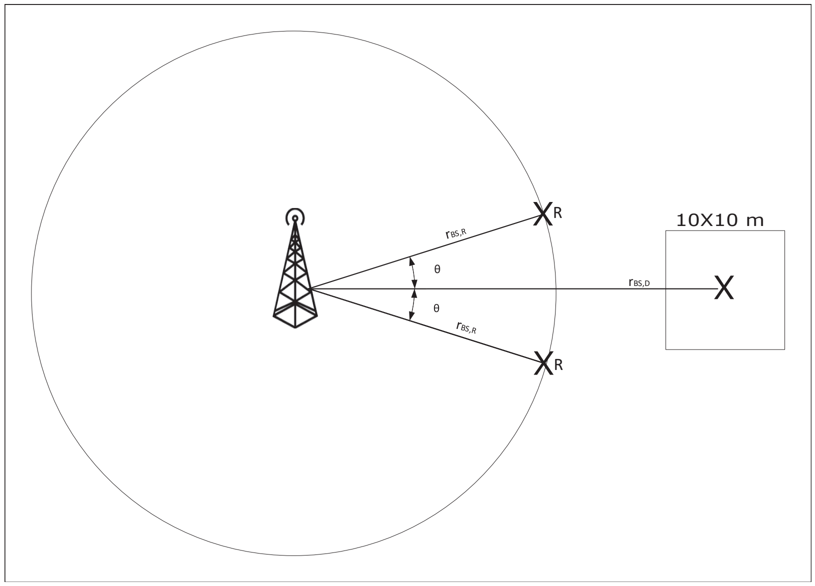

Figure 2.2: Cooperative multicast relaying.

have the same transmit power as the BS. Based on assumptions, the path loss exponent is assumed to be 2 between the BS and the relays, between the relays and the destination, and between the relays. Meanwhile, the path loss between the BS and destination is assumed to be 3. Therefore, the value of SNR at D accounting for signals transmitted by the relays is obtained by calculating $SNR_D^R = SNR_R^{BS}(r_{BS,R}/r_{R,D})^2$ where SNR_R^{BS} represents the value of SNR at D accounting for the signal transmitted by the BS.

Figure 2.3 shows the values of SNR received at relays and destination when the BS superimposes s_1 and s_2 by allocating 0.95 of the transmitted power for s_1 and 0.05 for s_2. We show: (i) in red curves, the SNR at node D for the recovery of $s_2(n-1)$ transmitted by the relays (before SIC decoding); (ii) in black curves, the SNR value at R_i $(i = 1, 2)$ for the recovery of $s_1(n)$ transmitted by the BS and (iii) in blue curves, the SNR value at R_i for the recovery of $s_2(n)$ transmitted by the BS. Triangle and circular shapes are used to differentiate between the two different values of SNR,

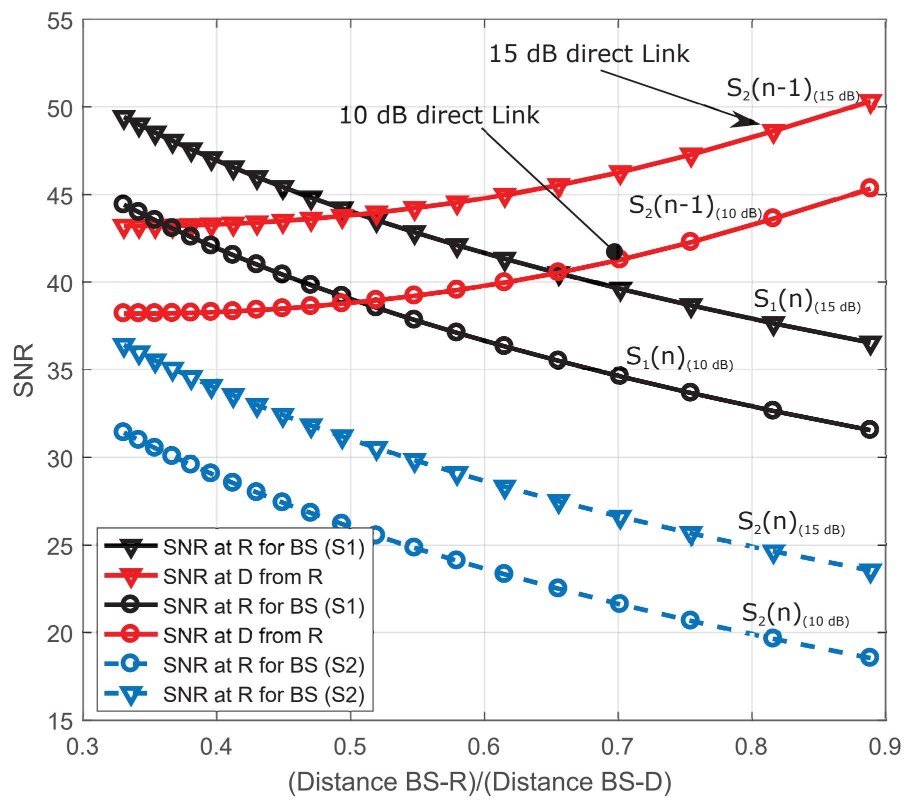

Figure 2.3: SNR values at relays and D when decoding different layers.

10 dB and 15 dB, of the direct transmission from the BS to node D.

It can be seen that when the fraction of the total power assigned to s_1 is 0.95 and 0.05 of the total power is assigned to s_2, the SINR to for s_1 is equal to 12.8 dB because s_1 is treated as an interferer, i.e., providing $10 \log_{10} \frac{0.95}{0.05} = 12.8\, dB$ SINR when decoding of $s_1(n)$ in the presence of $s_2(n)$ as interference at R_i. The maximum data rate that can be transmitted by s_1 is therefore limited by its SINR. After successfully decoding s_1 and subtracting it from the received signal, the relay decode s_2 with an SNR threshold ranging from $36\, dB$ to $24\, dB$, and $31\, dB$ to $19\, dB$ when the SNR value of the direct link between BS and D is $15\, dB$ and $10\, dB$, respectively.

From figure 2.3, the relays can transmit $s_2(n-1)$ to node D with an SNR threshold ranging from $43\,dB$ to $51\,dB$, and $38\,dB$ to $46\,dB$ when the SNR value of the direct link between BS and D is $15\,dB$ and $10\,dB$, respectively. It is worth mentioning here that the signal transmitted by the BS is considered as interference when decoding $s_2(n-1)$ transmitted by the relays at node D.

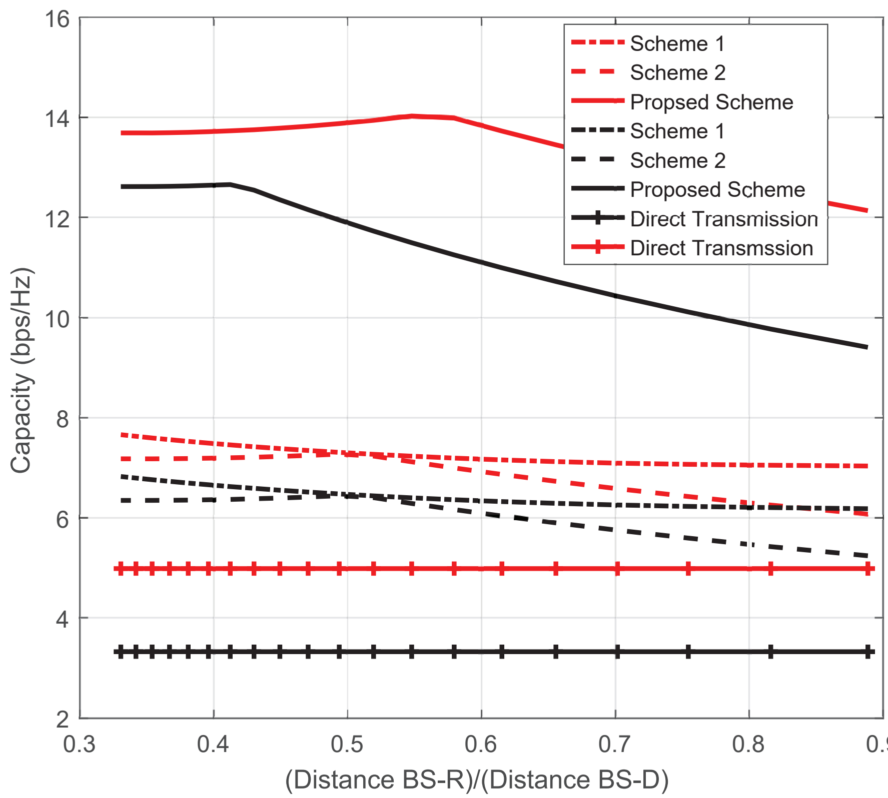

Figure 2.4: Capacity comparisons at the destination when using different schemes.

The values of SNRs are used to obtain the average capacities of the proposed schemes by calculating the the capacity expressions developed in the previous section. Figure 2.4 shows a comparison of maximum achieved capacities by the proposed and alternative schemes as a function of the normalized distance between the relays and destination. As can be seen, adaptive relaying scheme performs better than traditional relaying scheme in terms of capacity. However, adaptive relaying is more complicated than transitional as it requires buffering the messages. Traditional relaying can be considered a special case of the adaptive relaying as it can be noticed from figure 2.4 that both schemes have same performance when the relay is located in the middle between the BS and the destination. The proposed scheme out-preforms all alternate schemes in terms of capacity because it is efficient use of both time and power resources through the re-use of TSs and power adaptation, resulting in recovery of both layers by using SIC. The capacity of the proposed scheme is constant when the relay is located between 0.3-0.6, and that is because the limitation of the power at relays which is similar to the power at the BS.

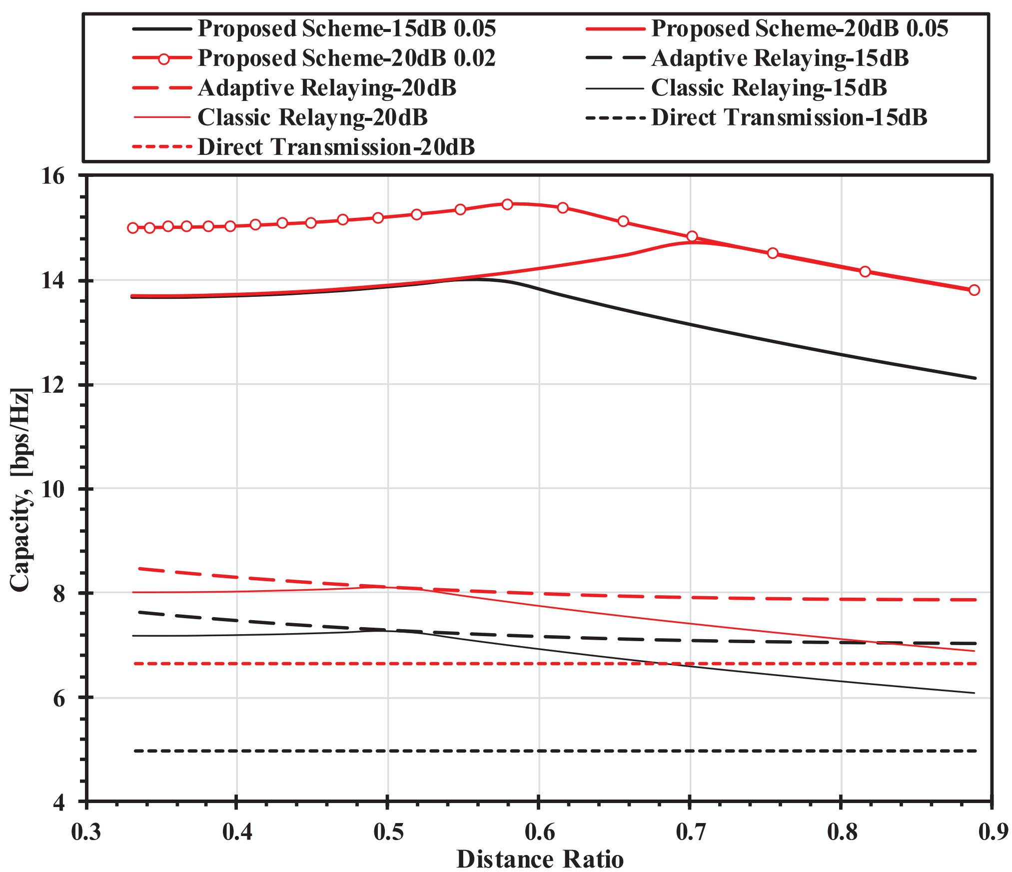

Figure 2.5: Capacity comparisons at the destination when using different schemes.

Figure 2.5 illustrates that the proposed scheme (layered transmission) with allocation factor 0.02 when the SNR of the direct like is 20 dB achieves about 1 bps/Hz performance gain over the proposed scheme when the allocation factor is 0.05. The power allocation factor plays an important rule in the proposed scheme as illustrated in figure 2.5 based on the direct link SNR.

2.1.6 Summary

This section introduced a relay-based transmission in a downlink wireless system using layered transmission. Through superimposing transmitted messages, two layers, into one signal, this scheme could be used to enable users with bad channel conditions to improve their data rate by using alternate-relaying protocol. However, one major issue of the alternate relaying is the inter-relay interference occurring during the successive relaying process, which compromises the overall system performance if not mitigated correctly. In our work, the IRI is dealt with at the destination and the relays using DF strategies. The two main components of the scheme are (i) superimposing the base and enhancement layers at the BS and multiplexed coding at the relays and (ii) power allocation to enable SIC decoding. So, by superimposing the base and enhancement layers at the source and using multiplexed coding through over-the-air signal summation, the destination is able to utilize all system resources for decoding the enhancement layer. Numerical results are presented to compare the average capacity with that of corresponding currently used schemes. The proposed scheme can provide a notable gain in terms of capacity, while successfully mitigating IRI between the relays.

2.2 Superposition Coding in Alternate DF Relaying Systems with Virtual MIMO IRI Cancellation in $1 \times 2 \times 2$ Antenna Configuration

As demonstrated in previous section, a combination of SC and AR system enables the IRI mitigation using SIC principles based on power profiles of different transmitted layers. The system showed a notable gain in terms of capacity. However, because it was assumed that all transmission nodes have the same power constraints, there were some locations for the relays when the BS was forced to lower the bit rate because the relays do not have enough power to deliver the enhancement layer in the presence of the interference from the BS at the destination. In this section, we consider the problem of canceling the IRI while using SC in the AR cooperative communication setup by using two receiving antennas at the destination and relays to efficiently cancel the IRI.

The purpose of this section is to present a layered transmission between the BS and the destination with the aid of two HD relays. The single antenna source continuously transmits the superimposed signal of base and enhancement layers, while relays alternately retransmit the enhancement layer utilizing only one antenna. So, by deploying both iterative SIC decoding and virtual MIMO, IRI will be successfully canceled. In detail, the receiving relay decouples two spatial streams representing data of interest using the virtual MIMO channel from the source and the transmitting relay, and then decode the base layer using SIC.

The rest of this section is organized as follows. In subsection 2.2.1 , we introduce the model of SC DF-SIMO alternate-relays communication system. The description of the proposed scheme is introduced in subsection 2.2.2. Subsections 2.2.3 and 2.2.4 present the the combination of semi-analytical capacity and simulated SNR results to demonstrate the performance of the scheme developed. Finally, we provide concluding remarks in subsection 2.2.5 .

2.2.1 Two-hop DF-MIMO Alternate-relaying Systems

We consider a wireless multicast communication system with one BS, two half-duplex relays (R_1,R_2), and a number of users in a multicast group as shown in figure 2.6. We assume that each node (relays and users) is equipped with two antennas, and the BS is equipped with only one antenna. Moreover, we study the systems where the channel conditions between (i) BS and relays, (ii) relays and users, and (iii) between relays are stronger than channel conditions between the BS and node D, which is an interesting case for relay-based systems. Single-carrier transmission under Rayleigh flat fading channel model is considered.

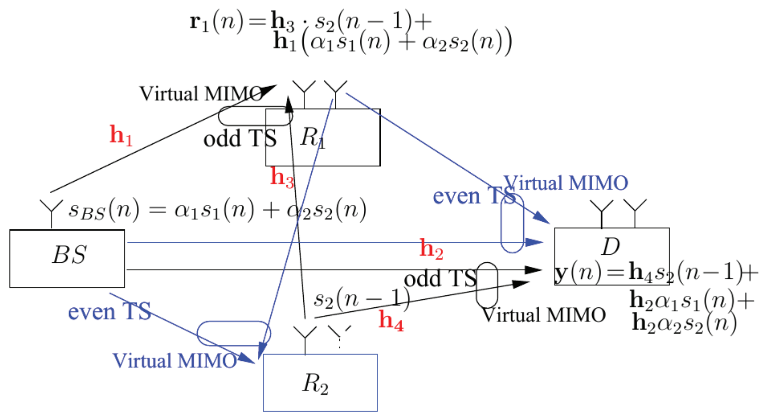

Figure 2.6: Two-path alternate relaying on a downlink with received asymmetric constellations

All wireless links exhibit independent and identically distributed (i.i.d.) block Rayleigh fading and are affected by AWGN. This means that the fading coefficients (channel gains) remain constant during one TS, but change independently from one slot to another according to a complex Gaussian distribution with zero mean and variance σ_h^2; the variance captures path-loss effects. The channel gains between all nodes are given by the column of vectors of size 2×1 (SIMO) as follows: (i) between BS and relays – $\mathbf{h}_1$ (in TS n); (ii) between the relays and D – $\mathbf{h}_4$; (iii) between BS and D – $\mathbf{h}_2$ and (iiii) between relays (represents the channel gains from the single transmit antenna at the transmitting relay to two receive antennas of the receiving relay)– $\mathbf{h}_3$. To simplify presentation, channel reciprocity between R_1 and R_2 is assumed. All the channel column vector gains are shown in Figure 2.6 using arrows, with the beginning of the arrow representing a single transmit antenna and the tip of the arrow representing two receive antennas.

In every time slot (TS) n, the BS multicasts its information-carrying symbols using SC where two layers, basic and enhancement, are linearly superimposed and transmitted as one symbol . The transmitted signal by the BS is generated as follows:

$$s_{BS}(n) = \sqrt{P_T} \cdot \Big(\alpha_1 s_1(n) + \alpha_2 s_2(n)\Big) \tag{2.15}$$

where $s_1(n)$ represents the base layer, $s_2(n)$ represents the enhancement layer. The average transmit power is assumed to be unity for s_1 and s_2,that is, $\mathbf{E}\{|s_1|^2\} = \mathbf{E}\{|s_2|^2\} = 1$, where $\mathbf{E}(\cdot)$ is the expectation operator and $0 < \alpha_i^2 \leq 1$ ($i \in \{0, 1\}$) are the power allocation coefficients between the two layers subject to constraint on $\alpha_1^2 + \alpha_2^2 = 1$, and P_T is the transmit power at the BS. Without loss of generality, the special cases of the system model, namely, signal transmissions in the first and the last time, are ignored in the sequel.

From Figure 2.6, transmissions/receptions in odd TSs are represented with black arrows, and the transmissions/receptions in even TSs are represented with blue arrows. During the odd time slot, R_1 receives the superimposed signal from the BS, while R_2 relays the enhancement layer (received in previous TS) to node D. Because of the simultaneous transmission of the BS and R_1, the received signal at relay node

R_1 can be given by:

$$\mathbf{r}_1(n) = \left[\mathbf{h}_1|\mathbf{h}_3\right] \begin{bmatrix} \sqrt{P_T} \cdot \left(\alpha_1 s_1(n) + \alpha_2 s_2(n)\right) \\ \sqrt{P_T} \cdot s_2(n-1) \end{bmatrix} + \mathbf{n}_{R_1}(n) \tag{2.16}$$

where (i) $\mathbf{n}_{R_1}$ stands for the circular AWGN (2-D $\mathcal{CN}(0, \sigma^2)$) at the relay R_1 and (ii) $\left[\mathbf{h}_1|\mathbf{h}_3\right]$ represents the virtual channel gain matrix of size 2×2 from the BS and R_2 to R_1.

On other hand, the received signal at node D is give by:

$$\mathbf{y}(n) = \left[\mathbf{h}_2|\mathbf{h}_4\right] \begin{bmatrix} \sqrt{P_T} \cdot \left(\alpha_1 s_1(n) + \alpha_2 s_2(n)\right) \\ \sqrt{P_T} \cdot s_2(n-1) \end{bmatrix} + \mathbf{n}_D(n) \tag{2.17}$$

where again (i) $\mathbf{n}_D(n)$ represents the circular AWGN at the destination and (ii) $\left[\mathbf{h}_2|\mathbf{h}_4\right]$ forms the virtual channel gain matrix of size 2×2 from the BS and R_2 to node D. The focus of our work is on the signal flow in the odd TSs because of the system symmetry; the signal analysis is the same in even TSs.

From equations 2.16) and 2.17, the power allocation for $s_1(n)$ and $s_2(n)$ is adjusted so that the destination can recover the $s_1(n)$ using iterative SIC process as proposed in the next sections. The SIC process at node D comes after the initial ZF post-processing of the received signal by exploiting the virtual MIMO created by the transmitting relay and the BS.

In the proposed scheme, node D can decode both signals $s_1(n)$ and $s_2(n)$ with the relay assistance in forwarding $s_2(n-1)$ to node D. The reason behind that is with the help of R_i by forwarding $s_2(n-1)$, node D will be able to recover it on the stronger spatial link between R_i and node D using virtual MIMO processing. Then, node D can suppress the effect of $s_2(n-1)$ from $S_{BS}(n-1)$. Hence, node D will have SNR to decode $s_1(n-1)$ and has already received $s_2(n-1)$.

In the next section, we demonstrate the IRI cancellation process using virtual

MIMO channels at the relays and destination shown in Figure 2.6 with black and blue ovals the relays during odd and even TSs, respectively.

2.2.2 Virtual SIMO and SIC Detection with IRI Cancellation

Assuming an odd TS n, as shown in Figure 2.6, the signal received at R_1 from R_2 and the BS creates a virtual MIMO channels, where transmissions/receptions in even TSs are represented with blue arrows with the conventional interpretation for the beginning and the tip of the arrow representing transmission and reception. In order for R_1 to relay $s_2(n)$ in the next $(n+1)$ TS, R_1 should decode $s_1(n)$, $s_2(n)$, and $s_2(n-1)$, i.e., the current base and enhancement layers and the delayed enhancement layer. R_1 can decode $s_2(n-1)$ by using ZF processing by inverting the virtual MIMO channel gain matrix $\left[\mathbf{h}_1|\mathbf{h}_3\right]$ (or pseudo-inversion). It is assumed that CSI for the BS-R_i,$R_i - D$, BS-D links, and between relays are available at the relays and node D. By ZF processing recovery, the relay can recover $s_{BS}(n)$ from $\mathbf{r}_1(n)$ as given in equation 2.16 as one of the spatial streams, then detect $s_1(n)$ as the second spatial stream by deploying SIC processing, then recovering $s_2(n)$ from the modified copy of $S_{BS}(n)$.

By employing more antennas at the relays, the associated inter-relay channel provides MIMO related degree of freedom in comparison with the conventional SISO relay system, hence offering the opportunity to eliminate the IRI in alternate relaying system. In other words, one degree of freedom provided by the inter-relay channel is used to efficiently remove the IRI.

Decoding at R_i, $s_1(n)$ and $s_2(n)$ can be recovered using SIC process only by assigning a certain α satisfying the following two conditions: (1) $\log_2\left(1+\frac{|h_1|^2P_1}{|h_1|^2P_2+N_0}\right) > R_{d1}$, and (2)$\log_2\left(1+\frac{|h_1|^2P_2}{N_0}\right) > R_{d2}$, where $> R_{d1}$ denotes the targeted data for $s_1(n)$, and $> R_{d2}$ is the targeted data rate for $S_2(n)$.

The decoding process at node D benefits also for the iterative SIC if we relax time constraints when recovering both layers. $s_2(n-1)$ is first recovered by using ZF processing because i) it is received with the highest power and ii) it is received on a spatial stream from the relay. For the recovery of $s_1(n-1)$, we revisit $s_{BS}(n-1)$ stored at node D, we remove the effect of reliably recovered $s_2(n-1)$ in the first phase, then we decode $s_1(n-1)$. The recovery of $s_1(n-1)$ can be achieved by satisfying the following condition: $\log_2\left(1+\frac{|h_2|^2P_1}{N_0}\right) > R_{d1}$.

When comparing data recovery at node D using virtual MIMO with SIC cancellation process of signals through power level separation of different layers as presented in [10], the current method benefits from decoding only 16-ary asymmetric QAM and QPSK on two spatial streams. In fact, in [10], we had to work with the 64 asymmetric QAM when we considered the case of the practical modulation at the BS.

The operation of the proposed SC alternate relaying scheme is summarized in Table 2.3.

2.2.3 Capacity Considerations

The performance of the proposed scheme is evaluated by means of a capacity analysis. Here we consider three alternative schemes to better understand the results presented in the numerical results section.

In the first scheme, we consider the direct transmission from the BS to node D. Here, node D implements the maximum ratio combining (MRC) for receiving the signal transmitted by the BS. The second scheme is the conventional two time slot relaying system with the aid of one of the relays (the one with the better channel conditions). In the first time slot, the BS transmits the data to one of the relays in a 1×2 SIMO mode; in the second time slot, the relay forwards the received data to node D by using two antennas leading to a 2×2 MIMO system. Here, the BS selects the relay with purpose to maximize the capacity by exploiting the channel diversity of the two links. The capacity of of this scheme, denoted by C_{CL}, can be found by:

Table 2.3: Summary of transmissions and decoding in the SC alternate relaying scheme with Virtual MIMO

	Transmission Phase
BS	In every TS n, BS sends $\sqrt{P_T} \cdot \Big(\alpha_1 s_1(n) + \alpha_2 s_2(n)\Big)$ with (i) P_T adjusted to meet the targeted SNR using a single antenna and (ii) α_1 and α_2 adjusted to decode s_1 and s_2 at R_1/R_2.
R_1/R_2	R_1 transmits in even and R_2 in odd TSs, both using single antenna out of the two available at any relay. In the designated TS n, using the maximum power P_T, the relay sends $s_2(n-1)$ decoded in the previous TS.
	Receiving Phase
R_1/R_2	In their respective TSs, e.g., n, based on virtual MIMO decoding, and SIC, R_1/R_2 decode in order: $s_{BS}(n)$ and $s_2(n)$
D	In every TS n, using virtual MIMO, D decodes in the first step $s_2(n-1)$ and stores $s_{BS}(n)$ for the use in the next TS. Then in the second step, based on $s_{BS}(n-1)$ recovered using virtual MIMO in the previous TS, uses $s_2(n-1)$ from the first step to decode $s_1(n-1)$.

$$C_{CL} = \frac{1}{2}\Big(\min\big(\max_{i\in\{1,2\}}(C_{BS-R_i}), \max_{i\in\{1,2\}}(C_{R_i-D})\big)\Big) \tag{2.18}$$

where C_{BS-R_i} is the capacity of the link from the BS to relay R_i, C_{R_i-D} is the capacity from relay R_i to D ($i = 1$ or $i = 2$).

In the third scheme, we adopt an adaptive relaying system in which relays receive and forward adaptively based on the channel conditions between the BS and relays, and between relays and node D. In the adaptive relaying scheme, the maximum achieved capacity, denoted by C_{AD}, is given by:

$$C_{AD} = \frac{(C_{BS-R_i}) \times (C_{R_i-D})}{(C_{BS-R_i}) + (C_{R_i-D})} \tag{2.19}$$

The capacity of the proposed scheme is obtained by summing the capacities associated with the delivery of the two layers as discussed initially in previous sections. We start by finding the capacity associated with the delivery of $s_1(n)$ and $s_2(n)$ at the relays from the BS which can be calculated from:

$$\log_2\Big(1 + \frac{|h_1|^2 P_1}{|h_1|^2 P_2 + N_0}\Big) + \log_2\Big(1 + \frac{|h_1|^2 P_2}{N_0}\Big) \tag{2.20}$$

We assume that $s_1(n)$ is reliably decoded so that its effect could be perfectly removed from $s_{BS}(n)$ to enable the decoding of $s_2(n)$. It is worth mentioning that the second term in equation 2.20 sets the limit on the capacity when recovering $s_2(n)$ and $s_2(n-1)$ at node D. Here, the first term is neglected from capacity considerations at node D due to the fact that $s_2(n)$ is the only signal to be forwarded to node D.

At the destination, the capacity is obtained by summing the capacities associated with the delivery of $s_1(n)$ from the BS and $s_2(n)$ for the relays. We first start by considering the capacity of the delayed enhancement layer($s_2(n-1)$) which can be obtained by:

$$C_{D,s_2} = \min\Big(\log_2\big(1+\frac{|h_1|^2 P_2}{N_0}\big), \log_2\big(1+\frac{|h_4|^2 P_T}{N_0}\big)\Big) \tag{2.21}$$

where the first term represents the capacity of delivering $s_2(n)$ from the BS to the relays, and the second term is the capacity of delivering $s_2(n)$ from the relays to node D. The capacity of $s_2(n)$ is, therefore, limited with minimum capacity associated with

BS-R_i, and R_i-D links.

We consider now the capacity associated with delivering the base layer($s_1(n)$), as proposed in previous sections, the recovery of $s_1(n)$ is not affected by any interference except for AWGN, i.e., $s_2(n)$ was perfectly suppressed from $s_{BS}(n)$ by using SIC process. Therefore, the capacity of $s_2(n)$ at node D is given by:

$$C_{D,s_1} = \log_2 \big(1+\frac{|h_2|^2 P_1}{N_0}) \tag{2.22}$$

From equation 2.21 and equation 2.22, the achievable data rate at the node D is: $C_{D,s_1} + C_{D,s_2}$.

In MIMO networks, ZF technique is frequently adopted to decode the received data as it requires the CSI only at the receiver. ZF detection is based on completely eliminating interference from other symbol layers when detecting a layer of interest. However, this results in suboptimal performance due to noise enhancement. In the developed work, we study the upper bound on the capacity of the underlying network topology and there was no practical implementation of the setup to study the noise enhancement effect.

2.2.4 Numerical Results

In this section, the proposed cooperative scheme is evaluated in terms of capacity through Monte-Carlo simulations based on averaging SNRs and capacities over 10^6 independent channel realizations for different positions of the relays and the destination. We consider a system with one BS, two relays, and one destination. The two relays and destination are all equipped with two antennas, while the BS is equipped with only one antenna. The distance between the BS and relays is $\hat{r}_{BS,D}, = 100\,m$ with an angle of $-18/18$ as shown in figure 2.2. The simulated end users are positioned randomly within a 10×10 m^2 square where the center of the square is at distance

$\hat{r}_{BS,D}$. For the performance evaluation, three scenarios are considered: (i) operation of the scheme when the distance between the BS and relays is 100 m, and the distance between the relays and destination ranges from $(20-100\ m)$; (ii) operation of the scheme when r_{BS-D} is 100 m and r_{BS-R_i} ranges from $(20-80\ m)$; (iii) the operation of the scheme when r_{BS-R_i} is 100 m and r_{R_i-D} is 100 m with SNR between the BS and D varying from 14 dB to 28 dB. For scenarios (i) and (ii), two values of SNR for the direct transmission link between the BS and node are considered, 15 dB and 20 dB.

The path loss exponent is assumed to be 2 between the BS and the relays, between the relays and the destination, and between the relays. The path loss between the BS and destination is assumed to be 3. Therefore, the value of SNR at D accounting for signals transmitted by the relays is obtained by calculating $SNR_D^R = SNR_R^{BS}(r_{BS,R}/r_{R,D})^2$ where SNR_R^{BS} represents the value of SNR at D accounting for the signal transmitted by the BS.

Figure 2.7 shows the SNR values at both the relays and node D when s_1 and s_2 are superimposed by allocating 95% of P_T for s_1, and 0.05 of P_T for s_2. The red line with bold circles shows the value of SNR for s_1, while the red line with empty circles shows the values of SNR for s_2 at the relays. From figure 2.7, the SINR for s_1 is equal to 12.8 dB because of the interference caused by s_2 signal. After successfully decoding s_1, s_2 then is decoded by using SIC with a SNR ranging from 24 dB to 31 dB.

In figure 2.8 , we compare the maximum achievable capacities of the four different schemes under scenario (ii) as a function of the distance between the BS and relays. The curves with bold circles shows that capacity when the SNR of the direct transmission from the BS to node D is 20 dB, while curves with no circles show the maximum achieved capacity when the SNR of the direct transmission from the BS to node D is 15 dB. It can be observed from figure 2.8 that the capacity of the

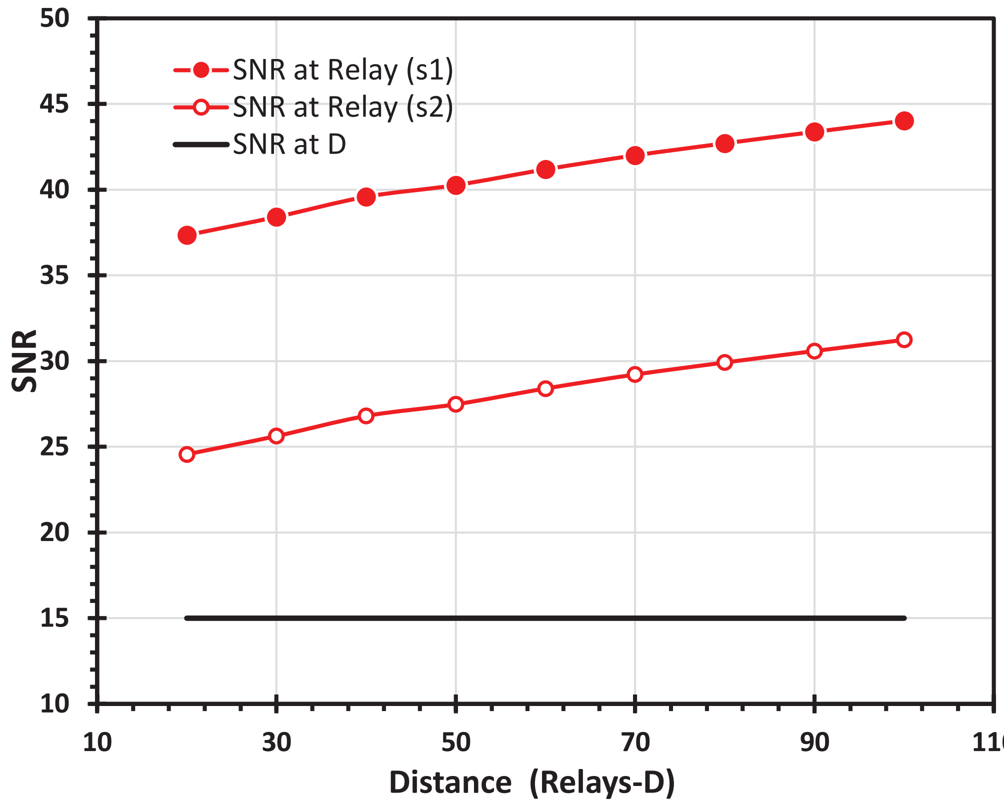

Figure 2.7: SNR values at relays and D when decoding different layers.

direct transmission from the BS to node D (black dashed lines) is 5.5 bps/Hz and 7.27 bps/Hz then the SNR values of the direct transmission are 15 dB and 20 dB, respectively. Direct transmission scheme provides lower capacity among all schemes due the poor channel conditions between the BS and node D. Classic relaying scheme (solid red curves) performs slightly better than the direct transmission scheme. The adaptive relaying scheme, depicted by red dashed lines, can achieve slightly higher capacity compared to direct transmission and classic relaying. The solid black line shows the maximum achieved capacity of the proposed scheme. It can be noticed that the proposed scheme outperforms all considered schemes by at least 2 bps/Hz when r_{BS-R_i} is 80 m .

Figure 2.9 shows a comparison of maximum achieved capacity between all schemes

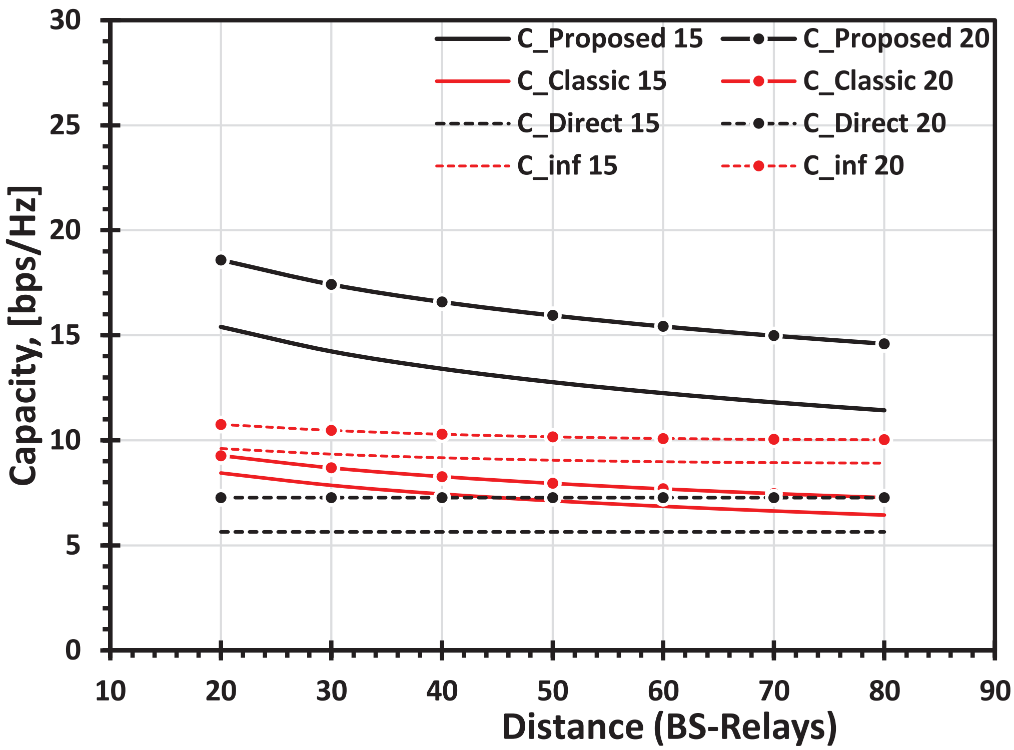

Figure 2.8: Capacity comparisons at the destination when using different schemes.

under scenario (i). It can be observed that the proposed scheme has the best performance among all schemes due to its ability to successfully combine the signals from both the relays and BS and use them efficiently. It can be noticed from figure 2.9 that the capacity of of all schemes increases with the increase of the distance between the relays and node D. The reason behind this increase is that SNR values at the relays increase when r_{R_i-D} increases, since the r_{BS-R_i} is set to be 100 m, as demonstrated in figure 2.7.

Figure 2.10 depicts the maximum achieved capacity by the different transmission schemes versus the SNR values of the direct link between the BS and node D. It can be seen that the proposed scheme provides good performance at all SNR values, benefiting from using signals from both the BS and the relays.

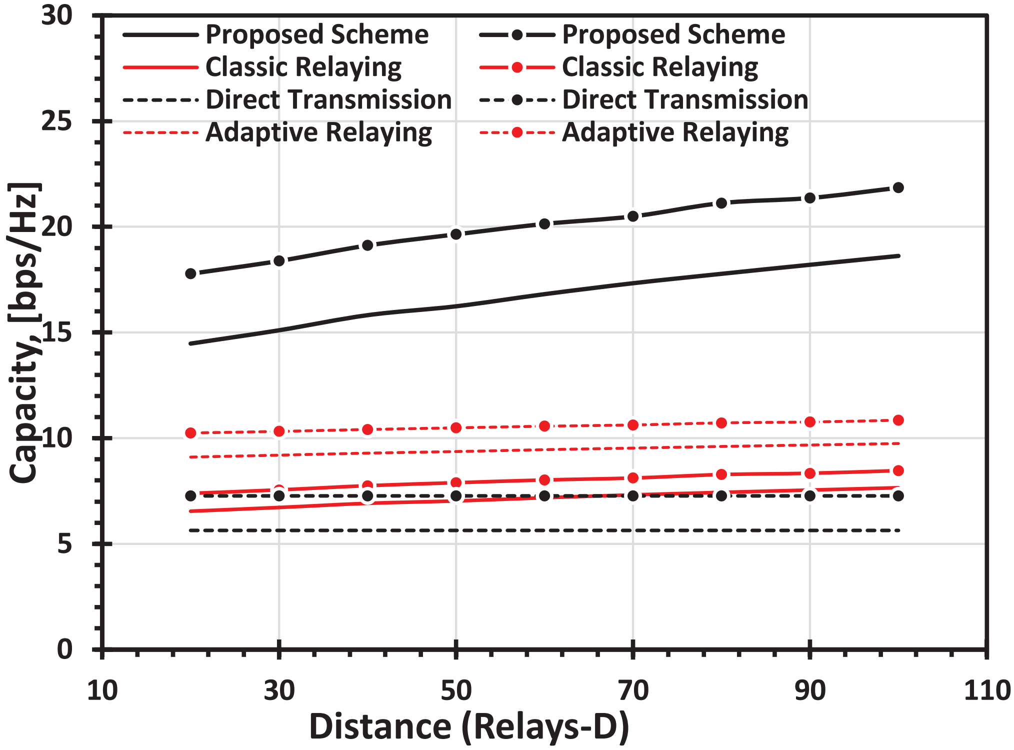

Figure 2.9: Capacity comparisons at the destination when using different schemes.

By comparing the achieved capacities between the system in section 1.1 and 1.2, it can be noticed that in system 1.1 there was a region of values of SNR for which the capacity of system was saturated because of the power constraint at the relays, whereas in system 1.2 the power constraint was relaxed, hence the proposed system in 1.2 achieved higher capacity than proposed system in section 1.1.

2.2.5 Summary

This section introduced an alternate relay-based transmission in a downlink wireless system using superposition coding. The scheme enables cell-edge users (destination) with bad channel conditions to improve their data rate by using SC at the BS with the aid of the relays. To remove the IRI and efficiently decode the two data layers at

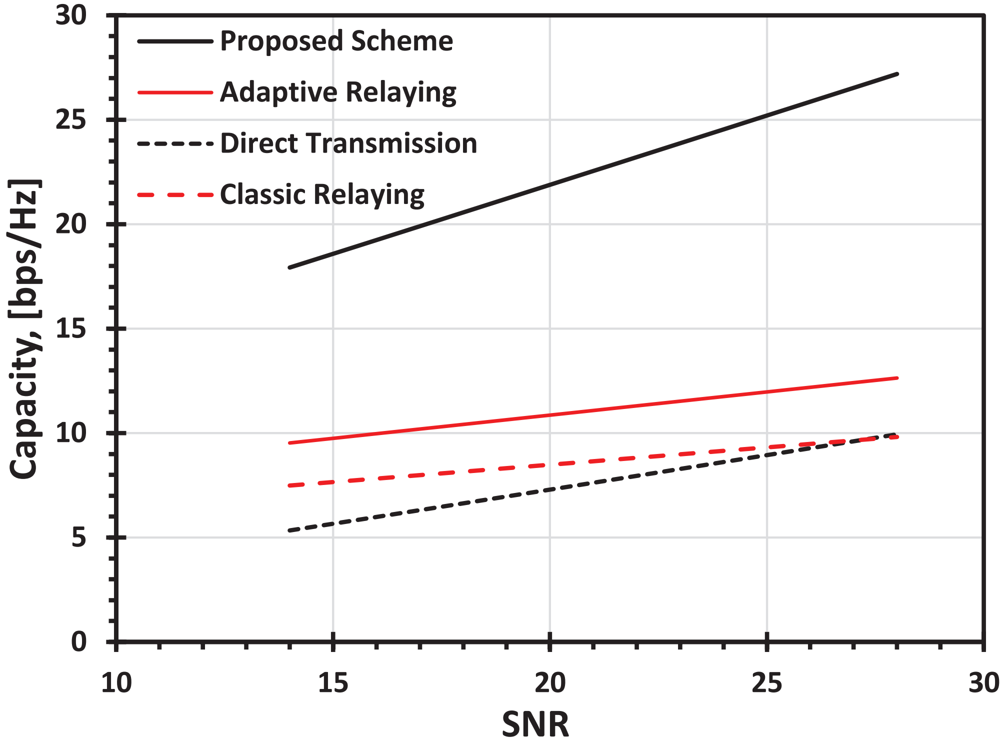

Figure 2.10: Capacity comparisons at the destination when using different schemes.

the destination, we created the opportunities for virtual MIMO decoding at both the relays and the destination. This was accomplished by using the relays (i) to transmit using a single antenna, and (ii) to receive using two antennas. The virtual MIMO channels created by the BS (one antenna) and relays (using a single antenna) enabled MIMO decoding at the destination by using two antennas and iterative SIC detection of both layers. The proposed scheme can provide a notable gain in terms of capacity, while successfully mitigating IRI between the relays.

2.3 Superposition Coding in Alternate DF Relaying Systems in a Generalized Antenna Configuration

This section establishes a general model for the system discussed in section 2.2. More specifically, a downlink multicast network is considered, where the BS is equipped with multiple antennas to communicate with the destination node, which represents the user with the worst channel conditions within a MG. Two relay nodes equipped with multiple antennas are used.

The BS continuously transmits the superimposed signal of the base and enhancement layers, while relays alternately retransmit the enhancement layer by utilizing half of the antennas.

The rest of this section is organized as follows. Subsection 2.3.1 introduces the model for a SC DF-MIMO alternate-relay communication system. The proposed scheme is described in subsection 2.3.2. Subsections 2.3.3 and 2.3.4 present a combination of semi-analytical capacity and simulated SNR results to demonstrate the performance of the proposed scheme. Finally, conclusions are provided in subsection 2.3.5.

2.3.1 Two-hop DF-MIMO Alternate-Relay Systems

A wireless multicast communication system is considered with one BS, two half-duplex relays (R_1,R_2), and a number of users in a multicast group, as shown in figure 2.11. It is assumed that R_1 and R_2 are equipped with M, the BS with $\frac{1}{2}M$, and node D with M antennas. Moreover, a practical network scenario is studied where the channel conditions between the BS and relays, between the relays and users, and between relays are better than the channel conditions between the BS and node D. This is an interesting case for relay-based systems. In the simulation section, these observations are used by working with different deterministic path loss laws for the power decay at different links. The power decay at distance r is given by $\frac{1}{r^\beta}$, where the path loss exponent β is varied based on the channel conditions. Single-carrier transmission with a Rayleigh flat fading channel model is considered.

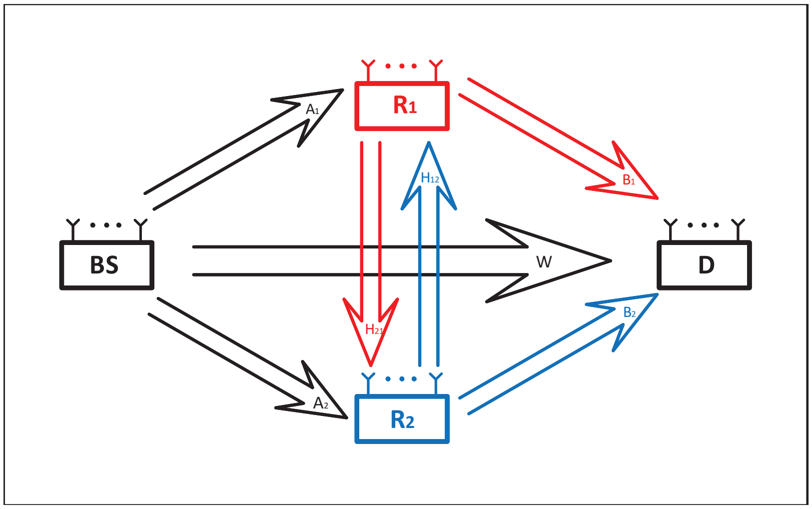

Figure 2.11: Two-path alternate relaying on a downlink with asymmetric constellations received

All wireless links exhibit independent and identically distributed (i.i.d.) block (vector) Rayleigh fading and are affected by circular AWGN, i.i.d., at each antenna. This means that the fading coefficients (channel gains) remain constant during one TS, but change independently from one slot to another according to a complex Gaussian distribution, with zero mean and variance σ_h^2. For instance, the $(M \times M)$-dimensional channel matrices between the two relay nodes are considered constant with one TS. The channel gains are given by the column vectors of size $M \times \frac{1}{2}M$ as follows: (i) between BS and relays – $\mathbf{A}_i$ (in TS n); (ii) between the relays and D (where relays transmit by using only $\frac{1}{2}M$ antennas) – $\mathbf{B}_i$ and (iii) between BS and D – $\mathbf{W}$, i.e., there is a direct link between the BS and D. The column vector $\mathbf{H}_{ij}$ (also of size $M \times \frac{1}{2}M$) represents the channel gains from the transmit antennas at the transmitting relay to the receive antennas of the receiving relay. To simplify the presentation, channel reciprocity between R_1 and R_2 is assumed. Here, the channels specified in bold letters, e.g., $\mathbf{A_2}$, should be interpreted as channel gain matrices with different subscripts to differentiate between various links, as shown in figure 2.11 and summarized in Table 2.4.

Table 2.4: Network parameter definitions

Symbol	Definition
BS	The source
R_1	Relay number 1
R_2	Relay number 2
D	The destination
$\mathbf{A_1}$	The channel from BS to R_1
$\mathbf{A_2}$	The channel from BS to R_2
$\mathbf{H_{21}}$	The channel from R_1 to R_2
$\mathbf{H_{12}}$	The channel from R_2 to R_1
$\mathbf{B_1}$	The channel from R_1 to D
$\mathbf{B_2}$	The channel from R_2 to D
$\mathbf{W}$	The channel from BS to D
P	S signal power

All the channel column vector gains are indicated in figure 2.11 by using arrows, where the beginnings of the arrows represent transmit antennas and the tips of the arrows represent receive antennas.

At $TS = 0$, the communication process is initiated. The BS continuously transmits the source message, denoted by s_{BS}, which contains two layers:

$$s_{BS}(n) = \sqrt{P_T} \cdot \Big(\alpha_1 s_1(n) + \alpha_2 s_2(n)\Big) \tag{2.23}$$

where $s_1(n)$ and $s_2(n)$ represent the base and enhancement layers respectively. The power allocation parameters ($0 < \alpha_i^2 \leq 1$ ($i \in \{0, 1\}$) between two layers are such that $\alpha_1^2 + \alpha_2^2 = 1$, and P_T is the total transmit power at the BS. It is further assumed that $P_1 = \alpha_1^2.P_T$ and $P_2 = \alpha_2^2.P_T$. Since the BS has $\frac{1}{2}M$ antennas, the number of messages being transmitted by the BS in each TS is equal to $\frac{1}{2}M$.

During the odd TS, as shown in figure 2.11, R_1 receives $\frac{1}{2}M$ superimposed signals from the BS, while R_2 transmits to node D $\frac{1}{2}M$ enhancement layers recovered at the previous TS. Because of the simultaneous transmission of the BS and R_1, the signals received at R_1 are not only the pure signals from the BS but also interference from the signals transmitted by R_2. The signals received at R_1 can thus be expressed as:

$$\mathbf{r}_1(n)= \mathbf{A}_1 s_{BS}(n) + \mathbf{H}_{21} s_2(n-1) + \mathbf{n}_{R1}(n) \tag{2.24}$$

where $\mathbf{n}_{R_1}$ stands for the circular AWGN (2-D $\mathcal{CN}(0, \sigma^2)$) at relay R_1.

In this section the task of the relays is to decode the signal received and to forward only $\frac{1}{2}M$ $s_2(n-1)$ messages by using $\frac{1}{2}M$ of the M antennas.

The signal received at node D is a composite signal from the BS and R_2. Thus, the following signals received at node D during the odd TS can be expressed as:

$$\mathbf{y}(n)= \mathbf{W} s_{BS}(n) + \mathbf{B}_2 s_2(n-1) + \mathbf{n}_D(n) \tag{2.25}$$

where $\mathbf{n}_D$ denotes the circular AWGN (2-D $\mathcal{CN}(0, \sigma^2)$) at node D. The signals received at the destination are composed of desired signals, interference, and AWGN from the destination.

In the proposed system, node D can decode both streams with layers s_1 and s_2 with the aid of the relays. Node D benefits from the relay transmission so that node D is reliably able to recover the signals of the usually stronger spatial streams from the relays. Then, by using SIC, node D can recover signals received in the spatial stream from the BS (direct link). The following section demonstrates the proposed efficient removal of IRI by using virtual MIMO channels created at the destination and relays, as shown in figure 2.11.

2.3.2 Virtual MIMO and SIC Detection

Thanks to the use of MIMO at the two intermediate relays, the inter-relay channel gives a MIMO-related (spatial) additional degree of freedom in comparison to the SISO system described in section 2.1. This additional degree of freedom is to be used for the efficient removal of IRI.

For a signal flow as shown in Fig. 2.11, in the odd TS n the signal received at R_1 is a composite of signals transmitted by the BS and R_2, as expressed in equation 2.24. Due to the system symmetry, the focus in this section is on the signal flow in the odd TSs, as in the preceding section.

Careful examination of equation 2.24 shows that $\mathbf{A}_1$ and $\mathbf{H}_{21}$ form a matrix with a size of $M \times M$. Moreover, $s_{BS}(n-1)$ and $s_2(n)$ form a M-dimensional vector with size $M \times 1$. The signal received at R_1 from the BS and R_2 can be expressed as a $M \times 1$ column vector:

$$\mathbf{r}_1(n) = \left[\mathbf{A}_1 | \mathbf{H}_{21}\right] \begin{bmatrix} \sqrt{P_T} \cdot \left(\alpha_1 s_1(n) + \alpha_2 s_2(n)\right) \\ \sqrt{P_T} \cdot s_2(n-1) \end{bmatrix} + \mathbf{n}_{R_1}(n) \tag{2.26}$$

The decoding process at R_1 is accomplished in two steps. In step one, a ZF-type processing technique is employed by inversion (or pseudo-inversion) of the virtual MIMO channel gain matrix $\left[\mathbf{h}_1 | \mathbf{h}_3\right]$ to recover the spatial stream with a higher power $s_2(n-1)$. After deployment of the ZF technique, M signals are successfully recovered as follows: (i) $\frac{M}{2}$ $s_2(n-1)$ messages are recovered from the spatial stream with a higher power (transmitted by the relays); and (ii) $\frac{M}{2}$ $s_{BS}(n)$ messages are recovered from the second spatial stream (transmitted by the BS). However, $s_{BS}(n)$ has two superimposed signals, so further processing is needed to recover both superimposed signals.

In step two, $s_1(n)$ is first decoded from $s_{BS}(n)$ by using SIC with a maximum achievable rate $R1 \le C = \log_2(1 + \frac{|h_1|^2 P_1}{|h_1|^2 P_2 + 1})$, i.e., $s_1(n)$ is limited by the power allocated to the second layer (interference/noise power). After the decoding of $s_1(n)$, the effect of R_1 is canceled from $s_{BS}(n)$ and recovered by $s_2(n)$ with an achievable rate of $R2 \leq C = \log_2(1 + |h_1|^2 P_2)$.

The main idea of the proposed IRI cancellation method is that the relays use $\frac{M}{2}$ degrees of freedom for IRI cancellation, with the remainder going to the signal transmitted by the BS. Thus, here the spatial streams from the transmitting relay are treated as a signal of interest.

Table 2.5: Superimposed signals at the destination

TS n	Relay streams $s_2(n-1)$	Mid Power $s_1(n)$(BS streams)	Low power $s_2(n)$(BS streams)	Low power $n_D(n)$
1	$s_2(0)$	$s_1(1)$	$s_2(1)$	$n_D(1)$
2	$s_2(1)$	$s_1(2)$	$s_2(2)$	$n_D(2)$
3	$s_2(2)$	$s_1(3)$	$s_2(3)$	$n_D(3)$

The decoding process at node D is similar to that at the relays, and also benefits from SIC if time constraints are relaxed when recovering s_1 and s_2.

Examination of equation 2.25 shows that $\mathbf{W}$ and $\mathbf{B}_2$ form a matrix with a size of $M \times M$. Moreover, $s_{BS}(n)$ and $s_2(n-1)$ form a M-dimensional vector with a size of $M \times 1$. Therefore, by combining both signals, a virtual MIMO of size $M \times M$ is created. The signal received at R_1 from the BS and R_2 can be expressed as a $M \times 1$ column vector:

$$\mathbf{y}(n) = \left[\mathbf{W}|\mathbf{B}_2\right] \begin{bmatrix} \sqrt{P_T} \cdot \left(\alpha_1 s_1(n) + \alpha_2 s_2(n)\right) \\ \sqrt{P_T} \cdot s_2(n-1) \end{bmatrix} + \mathbf{n}_D(n) \tag{2.27}$$

The decoding process at node D consists of two phases. In phase one, because of good channel conditions between the relay(s) and node D (with high channel gain values in $\mathbf{B}_2$), $s_2(n-1)$ is received with the highest power and is recovered first by applying ZF matrix inversion of the virtual MIMO channel matrix $\left[\mathbf{h}_2|\mathbf{h}_4\right]$. Phase two revisits $s_{BS}(n-1)$, recovered in the previous TS through ZF processing. Knowledge of the power allocation coefficients then makes it possible to remove the effects of the reliably recovered $s_2(n-1)$ in the current TS (phase one). This type of SIC-based detection of data encoded in $s_1(n-1)$ and $s_2(n-1)$ can be iterated a few times until the noise effects are stabilized.

The operation of the proposed SC in alternate DF relaying systems with multiple antenna relay nodes is summarized in Table 2.6.

2.3.3 Capacity Considerations

In general, the main focus of using alternate relays is to increase the throughput of the system by reducing the effect of the prelog factor. Thus, the primary motivation for studying alternate relaying schemes is to improve the capacity of the network by multiplexing more signals in the available resources. This subsection analyzes the performance of the proposed scheme by using capacity considerations. The capacity of the proposed scheme is obtained by summing the capacities associated with the delivery of the two messages transmitted from the BS and relays.

First the capacity provided by only one antenna is calculated. This is then generalized for multiple antennas. When SC is used with SIC in the SISO scenario, every receiver attempts first to decode the stronger signal ($s_1(n)$), and then subtract it from the superimposed signal. The message with the lower allocated power ($s_2(n)$) is then decoded. Without considering the IRI and MIMO system, the maximum achievable data rate for the relays, denoted by R_1, is $R1 \leq C = \log_2(1+\frac{|h_1|^2P_1}{|h_1|^2P_2+1})+\log_2(1+|h_1|^2P_2)$, i.e., the relays decode s_1 and s_2, where decoding of s_1 is limited by the power allocated to s_2 (interference power). When SC and SIC are used in MIMO networks, first antennas are selected with a maximum channel gain between the BS/relays and node D, to exploit the antenna selection diversity. The achievable data rate at node D comes from the direct transmission from the BS ($\frac{M}{2}$ spatial streams) and the transmission from the relays after antenna selection ($\frac{M}{2}$),

Table 2.6: Summary of transmission and decoding phases in the SC in alternate DF relaying systems with multiple antenna relay nodes

	Transmission Phase
BS	The BS sends $\frac{M}{2}$ $\sqrt{P_T} \cdot \Big(\alpha_1 s_1(n) + \alpha_2 s_2(n)\Big)$ messages by using $\frac{M}{2}$ antennas with (i) P_T adjusted to meet the targeted SNR for decoding R_1 at D (without the noise of s_2), (ii) α_1 and α_2 adjusted to decode s_1 and s_2 at R_1/R_2.
$R_1/\ R_2$	R_1 transmits in even and R_2 in odd TSs by using $\frac{M}{2}$ of the available M antennas. In the designated TS n, the relay sends $\frac{M}{2}$ $s_2(n-1)$, decoded in the previous TS.
	Receiving Phase
$R_1/\ R_2$	In their respective TSs, e.g., n, by using M antennas, $R_1/\ R_2$ decode $\frac{M}{2}$ $s_2(n-1)$, $\frac{M}{2}$ $s_1(n)$ and $\frac{M}{2}$ $s_2(n)$ messages transmitted by the BS and the other relay.
D	In every TS n, by using virtual MIMO and SIC, node D decodes $\frac{M}{2}$ $s_2(n-1)$ and uses this to cancel the effect of the stored $s_{BS}(n-1)$. Then $\frac{M}{2}$ $s_1(n-1)$ is decoded and $\frac{M}{2}$ $s_{BS}(n)$ is stored for the next time slot.

resulting in a $M \times M$ MIMO system, as shown in equation 2.27. However, because of the network topology, the data rates for each layer have to be chosen carefully according to SNR values associated with the power allocation coefficients, the direct link SNR value, and the channel conditions at node D. In the proposed scheme, s_1

is transmitted directly to node D with the aid of the relays. The data rate of s_1 is limited by the power allocation coefficients (α_1 and α_2) and the SNR value from the BS at node D (the direct link). s_2 is forwarded by the relays to the destination, and since s_2 has to be decoded at the relays first, the data rate is limited by α_2 (SNR from the BS at the relays). Following assignment of the power coefficients at the BS and the power value for s_2 at the relays, the capacity can then be obtained by using the $M \times M$ MIMO capacity system. It should be mentioned here that it is not necessary for the relays in this scheme to use the maximum total power.

To measure the efficiency of the proposed scheme, its capacity is compared with that of three alternative schemes: The first scheme involves direct transmission from the BS to node D, where a MRC technique is employed by node D. In the second scheme, a classic two-time-slot relay system is considered, i.e., the BS transmits the information to a selected relay in a $\frac{M}{2} \times M$ MIMO mode, and the relay transmits the received data to node D in a $M \times M$ MIMO mode. The capacity of this scheme, denoted by C_{CL}, is given by:

$$C_{CL}=\frac{1}{2}\Big(\min\big(\max_{i\in\{1,2\}}(C_{BS-R_i}), \max_{i\in\{1,2\}}(C_{R_i-D})\big)\Big) \tag{2.28}$$

where C_{BS-R_i} is the capacity of the link from the BS to relay R_i, and C_{R_i-D} is the capacity from relay R_i to D ($i = 1$ or $i = 2$). In the third scheme, an adaptive relaying system is adopted in which relays receive and forward adaptively based on the channel conditions of the BS-$R_(i)$ and $R_(i)$-D links. In the adaptive relaying scheme, the maximum capacity achieved, denoted by C_{AD_3}, is given by:

$$C_{AD} = \frac{(C_{BS-R_i}) \times (C_{R_i-D})}{(C_{BS-R_i}) + (C_{R_i-D})} \tag{2.29}$$

In MIMO networks, ZF technique is frequently adopted to decode the received data as it requires the CSI only at the receiver. ZF detection is based on completely eliminating interference from other symbol layers when detecting a layer of interest. However, this results in suboptimal performance due to noise enhancement. In the developed work, we study the upper bound on the capacity of the underlying network topology and there was no practical implementation of the setup to study the noise enhancement effect.

2.3.4 Numerical Results

In this section, the proposed cooperative scheme is evaluated in terms of its capacity through Monte-Carlo simulations based on averaging SNRs and capacities over 10^6 independent channel realizations for different positions of the relays and the destination. A system is considered which has one BS, two relays, and one destination. The two relays and the destination are all equipped with M antennas, while the BS is equipped with only $\frac{M}{2}$ antennas. The distance from the BS to the relays is $\hat{r}_{BS,D}, = 100\ m$, with an angle of $-18/18$ as shown in figure 2.2. The simulated end users are positioned randomly within a $10 \times 10\ m^2$ square, where the center of the square is located at a distance of $\hat{r}_{BS,D}$. For the performance evaluation, two scenarios are considered: (i) operation of the scheme where the distance from the BS to the relays is 100 m, and the distance from the relays to the destination ranges from $20 - 100\ m$; and (ii) operation of the scheme where r_{BS-R_i} is 100 m and r_{R_i-D} is 100 m, with the SNR between the BS and D ranging from 15 dB to 35 dB. In scenario (i), it is assumed that the value of the SNR for the direct transmission link between the BS and node D is 15 dB.

The path loss exponent is assumed to be 2 between the BS and the relays, between the relays and the destination, and among the relays. The path loss value between the BS and the destination is assumed to be 3.

Figure 2.12 compares the maximum capacity achieved by the four schemes under

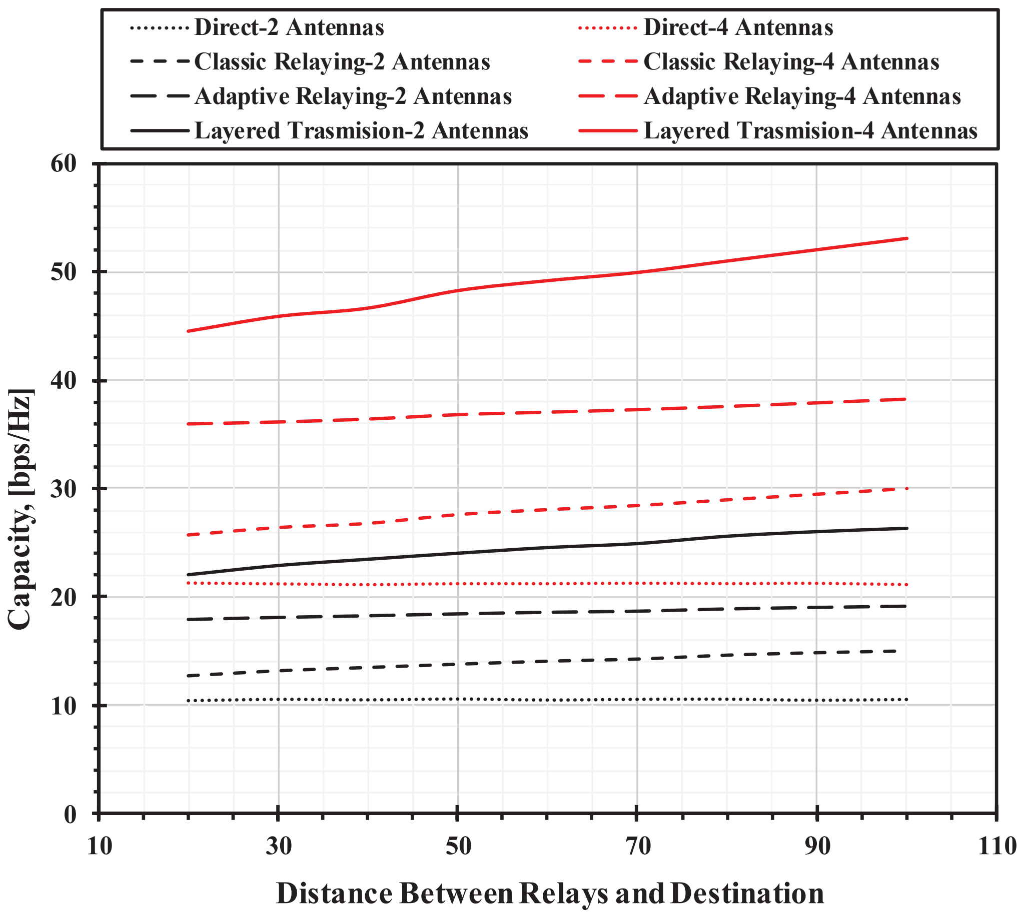

Figure 2.12: Comparison of capacity at the destination for different schemes under scenario (i).

scenario (i), as a function of the distance between the relays and the destination. Figure 2.12 shows the capacities achieved when the BS is equipped with two antennas (indicated in black), and with four antennas (indicated in red). The capacities for direct transmission from the BS to node D (indicated by dotted lines) are 10.5 bps/Hz when the BS has 2 antennas and 21.5 bps/Hz when the BS has 4 antennas. The direct transmission scheme has the lowest capacity of all the schemes considered, due the poor channel conditions between the BS and node D. The classic relaying scheme (indicated with short dashes) performs slightly better than the direct transmission scheme. The adaptive relaying scheme (indicated with longer dashes) achieves a slightly higher capacity in comparison to direct transmission and classic relaying. The maximum capacities achieved by the proposed scheme are indicated by solid lines. It is notable that the proposed scheme outperforms all the other schemes considered, due to its ability to benefit from the continuous direct transmission of the BS and relays, and due to the diversity improvement of the relays.

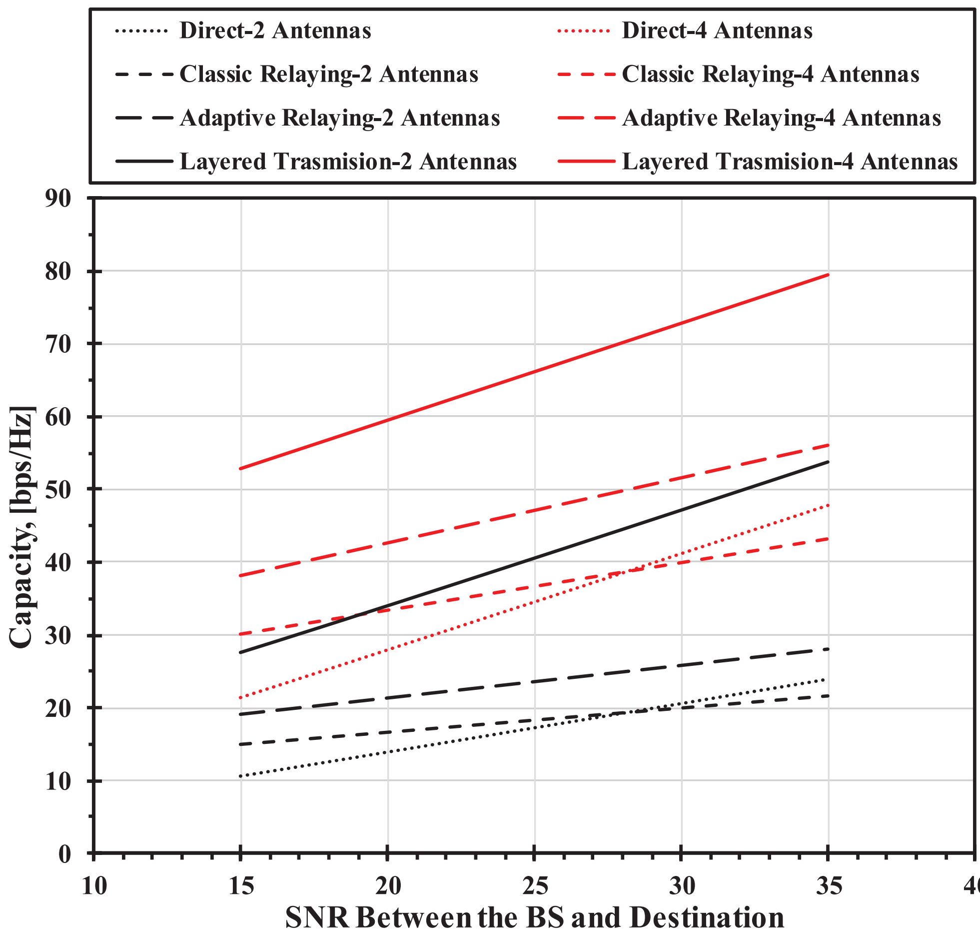

Figure 2.13: Comparison of capacity at the destination for different schemes under scenario (ii).

Figure 2.13 compares the maximum capacity achieved by the four schemes under scenario (ii), as a function of the SNR value between the BS and the destination. From figure 2.13, it can be seen that the direct transmission scheme has the lowest capacity of all the schemes considered. The classic relaying scheme performs slightly better than the direct transmission scheme. The adaptive relaying scheme (indicated with longer dashes) achieves a slightly higher capacity in comparison to direct transmission and classic relaying. The maximum capacities achieved by the proposed scheme are indicated by solid lines. It is clear that the proposed scheme outperforms all the other schemes considered, due to its ability to benefit from direct transmission from the BS and the relays simultaneously, with the use of multiple antennas and SIC-based detection.

2.3.5 Summary

This section introduces an alternate relay-based transmission in a downlink wireless system by using superposition coding, where all nodes are equipped with multiple antennas. The scheme enables cell-edge (destination) users with poor channel conditions to improve their data rates by using layered transmissions at the BS, SIC-based detection at the destination, and alternate relaying. For mitigation of the IRI and efficient decoding of the two data layers at the destination, opportunities for virtual MIMO decoding are created both at the relays and at the destination. This is accomplished by using the relays to transmit via $\frac{M}{2}$ antennas, and to receive via M antennas. The virtual MIMO channels created by the BS $\frac{M}{2}$ and relays enable MIMO $(M \times M)$ decoding at the destination, through the use of M antennas and iterative SIC detection of both layers. The proposed scheme provides a notable gain in terms of capacity, while successfully mitigating IRI between the relays.

Chapter 3

Increasing Data Rates in Relay-assisted Wireless Multicast Networks with Single Antenna Receiver

The increasing demand for mobile applications such as software updates, streaming media, traffic updates, location-based services, and certain applications that require the transmission of the same data to a selected group has triggered the need for wireless multicast technologies. The broadcasting nature of the wireless channel makes it naturally suitable for multicasting application, since every transmission by a node can be received by all nodes that lie within its communication range. However, the negative consequence of wireless broadcast is that serving different multicast groups simultaneously may cause MAI, and wireless channel is subject to signal fading. Various approaches have been proposed to ensure reliable delivery of data to all users in a MG. One of the most effective approaches to combat path loss and fading is to use intermediate relays between the source and destination due to its ability to extend the range and to improve the channel diversity.

In this chapter, we focus on using relays in a cooperative multicast network to benefit from the multiplexing gain. Specifically, spatial spectrum reuse is adopted in order to increase the downlink throughput of a multi-group multicast MISO cellular system, where inactive users may serve as cooperative relays to assist the operation of the centrally located BS.

The first section of this chapter presents the underlying system model and transmission methodology. In section 3.2, the principle behind the proposed data recovery algorithms is presented. In section 3.3, various configurations of relays selections are introduced. The numerical results are presented in section 3.4. And finally, the conclusions are presented.

3.1 System Model and Two-stage Relaying

In this chapter, we consider a multi-group multicast system which consists of a centrally located BS equipped with M antennas multicasts data to N MGs. We assume that each MG has L active users(MSs), and every user is equipped with a single antenna. All antennas are assumed to be omnidirectional. We initially assume that there are $M-1$ MSs in addition to the L MSs in each MG. The $M-1$ MSs will be serving as cooperative relays for MSs in the designated MG. We assume further that MSs in each MG are close to one another. Figure 3.1 shows the proposed system in the case when there are two MGs served by the BS.

In the figure, MG1 is located on the right side of the BS in the orange shaded area, while MG2 is located to the right side of the BS in the green shaded area. It is assumed that all wireless links exhibit flat fading channel which means that the fading coefficients, channel gains, remain constant during one TS, but change

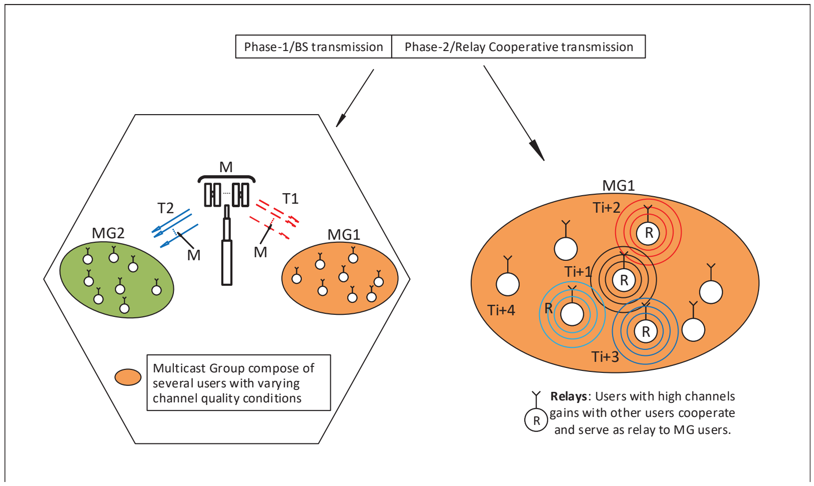

Figure 3.1: BS serving two MGs with the assistance of $M-1$ MSs acting as relays.

independently from one slot to another. AF approach is adopted by relays since MSs are interested in the raw signals. It is assumed that each receiving node (relays nodes and MSs) knows its receiving instantaneous CSI. Every relay must forward its CSI to the assisted MS. It is essential for the receiving MS to have the CSI of the forwarding relay to be able to decode the desired messages successfully. This step is usually done by using control pilots or frames between the relays and the receiver.

In the the proposed scheme, direct broadcasting to MSs in a MG doesn't take advantage of the M antennas at the BS to increase the multiplexing gain because the number of parallel streams that can be transmitted simultaneously is limited with the number of antennas at the receiving MS. In this case the number of antennas at the receiving MS is one, offering multiplexing gain of 1.

In this scheme, messages from the BS to the MG are delivered in two different stages. The first stage is allocated for the BS to multicast its messages. The second stage is used by the relays to forward the received signals to MSs within their designated MG. The next two subsections present the procedures followed in stage one and stage two of the proposed communication process/cycle.

3.1.1 Transmitter Stage

In this stage, the BS transmits M messages x_m^n through the M transmit antennas. The BS follows an orthogonal division protocol by allocating one time slot for each MG, as illustrated in Table 3.1 . This stage is comprised with N TS equal to the number of served MG.

Table 3.1: Stage one: Time slots allocated by the BS for every MS and its assisting relays.

Time slot	1	2	.	.	N
Served MG	MG^1	MG^2	.	.	MG^N

In TS n, the BS transmits M messages to the $n-th$ group. All MSs (Including assisting relays) within the $n-th$ MG can hear the transmitted signal with different scaled version based on their channel conditions. The variation in the received signals is due to variations of the CSI between each receiving MS and the BS antennas. The signal received by the $l-th$ MS in the $n-th$ MG can be expressed as follows:

$$Y_l^n(n) = \sum_{m=1}^{M} \sqrt{P} g_{lm}^n(n) x_m^n(n) + w_l^n(n) \tag{3.1}$$

where $\sqrt{P}$ represents the transmit power from the BS and g_{lm}^n represents the channel gain between the BS $m-th$ antenna and the $l-th$ receiving MS over the $n-th$ time slot. The notation $x_m^n(n)$ represents the message intended to be transmitted from BS antenna m to the n^{th} MG. The term w_l^n is the AWGN at the $l-th$ receiving node in the $n-th$ MG.

In equation 3.1, the time index n also identifies the MG. The indices l identifying the receiving MSs in the first stage are allocated such that the first L indices identify MSs actively receiving and recovering MG data in the second stage; the indices higher than L identify MSs that are serving as relays and initially (in this section) do not recover the MG data in the second stage. The sole purpose of MSs acting as relays is to forward signals enabling data recovery for the first L active MSs. It should be noted that MSs in each MG are concerned (active) only with the signal transmitted within their designated time slot. They remain inactive in any time slot allocated for other MG, and their received signals in such time slots are assumed to be zero.

3.1.2 Relay Stage

At the end of stage one, the BS turns off its transmission and remains inactive. The second stage begins then by splitting the entire network into N subnetworks corresponding to the N MGs. All MGs, in this stage, work concurrently and independently, using the same radio resource (frequency) over a period of $M-1$ TS, as illustrated in Table 3.2.

For the n^{th} MG, the generic MS is identified as MS_l^n and the MS serving as relay is identified as MSR_p^n. Every MG has L generic MSs and $M-1$ Relays. Within one MG, and in their designated TSs, the cooperating relays transmit in sequence the signals received in the first stage. The overlapping of transmissions in this stage is allowed because MGs and their corresponding MSs are assumed by sufficiently separated in space to allow frequency reuse. The index p of the transmitting relay MSR_p^n identifies also the time slot in stage two (p). It can be noticed that this stage is compromised with N TSs because of the overlapping in transmission between all MGs. By the end of stage two, the full transmission cycle is compromised $N+M-1$ TSs

3.2 Data Recovery

The AWGN is assumed to be small, so that neglecting it would not affect the analysis of the system. In fact, the MAI in the second stage dominates the SINR, and as a

Table 3.2: Stage two: Concurrent transmissions from relays to the desired MSs.

Time slot	N+1	N+2	.	.	N+M-1
from $\rightarrow$ to	$MS_1^1 \rightarrow MS_0^1$	$MS_2^1 \rightarrow MS_0^1$	.	.	$MS_{M-1}^1 \rightarrow MS_0^1$
	$MS_1^2 \rightarrow MS_0^2$	$MS_2^2 \rightarrow MS_0^2$	.	.	$MS_{M-1}^2 \rightarrow MS_0^2$
	.	.	.	.	.
	.	.	.	.	.
	$MS_1^N \rightarrow MS_0^N$	$MS_2^N \rightarrow MS_0^N$	.	.	$MS_{M-1}^N \rightarrow MS_0^N$

result, if not mitigated efficiently, may limit the performance of the system. Thus, the AWGN term will be omitted in this section derivations, though the AWGN impact is examined in the simulation section.

In the first stage of the proposed scheme, the BS allocates one TS per MG. The received signal at MS_l^n over the time slot n is given by equation 3.1. In the second stage, relay nodes starts working concurrently and independently. The signal received at MS_l^n over the time slot p is given by:

$$Z_p^{n,l} = d_p^n h_{lp}^n Y_l^n + \sum_{j=1, j\neq p}^{N} I_{jp}^n \tag{3.2}$$

where d_p^n is the amplifying gain at the relay (MSR_p^n), and h_{lp}^n is the channel gain between MSR_p^n and l^th user in n^th MG (MSR_l^n). The signal Y_l^n is the signal of interest which was received by MSR_p^n in the first stage as in equation 3.1 and is forwarded to the MSs in MG.

For simplicity, let $a_{lm}^n = d_l^n h_{0l}^n \sqrt{P} h_{lm}^n$. The $M-1$ signal received at MSR_l^n by the end of stage two can be represented by:

$$\begin{aligned} Z_1^{n,l} &= \sum_{m=1}^{M} a_{1m}^n x_m^n + \sum_j I_{j1}^n \\ &\vdots \\ Z_{M-1}^{n,l} &= \sum_{m=1}^{M} a_{M-1m}^n x_m^n + \sum_j I_{jM-1}^n \end{aligned} \tag{3.3}$$

In addition to the received signals in stage two, MS_l^n has also received a signal directly from the BS as given in equation 3.1 which can be expressed as (when disregarding AWGN):

$$Z_0^n = Y_0^n = \sum_{m=1}^{M} a_{0m}^n x_m^n \tag{3.4}$$

where $a_{0m}^n = \sqrt{P} h_{lm}^n$. Thus, by the end of this stage, every receiver MS_l^n, where $n \in \{1, \cdots, N\}$ and $l \in \{1, \cdots, L\}$, in each MG has to solve a system of M equations and M unknowns as shown by equation 3.3 and equation 3.4 in order to decode the desired M messages. The total transmitted messages to all N MGs by the BS is equal to MN. This is accomplished with the support of the designated relays. The aggregate efficiency which is measured by the total number of messages delivered per TS can be calculated by:

$$\eta_{designated} MSR = L.\frac{MN}{M+N-1} \tag{3.5}$$

where the $M + N - 1$ is the denominator represents the duration of full cycle of transmission. The actual savings in the proposed system depend on the channel conditions between BS and MG's users and relay nodes to allow reliable recovery of data at the maximized number of users in each MG. ZF precessing technique is adopted to decode the transmitted messages. However, ZF processing sufferers from noise enhancement problems and may limit the system capacity as presented in the simulation section.

In order to the solve the system of equations at every MS, it is assumed that each receiving side (relays nodes and active MSs) knows the receiving instantaneous CSI. Global availability of the CSI is not required, but every relay node must forward its CSI (g_{lm}^n) to the assisted MG. The virtual MIMO channel gain matrices created by equations 3.4 and 3.3 is defined by the CSI available at each receiving MS which is essential to decode the desired messages.

3.3 Relay Selection and Interference Management

In this section, we examine in detail the operation of the transmission scheme developed in the previous section. Alternative operations of relays will be considered.

We assumed initially, in section 3.2, that the assisting relays are not concerned with decoding all messages and their role was only to support the MG's users. The relays (MSR_p^n) supporting group n were receiving the signal in stage one in the designated n^{th} TS along with MSs in MG n. Then, in stage two, relays were forwarding the received signals in their designated TS p ($p \in \{1, \cdots, M-1\}$); they didn't have to hear the signal transmitted by other assisting relays (MSR_p^n $(p \neq k)$) because they were not interested in the the transmitted data. Even if a specific relay received signals from other $M-2$ assisting relays, the specific relay would only have $M-1$ copies (including the one received in the first stage). This means that it could not recover the M transmitted messages as it requires M equations to recover M unknowns. The issue with this scheme is that when MSR_n^p is transmitting, it cannot add one linear relation into its own system of equations. In order for these relays to be able to decode the transmitted messages, additional resources must be added. Our proposed solution in this section is to add additional time slot in this stage so that relays can have a number of equations equal to the number of unknowns. Therefore, we now have M MSR_p^n ($p \in \{1, \cdots, M\}$ allowing $L+M$ MSs (in each MG) to actively recover data designated for the n^{th} MG. With the addition of one TS this problem is rectified for all MSRs. This additional TS will create the redundancy that could result in more reliable data recovery but it may also reduce the throughput by increasing the time of each full transmission cycle. In this scheme, when relays are selected among active users, the M assisting relays do not have the overdetermined system of equation when recovering data. However, because they are selected to act as relays, they probably already have the favorable channel conditions.

The aggregate data rate of the proposed schemes is considered. In the proposed scheme, in this section, the total number of messages transmitted over $N+M$ time slots in full transmission cycle to all N MGs is equal to MN and in one MG there is now $L+M$ active MSs. This is achieved by the aid of M MSs acting as relays and actively recovering the data transmitted by the BS. The aggregate efficiency of this scheme is given by:

$$\eta_{designated} MSR = (L+M).\frac{MN}{M+N} \tag{3.6}$$

where $(L+M)$ represents here the total number of MSs in each MG. Comparing the aggregate efficiency of this scheme with that in the dedicated relays calculated by equation 3.5 in case of $M = 4$ and $N = 7$ reveals that the scheme with dedicated relays is more efficient for bigger MGs ($L \geq 35$).

Once the strategy for working with dedicated or active relays is decided; the relay selection could then be pursued. The strategy for working with either dedicated or active relays is usually based on the number of users per MG or the channel conditions between the active MSs and the BS. Because the transmit power of the BS, and the orthogonal TDMA transmission scheme for not causing any MAI, it can be safely assumed that all the M MISO signals received in first stage ,as in equation 3.1, are potentially useful for re-broadcasting in the relay stage.

The relay selection criterion in this scheme is dictated by power saving and MAI control requirements in the relay stage. So, based on the available CSI for a given MG, the BS selects potential relays with the strongest channel conditions that meet threshold requirements for stage two as follows:

$$\max\{\min |h_{lp}^{n}|^{2}\} \forall l \in \{1, \cdots, L\} \text{ and } |h_{lp}^{n}|^{2} \geq \gamma^{th} \tag{3.7}$$

To address the issues of virtual MIMO channel gain matrix invertibility, one option could be to reduce the number of transmitted messages from the BS. In specific, based on the channel conditions between the BS and the MG's users, a subset of M antennas at the BS will be used in stage one that offers better chances of inverting virtual channel gain matrix at MSs. However, by lowering the number of used antennas, the aggregate data rate will be reduced as measured in equations 3.5 and 3.6) but the latter do not account for the system reliability as elaborated in the simulation section.

Finally, the following observations about the MAI are in place. The power at the relays (MSR_p^n) should be carefully allocated ensuring that all the active MSs within a MG can receive signals at acceptable level of SNR. However, the allocated power used by MSR_p^n may contribute to MAI if it is allocated carefully. Specifically, in stage two, MSR_p^n forwards the signals received in stage one to all MSs within a MG concurrently, so controlling interference represented by the second term in equation 3.3 becomes necessity. The option of transmitting by a higher power by relays in the second stage may not be productive because of the "cocktail party effect" created by the simultaneous transmission by relays. Here, we target an acceptable level of SINR for each MG which means that power used by relays has to be adjusted so that all MGs have an acceptable SINR. To efficiently mitigate the impact of the MAI interference between different MGs in stage two, the inter distances between MGs have to be sufficiently large so that the leakage of the signals between groups is negligible.

The clustering concept of cellular systems can be adopted where the cell corresponds to an area of MG, hence, a larger number of MGs that are not necessarily separated sufficiently in space can be accommodated. The solution to control MAI would be to serve MGs in the same type of cells(one out of three in the 3-cell reuse pattern) in one communication cycle which allows for increased distance separation between active MGs and would extend the time to service all MGs in the system. The reuse pattern, as in a cellular system, is determined based on acceptable SINR levels between MGs versus the number of MGs in the system per single cycle. a Similar approach of separating the MGs in time that otherwise would contribute to increasing levels of MAI could be adopted in case the MGs overlap in space.

3.4 Result Analysis

In this section, the proposed two-stage relay-assisted multicasts is evaluated in terms of aggregate data rate and capacity. Three cooperative transmission scenarios are considered: (i) operation of the system with $M - 1$ dedicated relays; (ii) operation of the proposed system with active user/relays decoding the data designated for a specific MG; and (iii) operation of the system with dedicated relays and with a reduced number of active antennas. In this section, these scenarios are referred to as Scheme 1(dedicated), scheme 2 (4 antennas), and scheme 3 (3 atennas), respectively.

A cooperative transmission system consisting of (i) the BS with maximum number of antennas $M = 4$; (ii) N=2 MGs and (iii) number of MSs in every MG $L = 40$. It is assumed that MGs are far apart and clustering has been used with high reuse factor, so that the MAI is negligible. The MSs within the MG are distributed uniformly according to the Poisson point process. The channel between the BS antennas and MSs experience independent flat Rayleigh fading. It is assumed that the average SNR at all the MSs when receiving from the BSs in the transmit stage is the same, because of distances from the BS to MGs being much larger than the distances between MSs within the MGs. Relay selection is based on the criterion equation 3.7. In our simulation, we worked with different distances between the $MSRs$ and active MSs by working with the deterministic power attenuation in free space. Moreover, the same average of AWGN is assumed for MSs in stage two.

Figure 3.2 shows the minimum achievable capacity for all three considered schemes with SNR at MSs. It can be noticed that, in the presented network setup, scheme 1 performs better than the other two schemes. The full transmission cycle of scheme 1 and 3 consist of 5 TSs. In scheme 2, six TSs are needed to transmit 4 messages

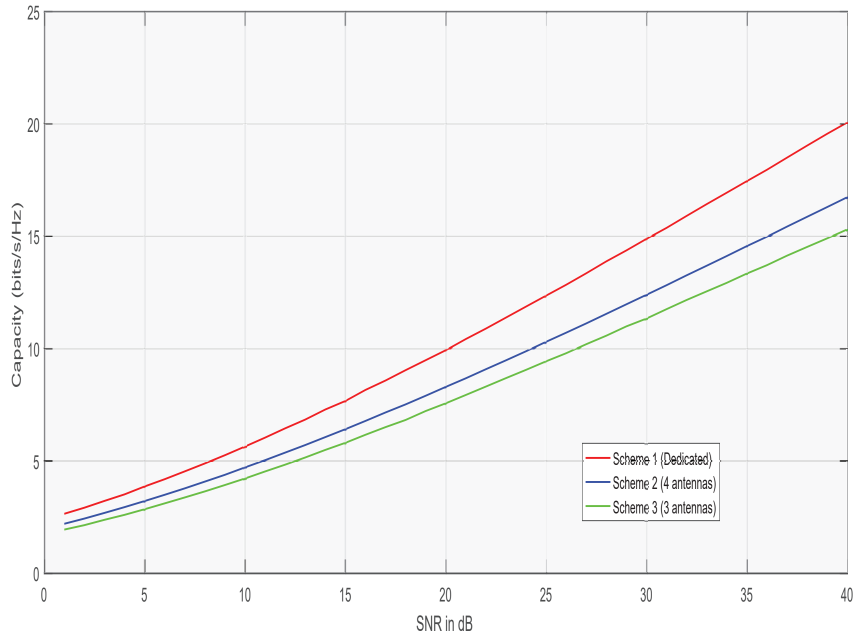

Figure 3.2: Capacity comparison between different schemes.

which means that each MG receives $\frac{4}{6}$ messages per TS. Scheme 1 and 3 can achieve $\frac{4}{5}$ and $\frac{3}{5}$ messages per TS, respectively. Scheme 3 might have worst performance in terms of capacity among the considered scheme2, however, it performs better in cases when the channel conditions between the BS and MG's users are not good. So, the advantage of adopting scheme 3 is that it is easier to find MSs that can serve as relays than in Schemes 1 and 2 as the optimization space is smaller.

In figure 3.3, the performance of the proposed schemes is evaluated in terms of aggregate data rate. The number of users in the system plays an important factor on the choice of the optimum strategy maximizing the data rate. When the number of users in the system is less than 12, scheme 1 has the worst performance among considered schemes, where in this region both schemes 2 and 3 provide higher aggregate data rate. When the number of users is greater than 18, scheme 3 outperforms all considered schemes.

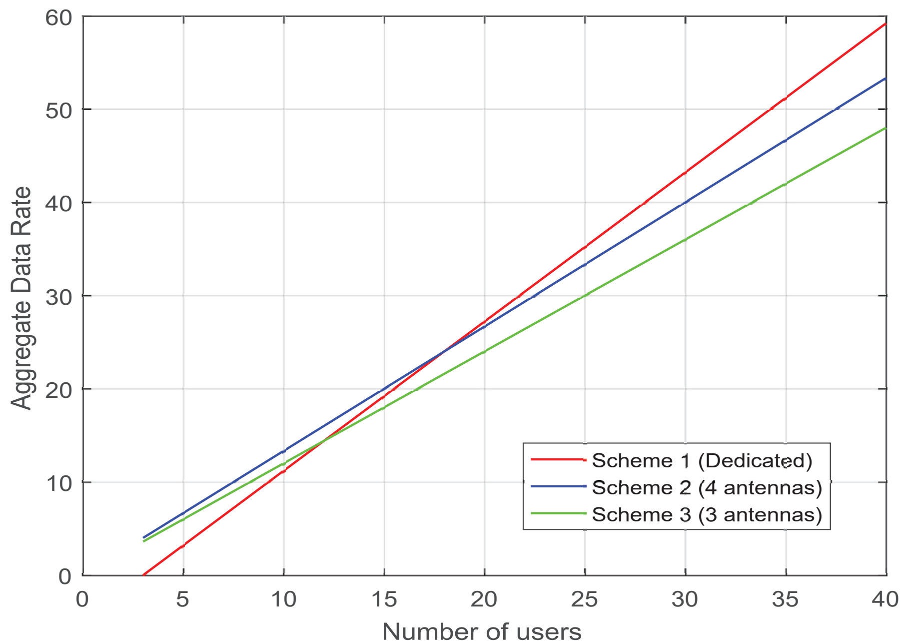

Figure 3.3: Aggregate data rate vs number of MSs in the MG.

3.5 Conclusion

In this chapter, we considered a two-stage virtual MIMO scheme to improve the data rates in wireless multicast networks. In our presented work, the M-antenna BS sends to single-antenna users within different MGs. The relays aid each MS to recover the transmitted messages by using AF technique so that a linear system of equations can be constructed at each user to decoded the M messages. Relays communicating in the same TS, through time scheduling, are spaced apart from each other to minimize the effect of MAI. Considering the aggregate data rate and capacity as performance measures, the proposed scheme improved the both aggregate data rate and capacity of the system. The proposed scheme benefits from the broadcast characteristic of the wireless medium, reusing access to the channel and opportunistic listening.

Chapter 4

Pattern-Based Multiple Access for Multigroup Multicast Systems

In this chapter, a pattern-based multicast transmission scheme is proposed for a multigroup environment. Traditionally, PDMA has been used to map the data of different users directly to radio resources, such as code, power, time, frequency, and spatial resources. However, in this chapter, a scheme is developed to assign unique patterns or signatures to MGs sharing the same radio resources. In the allocation of radio resources to a MG, the CSI is considered for all users within the MG, so as to optimize the fairness parameters representing the achievable data rates of users within each MG.

In section 4.1, an overview of PDMA and the system model is provided. This section presents the principles of assigning the patterns for each MG, as well as the principles of the transmitting and receiving processes. Fairness considerations are developed in section 4.2. Section 4.3 presents and discusses the simulation results, and section 4.4 provides the conclusions for the chapter.

4.1 System Model

This section considers a multigroup multicast system consisting of a BS and M MGs in a downlink transmission. Each MG has K users. The BS simultaneously multicasts the information to the MGs over L resource elements (REs), with a unique signature (pattern) for each MG, based on the CSI of users within the MG. It is assumed that CSI information is available at the BS for all users. It is also assumed that each transceiver is equipped with a single antenna. In subsection 4.1.1, a brief overview of PDMA pattern assigning and PDMA pattern matrix construction principles is presented. Subsection 4.1.2 describes the transmitting and receiving process.

4.1.1 Pattern Matrix Design

For each MG indexed by $m \in \{1, \cdots, M\}$, the BS maps the transmitted data onto the L REs by using a distinguishing PDMA pattern vector $\mathbf{v}_m$ of size $L \times 1$. The elements of $\mathbf{v}_m$ are either "0" or "1", where "1" means that the MG data are mapped to the corresponding RE, and "0" means they are not.

The PDMA matrix, $\mathbf{P}$, is generated by concatenating the M pattern vectors $\mathbf{v}_m$ for each MG to form an $L \times M$ matrix. An example of a pattern matrix, in the case of $L = 3$ REs and $M = 6$ MGs, can be expressed as:

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix} \tag{4.1}$$

In equation 4.1, it can be seen that several MGs share the same RE. This is because the positions of the MGs in the l-th row of $\mathbf{P}$ identify the set of MGs that map their data at the l-th RE, e.g., the frequency subcarrier, spatial stream. Conventionally, in

PDMA, the overload factor is defined as the ratio between the number of users and the number of available REs in the system. This study defines an overload factor of $\frac{M}{L}$, which reflects the multiplexing benefits of PDMA over orthogonal schemes, which use orthogonal REs with one-to-one mapping between users and REs. For $M = 6$ and $L = 3$, as in the case analyzed here, the overload factor is equal to 200%, which means that PDMA supports twice as many users as OMA. In addition to the PDMA implementation, the MGs are also separated in terms of the code domain resource. Specifically, within one RE, substreams from multiple description coding (MDC) for each MG are encoded by using a distinguishing code.

Once $\mathbf{v}_m$ has been determined for each MG, the BS begins the process of associating codes with each MG. An example of a pattern matrix following the assigment of codes is:

$$\mathbf{P_c} = \begin{bmatrix} C1 & C2 & C1 & C3 & 0 & C1 \\ C2 & 0 & C2 & C2 & C2 & C3 \\ 0 & C1 & C3 & C1 & C3 & 0 \end{bmatrix} \tag{4.2}$$

To resolve the non-orthogonal mapping of users and codewords, non-binary codewords are employed. This is equivalent to separating MGs within a given RE in both the code and power domains. In this work, based on the channel conditions of users within a MG, both orthogonal and non-orthogonal codes are used. Specifically, orthogonal codes are utilized when the MGs have similar channel conditions. Orthogonal variable spreading factor (OVSF) codes are adopted as orthogonal codes. Non-orthogonal codes are used when the channel conditions between the MGs are dissimilar. The adoption of non-orthogonal codes allows MGs to use the same code. Because the data are separated in the power domain, there are different power levels for every MG using the same code, for a given RE. The variation of channel conditions between MGs makes it possible to use SC (power scaling). In particular, before the data of two MGs are mixed, a power scaling factor is applied:

$$s_{BS} = P_1 s_1 + P_2 s_2 \tag{4.3}$$

where P_i is the power allocated to MG i. The total transmit power of the BS is P_T, where $P_T = P_1 + P_2$. Users in the MG with good channel conditions have to use SIC to recover the data of interest, whereas users with poor channel conditions must decode the data of interest in the presence of interference caused by data transmitted to other groups.

4.1.2 Transmitting and Receiving

On the transmitter side (BS), the data bits x_m for the m-th MG are encoded by using a unique code, and are then arranged in active REs in a column vector of size $L \times 1$. This creates a signal $\mathbf{s}_m = \mathbf{x}_m \bullet \mathbf{P}^c_m$, where $\mathbf{P}^c_m$ is the m-th column of $\mathbf{P}_c$ and $\bullet$ indicates the element-wise product of two matrices (columns). Thus, the signal $\mathbf{S}$ of size $L \times 1$ transmitted from the BS to all M MGs for the duration of codes Cs can be expressed as follows:

$$\mathbf{S} = \sum_{m=1}^{M} \mathbf{s}_m \tag{4.4}$$

On the receiver side, the signal received at user k in MG m can be expressed as:

$$\mathbf{y}^k_m = \mathbf{h}^k_m \bullet \mathbf{S} + n \tag{4.5}$$

where $\mathbf{h}^k_m$ denotes the channel gain column between user k in MG m and the BS, and the individual entry $\mathbf{h}^k_m$ captures the scalar channel gain of the particular RE (subcarrier). The scalar AWGN is represented by n, and the superimposed signal transmitted from the BS is represented by S.

The deterministic path loss of the signal as a function of distance determines the average channel conditions across all components in $\mathbf{h}_m^k$. In the simulation section of this study, this is reflected by working with a power decay with a distance r in free space, given by $\frac{1}{r^2}$. It is also assumed that the i.i.d. block Rayleigh fading in $\mathbf{h}_m^k$ can capture micro-scale signal variations.

In accordance with spread spectrum decoding principles, each user in a MG uses a unique pattern and code to recover the data of interest, as follows:

$$\hat{\mathbf{x}}_m = \mathbf{y}_m^k \bullet \mathbf{P}_m^c \tag{4.6}$$

In conventional PDMA with unicast-type traffic to individual users, the pattern assigned for a user depends mainly on the channel conditions of the users. However, in the present setup, users correspond to MGs. In multigroup multicast systems, fairness among users or groups plays an important role in resource allocation. Hence, it is crucial to consider both inter- and intra-group fairness when assigning codes for each MG. In the next section, suitable parameters are developed to measure both inter- and intra-group fairness, to maintain the fairness between users in a MG at an acceptable level while maximizing fairness among the MGs.

4.2 Fairness Considerations

In a multicast transmission environment, both fairness and data rates should be considered in resource allocation. However, balancing fairness and data rates poses a serious challenge due to the variations in channel conditions between users within the same MG. Here, this study focuses on (i) maximizing the intra-group fairness,(ii) maintaining an acceptable level of inter-group fairness, and (iii) maintaining acceptable data rates for each MG. First, various fairness indices are introduced to optimize the inter- and intra-group fairness. Then resource decisions are described for allocation of the PDMA patterns.

The intra-group fairness index (FI) is first defined for user k in MG m as follows:

$$FI_k^m = \frac{C_A}{C_{MAX}} \tag{4.7}$$

where C_A is the data rate assigned by the BS for MG m, and C_{MAX} is the maximum data rate achievable by user k in MG m. Thus, the optimum fairness and data rate utilization are achieved when $FI = 1$. This means that the MG-assigned data rate is equal to the maximum data rate achievable by all users within the MG. The objective of the assignment of the data rate C_A^m is to ensure that $C_A^m \leq C_{MAX}^k$. If $C_A^m > C_{MAX}^k$, the user cannot recover the data, hence, $FI_k^m = 0$.

The average of the FI for users within the m-th MG is referred to as the group fairness index (GFI^m), which is calculated as:

$$GFI^m = \frac{1}{K}\sum_{k=1}^{K} FI_k^m \tag{4.8}$$

Maximum satisfaction for the m-th group is achieved when GFI^m is equal to one. However, in a multigroup environment, it is essential to consider inter-group fairness to ensure that all MGs have an acceptable QoS.

The inter-group fairness index is measured based on the assigned data rate for each MG. The assigned data rate for MG m (R_m) can be calculated as:

$$R_m = \sum_{l=1}^{L} R_l^C U_l \tag{4.9}$$

where R_l^C is the spreading code rate assigned to the m-th MG on the l-th RE, and U_l (with values of 0 or 1) is an indicator of whether the RE is being utilized by this MG. The group fairness index for the m-th MG in the system (IGF^m) can be calculated as:

$$IGF_m = \frac{R_m}{\sum_{m=1}^{M} R_m} \tag{4.10}$$

From equation 4.10, the inter-group fairness among groups can be obtained as follows:

$$IGF_m^n = \frac{IGF_m}{IGF_n} \tag{4.11}$$

From equation 4.11, it can be seen that the maximum inter-group fairness between MGs can be achieved when $IGF_m^n = 1$. A higher fairness index value means that group m has a higher data rate and group n has a lower rate.

The proposed scheme aims to select RE patterns and code rates for each MG to maximize inter-group fairness based on equation 4.11, while achieving an acceptable level of intra-group fairness among users, based on equation 4.8. This should lead to better performance and greater satisfaction for users and MGs. For example, if the channel conditions between the users within each MG are similar, orthogonal codes are used to serve MGs in the same time and frequency domains, which leads to better inter-group fairness at a small cost in terms of data rate. If MGs have a high variation in channel conditions, multi-level spreading codes are used to differentiate the MG data so that the data are separated in the power domain. In this case, users need to use SIC to recover the data of interest for their corresponding MGs. Thus, it can be seen that the use of multilevel codes improves both the data rate and inter-group fairness.

For resource allocation, first the channel gains for MG m are represented by a single column vector:

$$\mathbf{h}_m^{\min} = \min_{\text{column-wise}} [\mathbf{h}_m^1 | \mathbf{h}_m^2 | \cdots | \mathbf{h}_m^K] \tag{4.12}$$

where for the CMS scheme, multiple minimizations are implemented across all the rows of concatenated individual channel gains to capture the worst user channel conditions on REs within the group. For the AVG scheme, the average of concatenated individual channel gains is calculated for each MG. Based on the targeted SNR, the pattern vectors $\mathbf{v}_m$ for each MG are determined. Then, to equalize fairness for each group in the system (as calculated in equation 4.10) and intra-group fairness (as calculated in equation 4.8), different codes and corresponding data rates are allocated to the MGs on their assigned REs. The final step is an iterative process that involves optimizing equation 4.11 and then checking for acceptable performance by using equation 4.8.

4.3 Numerical Results

This section provides a numerical performance evaluation for the proposed PDMA scheme in a multigroup multicast environment by using Monte-Carlo simulations. The system performance is evaluated by averaging the aggregate data for each MG, and inter-group fairness between MGs over 10^5 independent channel realizations for a given number of users in each MG. The users are distributed within a cell by using a Poisson point process.

In the simulation, two MGs are considered ($M = 2$) with a variable number of users, ranging from $K = 10$ to $K = 100$, distributed within a cell with radius $100m$ with pedestrian user mobility. The number of REs is $L = 4$. In the evaluation, the performance of the proposed scheme is compared with that of the CMS and AVG schemes.

The first evaluation is in terms of the aggregate data rate, which is the sum of data rates obtained by all users in the network, defined as:

$$ADR = \sum_{m=1}^{M} \sum_{k=1}^{K} R_k^m \tag{4.13}$$

where R_k^m is the data rate for user k in group m.

Figure 4.1 shows the aggregate data rate for the proposed PDMA scheme and the CMS scheme, with various group sizes, where the number of users K per group ranges from 10 to 100. From Figure 4.13, it can be seen that the CMS scheme and the

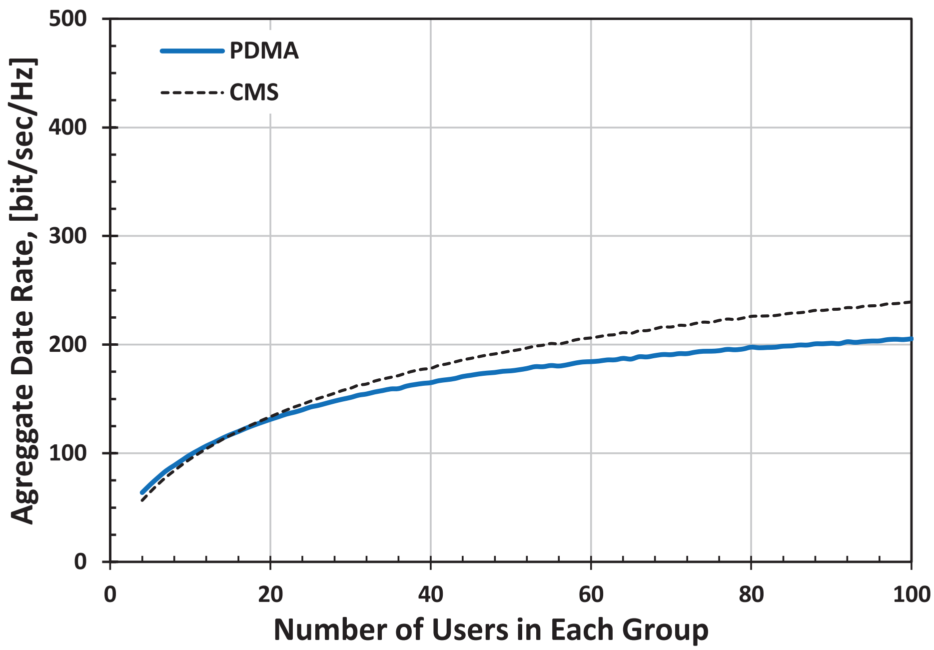

Figure 4.1: Aggregate data rate for the entire system.

proposed PDMA scheme have a comparable performance when the number of users in a MG is less than 20. The CMS scheme gradually begins to perform better than the PDMA scheme when the number of users is more than 20. When $K = 100$, the CMS scheme outperforms the PDMA scheme by 10% in terms of the aggregate data rate.

In multigroup multicast transmission systems, the main objective is to maximize inter-group fairness while maintaining an acceptable level of intra-group fairness. This is achieved by using the fairness indices introduced in the previous section to ensure an acceptable level of fairness in the system. The fairness for each MG in the system is measured by equation 4.10, where IGF_m represents the fairness index of group m. For example, if IGF_m is equal to 0.25, assuming that the system has only two MGs (as in these simulations), the received aggregate data rate of group m is 25% of the total rate, and the received aggregate data rate of the other group is 75% of the total rate.

Figure 4.2 provides a comparison between the CMS, PDMA, and AVG schemes in terms of inter-group fairness indices, with different numbers of REs.

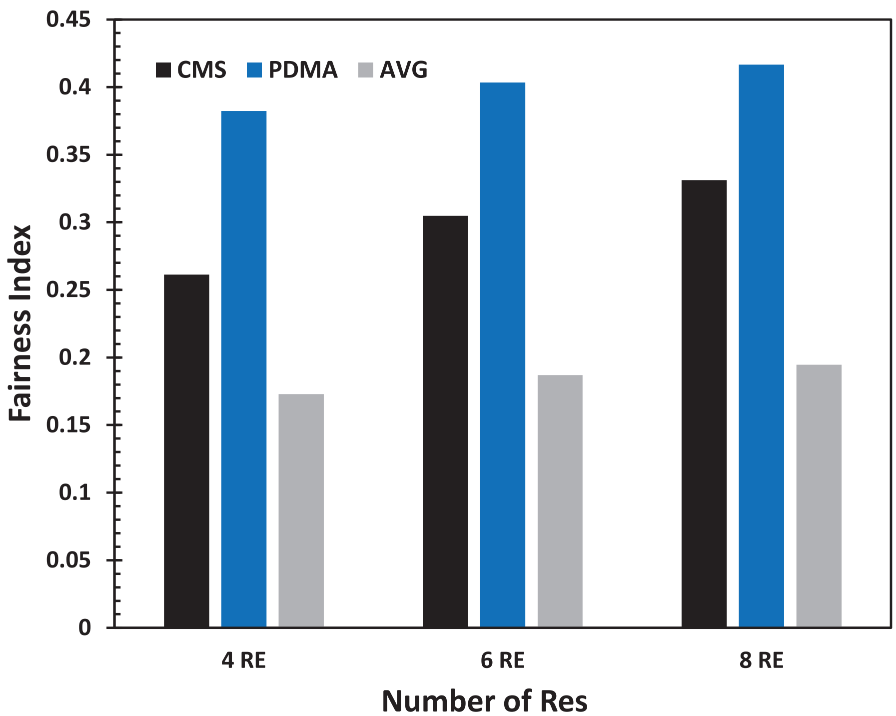

Figure 4.2: Fairness index comparison between CMS, PDMA, and AVG schemes.

When the number of available REs is 4, IGF_1 is equal to 0.25 when CMS is implemented, which means that group 1 can receive only 25% of the total aggregate data rate. In contrast, group 2 can receive 75% of the total aggregate data rate ($IGF_2 = 75\%$), which means that IGF_2^1 is equal to 33.33%. IGF_1 is equal to 0.39 when PDMA is implemented, which means that group 1 can receive only 39% of the total aggregate data rate, while group 2 can receive 61% of the total aggregate data rate, which means that IGF_2^1 is equal to 63.9%. The AVG scheme performs poorly in terms of inter-group fairness, with IGF_2^1 equal to 20%. In a comparison of IGF_2^1 for all three schemes, it can be seen that the PDMA scheme outperforms the CMS scheme by 30.57% and outperforms the AVG scheme by 43.9% in terms of inter-group fairness.

Inter-group fairness improves as the number of REs increases. For example, when the number of REs is 6, the IGF_1 achieved by the PDMA scheme is 0.40, whereas the IGF_1 achieved by the CMS scheme is 0.30. An IGF_1 of 0.3 results in an IGF_2^1 of 42.8%, while an IGF_1 of 0.4 results in an IGF_2^1 of 66.6%. The PDMA scheme still outperforms the CMS scheme by 23.6% and the AVG scheme by 45.5% in terms of inter-group fairness. Increasing the number of REs to 8 leads to an IGF_1 equal to 0.33 for the CMS scheme, and an IGF_1 equal to 0.42 for the PDMA scheme. When 8 REs are used, the PDMA scheme outperforms the CMS and AVG schemes by 23.2% and 50%, respectively.

4.4 Conclusions

This chapter presents a PDMA-type scheme for a multigroup multicast transmission system. For the scheme developed here, the PDMA pattern assignment is analyzed to maximize inter-group fairness while providing MGs with acceptable aggregate data rates. A consideration of aggregate data rate and inter-group fairness index performance measures shows that the PDMA scheme represents an improvement in terms of fairness, and performs comparably to the CMS scheme in terms of aggregate data rates.

Chapter 5

Conclusions and Future Work

The primary objective of the research in this dissertation was to design and implement radio resource management algorithms for downlink multicast systems by deploying layered transmission protocols at the BS, with the aim of improving both the system capacity and fairness. In this chapter, the dissertation contributions and conclusions are presented in section 5.1, followed in section 5.2 by suggestions for possible future work based on the results of this research.

5.1 Dissertation Summary

One of the fundamental drivers of wireless communications and network research is the goal of increasing communication rates over limited system bandwidth. The broadcast nature of the wireless channel makes it suitable for multicasting applications, since a single transmission can be received simultaneously by a number of users. The field of multicast transmissions considered in this dissertation is a special case of broadcast transmissions. It is emerging as a promising technology, offering significant improvement in spectral efficiency with the potential to extend the capabilities of next-generation communication systems. Layered transmissions are of particular interest, providing flexible use of the broadcast spectrum by transmitting several data streams over the same radio resources. The primary focus of this dissertation is the implementation of layered transmission protocols in a mutlicast system to improve network data rates and reliability.

The main contributions of this work are as follows:

1. A framework is provided for combining layered transmissions and successive relaying to improve throughput and data reliability for cell-edge users. A scheme using superposition coding at the BS with SIC for both relays and cell-edge users was implemented and a performance analysis was undertaken. Several schemes are developed which provide some improvement over schemes currently being used. In the decode-and-forward (DF) transmission schemes developed, transmissions from the BS and the relays are aligned in the time, power, and spatial domains to improve the system throughput and reliability. The approach developed here benefits from layered transmissions, time scheduling of transmissions, and spatial separation between relays, with cell-edge users and BSs reusing the same spectrum.

2. Three methods are developed to cancel the IRI efficiently in two-path successive relaying, based on iterative SIC and zero-forcing techniques. The first method is based on applying SIC at both relays and destination. It considers different power allocations to the transmitted layers, so that IRI is dealt with as a different layer (signal of interest). In the second method, the IRI is cancelled by using an additional spatial stream, where the relays and destination are equipped with two antennas. In the third approach, a generalized system is developed by using M multiple antennas at the relays and destination. The associated inter-relay channel provides a MIMO-related additional degree of freedom in comparison to conventional SISO relay systems, thus permitting the efficient removal of IRI.

3. A transmission scheduling system for downlink cellular networks which incorporates cooperative relaying and a virtual MIMO method is presented for a multi-group multicast MISO system. Cooperative relaying is used to recover independent messages for the desired users in a MG, by taking advantage of independent signal replicas and channel reuse at different locations, with reduced co-channel interference.

4. A PDMA-based multicast transmission system to allocate available radio resources fairly among users/groups is developed and analyzed. A joint design for both transmitter and receiver is considered. At the receiver, multiple groups are detected by using a successive interference cancellation (SIC) detection method. In the work developed here, several indices for both intra- and inter-group fairness are introduced, in order to maximize inter-group fairness among MGs while maintaining an acceptable level of intra-group fairness.

The contents of the individual chapters of this dissertation and the associated research results can be summarized as follows:

Chapter 2 proposes a layered transmission between the BS and the destination, with the aid of two HD relays (successive relaying) and superposition coding. In this two-hop decode-and-forward (DF) relaying system, because of the network topology, the broadcast spectrum is utilized simultaneously by the BS and one of the relays, causing what is known as inter-relay interference (IRI). In this chapter different signal processing schemes are developed that fully cancel the inter-relay interference at the relays by using superposition coding (SC) and MIMO. The first IRI mitigation method is based on a SIC scheme that considers different power profiles for the superimposed layers. Specifically, based on the channel gains between the BS and the destination, the transmissions from the BS and the relays are aligned in the time and power domains so as to optimize the system throughput and cancel the IRI. In the second method, IRI is cancelled by deploying additional antennas at the relays and destination. This offers an additional spatial stream that is exploited to cancel the IRI by using a zero-forcing processing technique. The third method of cancelling the IRI is a generalization of method two. This scheme benefits from using M multiple antennas at the relays and destination. The associated inter-relay channel provides a MIMO-related additional degree of freedom in comparison to conventional SISO relay systems and permits the efficient removal of IRI.

Chapter 3 deals with single-antenna users in a multi-group multicast cooperative network, where the BS is equipped with M antennas. In this network, the users within a MG are assisted by a group of single-antenna relays. The messages from the BS to the active users are delivered in two stages. In the first stage, the BS uses multiple antennas with time division multiplexing and transmits to different MGs in their allocated time slots. The number of messages sent in one time slot is equal to the number of transmit antennas at the BS. In the next stage, the multicast group of co-located users and the assigned relays form an independent network, where amplify-and-forward (AF) relays aid the recovery of the BS messages, with users solving a linear system of equations at the end of the stage. All these subnetworks utilize the available channel concurrently, producing an acceptable level of multiple access interference (MAI). The MAI is controlled by using a clustering technique tailored to the location of the MGs.

Chapter 4 proposes a pattern-based multicast transmissions in a multi-group environment, with fairness-aware resource allocation. A scheme is developed to assign unique patterns or signatures to MGs sharing the same radio resources, to permit a more efficient and flexible exploitation of multi-group diversity. For radio resources allocated to a MG, channel state information (CSI) for all users in the MG is used to optimize the fairness parameters which represent the achievable data rates for users within each MG. In the scheme developed, to differentiate signals from various MGs, resource blocks are considered that span multiple aspects of signal separation. The proposed scheme is developed to achieve an acceptable throughput for each multicast group while maximizing fairness among the groups.

5.2 Suggested Future Work

Many unresolved problems remain to be addressed with regard to layered transmissions in multicast networks. Some of these problems arise in connection with proposed new solutions that use cooperative communications, superposition coding, and non-orthogonal multiple access techniques, while other problems, such as channel estimation and synchronization, are generic to all signal processing algorithms related to relaying systems. In this dissertation, all algorithms and methods were first analyzed by using mathematical tools and then verified through computer simulations. The main challenge facing new theoretical results in wireless network research is to find a path to apply these designs in practical systems.

The following topics related to the present study are suggested for future work.

1. Effects of Imperfect Channel Information

In the research presented in this dissertation, it has been assumed that perfect CSI is available at one node at least in the network, when applying relay interference cancellation. However, the availability of perfect CSI may be difficult to realize in practice, since knowledge about the channel is obtained through statistical estimation and feedback between the nodes. The problem of imperfect CSI in different network settings has been extensively studied and many algorithms to address this issue have been proposed in the literature. Possible performance deterioration due to imperfect CSI in the algorithms developed could be investigated.

2. Synchronized Transmissions and Reception

An important challenge for IRI cancellation at the relays is the difficulty of synchronizing the arrival of the signals transmitted from the BS and the interference signals from the interfering relays. There are methods of achieving synchronous transmissions of the network nodes, however in the work developed in this dissertation the synchronous reception of signals from terminals at different locations is required. Because of the variability in the propagation delays of the signals, this may present a challenge, especially at high data rates. A potential solution is to deploy orthogonal frequency-division multiplexing (OFDM), which extends the duration of transmitted symbols and may help to alleviate the impact of signals arriving with different delays. This could be a subject of further study.

3. IRI Cancellation at Multiple-antenna Relays

In the parts of this dissertation that consider successive relaying, it is assumed that the relays and destination are always equipped with twice as many antennas as the BS. The analysis of networks where relays are equipped with same number of antennas, and of the impact of using multiple relays equipped with different numbers of antennas would be of interest, since this could contribute to making the IRI cancellation scheme more flexible to benefit from relay diversity.

4. Use of a PDMA-based System for Multicast Subgroup Formation

The sections of this dissertation which examine PDMA-based transmissions are based on serving multiple MGs. In the case of a single MG, an interesting solution is to split any MG into subgroups and to apply subgroup-based adaptive modulation and coding schemes, which permit a more efficient exploitation of multi-user diversity. Moreover, with regard to PDMA implementation, low complexity implementations and the impact of error cancellation on the overlaid layers should also be studied.

5. Improving SNIR for Single-Antenna Users

In the investigation of relay-assisted downlink transmissions to support increased data rates for single-antenna users in Chapter 3, one factor limiting the scheme performance is interference from simultaneously transmitting "adjacent" relays. However, successive interference cancellation in a comparable system in Chapter 2 achieves further improvement in bandwidth efficiency. The merging of both approaches to reduce the SNIR through spatial separation of interfering relays is a topic worthy of further investigation.

Bibliography

[1] M. M. da Silva, A. Correia, R. Dinis, N. Souto, and J. C. Silva, *Transmission Techniques for Emergent Multicast and Broadcast Systems*, 1st ed. Boston, MA, USA: Auerbach Publications, 2010.

[2] H. Won, H. Cai, D. Y. Eun, K. Guo, A. Netravali, I. Rhee, and K. Sabnani, "Multicast scheduling in cellular data networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 9, pp. 4540–4549, Sep. 2009.

[3] H. Wang, X. Chu, and H. Xiang, "Capacity maximization and transmission rates optimization for relay assisted wireless multicast," *IEEE Communications Letters*, vol. 16, no. 10, pp. 1592–1595, October 2012.

[4] H. Wang, X. Chu, and H. Xiang, "Capacity maximization and transmission rates optimization for relay assisted wireless multicast," *IEEE Communications Letters*, vol. 16, no. 10, pp. 1592–1595, October 2012.

[5] A. M. C. Correia, J. C. M. Silva, N. M. B. Souto, L. A. C. Silva, A. B. Boal, and A. B. Soares, "Multi-resolution broadcast/multicast systems for mbms," *IEEE Transactions on Broadcasting*, vol. 53, no. 1, pp. 224–234, March 2007.

[6] H. Haas and S. McLaughlin, *Next Generation Mobile Access Technologies: Implementing TDD*. New York, NY, USA: Cambridge University Press, 2008.

[7] H. C. Wu, C. Akamine, B. Rong, M. Velez, C. Wang, and J. Wang, "Point-to-multipoint communications and broadcasting in 5G," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 72–73, Mar 2018.

[8] L. Zhang, W. Li, Y. Wu, X. Wang, S. Park, H. M. Kim, J. Lee, P. Angueira, and J. Montalban, "Layered-division-multiplexing: Theory and practice," *IEEE Transactions on Broadcasting*, vol. 62, no. 1, pp. 216–232, March 2016.

[9] R. Zhang and L. Hanzo, "A unified treatment of superposition coding aided communications: theory and practice," *IEEE Commun. Surv. Tutor.*, vol. 13, no. 3, pp. 503–520, 2011.

[10] L. Wang, E. Sasoglu, B. Bandemer, and Y. Kim, "A comparison of superposition coding schemes," *CoRR*, vol. abs/1302.1258, 2013. [Online]. Available: http://arxiv.org/abs/1302.1258

[11] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (noma)," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sep. 2013, pp. 611–615.

[12] S. Timotheou and I. Krikidis, "Fairness for non-orthogonal multiple access in 5g systems," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1647–1651, Oct 2015.

[13] L. Dai, B. Wang, Y. Yuan, S. Han, C. l. I, and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, September 2015.

[14] A. Abrardo, G. Fodor, and B. Tola, "Network coding schemes for device-to-device communications based relaying for cellular coverage extension," in *2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2015, pp. 670–674.

[15] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, Dec 2004.

[16] V. Havary-Nassab, S. Shahbazpanahi, and A. Grami, "Optimal distributed beamforming for two-way relay networks," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1238–1250, March 2010.

[17] N. Zlatanov, V. Jamali, and R. Schober, "On the capacity of the two-hop half-duplex relay channel," in *2015 IEEE Global Communications Conference (GLOBECOM)*, Dec 2015, pp. 1–7.

[18] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. New York, NY, USA: Cambridge University Press, 2005.

[19] D. Bliss, K. Forsythe, and A. Chan, "MIMO wireless communication," *Linconlan Laboratory Journal*, vol. 15, no. 1, pp. 97–126, 2005.

[20] C. Lim, T. Yoo, B. Clerckx, B. Lee, and B. Shim, "Recent trend of multiuser MIMO in LTE-advanced," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 127–135, 2013.

[21] Z. Ding, S. Dasgupta, and R. Lopez, "Interference cancellation and blind equalization for linear multi-user systems," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process., ICASSP '00*, vol. 1, 2000, pp. 145–148.

[22] S. K. S. Yoshioka and F. Adachi, "Joint Tx/Rx signal processing for uplink distributed antenna network using single-carrier MU-MIMO," in *2015 10th Int. Conf. on Inform., Commun. and Signal Process. (ICICS)*, Dec 2015, pp. 1–5.

[23] X. Dai, Z. Zhang, B. Bai, S. Chen, and S. Sun, "Pattern division multiple access: A new multiple access technology for 5G," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 54–60, April 2018.

[24] R. Alsakarnah and J. Ilow, "Superposition coding in alternate DF relaying systems with virtual MIMO IRI cancellation," in *29th IEEE Annual International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2018, Bologna, Italy, September 9-12, 2018*, 2018, pp. 1–6.

[25] R. Alsakarnah, "Fairness-aware pattern-based multiple access for multi-group multicast systems," in *2018 11th International Symposium on Communication Systems, Networks Digital Signal Processing (CSNDSP)*, July 2018, pp. 1–5.

[26] R. Alsakarnah, F. Alhumaidi, and J. Ilow, "Increasing data rates in relay-assisted wireless multicast networks with single antenna receivers," in *2018 Wireless Days (WD)*, April 2018, pp. 202–206.

[27] R. Alsakarnah and J. Ilow, "Superposition coding in alternate DF relaying systems with inter-relay interference cancellation," in *13th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, WisaMob 2017, Rome, Italy, October 9-11, 2017*, 2017, pp. 104–109.

[28] J. Montalbán, L. Zhang, U. Gil, Y. Wu, I. Angulo, K. Salehian, S. Park, B. Rong, W. Li, H. M. Kim, P. Angueira, and M. Vélez, "Cloud transmission: System performance and application scenarios," *IEEE Transactions on Broadcasting*, vol. 60, no. 2, pp. 170–184, June 2014.

[29] S. M. R. Islam, N. Avazov, O. A. Dobre, and K. Kwak, "Power-domain non-orthogonal multiple access (noma) in 5g systems: Potentials and challenges," *IEEE Communications Surveys Tutorials*, vol. 19, no. 2, pp. 721–742, Secondquarter 2017.

[30] S. Vanka, S. Srinivasa, Z. Gong, P. Vizi, K. Stamatiou, and M. Haenggi, "Superposition coding strategies: Design and experimental evaluation," *IEEE Transactions on Wireless Communications*, vol. 11, no. 7, pp. 2628–2639, July 2012.

[31] J. Kim and I. Lee, "Capacity analysis of cooperative relaying systems using non-orthogonal multiple access," *IEEE Communications Letters*, vol. 19, no. 11, pp. 1949–1952, Nov 2015.

[32] L. Wang, E. Sasoglu, B. Bandemer, and Y. Kim, "A comparison of superposition coding schemes," in *2013 IEEE International Symposium on Information Theory*, July 2013, pp. 2970–2974.

[33] J. Kim and I. Lee, "Capacity analysis of cooperative relaying systems using non-orthogonal multiple access," *IEEE Communications Letters*, vol. 19, no. 11, pp. 1949–1952, Nov 2015.

[34] H.-W. Chang and C.-W. Yu, "Optimizing one-to-one data communication on double-loop networks," in *Circuits and Systems, 2004. Proceedings. The 2004 IEEE Asia-Pacific Conference on*, vol. 1, Dec 2004, pp. 473–476 vol.1.

[35] C. Huang and S. Cui, "On the alternative relaying gaussian diamond channel with conferencing links," *IEEE Transactions on Wireless Communications*, vol. 12, no. 2, pp. 758–768, February 2013.

[36] J. Pope and R. Simon, "Efficient one-to-many broadcasting for resource-constrained wireless networks," in *Local Computer Networks (LCN), 2015 IEEE 40th Conference on*, Oct 2015, pp. 518–525.

[37] P. Agashe, R. Rezaiifar, and P. Bender, "cdma2000/sup /spl reg// high rate broadcast packet data air interface design," *IEEE Communications Magazine*, vol. 42, no. 2, pp. 83–89, Feb 2004.

[38] K. Bakanoglu, W. Mingquan, L. Hang, and M. Saurabh, "Adaptive resource allocation in multicast ofdma systems," in *2010 IEEE Wireless Communication and Networking Conference*, April 2010, pp. 1–6.

[39] H. Won, H. Cai, D. Y. Eun, K. Guo, A. Netravali, I. Rhee, and K. Sabnani, "Multicast scheduling in cellular data networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 9, pp. 4540–4549, Sep. 2009.

[40] G. Araniti, M. Condoluci, L. Militano, and A. Iera, "Adaptive resource allocation to multicast services in lte systems," *IEEE Transactions on Broadcasting*, vol. 59, no. 4, pp. 658–664, Dec 2013.

[41] T. Han and N. Ansari, "Energy efficient wireless multicasting," *IEEE Communications Letters*, vol. 15, no. 6, pp. 620–622, June 2011.

[42] J. Montalban, P. Scopelliti, M. Fadda, E. Iradier, C. Desogus, P. Angueira, M. Murroni, and G. Araniti, "Multimedia multicast services in 5g networks: Subgrouping and non-orthogonal multiple access techniques," *IEEE Communications Magazine*, vol. 56, pp. 91–95, 2018.

[43] M. Condoluci, G. Araniti, A. Molinaro, and A. Iera, "Multicast resource allocation enhanced by channel state feedbacks for multiple scalable video coding streams in lte networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 2907–2921, May 2016.

[44] V. K. Goyal, "Multiple description coding: compression meets the network," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 74–93, Sep. 2001.

[45] Y. Jing, "A relay selection scheme for two-way amplify-and-forward relay networks," in *Int. Conf. on Wireless Commun. Signal Processing, WCSP*, Nov 2009, pp. 1–5.

[46] M. Król, Y. Ji, S. Yamada, C. Borcea, L. Zhong, and K. Takano, "Extending network coverage by using static and mobile relays during natural disasters," in *2016 30th International Conference on Advanced Information Networking and Applications Workshops (WAINA)*, March 2016, pp. 681–686.

[47] S. Zhang, X. G. Xia, and J. Wang, "Cooperative performance and diversity gain of wireless relay networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 9, pp. 1623–1632, October 2012.

[48] S. Zhao, M. Peng, and W. Wang, "An adaptive relay selection algorithm for optimizing diversity and multiplexing gains in wireless cooperative networks," in *2010 IEEE Wireless Communication and Networking Conference*, April 2010, pp. 1–6.

[49] D. Hwang, D. I. Kim, S. K. Choi, and T. J. Lee, "UE relaying cooperation over D2D uplink in heterogeneous cellular networks," *IEEE Transactions on Communications*, vol. 63, no. 12, pp. 4784–4796, Dec 2015.

[50] A. S. Ibrahim, A. K. Sadek, W. Su, and K. J. R. Liu, "Cooperative communications with relay-selection: when to cooperate and whom to cooperate with?" *IEEE Transactions on Wireless Communications*, vol. 7, no. 7, pp. 2814–2827, July 2008.

[51] H. J. Lee, S. H. Kang, and J. H. Kim, "A cooperative scanning mechanism for the mobile relay in the moving network environment," in *2010 7th IEEE Consumer Communications and Networking Conference*, Jan 2010, pp. 1–5.

[52] W. Sang, D. Shen, W. Ren, and X. Shuai, "A survey of capacity in cooperative relay networks," in *Mobile Congress (GMC), 2011 Global*, Oct 2011, pp. 1–8.

[53] Q. T. Vien, H. X. Nguyen, and W. Tu, "Optimal relay positioning for green wireless network-coded butterfly networks," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sept 2013, pp. 286–290.

[54] E. Yilmaz, R. Knopp, and D. Gesbert, "On the gains of fixed relays in cellular networks with intercell interference," in *2009 IEEE 10th Workshop on Signal Processing Advances in Wireless Communications*, June 2009, pp. 603–607.

[55] T. Kwon, "Design and implementation of a simulator based on a cross-layer protocol between mac and phy layers in a wibro compatible.ieee 802.16e ofdma system," *IEEE Communications Magazine*, vol. 43, no. 12, pp. 136–146, Dec 2005.

[56] H. Yomo and E. D. Carvalho, "A csi estimation method for wireless relay network," *IEEE Communications Letters*, vol. 11, no. 6, pp. 480–482, June 2007.

[57] B. Sunil and N. B. T, "Performance analysis and comparison of AF and DF relaying systems in rayleigh fading channel considering poisson interference field," in *Innovations in Information, Embedded and Communication Systems (ICIIECS), 2015 International Conference on*, March 2015, pp. 1–5.

[58] M. Yu, J. Li, and H. Sadjadpour, "Amplify-forward and decode-forward: the impact of location and capacity contour," in *MILCOM 2005 - 2005 IEEE Military Communications Conference*, Oct 2005, pp. 1609–1615 Vol. 3.

[59] B. D. Moubagou and Y. Y. Chang, "Performance of AF and DF relays by employing half duplex and full duplex mode," in *Measurement, Information and Control (ICMIC), 2013 International Conference on*, vol. 02, Aug 2013, pp. 1379–1383.

[60] O. Munoz, A. Agustin, and J. Vidal, "Cellular capacity gains of cooperative MIMO transmission in the downlink," in *Communications, 2004 International Zurich Seminar on*, 2004, pp. 22–26.

[61] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 379–389, February 2007.

[62] H. Wicaksana, S. H. Ting, C. K. Ho, W. H. Chin, and Y. L. Guan, "Af two-path half duplex relaying with inter-relay self interference cancellation: diversity

analysis and its improvement," *IEEE Transactions on Wireless Communications*, vol. 8, no. 9, pp. 4720–4729, September 2009.

[63] H. Park and J. Chun, "Inter-relay interference cancellation for AF MIMO two-path relay systems," in *Vehicular Technol. Conf. (VTC Fall)*, Sept 2012, pp. 1–5.

[64] K.-H. Park and M.-S. Alouini, "Alternate MIMO relaying with three AF relays using interference alignment," in *IEEE Int. Conf. on Commun. (ICC)*, June 2012, pp. 3526–3531.

[65] Z. Ding, I. Krikidis, B. Rong, J. S. Thompson, C. Wang, and S. Yang, "On combating the half-duplex constraint in modern cooperative networks: protocols and techniques," *IEEE Wireless Communications*, vol. 19, no. 6, pp. 20–27, December 2012.

[66] R. Zhang, "On achievable rates of two-path successive relaying," *IEEE Transactions on Communications*, vol. 57, no. 10, pp. 2914–2917, October 2009.

[67] W. Chen, K. B. Letaief, and Z. Cao, "Achieving spectral efficient cooperative diversity with network interference cancellation," in *2011 IEEE International Conference on Communications (ICC)*, June 2011, pp. 1–5.

[68] W. Hou, M. Jiang, and S. Feng, "Beamforming for inter-relay interference reduction in mimo-aided two-path successive relaying," *Wireless Communications and Mobile Computing*, vol. 16, no. 15, pp. 2274–2286, 2016.

[69] X. Dai, S. Chen, S. Sun, S. Kang, Y. Wang, Z. Shen, and J. Xu, "Successive interference cancelation amenable multiple access (sama) for future wireless communications," in *2014 IEEE International Conference on Communication Systems*, Nov 2014, pp. 222–226.

[70] S. Chen, B. Ren, Q. Gao, S. Kang, S. Sun, and K. Niu, "Pattern division multiple access-a novel nonorthogonal multiple access for fifth-generation radio networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3185–3196, April 2017.

[71] W. Tang, S. Kang, and B. Ren, "Performance analysis of cooperative pattern division multiple access (co-pdma) in uplink network," *IEEE Access*, vol. 5, pp. 3860–3868, 2017.