

ABSTRACT

Title of Dissertation: CROSS-LAYER DESIGN FOR
 COOPERATIVE COMMUNICATIONS
 AND NETWORKING

Ahmed K. Sadek, Doctor of Philosophy, 2007

Dissertation directed by: Professor K. J. Ray Liu
 Department of Electrical and Computer Engineering

Cooperative communications is a new communication paradigm in which different terminals in the wireless network share their antennas and resources for distributed transmission and processing. Recent studies have shown that cooperative communications can yield significant performance improvement due to spatial diversity gains. The theory of cooperative communications is however still immature to fully understand its broader impacts on the design of future wireless networks. This thesis contributes to the advancement of cooperative communications by developing and analyzing cooperation protocols at different network levels, with the goal to provide significant improvements in signal reliability, coverage area, network throughput, and energy efficiency with respect to other existing alternatives.

We first propose a family of cooperative protocols for multi-node cooperative communications. We demonstrate that full diversity gains is achieved, which yields

a significant improvement in the error performance. Based on the derived symbol-error-rate expressions, we characterize the optimal power allocation strategy among the relays and the source to further improve the performance of the system.

We develop distributed relay assignment protocols, and analyze their outage performance. We derive lower bounds on any relay-assignment scheme to benchmark the performance of our proposed schemes. We study the impact of our proposed protocols on increasing the coverage area of cellular networks without increasing the transmit power or adding extra base-stations.

We demonstrate that the gains promised by cooperation can be leveraged to the multiple-access layer. We propose the deployment of cognitive relays to utilize the periods of silence of the terminals to enable cooperation. This alleviates the spectral inefficiency problems inherent in conventional cooperation protocols. Our results reveal significant improvements in the maximum stable throughput region and delay performance of the network.

Finally, an analytical framework for studying the energy efficiency of cooperation in wireless networks is presented. This framework considers the overhead in the processing and receiving powers introduced by cooperation. The results characterize the regions where cooperation is more energy efficient than direct transmission. The results also provide guidelines for the design of power allocation strategies, relay-assignment algorithms and the selection of the optimal number of relays to help the source.

CROSS-LAYER DESIGN FOR COOPERATIVE
COMMUNICATIONS AND NETWORKING

by

Ahmed K. Sadek

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2007

Advisory Committee:

Professor K. J. Ray Liu, Chairman
Professor Anthony Ephremides
Professor Prakash Narayan
Professor Sennur Ulukus
Professor Amr Baz

©Copyright by
Ahmed K. Sadek
2007

PREFACE

All praise is due to Allah

“And of Knowledge, you (mankind) have been given only a little”,
translation of the meaning of Verse 85, Chapter 17, Quran.

DEDICATION

To my Parents and my beloved wife Dina

ACKNOWLEDGEMENTS

Completing my Ph.D. thesis is a milestone in my career that would not have been possible without the help and support of special individuals to whom I dedicate the following few words that can not capture my appreciation for them.

First of all, I would like to express my deep gratitude to my advisor, Prof. Ray Liu, who provided me the opportunity to be the researcher I am today. Dr. Liu has been my inspiration for doing exciting world-class research over the past few years. He was always available whenever I needed his guidance through problems in my research or even in my life. Most important of all, he made me and other members of our group feel as one family to which he always offered his countless efforts and dedication. Thank you Dr. Liu for offering me your friendship.

I would like to thank the other members in my dissertation committee. Special thanks are due to Prof. Anthony Ephremides for the the opportunity of doing a joint work with him, and for our several discussions that led to fruitful research. I would also like to thank Prof. Prakash Narayan not only for accepting to be in my thesis committee, but for his continuous support and kind guidance. I also want to thank Prof. Sennur Ulukus and Prof. Amr Baz for agreeing to serve in my thesis committee.

I would like to thank every member in our Signals and Information group for the wonderful time we spent together and for the very nice discussions we had during our weekly group meetings. I have learned a lot from each one of you, and

for that I am deeply grateful. I want to specially thank Dr. Weifeng Su and Karim Seddik for the research discussions that we had that were all inspiring for me.

I would like to thank my parents and brothers for their endless support and love. Last, my most sincere thanks go to my beloved wife, Dina, for her love, support, and patience specially during the last year. Without you, I would not have been able to accomplish this achievement, and to you I do not only offer my thanks but my love!

TABLE OF CONTENTS

List of Tables	ix
List of Figures	x
1 Introduction	1
1.1 The Wireless Channel	1
1.1.1 Performance Characterization of Fading Channels	3
1.2 Combating Channel Fading: Diversity	5
1.3 Cooperative Diversity	7
1.4 Dissertation Outline	11
1.4.1 Multi-node Cooperative Communications (Chapter 2)	11
1.4.2 Relay-Assignment Protocols for Coverage Extension (Chapter 3)	12
1.4.3 Cognitive Cooperative Multiple Access (Chapter 4)	12
1.4.4 Energy Efficiency of Cooperative Communications (Chapter 5)	14
2 Multinode Cooperative Communications	15
2.1 System Model and Protocol Description	16
2.2 Exact SER Performance Analysis	20
2.2.1 Exact SER for General Cooperation Scheme	21
2.2.2 Verifying the Validity of our Theoretical Model for Selective Relaying	25
2.3 Approximate SER Expression	28
2.3.1 SER Approximation for General Cooperative Protocol	28
2.3.2 Diversity order and Cooperation Gain	34
2.3.3 Bandwidth efficiency versus Diversity Gain	36
2.4 Optimal Power Allocation	39
2.4.1 Single-Relay Scenario	44
2.4.2 Networks with linear topologies	45
2.4.3 Relays located near the source or the destination	47
2.4.4 Numerical Examples	49

3	Relay Assignment Protocols for Coverage Expansion	53
3.1	System Model	54
3.2	Relay Assignment: Protocols and Analysis	57
3.2.1	Direct Transmission	58
3.2.2	Cooperative Transmission: Conditional Outage Probability	58
3.2.3	Optimal Relay Position	60
3.3	Relay Assignment Algorithms	63
3.3.1	Nearest-Neighbor Protocol	63
3.3.2	Fixed Relays Strategy	66
3.3.3	Lower Bound on the Performance: The Genie-aided algorithm:	69
3.4	Numerical Results	71
4	Cognitive Cooperative Multiple-Access	76
4.1	System Model	81
4.2	Cooperative Cognitive Multiple Access (CCMA) Protocols	85
4.2.1	CCMA-Single frame (CCMA-S)	86
4.2.2	CCMA-Multiple frames (CCMA-M)	87
4.3	Stability Analysis	88
4.3.1	Stability Analysis of CCMA-S	89
4.3.2	Stability Analysis of CCMA with multiple frames (CCMA-M)	99
4.3.3	Existing Cooperation Protocols: Stability Analysis	105
4.3.4	Numerical Results	110
4.4	Throughput Region	117
4.5	Delay Analysis	118
4.5.1	Delay Performance for CCMA-S	118
4.5.2	Delay Performance of CCMA-Me	121
4.5.3	Numerical Results for Delay Performance	122
5	On the Energy Efficiency of Cooperative Communications	133
5.1	System Model	135
5.2	Performance Analysis and Optimum Power Allocation	138
5.2.1	Direct Transmission	139
5.2.2	Cooperative Transmission	140
5.3	Multi-Relay Scenario	145
5.4	Experimental and Simulation Results	148
5.4.1	Experimental Results	148
5.4.2	Simulation Results	151
6	Conclusions and Future Work	158
6.1	Conclusions	158
6.2	Future Work	161
6.2.1	Cooperation among Correlated Sources	161

6.2.2 Cognitive Cooperative Communications	162
Bibliography	162

LIST OF TABLES

2.1	Comparison between optimal power allocation via exhaustive search and analytical results. $N = 3$ relays, uniform network topology. . .	49
2.2	Comparison between optimal power allocation via exhaustive search and analytical results. $N = 3$ relays: (a)all relays near the source; (b) all relays near the destination.	50

LIST OF FIGURES

1.1	Probability of error performance for a BPSK system for AWGN and fading channels.	4
2.1	Illustrating cooperation under $\mathcal{C}(1)$: The $(k + 1)$ -th relay combines the signals received from the source and the k -th relay.	18
2.2	Illustrating cooperation under $\mathcal{C}(N - 1)$: The $(k + 1)$ -th relay combines the signals received from the source and all of the previous relays.	20
2.3	SER vs. SNR for two different scenarios. The first is the simulated SER for the model described in this chapter in which the relays know whether each symbol is decoded correctly or not. The second is the simulated SER for a practical scenario in which the relay forwards the decoded symbol based on comparing the received SNR with a threshold. Also the exact SER expression in (2.22) is plotted as '+'. The cooperation protocol utilized is $\mathcal{C}(1)$ and the modulation scheme is QPSK.	27
2.4	Comparison between the approx. SER in (2.37), and the simulated SER for different number of relays. The cooperation protocol utilized is $\mathcal{C}(1)$ and the modulation scheme is QPSK.	36
2.5	Comparison between the performance of schemes $\mathcal{C}(1)$ and $\mathcal{C}(N - 1)$ for both QPSK and 16QAM modulation, $N = 3$	37
2.6	BER performance comparison between different numbers of cooperating relays taking into account the BW efficiency, $\mathcal{C}(1)$	38
2.7	Comparison between the SER performance of time diversity without any relaying and cooperative diversity. Two relays are utilized for cooperation and correspondingly three time slots for time diversity. The first order Markov model is utilized to account for time correlation, and different relays' positions are depicted with the source-destination distance normalized.	40
2.8	Optimal power allocation for $N = 2$ relays under different relays positions.	44
2.9	SER vs. power allocation ratio at the source node for different relay positions.	45

3.1	Illustrating the difference between the direct and cooperative transmission schemes, and the coverage extension prospected by cooperative transmission.	57
3.2	The effect of the relay location on the outage probability.	62
3.3	Illustrating cooperation under nearest neighbor protocol: The nearest neighbor is at a distance r_{sl} from the source. Therefore, the shaded area should be empty from any users.	64
3.4	Illustrating cooperation under nearest neighbor protocol: The nearest neighbor is at a distance r_{sl} from the source. Therefore, the shaded area should be empty from any users.	68
3.5	Average outage probability versus the number of relays in fixed relaying.	69
3.6	Average outage probability versus the transmit power.	74
3.7	Average outage probability versus the cell radius.	74
3.8	Average outage probability versus the spectral efficiency.	75
4.1	Network and channel Model.	81
4.2	Stable throughput region for system CCMA-S for a fixed resource-sharing vector (ω_1, ω_2) given by $\mathcal{R}(S_1) \cup \mathcal{R}(S_2)$	94
4.3	Stability regions for the different considered protocols at a SNR threshold equal to $\beta = 35$. For this value of β CCMA-S is equivalent to TDMA as depicted. CCMA-Me has the largest throughput region.	113
4.4	Stability regions for the different considered protocols at a SNR threshold equal to $\beta = 64$. TDMA is contained in CCMA-S, and the gap between SDF and CCMA-Me increases in this case.	113
4.5	Aggregate maximum stable throughput versus SNR threshold β in dB. The propagation path loss is set to $\gamma = 3.5$. First scenario is used for SDF and SIDF. CCMA-Me has the best tradeoff curve among all the other protocols.	114
4.6	Aggregate maximum stable throughput versus spectral efficiency R in b/s/Hz. The propagation path loss is set to $\gamma = 3.5$. Second scenario is used for SDF and SIDF. SIDF has the best performance for low spectral efficiency but it suffers from a catastrophic degradation when increasing the spectral efficiency R . CCMA-Me has a graceful degradation due to its bandwidth efficiency, and it has the best tradeoff for medium to high spectral efficiency.	115
4.7	Average queueing delay per terminal versus the arrival rate for a SNR threshold of $\beta = 15$	123
4.8	Average queueing delay per terminal versus the arrival rate for a SNR threshold of $\beta = 64$	124
4.9	Envelopes for the stability region of CCMA-S.	130
4.10	Numerical evaluation of δ for $f_{1l} = 1$	132

5.1	System Model	135
5.2	Sequence of packet errors at the two utilized wireless cards.	149
5.3	Modelling the channel by a two (on-off) state Markov chain to study the time correlation.	150
5.4	Cooperation gain versus the source-destination distance for different values of received power consumption	154
5.5	Cooperation gain versus the SNR threshold β	155
5.6	Optimal power consumption for both cooperation and direct transmission scenarios for different values of power amplifier loss α	155
5.7	Cooperation gain versus the source-destination distance for different values of QOS.	156
5.8	Optimal Consumed Power versus distance for different relay locations for equal power allocation at source and relay.	156
5.9	Optimal Consumed Power versus distance for different relay locations for optimal power allocation at source and relay.	157
5.10	Optimal Consumed Power versus number of relays for different values of required outage probability.	157

Chapter 1

Introduction

1.1 The Wireless Channel

Although the study of wireless communications and networking has started many decades ago and resulted in a large body of work, there is still an increasing thrill from the research community in further exploring this field. This is attributed to the proliferation of wireless applications with high demands in terms of signal quality, data rates, and coverage. Various wireless devices are becoming an integral part of our everyday life. Many emerging applications require a collection of freely and probably dynamic wireless nodes to communicate with each other without the existence of an infrastructure. The main challenge facing designers in achieving the demands of these future wireless applications is the unpredictability associated with the wireless channel. Wireless channels feature fading, shadowing, interference, and other impairments that affect the performance of communications. Another major challenge is the scarcity of the two fundamental resources for communications, namely, energy and bandwidth.

Among the most severe impairments to wireless communications is signal fad-

ing. Fading results in random fluctuations in the amplitude of the received signals that can result in the received signal amplitude being very low to the extent that the receiver may not be able to distinguish the signal from thermal noise [1], [2]. Fading is the result of the random scattering from reflectors with different attenuation coefficients that results in multiple copies of the signal arriving at the receiver with different gains, phase shifts and delays. These multiple signal replicas can add together in constructive or destructive ways resulting in the fading phenomenon [1], [2].

If we denote the transmitted signal by $x(t)$ and the received signal by $y(t)$, we can model the multipath channel model as follows

$$y(t) = \sum_{i=1}^L h_i(t)x(t - \tau_i(t)) + n(t), \quad (1.1)$$

where $h_i(t)$ is the channel coefficient for the i -th path at time t , $\tau_i(t)$ is the corresponding path delay, L is the number of paths, and $n(t)$ is the additive noise. This model implicitly assumes the channel to be linear. The channel delay spread is defined as the time difference between the first received path and the last received path, i.e., $\max_{i,j \in \{1, \dots, L\}} \tau_i - \tau_j$ [1], [2]. If the channel delay spread is very small compared to the symbol duration, then we can consider that all of the paths are received simultaneously and the corresponding channel gain is the sum of the path gains in (1.1), i.e.,

$$y(t) = x(t - \tau) \sum_{i=1}^L h_i(t) + n(t). \quad (1.2)$$

In the frequency domain, this is equivalent to looking at the channel in the time domain as a single impulse, which will have a flat spectrum in the frequency domain. The channel is said to be a flat fading channel under such scenario because all of the signal components in the frequency domain are affected by the same fade value. Furthermore, if the number of paths L are large enough, we can use the

central limit theorem to approximate the distribution of the channel gain by a Gaussian distribution. If the channel gains are zero-mean, e.g. there is no line of sight, then this fading scenario is known as Rayleigh flat fading [1], [2].

1.1.1 Performance Characterization of Fading Channels

To better understand the severe effects of fading on the performance of wireless communication systems, we are going to consider the performance of a simple channel under two scenarios: additive-white-Gaussian noise (AWGN), and Rayleigh flat fading. We will illustrate the probability of error performance of transmitting binary phase-shift-keying (BPSK) under these two scenarios.

Consider an AWGN channel modeled as follows

$$y = hx + n, \tag{1.3}$$

where h is a deterministic known channel gain at the receiver, x is the transmitted binary signal that takes values $+\sqrt{P}$ or $-\sqrt{P}$, and the noise n is a circular symmetric zero-mean complex Gaussian random variable with variance N_o . The probability of error for such a communication system is given by [1], [2]

$$P_e = Q\left(\sqrt{2|h|^2\text{SNR}}\right) \tag{1.4}$$

where $\text{SNR} = \frac{P}{N_o}$, and $Q(x)$ is the complementary distribution function (cdf) of a zero-mean unit variance Gaussian random variable. It is known that $Q(x)$ decays exponentially with x as [2]

$$Q(x) \leq \exp\left(-\frac{x^2}{2}\right). \tag{1.5}$$

For the Rayleigh flat fading case, h is a random variable that follows a circular symmetric zero mean complex Gaussian process. Assuming the channel fading

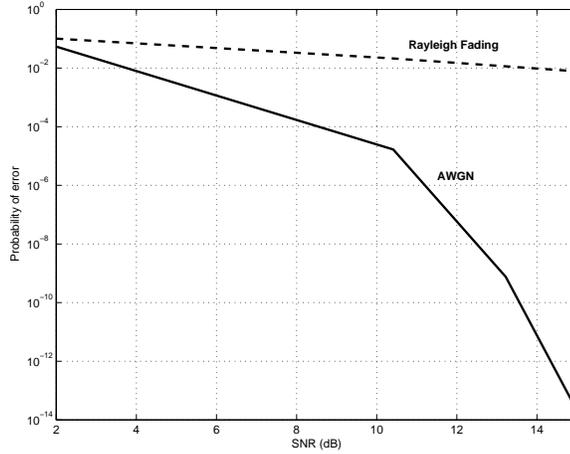


Figure 1.1: Probability of error performance for a BPSK system for AWGN and fading channels.

to have variance one, and that the receiver has perfect channel state information (coherent detection), the probability of error for such scenario can be shown to be equal to [1], [2]

$$P_e = \frac{1}{2} \left(1 - \sqrt{\frac{\text{SNR}}{1 + \text{SNR}}} \right) \simeq \frac{1}{4\text{SNR}}, \quad (1.6)$$

where the approximation is tight at high SNR. From the above, we can see that while the error performance decays exponentially fast with the signal-to-noise ratio for AWGN channels, it decays only with the inverse of the SNR for the Rayleigh fading channel. This shows that the performance of wireless channels under fading is very poor. Figure 1.1 depicts the probability of error for BPSK signalling discussed above under AWGN and Rayleigh flat fading. It is clear from the figure the severe losses in the performance due to fading. This emphasizes the fact that fading is one of the major challenges facing wireless communications systems.

1.2 Combating Channel Fading: Diversity

An effective and well established technique to combat fading is diversity. Diversity can be defined as any technique by which multiple copies of the signal are delivered to the receiver via independently fading channels [2]. The diversity order of a system can be loosely defined as the number of independent channels over which information is being sent. There are three physical domains in which we can generate independent channels: time, frequency and space. The use of spatial diversity has gained a lot of interest in the recent years as time diversity, e.g. channel coding and interleaving, can result in excessive system delays, and frequency diversity can result in high bandwidth losses. Diversity order was rigorously defined in [2] as the rate of decay of the probability of error with the SNR when using log – log scale, i.e.,

$$\text{Diversity gain} = - \lim_{\text{SNR} \rightarrow \infty} \frac{\log P_e}{\log \text{SNR}}. \quad (1.7)$$

For example, for the flat fading channel in (1.6) the diversity order is one.

To achieve spatial diversity, multiple-input-multiple-output (MIMO) systems have been introduced in the last decade, and a large body of work has been established to study the performance of these systems [3, 4]. A MIMO system is simply one where both the transmitter and the receiver have multiple antennas. This implies that the transmitter has the capability of transmitting a different signal from each antenna and the receiver has as input different signals from each antenna. The signal present at each receive antenna is the combination of signals from the transmit antennas after each having traveled through their different fading channels. The MIMO configuration can be exploited through different designs that differ, among other factors, in the particular form of performance improvement that it is intended to obtain. One of these possible design approaches may aim at

obtaining array gain, which is an increase in average received SNR by processing the signals at each transmit and receive antenna in such a way that the received signals are coherently combined. A similar technique can be applied to achieve interference reduction by shaping the energy emitted to each receiver in such a way that most of the energy is useful and not interference.

MIMO systems can provide performance improvement through diversity gain. For example, if the number of antennas at the transmitter and receiver are M and N , respectively, and assuming independent fading between all antenna element pairs, the probability of error at the receiver side can be shown to decay with the SNR as SNR^{-MN} . Code design to achieve diversity in flat fading MIMO systems, also known as space-time codes, has been the focus of many researchers in the last decade [5–7]. The analysis and design of codes for MIMO systems was not limited to the flat fading scenario, but further extended to frequency selective fading channel models. Orthogonal-frequency-division-multiplexing (OFDM) is usually utilized with such frequency selective fading channels to overcome the intersymbol interference in the channel and reduce the complexity of equalizer design at the receiver side [8]. Space-frequency and space-time-frequency codes were designed to achieve diversity in space, time, and frequency in MIMO-OFDM systems [9–12].

The gains of MIMO systems in terms of increasing the channel capacity, higher throughput, improved error performance, and better energy efficiency are well established by now. In practice, however, installing multiple antennas on a device might not be feasible because of space or cost limitations. Besides, to achieve full diversity gains in MIMO, there must be sufficient separation between the antenna elements at the transmitter and receiver sides, which is difficult to achieve in

practice. This will result in the fades of the channels between different antenna pairs to be correlated which can reduce the diversity gains of the system [13–16].

To avoid such problems in MIMO, cooperative diversity has been recently introduced [17–23]. In cooperation, different nodes in the network can share their antennas and resources for distributed transmission and processing. In other words, cooperative communications benefit from the broadcast nature of the wireless channel to form a distributed MIMO system via relaying. In the following, we further discuss the relay channel and the new term cooperative diversity.

1.3 Cooperative Diversity

The classical relay channel introduced by Van der Meulen [24] models a three terminal communication channel. A relay channel contains a terminal, called a relay, that listens to the signal transmitted by the source, processes it, and then transmits it to the destination to improve the system performance. Later, Cover and El Gamal [25] developed lower and upper bounds on the channel capacity for specific non-faded relay channel models. The lower and upper bounds do not coincide in general except for some special cases as in the degraded relay channel [26]. Later, several works have studied the capacity of the relay channels and developed coding strategies that can achieve the ergodic capacity of the channel under certain scenarios, see [27] and the references there in. Capacity of multiple-access relay channels was studied in [28], and capacity theorems for transmitter receiver cooperation was developed in [29].

User-cooperation has been first introduced and studied in [18,19]. In these two-part paper, a specific cooperation model was developed for code-division-multiple-access (CDMA) systems in which each two users in the system are coupled to

help each other. The two users use orthogonal codes to avoid multiple-access interference. Assuming the knowledge of channel phases at the transmitter sides, increased data rates for the cooperating users have been demonstrated. In another work [30], error-control-coding was incorporated into cooperation and this technique was termed as coded-cooperation.

In [22], the term cooperative diversity was introduced. Several cooperation protocols were presented and their outage capacity was analyzed. Outage capacity can be defined as the probability that the mutual information of a channel falls below a certain required rate [31]. Half-duplex constraints were assumed, meaning that the relay can not transmit and receive at the same time. Two main categories of cooperative diversity protocols were proposed: fixed relaying and adaptive relaying.

In fixed relaying, the channel resources are divided between the source and the relay in a fixed (deterministic) manner. Cooperation is generally done in two phases. In the first phase the transmitter sends a message and both the relay and the destination try to receive. In the second phase, the relay transmits a processed version of its received message to the destination. The destination then combines both copies from the source and the relay to form a detection statistic. The processing at the relay differs according to the employed protocol. In amplify-and-forward, the relay simply scales the received version and transmits an amplified version of it to the destination. Note that the amplified version is noisy because of the noise added at the relay. Despite of the noise propagation, it was shown in [22] that amplify-and-forward can achieve full diversity gain equal to two, the number of cooperating nodes in this case. Another possibility of processing at the relay node is for the relay to decode the received signal, re-encodes it and then retransmits

it to the receiver. This kind of relaying is termed as decode-and-forward. In [22], decode-and-forward requires the correct decoding at the relay node, otherwise the signal is considered decoded in error at the destination. It is clear that for such a scheme the diversity achieved is one only, because the performance of the system is limited by the worst link from the source-relay and source-destination.

Fixed relaying has the advantage of easy implementation, but the disadvantage of low bandwidth efficiency. This is because half of the channel resources are allocated to the relay for transmission, which reduces the overall rate. This is true especially when the source-destination channel is not very bad, because under such scenario a high percentage from the packets transmitted by the source to the destination can be received correctly by the destination and the relays transmissions are wasted. To overcome this problem, [22] proposed adaptive relaying protocols. The proposed adaptive relaying protocols comprise two strategies, selective and incremental relaying.

In selective relaying, the relay and the source are assumed to know the fade of the channel between them, and if the signal-to-noise ratio of the signal received at the relay exceeds a certain threshold, the relay performs decode-and-forward on the message. On the other hand, if the channel between the source and the relay falls below the threshold, the relay idles. Furthermore, assuming reciprocity in the channel, the source also knows that the relay idles, and the source transmits a copy of its signal to the destination instead. Selection relaying improves upon the performance of decode-and-forward, as the signal-to-noise ratio threshold at the relay can be designed to overcome the inherent problem in decode-and-forward that the relay is required to decode correctly. Selection relaying was shown in [22] to achieve diversity gain two.

For the second adaptive relaying protocol proposed in [22], namely, incremental relaying, it is assumed that there is a feedback channel from the destination to the relay. The destination feeds back an acknowledgement to the relay if it was able to receive the source's message correctly in the first transmission phase, and the relay does not need to transmit then. It was shown in [22], that this protocol has the best spectral efficiency among the proposed protocols because the relay does not need to transmit always, and hence, the second transmission phase becomes opportunistic depending on the channel fade of the direct channel between the source and the destination. Nevertheless, incremental relaying achieves diversity order two [22]. In [23], distributed space-time coding was proposed in which multiple relays receive the source's message and performs decode-and-forward with each relay assigned a unique codeword which is a column from a space-time code. The relays need to be synchronized, and the destination combines the signals received from the relays simultaneously in the second phase. Such a scheme was shown to provide full diversity gain equal to the number of cooperating terminals, and it still requires only two phases for transmission. Several works followed to study the implementation and performance of practical distributed space-time codes [32–36].

The symbol-error-rate performance for Rayleigh flat fading single-relay channel was analyzed in [37, 38] for both decode-and-forward and amplify-and-forward protocols, and in [39] for amplify-and-forward relaying. For both relaying strategies, exact symbol-error-rate expressions and outer bounds for the performance were derived. Furthermore, optimal power allocation between the source and the relay was studied in [37, 38], and it was shown that equal power allocation is not optimal in general. However, if the source-relay channel condition is good, then equal power allocation is close to optimal.

1.4 Dissertation Outline

From the discussion above, cooperative communications is a new communication paradigm that allows different users or nodes in a wireless network to share their resources and antennas for distributed transmission and/or processing. This thesis develops and analyzes a cross-layer framework for utilizing the cooperative communication paradigm in wireless networks with the goal to provide significant improvements in signal reliability, coverage area, network throughput, and energy efficiency with respect to other existing alternatives. We envision that cooperative communications will result in a paradigm shift in existing wireless network, a shift that will impact the design of future ad-hoc networks, sensor networks, as well as next generation cellular networks. The rest of the thesis is organized as follows.

1.4.1 Multi-node Cooperative Communications (Chapter 2)

Previous work on multiple relays cooperative networks [40] has only considered protocols in which each relay node needs to combine signals from all of the previous transmissions to achieve full diversity. In this thesis, we consider a more general setup and study a family of cooperative protocols in which each relay can combine an arbitrary subset from the previous transmission. We characterize the exact symbol-error-rate expressions for this family of protocols, besides deriving approximations for the performance at high SNR. We demonstrate that it is sufficient to combine signals only from the previous relay and the source in order to asymptotically achieve the same performance as combining all of the previous transmissions. Moreover, we characterize the optimal power allocation strategy

among the relays and the source to further improve the performance of the cooperation protocol [41–43].

1.4.2 Relay-Assignment Protocols for Coverage Extension (Chapter 3)

One important application of cooperative communications is to extend the coverage area in wireless networks without increasing infrastructure. However, a crucial challenge in implementing cooperation protocols is how to select relay-source pairs. In this thesis, we address this problem based on the knowledge of the users spatial distribution which determines the channel statistics. We consider two scenarios at the destination node, when the receiver uses maximal-ratio-combiner (MRC) and when no MRC is used. First we characterize the optimal relay location to minimize the outage probability, and we derive lower bounds on any relay-assignment scheme. Then, we develop and analyze the performance of two schemes: a distributed nearest neighbor relay-assignment protocol in which users can act as relays, and fixed-relay assignment where fixed relays are deployed in the network to help the users forward their data. The outage probability of these two schemes are derived, and the results reveal a significant gain in coverage area over direct-transmission under the same bandwidth efficiency and average transmitted power. The results also show that with increasing the cell-radius, the gap between the performance of direct and cooperative transmission diminishes [44–46].

1.4.3 Cognitive Cooperative Multiple Access (Chapter 4)

We propose a novel cognitive multiple-access strategy in the presence of a cooperating relay. Exploiting an important phenomenon in wireless networks, source

burstiness, the cognitive relay utilizes the periods of silence of the terminals to enable cooperation. Utilizing the unused channel resources in cooperation adds new dimensions to the problem as it alleviates the spectral inefficiency problems inherent in conventional cooperation protocols. Two protocols are developed to implement the proposed multiple-access strategy. The maximum stable throughput region and the delay performance of the proposed protocols are characterized. The results reveal that the proposed protocols provide significant performance gains over conventional relaying strategies as selection and incremental relaying, specially at high spectral efficiency regimes. The rationale is that the lossless bandwidth property of the proposed protocols results in a graceful degradation in the maximum stable throughput with increasing the required rate of communication. On the other hand, conventional relaying strategies suffer from catastrophic performance degradation because of their inherent bandwidth inefficiency that results from allocating specific channel resources for cooperation at the relay. Moreover, the analysis reveals an interesting result that the throughput region of the proposed strategy is a subset of its maximum stable throughput region, which is different from random-access ALOHA where both regions are conjectured to be identical. The proposed protocol provides a new view to the utilization of the unused channel resources; besides sharing the unused spectrum with cognitive secondary users, we show that the unused channel resources could be utilized to enhance the original system performance via cooperation [47–49].

1.4.4 Energy Efficiency of Cooperative Communications (Chapter 5)

An analytical framework for studying the energy efficiency of cooperation in wireless networks is presented. In this framework, we consider the overhead in the processing and receiving power introduced by cooperation. By taking into consideration such overhead, we study the tradeoff in the gains provided by cooperation in the form of a reduction in the transmit power, due to the spatial diversity gain, and the increase in the receiving and processing power that results from the operation of the relay. This tradeoff is shown to depend on many parameters such as the values of the receive and processing powers, the application, the power amplifier loss, and several other factors. The results reveal an interesting threshold behavior; below a certain threshold distance between the source and destination direct transmission becomes more energy efficient than cooperation. The results also provide guidelines for the design of power allocation strategies, relay-assignment algorithms and the selection of the optimal number of relays to help the source [50, 51].

Chapter 2

Multinode Cooperative Communications

As discussed in Chapter 1, cooperative communications allows nodes in the network to share their antennas and resources for distributed transmission and processing. The goal is to combat channel fading by generating independent paths between the source and the destination. The question that arises is whether such a distributed MIMO system can achieve the same diversity gains as MIMO systems where all the antennas are located at the same place and have access to the source information. The non-identical statistics of the channels between different pairs of nodes in the system renders the problem more challenging to analyze.

Furthermore, if multiple nodes are assigned to help a source node, several scenarios to implement cooperation between the nodes arise. For instance, there are different scenarios for how each node processes transmissions from previous nodes. Previous work on multi-node cooperation [40] considered one possible scenario in which each node forms a decision statistic based on all previous transmissions. How other scenarios compare to such a complicated scheme is not clear.

In this chapter we address the multinode cooperation problem, and we seek answers for the above posed questions. We start by proposing a class of cooperative decode-and-forward protocols for arbitrary N -relay wireless networks, in which each relay can combine the signal received from the source along with one or more of the signals transmitted by previous relays. We consider selective relaying in which each receiving relay can judge on the quality of the receiving signal and decide whether to forward the received signal or not, similar treatment for the amplify-and-forward multi-relay problem is considered in [52]. In our proposed protocols, we refer to the scenario in which each relay combines the signals received from the previous m relays along with that from the source as $\mathcal{C}(m)$, where $1 \leq m \leq N - 1$. Note that the multihop diversity scheme introduced in [40] is similar to the scheme $\mathcal{C}(N - 1)$ we are considering without selective relaying. We provide symbol-error-rate performance analysis for the class of proposed protocols. Finally, we analyze the optimal power allocation among the cooperating nodes assuming a fixed average power available for transmission.

2.1 System Model and Protocol Description

We consider an arbitrary N -relay wireless network, where information is to be transmitted from a source to a destination. Due to the broadcast nature of the wireless channel, some relays can overhear the transmitted information and thus can cooperate with the source to send its data. The wireless link between any two nodes in the network is modeled as a Rayleigh fading narrowband channel with additive white Gaussian noise (AWGN). The channel fades for different links are assumed to be statistically independent. This is a reasonable assumption as the relays are usually spatially well separated. The additive noise at all receiving

terminals is modeled as zero-mean complex Gaussian random variables with variance \mathcal{N}_o . For medium access, the relays are assumed to transmit over orthogonal channels, thus no inter-relay interference is considered in the signal model.

The cooperation strategy we are considering employs a selective decode-and-forward protocol at the relaying nodes. Each relay can measure the received SNR and forwards the received signal if the SNR is higher than some threshold. For mathematical tractability of symbol-error-rate calculations we assume the relays can judge whether the received symbols are decoded correctly or not and only forwards the signal if decoded correctly otherwise remains idle. This assumption will be shown via simulations to be very close to the performance of the practical scenario of comparing the received SNR to a threshold, specially when the relays operate in a high SNR regime, as for example when the relays are selected close to the source node. The rationale behind this is that when the relays are closer to the source node, or more generally operate in a high SNR regime, the channel fading (outage event defined in [53]) becomes the dominant source of error [53], and hence measuring the received SNR gives a very good judgement on whether the received symbol can be decoded correctly or not with high probability.

Various scenarios for the cooperation among the relays can be implemented. A general cooperation scenario, denoted as $\mathcal{C}(m)$ ($1 \leq m \leq N - 1$), can be implemented in which each relay combines the signals received from the m previous relays along with that received from the source. The simplest scenario $\mathcal{C}(1)$ among the class of proposed protocols is depicted in Fig. 2.1, in which each relay combines the signal received from the previous relay and the source. The most complicated scenario $\mathcal{C}(N - 1)$ is depicted in Fig. 2.2, in which each relay combines the signals received from all of the previous relays along with that from the source, and thus

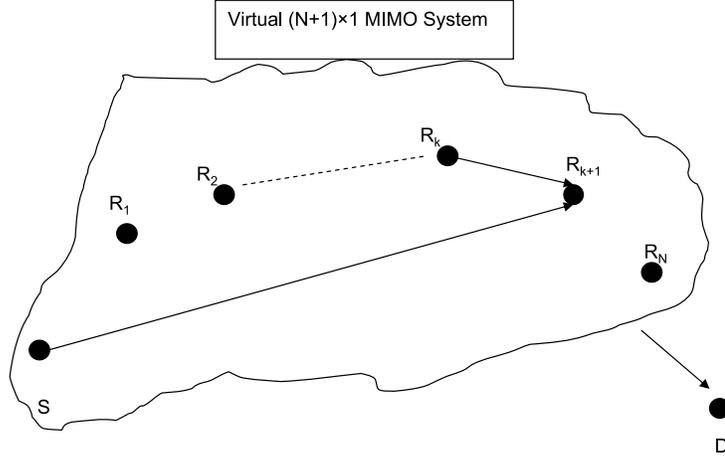


Figure 2.1: Illustrating cooperation under $\mathcal{C}(1)$: The $(k + 1)$ -th relay combines the signals received from the source and the k -th relay.

is similar to the scenario considered in [40]. This is the most sophisticated scenario and should provide the best performance in the class of proposed protocols $\{\mathcal{C}(m)\}_{m=1}^{N-1}$ as in this case each relay utilizes the information from all previous phases of the protocol. In all of the considered cooperation scenarios, the destination coherently combines the signals received from the source and all of the relays. In the sequel, we focus on presenting the system model for a general cooperative scheme $\mathcal{C}(m)$ for any $1 \leq m \leq N - 1$.

For a general scheme $\mathcal{C}(m)$, $1 \leq m \leq N - 1$, each relay decodes the information after combining the signals received from the source and the previous m relays. The cooperation protocol has $(N + 1)$ Phases. In Phase 1, the source transmits the information, and the received signal at the destination and the i -th relay can be modeled respectively as

$$y_{s,d} = \sqrt{P_0}h_{s,d}x + n_{s,d}, \quad y_{s,l_i} = \sqrt{P_0}h_{s,l_i}x + n_{s,l_i}, \quad 1 \leq i \leq N, \quad (2.1)$$

where P_0 is the power transmitted at the source, x is the transmitted symbol with

unit power, $h_{s,d} \sim CN(0, \sigma_{s,d}^2)$ and $h_{s,l_i} \sim CN(0, \sigma_{s,l_i}^2)$ are the channel fading coefficients between the source and the destination, and i -th relay, respectively, and $CN(\alpha, \sigma^2)$ denotes a circularly symmetric complex Gaussian random variable with mean α and variance σ^2 . The terms $n_{s,d}$ and n_{s,l_i} denote the AWGN. In Phase 2, if the first relay correctly decodes, it forwards the decoded symbol with power P_1 to the destination, otherwise it remains idle.

Generally in Phase k , $2 \leq k \leq N$, the k -th relay combines the received signals from the source and the previous $\min\{m, k-1\}$ relays using a maximal-ratio-combiner (MRC) as follows

$$y_{l_k} = \sqrt{P_0} h_{s,l_k}^* y_{s,l_k} + \sum_{i=\max(1, k-m)}^{k-1} \sqrt{\widehat{P}_i} h_{i,l_k}^* y_{l_i,l_k}, \quad (2.2)$$

where $h_{l_i,l_k} \sim CN(0, \sigma_{l_i,l_k}^2)$ is the channel fading coefficient between the i -th and the k -th relays. In (2.2), y_{l_i,l_k} denotes the signal received at the k -th relay from the i -th relay, and can be modeled as

$$y_{l_i,l_k} = \sqrt{\widehat{P}_i} h_{i,l_k} x + n_{l_i,l_k}, \quad (2.3)$$

where \widehat{P}_i is the power transmitted at relay i in Phase $(i+1)$, and $\widehat{P}_i = P_i$ if relay i correctly decodes the transmitted symbol, otherwise $\widehat{P}_i = 0$. The k -th relay uses y_{l_k} in (2.2) as the detection statistics. If relay k decodes correctly it transmits with power $\widehat{P}_l = P_l$ in Phase $(k+1)$, otherwise it remains idle. Finally, in Phase $(N+1)$, the destination coherently combines all of the received signals using an MRC as follows

$$y_d = \sqrt{P_0} h_{s,d}^* y_{s,d} + \sum_{i=1}^N \sqrt{\widehat{P}_i} h_{i,d}^* y_{l_i,d}. \quad (2.4)$$

In all the cooperation scenarios considered, the total transmitted power is fixed as $P_0 + \sum_{i=1}^N P_i = P$.

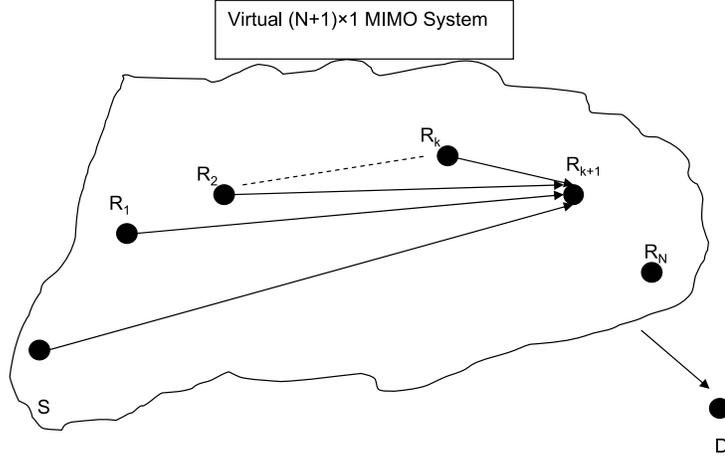


Figure 2.2: Illustrating cooperation under $\mathcal{C}(N-1)$: The $(k+1)$ -th relay combines the signals received from the source and all of the previous relays.

2.2 Exact SER Performance Analysis

In this section, we present SER performance analysis for a general cooperative scheme $\mathcal{C}(m)$ for any $1 \leq m \leq N-1$. Exact SER expressions of this general scheme is provided for systems utilizing either M-PSK or M-QAM modulation.

First, we introduce some terminologies that will be used throughout the analysis. For a given transmission, each relay can be in one of two states: either it decoded correctly or not. Let us define a $1 \times n$, $1 \leq n \leq N$, vector \mathbf{S}_n to represent the states of the first n relays for a given transmission. The k -th entry of the vector \mathbf{S}_n denotes the state of the k -th relay as follows

$$S_n[k] = \begin{cases} 1 & \text{if relay } k \text{ correctly decodes,} \\ 0 & \text{otherwise,} \end{cases} \quad 1 \leq k \leq n. \quad (2.5)$$

Since the decimal value of the binary vector \mathbf{S}_n can take on values from 0 to $2^n - 1$, for convenience we denote the state of the network by an integer decimal number.

Let $\mathbf{B}_{x,n} = (B_{x,n}[1], B_{x,n}[2], \dots, B_{x,n}[n])$ be the $1 \times n$ binary representation of a decimal number x , with $B_{x,n}[1]$ being the most significant bit. So, $\mathbf{S}_N = \mathbf{B}_{x,N}$ indicates that the k -th relay, $1 \leq k \leq N$, is in state $S_N[k] = B_{x,N}[k]$.

2.2.1 Exact SER for General Cooperation Scheme

We consider a general cooperation scheme $\mathcal{C}(m)$, $1 \leq m \leq N - 1$, in which the k -th ($1 \leq k \leq N$) relay coherently combines the signals received from the source along with the signals received from the previous $\min\{m, k - 1\}$ relays. The state of each relay in this scheme depends on the states of the previous m relays, i.e., whether these relays decoded correctly or not. This is due to the fact that the number of signals received at each relay depends on the number of relays that decoded correctly from the previous m relays. Hence, the joint probability of the states is given by

$$P(\mathbf{S}_N) = P(S_N[1])P(S_N[2] | S_N[1]) \cdots P(S_N[N] | S_N[N-1], \dots, S_N[N-m]). \quad (2.6)$$

Conditioning on the network state, which can take 2^N values, the probability of error at the destination given the channel state information (CSI) can be calculated using the law of total probability as follows

$$P_{e|CSI} = \sum_{i=0}^{2^N-1} Pr(e | \mathbf{S}_N = \mathbf{B}_{i,N})Pr(\mathbf{S}_N = \mathbf{B}_{i,N}), \quad (2.7)$$

where e denotes the event that the destination decoded in error. The summation in the above equation is over all possible states of the network.

Now, let us compute the terms in (2.7). The destination collects the copies of the signal transmitted in the previous phases using a MRC (2.4). The resulting

SNR at the destination can be computed as

$$SNR_d = \frac{P_0 |h_{s,d}|^2 + \sum_{j=1}^N P_j B_{i,N}[j] |h_{l_j,d}|^2}{\mathcal{N}_o}, \quad (2.8)$$

where $B_{i,N}[j]$ takes value 1 or 0 and determines whether the j -th relay has decoded correctly or not. The k -th relay coherently combines the signals received from the source and the previous m relays. The resulting SNR can be calculated as

$$SNR_{l_k}^m = \frac{P_0 |h_{s,l_k}|^2 + \sum_{j=\max(1,k-m)}^{k-1} P_j B_{i,N}[j] |h_{l_j,l_k}|^2}{\mathcal{N}_o}. \quad (2.9)$$

If M-PSK modulation is used in the system, with instantaneous SNR γ , the SER given the channel state information is given by [54]

$$P_{CSI}^{PSK} = \Psi_{PSK}(\gamma) \triangleq \frac{1}{\pi} \int_0^{(M-1)\pi/M} \exp\left(-\frac{b_{PSK}\gamma}{\sin^2(\theta)}\right) d\theta, \quad (2.10)$$

where $b_{PSK} = \sin^2(\pi/M)$. If M-QAM ($M = 2^k$ with k even) modulation is used in the system, the corresponding conditional SER can be expressed as [54]

$$P_{CSI}^{QAM} = \Psi_{QAM}(\gamma) \triangleq 4CQ(\sqrt{b_{QAM}\gamma}) - 4C^2Q^2(\sqrt{b_{QAM}\gamma}), \quad (2.11)$$

in which $C = 1 - \frac{1}{\sqrt{M}}$, $b_{QAM} = 3/(M - 1)$, and $Q(x)$ is the complementary distribution function (CDF) of the Gaussian distribution, and is defined as $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-\frac{t^2}{2}) dt$.

Let us focus on computing the SER in the case of M-PSK modulation, and the same procedure is applicable for the case of M-QAM modulation. From (2.8), and for a given network state $\mathbf{S}_N = \mathbf{B}_{i,N}$, the conditional SER at the destination can be computed as

$$Pr(e|\mathbf{S}_N = \mathbf{B}_{i,N}) = \Psi_{PSK}(SNR_d). \quad (2.12)$$

Denote the conditional probability that the k -th relay is in state $B_{i,N}[k]$ given the states of the previous m relays by $P_{k,i}^m$. From (2.9), this probability can be

computed as follows

$$\begin{aligned}
P_{k,i}^m &\triangleq Pr(S_N[k] = B_{i,N}[k] | S_N[k-1] = B_{i,N}[k-1], \dots, S_N[k-m] = B_{i,N}[k-m]) \\
&= \begin{cases} \Psi_{PSK}(SNR_{l_k}^m), & \text{if } B_{i,N}[k] = 0, \\ 1 - \Psi_{PSK}(SNR_{l_k}^m), & \text{if } B_{i,N}[k] = 1. \end{cases}
\end{aligned} \tag{2.13}$$

To compute the average SER, we need to average the probability in (2.7) over all channel realizations, i.e., $P_{SER}(m) = E_{CSI} [P_{e|CSI}]$. Using (2.6), (2.12), and (2.13), $P_{SER}(m)$ can be expanded as follows

$$P_{SER}(m) = \sum_{i=0}^{2^N-1} E_{CSI} \left[\Psi_{PSK}(SNR_d) \prod_{k=1}^N P_{k,i}^m \right]. \tag{2.14}$$

Since the channel fades between different pairs of nodes in the network are statistically independent by the virtue that different nodes are not co-located, the quantities inside the expectation operator in the above equation are functions of independent random variables, and thus can be further decomposed as

$$P_{SER}(m) = \sum_{i=0}^{2^N-1} \left\{ E_{CSI} [\Psi_{PSK}(SNR_d)] \prod_{k=1}^N E_{CSI} [P_{k,i}^m] \right\}. \tag{2.15}$$

The above analysis is applicable to the M-QAM case by changing the function $\Psi_{PSK}(\cdot)$ into $\Psi_{QAM}(\cdot)$.

Since the channels between the nodes are modeled as Rayleigh fading channels, the absolute norm square of any channel realization $h_{i,j}$ between any two nodes i and j in the network has an exponential distribution with mean $\sigma_{i,j}^2$. Hence, $E_{CSI}[\Psi_q(\gamma)]$ can be expressed as

$$E_{CSI}[\Psi_q(\gamma)] = \int_{\gamma} \Psi_q(\gamma) f(\gamma) d\gamma \tag{2.16}$$

where $f(\gamma)$ is the probability density function of the random variable γ , and $q = 1$ ($q = 2$) correspond to M-PSK (M-QAM), respectively. If γ is an exponentially distributed random variable with mean $\bar{\gamma}$, then it can be shown [54] that $E_{CSI}[\Psi_q(\gamma)]$ is given by

$$E_{CSI}[\Psi_q(\gamma)] = F_q \left(1 + \frac{b_q \bar{\gamma}}{\sin^2(\theta)} \right), \quad (2.17)$$

where $F_q(\cdot)$ and the constant b_q are defined as

$$\begin{aligned} F_1(x(\theta)) &= \frac{1}{\pi} \int_0^{(M-1)\pi/M} \frac{1}{x(\theta)} d\theta, & b_1 &= b_{psk} \\ F_2(x(\theta)) &= \frac{4C}{\pi} \int_0^{\pi/2} \frac{1}{x(\theta)} d\theta - \frac{4C^2}{\pi} \int_0^{\pi/4} \frac{1}{x(\theta)} d\theta, & b_2 &= \frac{b_{QAM}}{2}. \end{aligned} \quad (2.18)$$

In order to get the above expressions, we use two special properties of the $Q(\cdot)$ function, specifically, $Q(x) = \frac{1}{\pi} \int_0^{\pi/2} \exp(-\frac{x^2}{2\sin^2(\theta)}) d\theta$, and $Q^2(x) = \frac{1}{\pi} \int_0^{\pi/4} \exp(-\frac{x^2}{2\sin^2(\theta)}) d\theta$ for $x \geq 0$ [54].

Averaging over all the Rayleigh fading channel realizations, the SER at the destination for a given network state $\mathbf{B}_{i,N}$ is given by

$$E_{CSI}(\Psi_q(SNR_d)) = F_q \left[\left(1 + \frac{b_q P_0 \sigma_{s,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \prod_{j=1}^N \left(1 + \frac{b_q B_{i,N}[j] P_j \sigma_{l_j,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \right]. \quad (2.19)$$

Similarly, the probability that the k -th relay is in state $B_{i,N}[k]$ given the states of the previous m relays is given by

$$E_{CSI}[P_{k,i}^m] = G_k^m(B_{i,N}[k]), \quad (2.20)$$

where $G_k^m(\cdot)$ is defined as

$$G_k^m(x) = \begin{cases} F_q \left[\left(1 + \frac{b_q P_0 \sigma_{s,l_k}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \prod_{j=\max(1,k-m)}^{k-1} \left(1 + \frac{b_q B_{i,N}[j] P_j \sigma_{l_j,l_k}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \right], & \text{if } x = 0, \\ 1 - F_q \left[\left(1 + \frac{b_q P_0 \sigma_{s,l_k}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \prod_{j=\max(1,k-m)}^{k-1} \left(1 + \frac{b_q B_{i,N}[j] P_j \sigma_{l_j,l_k}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \right], & \text{if } x = 1. \end{cases} \quad (2.21)$$

in which $F_q(\cdot)$ and the constant b_q are specified in (2.18). As a summary, the SER in (2.15) of the cooperative multi-node system employing scenario $\mathcal{C}(m)$ with M-PSK or M-QAM modulation can be determined from (2.19), (2.20), and (2.21) in the following Theorem.

Theorem 1 *The SER of an N -relay decode-and-forward cooperative diversity network utilizing protocol $\mathcal{C}(m)$, $1 \leq m \leq N - 1$, and M-PSK or M-QAM modulation is given by*

$$P_{SER}(m) = \sum_{i=0}^{2^N-1} F_q \left[\left(1 + \frac{b_q P_0 \sigma_{s,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \prod_{j=1}^N \left(1 + \frac{b_q B_{i,N}[j] P_j \sigma_{l_j,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \right] \prod_{k=1}^N G_k^m(B_{i,N}[k]), \quad (2.22)$$

where the functions $F_q(\cdot)$ and $G_k^m(\cdot)$ are defined in (2.18) and (2.21), respectively.

2.2.2 Verifying the Validity of our Theoretical Model for Selective Relaying

In this subsection, we will illustrate with some simulation experiments the validity of the theoretical results we obtained. In the simulations, we considered only cooperative protocol $\mathcal{C}(1)$. The number of relays is taken to be $N = 1, 2, 3$, in addition to the source and the destination nodes. We considered two simulation setups. In the first setup we simulate the SER performance under the assumption that the relay correctly judges whether the received signal is decoded correctly or not, i.e., no error propagation. In the second setup, we consider the more practical scenario in which each relay compares the instantaneous received SNR to a threshold and hence decides whether to forward the received signal or not, and thus error propagation is allowed (the threshold is taken equal to 3dB here and is selected by experiment). The relays are considered closer to the source than

the destination. The channel variance depends on the distance l and propagation path-loss α as follows $\sigma^2 \propto l^{-\alpha}$, and $\alpha = 3$ in our simulations. The channel gains are as follows: $\sigma_{s,l_i}^2 = 8\sigma_{s,d}^2$, and $\sigma_{l_i,d}^2 = \sigma_{s,d}^2$. The noise variance is taken to be $N_o = 1$. The total transmitted power in each case is considered fixed to P .

Fig. 2.3 depicts the SER vs. P/N_o performance of cooperation scenario $\mathcal{C}(1)$ with QPSK. As shown in the figure, the performance curves of the two previously described simulation setups are very close for different number of relays. This validates that our model for selective relaying assumed for mathematical tractability has close performance to that of practical selective relaying when comparing the SNR to a threshold. The intuition behind this is, as we illustrated before, that when the relays in general operate in a high SNR regime, in this case the relays are closer to the source node, the error propagation from the relays becomes negligible and this is due to the fact that the channel outage event (SNR less than the threshold) becomes the dominating error event as proved in [53].

The performance of direct transmission without any relaying is also shown in Fig. 2.3 as a benchmark for a no-diversity scheme. Moreover, the exact SER expression from Theorem 1 is depicted as a '+' mark. It is clear from the depicted figure that the analytical SER expression in (2.22) for scenario $\mathcal{C}(1)$ exactly matches the simulation results for each case. This confirms our theoretical analysis. The results also reveal that the cooperative diversity protocols can achieve full diversity gain in the number of cooperating terminals, which can be seen from the slopes of the performance curves which become more steeper with increasing the number of relays.

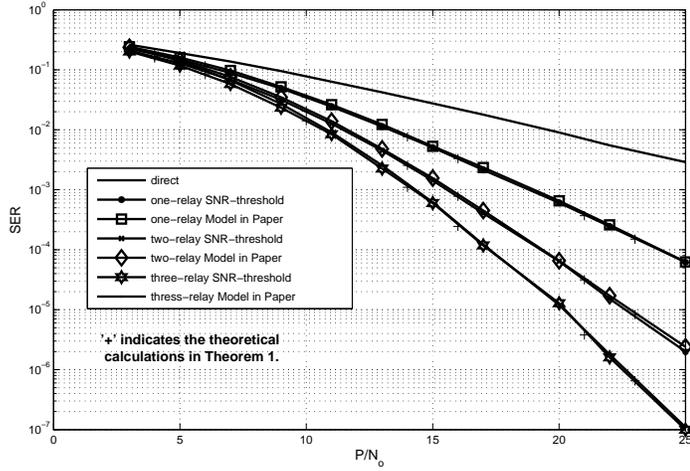


Figure 2.3: SER vs. SNR for two different scenarios. The first is the simulated SER for the model described in this chapter in which the relays know whether each symbol is decoded correctly or not. The second is the simulated SER for a practical scenario in which the relay forwards the decoded symbol based on comparing the received SNR with a threshold. Also the exact SER expression in (2.22) is plotted as '+'. The cooperation protocol utilized is $\mathcal{C}(1)$ and the modulation scheme is QPSK.

2.3 Approximate SER Expression

In the previous section, we provided exact expressions for the SER of a general cooperative scheme $\mathcal{C}(m)$, $1 \leq m \leq N - 1$, for arbitrary N -relay networks with either M-PSK or M-QAM modulation. The derived SER expressions, however, involve 2^N terms and integral functions. In this section, we provide approximate expressions for the SER performance of the proposed class of cooperative diversity schemes. The approximation is derived at high SNR and yields simple expressions that can provide insights to understanding the factors affecting the system performance, which helps in designing different network functions as power allocation, scheduling, routing, and node selection.

2.3.1 SER Approximation for General Cooperative Protocol

One can see that any term in the exact SER formulation (2.22) in Theorem 1 consists of the product of two quantities: i) The first is

$$F_q \left[\left(1 + \frac{b_q P_0 \sigma_{s,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \prod_{j=1}^N \left(1 + \frac{b_q B_{i,N}[j] P_j \sigma_{t_j,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \right] \quad (2.23)$$

, which corresponds to the conditional SER at the destination for a given network state $\mathbf{B}_{i,N}$; and ii) The second is the probability of the network being in that state, and is given by $\prod_{k=1}^N G_k^m(B_{i,N}[k])$. At high enough SNR, the probability of error $F_q(\cdot)$ is sufficiently small compared to 1, thus we can assume that $1 - F_q(\cdot) \simeq 1$. Hence, the only terms in the second quantity $\prod_{k=1}^N G_k^m(B_{i,N}[k])$ that will count are those corresponding to relays that have decoded in error. For convenience, we make the following definition: Let $\Omega_i(n, m)$ denote the subset of nodes that decode correctly from node $\max(1, n - m)$ till node $n - 1$, when the network was in state

$\mathbf{B}_{i,N}$. More specifically

$$\Omega_i(n, m) \triangleq \{\text{relay } j: \text{ s.t. } B_{i,N}[j] = 1, \max(1, n - m) \leq j \leq n - 1.\} \quad (2.24)$$

Then, the SER formulation (2.22) in Theorem 1 can be approximated as

$$P_{SER}(m) \simeq \sum_{i=0}^{2^N-1} F_q \left[\left(1 + \frac{b_q P_0 \sigma_{s,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \prod_{j=1}^N \left(1 + \frac{b_q B_{i,N}[j] P_j \sigma_{l_j,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \right] \quad (2.25)$$

$$\prod_{k \in \Omega^c(N+1, N)} G_k^m(B_{i,N}[k]),$$

where Ω^c is the complementary set of Ω , i.e., the set of nodes that decoded erroneously.

First, we simplify the first term corresponding to the SER at the destination. Using the definition of F_q in (2.18), and ignoring all the 1's¹ in $F_q(\cdot)$ in (2.25), the conditional SER at the destination for a given network state $\mathbf{B}_{i,N}$ can be approximated as

$$F_q \left[\left(1 + \frac{b_q P_0 \sigma_{s,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \prod_{j=1}^N \left(1 + \frac{b_q B_{i,N}[j] P_j \sigma_{l_j,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \right] \quad (2.26)$$

$$\simeq \frac{\mathcal{N}_o^{1+|\Omega_i(N+1, N)|} g_q(1 + |\Omega_i(N+1, N)|)}{b_q^{1+|\Omega_i(N+1, N)|} P_0 \sigma_{s,d}^2 \prod_{j \in \Omega_i(N+1, N)} P_j \sigma_{l_j,d}^2},$$

where $|\Omega_i(N+1, N)|$ denotes the cardinality of the set $\Omega_i(N+1, N)$, i.e. the number of nodes that decodes correctly, which also denotes the number of signal copies transmitted from the N relays to the destination at network state $\mathbf{B}_{i,N}$. The function $g_q(\cdot)$ in (2.26) is specified as

$$g_q(x) = \begin{cases} \frac{1}{\pi} \int_0^{(M-1)\pi/M} \sin^{2x}(\theta) d\theta, & \text{for M-PSK, } q = 1 \\ \frac{4C}{\pi} \left[\int_0^{\pi/2} \sin^{2x}(\theta) d\theta - C \int_0^{\pi/4} \sin^{2x}(\theta) d\theta \right], & \text{for M-QAM, } q = 2. \end{cases} \quad (2.27)$$

¹The tightness of these approximations can be proved easily by computing some limit functions for $F_q(x)$ and $1 - F_q(x)$ as x , which denotes an affine function of the power, goes to ∞ . For page limitations, we only include the proof for the single relay scenario using M-PSK in appendix 1.

Let us write the transmitter powers allocated at the source and different relays as a ratio of the total available power P as follows, $P_0 = a_0 P$, and $P_i = a_i P$, $1 \leq i \leq N$, in which the power ratios are normalized as $a_0 + \sum_{i=1}^N a_i = 1$. One can then rewrite (2.26) in terms of the power allocation ratios as follows

$$F_q \left[\left(1 + \frac{b_q P_0 \sigma_{s,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \prod_{j=1}^N \left(1 + \frac{b_q B_{i,N}[j] P_j \sigma_{l_j,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \right] \quad (2.28)$$

$$\simeq \frac{(\mathcal{N}_o/P)^{1+|\Omega_i(N+1,N)|} g_q (1 + |\Omega_i(N+1,N)|)}{b_q^{1+|\Omega_i(N+1,N)|} a_0 \sigma_{s,d}^2 \prod_{j \in \Omega_i(N+1,N)} a_j \sigma_{l_j,d}^2},$$

Note that the SNR term (\mathcal{N}_o/P) in (2.28) is of order $(1 + |\Omega_i(N+1,N)|)$. This is intuitively meaningful since the destination receives $(1 + |\Omega_i(N+1,N)|)$ copies of the signal, in which the term 1 is due to the copy from the source. Thus (2.28) decays as $SNR^{-(1+|\Omega_i(N+1,N)|)}$ at high SNR.

At the k -th relay, $1 \leq k \leq N$, the conditional SER for a given network state $\mathbf{B}_{i,N}$ can be similarly approximated as

$$F_q \left[\left(1 + \frac{b_q P_0 \sigma_{s,l_k}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \prod_{j=\max(1,k-m)}^{k-1} \left(1 + \frac{b_q B_{i,N}[j] P_j \sigma_{l_j,l_k}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \right] \quad (2.29)$$

$$\simeq \frac{(\mathcal{N}_o/P)^{1+|\Omega_i(k,m)|} g_q (1 + |\Omega_i(k,m)|)}{b_q^{1+|\Omega_i(k,m)|} a_0 \sigma_{s,l_k}^2 \prod_{j \in \Omega_i(k,m)} a_j \sigma_{l_j,l_k}^2}.$$

where $|\Omega_i(k,m)|$ is the number of relays that decodes correctly from the previous $\min(k-1, m)$ relays. The SNR in the above expression is of order $1 + |\Omega_i(k,m)|$.

From (2.29), the product $\prod_{k \in \Omega_i^c(N+1,N)} G_k^m(B_{i,N}[k])$ in (2.25) is given by

$$\prod_{k \in \Omega_i^c(N+1,N)} G_k^m(B_{i,N}[k]) = \prod_{k \in \Omega_i^c(N+1,N)} \frac{(\mathcal{N}_o/P)^{1+|\Omega_i(k,m)|} g_q (1 + |\Omega_i(k,m)|)}{b_q^{1+|\Omega_i(k,m)|} a_0 \sigma_{s,l_k}^2 \prod_{j \in \Omega_i(k,m)} a_j \sigma_{l_j,l_k}^2}, \quad (2.30)$$

in which the SNR is of order $\sum_{k \in \Omega_i^c(N+1,N)} (1 + |\Omega_i(k,m)|) = |\Omega_i^c(N+1,N)|$

+ $\sum_{k \in \Omega_i^c(N+1, N)} |\Omega_i(k, m)|$. Substituting (2.28) and (2.30) into (2.25), we get

$$P_{SER}(m) \simeq \sum_{i=0}^{2^N-1} \frac{(N_o/P)^{d_i} g_q(1 + |\Omega_i(N+1, N)|) \prod_{k \in \Omega_i^c(N+1, N)} g_q(1 + |\Omega_i(k, m)|)}{b_q^{d_i} a_0^{1+|\Omega_i^c(N+1, N)|} \sigma_{s,d}^2 \prod_{j \in \Omega_i(N+1, N)} a_j \sigma_{r_j,d}^2 \prod_{k \in \Omega_i^c(N+1, N)} \sigma_{s,l_k}^2 \prod_{l \in \Omega_i(k, m)} a_l \sigma_{l,l_k}^2} \quad (2.31)$$

where $d_i = 1 + |\Omega_i(N+1, N)| + |\Omega_i^c(N+1, N)| + \sum_{k \in \Omega_i^c(N+1, N)} |\Omega_i(k, m)|$. From (2.31), we can see that the SNR is of order d_i . Since $|\Omega_i(N+1, N)| + |\Omega_i^c(N+1, N)| = N$, the order d_i can be lower bounded as follows

$$d_i = 1 + N + \sum_{k \in \Omega_i^c(N+1, N)} |\Omega_i(k, m)| \geq N + 1, \quad (2.32)$$

in which the equality holds if and only if $\sum_{k \in \Omega_i^c(N+1, N)} |\Omega_i(k, m)| = 0$. Thus the smallest order of the SNR is $N + 1$.

The equality in (2.32) holds if and only if $|\Omega_i(k, m)| = 0$, for any $k \in \Omega_i^c(N+1, N)$, and $0 \leq i \leq 2^N - 1$. Essentially, this means that the equality in (2.32) is satisfied if and only if for each relay k that decodes erroneously, the m preceding relays also must have decoded erroneously. One can think of this condition as a chain rule, and this leads to the conclusion that the equality holds if and only if for each relay k that decodes in error all the previous relays must have decoded in error. As a result, the only network states that will contribute in the SER expression with terms of order $N + 1$ in the SNR are those of the form $\mathbf{S}_N = \mathbf{B}_{2^n-1, N}$, $0 \leq n \leq N$. For example a network state of the form $\mathbf{S}_N = [0, \dots, 0, 1, \dots, 1]$ will contribute to a term in the SER with SNR raised to the order $N + 1$, and a network state $\mathbf{S}_N = [0, \dots, 0, 1, 1, 0, 1, \dots, 1]$ will contribute to a term in the SER with SNR raised to an exponent larger than $(N + 1)$ depending on m . Therefore, only $N + 1$ states of the network have SER terms that decays as $1/SNR^{N+1}$ and the rest of the network states decay with faster rates, hence these $N + 1$ terms will dominant the SER expression at high enough SNR.

In order to write the approximate expression for the SER corresponding to these $N + 1$ terms, we need to note the following points that can be deduced from the above analysis. As described above, in order for the equality in (2.32) to hold, the following set of conditions must be satisfied. First, since for any relay that decodes erroneously all the previous m relays must have decoded in error, we have

$$\Omega_i(k, m) = \Phi, \quad (2.33)$$

for all $k \in \Omega_i^c(N + 1, N)$, where Φ is the empty set. Second, for these $N + 1$ states that satisfy the equality in (2.32) the set $\Omega_i^c(N + 1, N)$ takes one of the following forms

$$\Omega_i^c(N + 1, N) \in \{\Phi, \{1\}, \{1, 2\}, \dots, \{1, 2, \dots, N\}\}, \quad (2.34)$$

For example, $\Omega_i^c(N + 1, N) = \{1, 2, \dots, k\}$ denotes the state in which only the first k relays decoded erroneously. Accordingly, its cardinality, denoted by $|\Omega_i^c(N + 1, N)|$, takes one of the following values

$$|\Omega_i^c(N + 1, N)| \in \{0, 1, 2, \dots, N\}. \quad (2.35)$$

Thus, only the $N + 1$ states determined from the above conditions will contribute to the SER expression at high SNR because they decay as $1/SNR^{N+1}$, which is the slowest decaying rate as seen from (2.32). From (2.31), (2.33), (2.34), and (2.35), the conditional SER for any of these states, e.g., $\Omega_i^c(N + 1, N) = \{1, 2, \dots, k\}$, can be determined as follows

$$SER_k(m) = \frac{(\mathcal{N}_o/P)^{N+1} g_q(N - k + 1) g_q^k(1)}{b_q^{N+1} \sigma_{s,d}^2 a_o^{1+k} \prod_{j \in \Omega_i(N+1, N)} a_j \sigma_{l_j, d}^2 \prod_{t=1}^k \sigma_{s, l_t}^2}. \quad (2.36)$$

Summing the above expression over the $N + 1$ states in (2.34), we can further determine the approximate expression for the SER in the following Theorem.

Theorem 2 *At high enough SNR, the SER of an N relay decode-and-forward cooperative diversity network employing cooperation scheme $\mathcal{C}(m)$ and utilizing M -PSK or M -QAM modulation can be approximated by*

$$P_{SER}(m) \simeq \frac{(\mathcal{N}_o/P)^{N+1}}{b_q^{N+1}\sigma_{s,d}^2} \sum_{j=1}^{N+1} \frac{g_q(N-j+2)g_q^{j-1}(1)}{a_0^j \prod_{i=j}^N a_i \sigma_{i,d}^2 \prod_{t=1}^{j-1} \sigma_{s,l_t}^2}. \quad (2.37)$$

A very important point to be noticed from the above theorem is that the approximate SER expression in (2.37) does not depend on m , the class parameter. Hence, the whole class of cooperative diversity protocols $\{\mathcal{C}(m)\}_{m=1}^{N-1}$ shares the same asymptotic performance at high enough SNR. The results obtained in Theorem 2 illustrate that utilizing the simplest scheme, namely, scenario $\mathcal{C}(1)$, results in the same asymptotic SER performance as the most sophisticated scheme, namely, $\mathcal{C}(N-1)$. This motivates us to utilize scenario $\mathcal{C}(1)$ as a cooperative protocol for multi-node wireless networks employing decode-and-forward relaying. The simplicity behind scenario $\mathcal{C}(1)$ is due to the fact that it does not require each relay to estimate the CSI for all the previous relays as in scenario $\mathcal{C}(N-1)$. It only requires each relay to know the CSI to the previous relay and the destination thus saving a lot in the channel estimation computations.

In the following, we determine roughly the savings in the computations needed for channel estimation when using scenario $\mathcal{C}(1)$ as opposed to scenario $\mathcal{C}(N-1)$ by computing the number of channels needed to be estimated in each case. The number of channels needed to be estimated in scenario $\mathcal{C}(1)$ is given by

$$n_{h,1} = 3N, \quad (2.38)$$

where N is the number of relays forwarding for the source. This value accounts for the $N+1$ channels estimated at the destination and $2N-1$ channels estimated by the N relays; the first relay estimates only one channel. In scenario $\mathcal{C}(N-1)$,

the k -th relay estimates k channels, and thus the amount of computations for this case is given by

$$n_{h,N-1} = \frac{1}{2} [N^2 + 3N + 1]. \quad (2.39)$$

From (2.38) and (2.39), the savings in the computations needed for channel estimation when using scenario $\mathcal{C}(1)$ as opposed to scenario $\mathcal{C}(N-1)$ is given by

$$\frac{n_{h,1}}{n_{h,N-1}} = \frac{6N}{N^2 + 3N + 1}. \quad (2.40)$$

The above ratio approaches 0 in the limit as N tends to ∞ . Hence, utilizing scenario $\mathcal{C}(1)$ will reduce the protocol complexity while having the same asymptotic performance as the best possible scenario.

2.3.2 Diversity order and Cooperation Gain

The philosophy before employing cooperative diversity techniques in wireless networks is to form virtual MIMO systems from separated single-antenna terminals. The aim behind this is to emulate the performance gains that can be achieved in point-to-point communications when employing MIMO systems. Two well known factors that describe the performance of the system are the diversity order and coding gain of the transmit diversity scheme. To define these terms, the SER can be written in the following form

$$P_{SER} \sim (\Delta \cdot SNR)^{-d}. \quad (2.41)$$

The constant Δ which multiplies the SNR denotes the coding gain of the scheme, and the exponent d denotes the diversity order of the system.

In the cooperative diversity schemes considered in this chapter, the relays simply repeat the decoded information, and thus we do not really have the notion of

coding; although it can still be seen as a repetition coding scheme. Hence, we will donate the constant Δ that multiplies the SNR by the cooperation gain. From (2.37) in Theorem 2, the following observations can be deduced from the previous relation

- It is clear that the diversity order of the system is given by $d = N + 1$, which indicates that the proposed cooperative diversity schemes achieves full diversity order in the number of cooperating terminals; the source and the N relays.
- The cooperation gain of the system is given by

$$\Delta = \left[\frac{1}{b_q^{N+1} \sigma_{s,d}^2} \sum_{j=1}^{N+1} \frac{g_q(N-j+2)g_q^{j-1}(1)}{a_0^j \prod_{i=j}^N a_i \sigma_{i,d}^2 \prod_{t=1}^{j-1} \sigma_{s,l_t}^2} \right]^{-1/(N+1)}. \quad (2.42)$$

In order to validate the accuracy of the derived approximate SER we conducted some simulation experiments. Throughout all the simulations, and without loss of generality, the channel gains are assumed to be unity and the noise variance is taken to be $N_o = 1$. Figure 2.4 considers scenario $\mathcal{C}(1)$ and depicts the SER performance vs. P/N_o for QPSK signalling. The transmitting power P is fixed for different number of cooperating relays in the network. The results reveal that the derived approximations for the SER are tight at high enough SNR. Regarding scenario $\mathcal{C}(N-1)$, we considered the $N = 3$ relays case. Figure 2.5 depicts the SER performance for QPSK and 16QAM modulation. The results for scenario $\mathcal{C}(1)$ under the same simulation setup are included for comparison. It can be seen from the results that there is a very small gap between the SER performance of scenarios $\mathcal{C}(1)$ and $\mathcal{C}(N-1)$, and that they almost merge together at high enough SNR. This confirms our observations that utilizing scenario $\mathcal{C}(1)$ can deliver the required SER performance for a fairly wide range of SNR. Hence, saving a lot

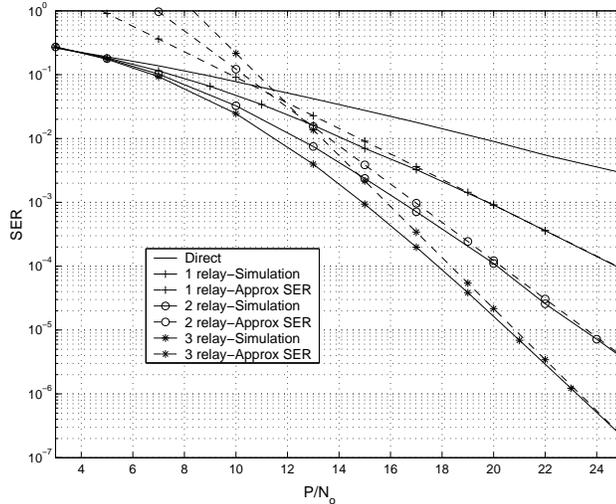


Figure 2.4: Comparison between the approx. SER in (2.37), and the simulated SER for different number of relays. The cooperation protocol utilized is $\mathcal{C}(1)$ and the modulation scheme is QPSK.

in terms of channel estimation, thus computational complexity, requirements to implement the protocol.

2.3.3 Bandwidth efficiency versus Diversity Gain

Up to this point, we did not take into account the bandwidth (BW) efficiency as another important factor to determine the performance besides the SER. Increasing the number of relays reduces the BW efficiency of the system, as the source uses only a fraction of the total available degrees of freedom to transmit the information. There is a tradeoff between the diversity gain and the BW efficiency of the system, as higher diversity gain is usually translated into utilizing the available degrees of freedom to transmit more copies of the same message which reduces the BW efficiency of the system. In order to have a fair comparison, we will fix the

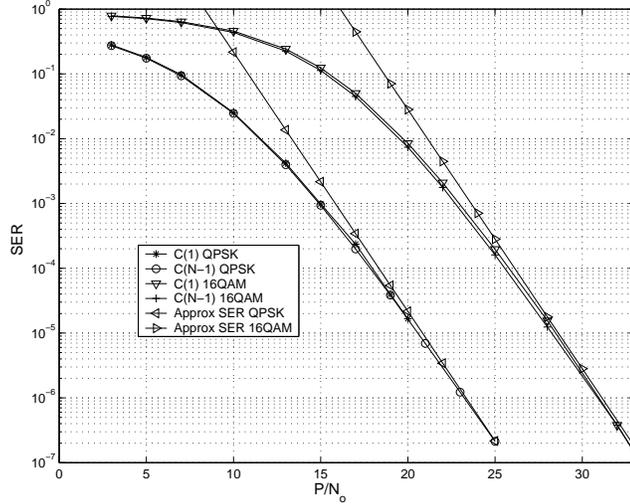


Figure 2.5: Comparison between the performance of schemes $\mathcal{C}(1)$ and $\mathcal{C}(N - 1)$ for both QPSK and 16QAM modulation, $N = 3$.

BW efficiency throughout the simulations. In order to achieve this, larger signal constellations are utilized with larger number of cooperating relays. For the direct transmission case, BPSK is used as a benchmark to achieve bandwidth efficiency of 1 bit/channel use. QPSK is used with the $N = 1$ relay case, 8PSK with $N = 2$ relays and 16QAM with $N = 3$ relays. In all of the aforementioned cases, the achieved BW efficiency is 1 bit/channel use. Fig. 2.6 depicts the BER vs. SNR per bit in dB for $N = 1, 2, 3$ relays along with the direct transmission case. The results reveal that at low SNR, lower number of nodes achieves better performance due to the BW efficiency loss incurred with utilizing larger number of cooperating nodes.

Another important point of concern is how the performance of cooperative diversity compare to that of time diversity without relaying under the same bandwidth efficiency. For example, if the target diversity gain is $N + 1$ then cooperation

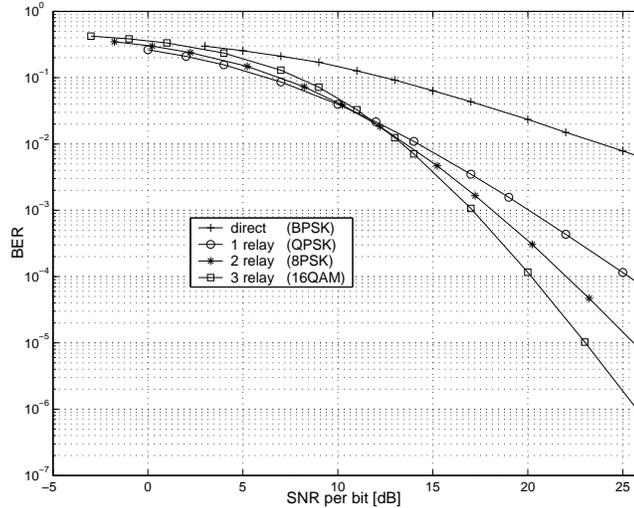


Figure 2.6: BER performance comparison between different numbers of cooperating relays taking into account the BW efficiency, $\mathcal{C}(1)$.

requires the employment of N relays, while in time diversity the source simply repeats the information for $N + 1$ successive time slots. Two factors can lead to cooperation yielding better performance than time diversity. The first is that the cooperation gain of cooperative diversity (2.42) can be considerably higher than that of time diversity if the propagation path loss is taken into account. This is because the relay nodes are usually closer to the destination node than the source itself, which results in less propagation path loss in the relays-destination links compared to the source-destination link. This is a natural gain offered by cooperation because of the distributed nature of the formed virtual array, and this is the same reason multihop communications offer more energy efficient transmission in general. The second factor which can lead to cooperation being a more attractive scheme than time diversity is that the spatial links between different nodes in the network fade independently, again because of the distributed nature

of the formed virtual array, which leads to full diversity gain. In time diversity, however, full diversity gain is not guaranteed as there might be time correlation between successive time slots. This correlation is well modeled by a first order Markov chain [55]. To illustrate more the above described factors we compare the SER performance of time and cooperative diversity in Fig. 2.7. The desired diversity gain is 3. The time correlation factor for the first order Markov model is taken equal to $\rho = 0.9, 0.7, 0.3, 0.1$. The two relays are taken in different positions as illustrated in the figure to illustrate different coding gains. It is clear from Fig. 2.7 that cooperative diversity can offer better performance than time diversity because of the higher possible coding gain that depends on the relay positions, and the degradation in the achieved performance of time diversity due to the correlation factor ρ .

2.4 Optimal Power Allocation

In this section, we try to find the optimal power allocation strategy for the multi-node cooperative scenarios considered in the previous sections. The approximate SER formula derived in (2.37) is a function of the power allocated at the source and the N relays. For a fixed transmission power budget P , the power should be allocated optimally at the different nodes in order to minimize the SER.

Since the approximation in (2.37) is tight at high enough SNR, we use it to determine the asymptotic optimum power allocation, also we drop the parameter m as the asymptotic SER performance is independent of it. The SER can be written in terms of the power ratios allocated at the transmitting nodes as follows

$$P_{SER} \simeq \left(\frac{P}{N_o} \right)^{-(N+1)} \frac{1}{b_q^{N+1} \sigma_{s,d}^2} \sum_{j=1}^{N+1} \frac{g_q(N-j+2)g_q^{j-1}(1)}{a_0^j \prod_{i=j}^N a_i \sigma_{i,d}^2 \prod_{k=1}^{j-1} \sigma_{s,l_k}^2}. \quad (2.43)$$

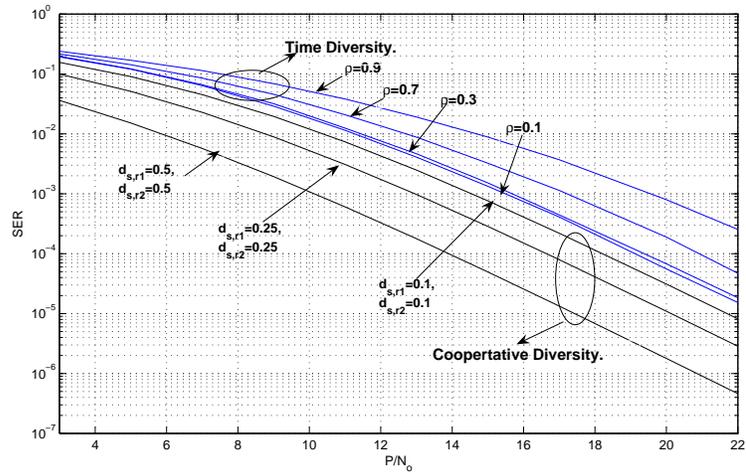


Figure 2.7: Comparison between the SER performance of time diversity without any relaying and cooperative diversity. Two relays are utilized for cooperation and correspondingly three time slots for time diversity. The first order Markov model is utilized to account for time correlation, and different relays' positions are depicted with the source-destination distance normalized.

The nonlinear optimization problem can be formulated as follows

$$\begin{aligned} \mathbf{a}_{opt} &= \arg \min_{\mathbf{a}} P_{SER} & (2.44) \\ \text{subject to} & \quad a_i \geq 0 \quad (0 \leq i \leq N), \quad \sum_{i=0}^N a_i = 1, \end{aligned}$$

where $\mathbf{a} = [a_0, a_1, \dots, a_N]$ is the power allocating vector. The Lagrangian of this problem can be written as

$$L = P_{SER} + \nu \left(\sum_{i=0}^N a_i - 1 \right) - \sum_{j=0}^N \beta_j a_j \quad (2.45)$$

where the β 's act as slack variables.

Although this nonlinear optimization problem should, in general, be solved numerically, there are some insights which can be drawn out of it. Applying first order optimality conditions, we can show that the optimum power allocation vector \mathbf{a}_{opt} must satisfy the following necessary conditions

$$\frac{\partial P_{SER}}{\partial a_i} = \frac{\partial P_{SER}}{\partial a_j}, \quad i, j \in \{0, 1, 2, \dots, N\}. \quad (2.46)$$

Next, we solve these equations simultaneously to get the relations between the optimal power allocations at different nodes. To simplify the notations, let μ_j denote the constant quantity inside the summation in (2.43), i.e.,

$$\mu_j = \left(\frac{P}{N_o} \right)^{-(N+1)} \frac{g_q(N-j+2)g_q^{j-1}(1)}{b_q^{N+1}\sigma_{s,d}^2 \prod_{i=j}^N \sigma_{l_i,d}^2 \prod_{k=1}^{j-1} \sigma_{s,l_k}^2}. \quad (2.47)$$

The derivative of the SER with respect to a_0 is given by

$$\frac{\partial P_{SER}}{\partial a_0} = \sum_{j=1}^{N+1} \frac{-j\mu_j}{a_0^{j+1} \prod_{i=j}^N a_i}, \quad (2.48)$$

while the derivative with respect to a_k , $1 \leq k \leq N$ is given by

$$\frac{\partial P_{SER}}{\partial a_k} = \sum_{j=1}^k \frac{-\mu_j}{a_0^j a_k \prod_{i=j}^N a_i}, \quad (2.49)$$

where the summation is to the k -th term only as a_k does not appear in the terms from $k + 1$ to N . Using (2.46), we equate the derivatives of the SER with respect to any two consecutive variables a_k and a_{k+1} , $1 \leq k \leq N - 1$, as follows

$$\sum_{j=1}^k \frac{-\mu_j}{a_0^j a_k \prod_{i=j}^N a_i} = \sum_{j=1}^{k+1} \frac{-\mu_j}{a_0^j a_{k+1} \prod_{i=j}^N a_i}. \quad (2.50)$$

Rearranging the terms in the above equation we get

$$\frac{a_{k+1} - a_k}{a_k a_{k+1}} \sum_{j=1}^k \frac{\mu_j}{a_0^j \prod_{i=j}^N a_i} = \frac{\mu_{k+1}}{a_0^j a_{k+1} \prod_{i=k+1}^N a_i}. \quad (2.51)$$

Since both sides of the above equation is positive, we conclude that $a_{k+1} \geq a_k$ for any $k = 1, 2, \dots, N$. Similarly, we can show that $a_o \geq a_k$ for all $1 \leq k \leq N$. Hence, solving the optimality conditions simultaneously we get the following relationships between the powers allocated at different nodes

$$P_0 \geq P_N \geq P_{N-1} \geq \dots \geq P_1. \quad (2.52)$$

The above set of inequalities demonstrates an important concept: Power is allocated at different nodes according to the received signal quality at these node. We refer to the quality of the signal copy at a node as the reliability of the node, thus the more reliable the node the more power allocated to this node. To further illustrate this concept, the $N + 1$ cooperative nodes form a virtual $(N + 1) \times 1$ MIMO system. The difference between this virtual array and a conventional point-to-point MIMO system is that in conventional point-to-point communications all the antenna elements at the transmitter are allocated at the same place and hence all the antenna elements can acquire the original signal. In a virtual array, the antenna elements constituting the array (the cooperating nodes) are not allocated at the same place and the channels among them are noisy. The source is the most reliable node as it has the original copy of the signal and thus it should be allocated

the highest share of the power. According to the described cooperation protocols, each relay combines the signal received from the source and the previous relays. As a result, each relay is more reliable than the previous relay, and hence the N -th relay is the most reliable node and is allocated the largest ratio of the power after the source, and the 1-st relay is the least reliable and is allocated the smallest ratio of the transmitted power. Another important point to notice is that the channel quality of the direct link between the source and the destination $\sigma_{s,d}^2$ is a common factor in the μ_j 's that appear in (2.51), hence the optimal power allocation does not depend on it.

To illustrate the effect of relay position on the values of the optimal power allocation ratios at the source and relay nodes, we consider a 2 relays scenario in Fig. 2.8. The two relays are taken in three different positions, close to the source, close to the destination, and in the middle between the source and the destination. In the first scenario, almost equal power allocation between the three nodes is optimal. When the relays are closer to the destination, more power is allocated to the source node, but still the second relay has a higher portion of the power relative to the first one. Similarly, in the last scenario the last relay has more power than the first one. These results reveal the fact that the further the relays from the source node the less power is allocated to the relays as they become less reliable, while as the relays become closer to the source, equal power allocation becomes near optimal. This is similar to the results of optimal power allocation for distributed space-time-coding in [34].

There are a few special cases of practical interest that permits a closed-form solution for the optimization problem in (2.44), and they are discussed in the sequel.

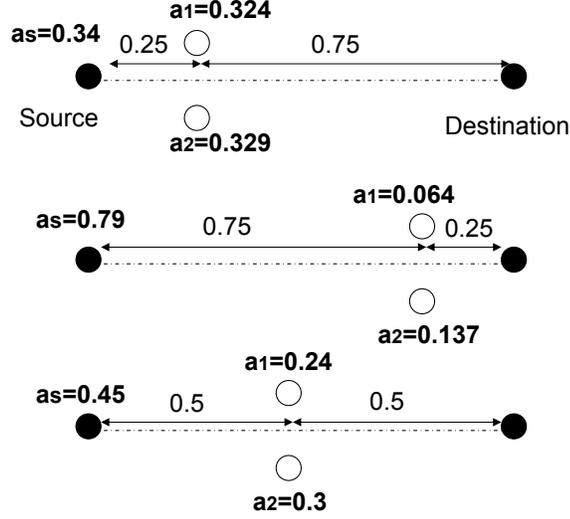


Figure 2.8: Optimal power allocation for $N = 2$ relays under different relay positions.

2.4.1 Single-Relay Scenario

For the $N = 1$ relay scenario [37], the optimization problem in (2.44) admits closed form expression. The SER for this case is simply given as

$$P_{SER} = \left(\frac{P}{N_o} \right)^{-2} \frac{1}{b_q^2 \sigma_{s,d}^2} \left(\frac{g_q(2)}{a_0 a_1 \sigma_{r_1,d}^2} + \frac{g_q^2(1)}{a_0^2 \sigma_{s,r_1}^2} \right) \quad (2.53)$$

Solving the optimization problem for this case leads the following solution for the optimal power allocation

$$a_0 = \frac{\sigma_{s,r_1} + \sqrt{\sigma_{s,r_1}^2 + 8 \frac{g_q^2(1)}{g_q(2)} \sigma_{r_1,d}^2}}{3\sigma_{s,r_1} + \sqrt{\sigma_{s,r_1}^2 + 8 \frac{g_q^2(1)}{g_q(2)} \sigma_{r_1,d}^2}}, \quad a_1 = \frac{2\sigma_{s,r_1}}{3\sigma_{s,r_1} + \sqrt{\sigma_{s,r_1}^2 + 8 \frac{g_q^2(1)}{g_q(2)} \sigma_{r_1,d}^2}}. \quad (2.54)$$

To study the effect of relay position on the optimal power allocation, we depict in Fig. 2.9 the SER performance of a single relay scenario versus the power allocation at the source node a_0 for different relay positions. The first observation that the figure reveals is that the SER performance is relatively flat around equal power

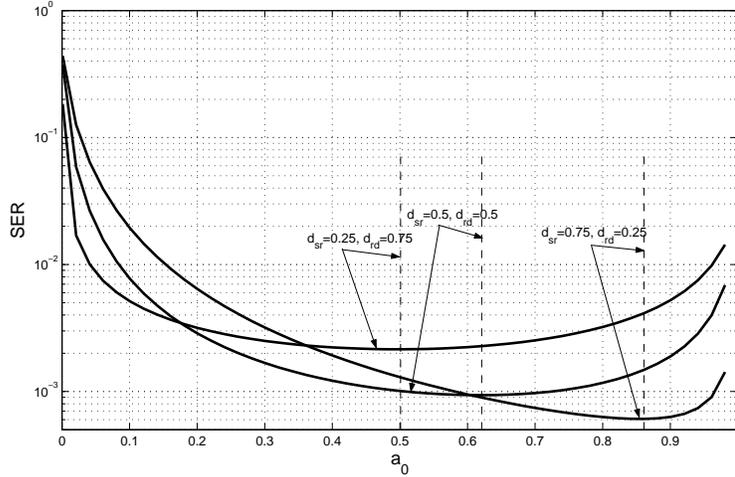


Figure 2.9: SER vs. power allocation ratio at the source node for different relay positions.

allocation when the relay is not very close to the destination, an observation that was also made for distributed space-time coding in [34]. Another observation to notice here is that as the relay becomes closer to the destination, the value of the optimal power allocation at the source node a_0 approaches 1, which means that as the relay node becomes less reliable more power should be allocated to the source node.

2.4.2 Networks with linear topologies

The propagation path-loss will be taken into account here. The channel attenuation between any two nodes $\sigma_{i,j}^2$ depends on the distance between these two nodes $d_{i,j}$ as follows: $\sigma_{i,j}^2 \propto d^{-\alpha}$, where α is the propagation constant. For a linear network topology, the most significant channel gains are for the channels between the source and the first relay σ_{s,l_1}^2 , and that between the last relay and the destination $\sigma_{l_N,d}^2$,

the other channel gains are considerably smaller than these two channels. In the SER expression in (2.37), these two terms appear as a product in all the terms except the first and the last terms. Hence these two terms dominate the SER expression, and we can further approximate the SER in this case as follows

$$P_{SER} \simeq \frac{(\mathcal{N}_o/P)^{N+1}}{b_q^{N+1} \sigma_{s,d}^2} \left[\frac{g_q(N+1)}{a_0 \prod_{i=1}^N a_i \sigma_{l_i,d}^2} + \frac{g_q^{N+1}(1)}{a_0^{N+1} \prod_{i=1}^N \sigma_{s,l_i}^2} \right]. \quad (2.55)$$

Taking the power constraint into consideration, the Lagrangian of the above problem can be written as

$$L(\mathbf{a}) = \frac{\mu_1}{a_o \prod_{i=1}^N a_i} + \frac{\mu_{N+1}}{a_o^{N+1}} + \lambda \left(\sum_{i=0}^N a_i - 1 \right), \quad (2.56)$$

where the constants μ_1 and μ_{N+1} are defined in (2.47).

Taking the partial derivatives of the Lagrangian with respect to a_j , $1 \leq j \leq N$, and equating with 0, we get

$$a_j = \frac{\mu_1}{\lambda a_o \prod_{i=1}^N a_i}. \quad (2.57)$$

Thus, we deduce that the power allocated to all of the relays are equal. Let the constant κ be defined as follows

$$\kappa = \frac{P_0 - P_j}{P_j}. \quad (2.58)$$

From the above definition, along with the power constraint we get

$$P_j = \frac{1}{1 + \kappa + N} P, \quad P_0 = \frac{1 + \kappa}{1 + \kappa + N} P. \quad (2.59)$$

To find the optimum value for κ , substitute (2.59) into the expression for the SER in (2.55) to get

$$P_{SER} \simeq \frac{\mu_1 (1 + \kappa)^N (1 + \kappa + N)^{N+1} + \mu_{N+1} (1 + \kappa + N)^{N+1}}{(1 + \kappa)^{N+1}}. \quad (2.60)$$

Differentiating (2.60) and equating to 0, we can find that the optimum κ satisfies the equation $\kappa(1 + \kappa)^N = A$, in which A is a constant given by $(N + 1) \frac{g_q^{N+1}(1) \prod_{i=1}^N \sigma_{i,d}^2}{g_q(N+1) \prod_{i=1}^N \sigma_{s,i}^2}$.

From the above analysis, the optimal power allocation for a linear network can be found in the following theorem:

Theorem 3 *The optimal power allocation for a linear network that minimizes the SER expression in (2.55) is as follows*

$$P_0 = \frac{1 + \kappa}{1 + \kappa + N} P, \quad P_i = \frac{1}{1 + \kappa + N} P, \quad 1 \leq i \leq N, \quad (2.61)$$

where κ is found through solving the equation $\kappa(1 + \kappa)^N = A$, in which A is a constant given by $(N + 1) \frac{g_q^{N+1}(1) \prod_{i=1}^N \sigma_{i,d}^2}{g_q(N+1) \prod_{i=1}^N \sigma_{s,i}^2}$.

Theorem 3 agrees with optimality conditions we found for the general problem in (2.52). Also, it shows an interesting property that in linear network topologies equal power allocation at the relays is asymptotically optimal.

2.4.3 Relays located near the source or the destination

The cooperating relays can be chosen to be closer to the source than to the destination, in order for the $N + 1$ cooperating nodes to mimic a multi-input-single-output (MISO) transmit antenna diversity system. This case is of special interest as it was shown in [27] that decode-and-forward relaying can be a capacity achieving scheme when the relays are taken to be closer to the source and it has the best performance compared to amplify-and-forward and compress-and-forward relaying in this case. In order to model this scenario in our SER formulation, we will consider the channel gains from the source to the relays to have higher gains than those from the relays to the destination, i.e., $\sigma_{s,r_i}^2 \gg \sigma_{r_i,d}^2$ for $1 \leq i \leq N$. Taking this into

account, the approximate SER expression in (2.37) can be further approximated as

$$P_{SER} \simeq \frac{N_o^{N+1} g_q(N+1)}{b_q^{N+1} \sigma_{s,d}^2 P^{N+1} a_0 \prod_{i=1}^N a_i \sigma_{l_i,d}^2}. \quad (2.62)$$

It is clear from the above equation that the SER depends equally on the power allocated to all nodes including the source, and thus the optimal power allocation strategy for this case is simply given by

$$P_0 = P_i = \frac{P}{N+1}, \quad 1 \leq i \leq N. \quad (2.63)$$

This result is intuitively meaningful as all the relays are located near to the source and thus they all have high reliability and are allocated equal power as if they form a conventional antenna array.

Now we consider the opposite scenario in which all the relays are located near the destination. In this case the channels between the relays and the destination are of a higher quality, higher gain, than those between the source and the relays, i.e., $\sigma_{l_i,d}^2 \gg \sigma_{s,l_i}^2$ for $1 \leq i \leq N$. In this case the SER can be approximated as

$$P_{SER} \simeq \frac{N_o^{N+1} g(1)^{N+1}}{b_q^{N+1} \sigma_{s,d}^2 P_0^{N+1} \prod_{k=1}^N \sigma_{s,l_k}^2}. \quad (2.64)$$

The SER in the above equation is not a function of the power allocated at the cooperating relays, and thus the optimal power allocation in this case is simply $P_0 = P$, i.e., allocating all the available power at the source. This result is very interesting as it reveals a very important concepts: If the relays are located closer to the destination than to the transmitter then direct transmission can lead better performance than decode-and-forward relaying. This is also consistent with the results in [27] in which it was shown that the performance of the decode-and-forward strategy degrades significantly when the relays get closer to the destination . This result can be intuitively interpreted as follows: The farther the relays from

Exhaustive Search	Analytical Results
$P_0 = 0.31P$	$P_0 = 0.31P$
$P_1 = 0.23P$	$P_1 = 0.23P$
$P_2 = 0.23P$	$P_2 = 0.23P$
$P_3 = 0.23P$	$P_3 = 0.23P$

Table 2.1: Comparison between optimal power allocation via exhaustive search and analytical results. $N = 3$ relays, uniform network topology.

the source the more noisy the channels between them and the less reliable the signals received by those relays to the extent that we can not rely on them on forwarding copies of the signal to the destination.

2.4.4 Numerical Examples

In this subsection, we present some numerical results to verify the analytical results for the optimal power allocation problem for the considered network topologies. The effect of the geometry on the channel links qualities is taken into consideration. We assume that the channel variance between any two nodes is proportional to the distance between them, more specifically $\sigma_{i,j}^2 \propto d_{i,j}^{-\alpha}$, where α is determined by the propagation environment is taken equal to 4 throughout our simulations. We provide comparisons between the optimal power allocation via exhaustive search to minimize the SER expression in (2.37), and optimal power allocation provided by the closed form expressions provided in this section.

First, for the linear network topology, we consider a uniform linear network, i.e., $d_{s,l_1} = d_{l_1,l_2} = \dots = d_{r_l,d}$. The variance of the direct link between the source and the destination is taken to be $\sigma_{s,d}^2 = 1$. Table 2.1 demonstrates the results

(A) Exhaustive Search	Analy. Results	(B) Exhaustive Search	Analy. Results
$P_0 = 0.25P$	$P_0 = 0.25P$	$P_0 = 0.875P$	$P_0 = P$
$P_1 = 0.25P$	$P_1 = 0.25P$	$P_1 = 0.015P$	$P_1 = 0$
$P_2 = 0.25P$	$P_2 = 0.25P$	$P_2 = 0.035P$	$P_2 = 0$
$P_3 = 0.25P$	$P_3 = 0.25P$	$P_3 = 0.075P$	$P_3 = 0$

Table 2.2: Comparison between optimal power allocation via exhaustive search and analytical results. $N = 3$ relays: (a) all relays near the source; (b) all relays near the destination.

for $N = 3$ relays. Second, for the case when all the relays are near the source, the channel links are taken to be: $\sigma_{s,l_i}^2 = \sigma_{l_i,l_j}^2 = 10$, while $\sigma_{s,d}^2 = \sigma_{l_i,d}^2 = 0.1$. Finally, for the case when all of the relays are near the destination, the channel link qualities are taken to be: $\sigma_{s,d}^2 = \sigma_{s,l_i}^2 = 0.1$, while $\sigma_{l_i,l_j}^2 = \sigma_{l_i,d}^2 = 10$. Table 2.2 illustrates the results for $N = 3$ relays for the two previous cases. In all of the provided numerical examples it is clear that the optimal power allocations obtained via exhaustive search agree with that via analytical results for all the considered scenarios. Also, the numerical results show that the optimal power allocation obtained via exhaustive search has the same ordering as the one we got in (2.52).

Appendix 1

In this appendix, we provide a proof for the tightness of the approximations we use to derive the asymptotic SER expressions at high SNR. For space limitations, we include only the proof for a single relay scenario using M-PSK modulations. The proof for the general scenario follows easily in the same footsteps. The purpose for

this proof is just to illustrate what we rigorously mean by ignoring the 1's in the $F_q(\cdot)$ functions in the SER expressions at high SNR.

For the single relay case the SER is given by,

$$\begin{aligned}
P_{SER} = & F_q \left(1 + \frac{b_q P_0 \sigma_{s,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) F_q \left(1 + \frac{b_q P_0 \sigma_{s,l_1}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \\
& + F_q \left[\left(1 + \frac{b_q P_0 \sigma_{s,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \left(1 + \frac{b_q P_1 \sigma_{l_1,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \right] \left[1 - F_q \left(1 + \frac{b_q P_0 \sigma_{s,l_1}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \right].
\end{aligned} \tag{2.65}$$

Define the functions $I_1(x)$ and $I_2(x, y)$ as follows

$$\begin{aligned}
I_1(x) &= F_1 \left(1 + \frac{b_1 x \sigma_{s,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) F_1 \left(1 + \frac{b_1 x \sigma_{s,l_1}^2}{\mathcal{N}_o \sin^2(\theta)} \right), \\
I_2(x, y) &= F_1 \left[\left(1 + \frac{b_1 x \sigma_{s,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \left(1 + \frac{b_1 y \sigma_{l_1,d}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \right] \left[1 - F_1 \left(1 + \frac{b_1 x \sigma_{s,l_1}^2}{\mathcal{N}_o \sin^2(\theta)} \right) \right].
\end{aligned} \tag{2.66}$$

where the function $F_q(\cdot)$ is defined for M-PSK ($q = 1$) in (2.18).

We are now going to prove that

$$\lim_{x \rightarrow \infty} x^2 I_1(x) = \frac{g_1^2(1)}{b_1^2 \sigma_{s,d}^2 \sigma_{s,l_1}^2}, \quad \lim_{x,y \rightarrow \infty} xy I_2(x, y) = \frac{g_1(2)}{b_1^2 \sigma_{s,d}^2 \sigma_{l_1,d}^2}. \tag{2.67}$$

The proofs are as follows.

$$\begin{aligned}
\lim_{x \rightarrow \infty} x^2 I_1(x) &= \lim_{x \rightarrow \infty} F_1 \left(\frac{1}{x} + \frac{b_1 \sigma_{s,d}^2}{\sin^2 \theta} \right) F_1 \left(\frac{1}{x} + \frac{b_1 \sigma_{s,l_1}^2}{\sin^2 \theta} \right) \\
&= \lim_{x \rightarrow \infty} \frac{1}{\pi^2} \int_0^{(M-1)\pi/M} \int_0^{(M-1)\pi/M} \frac{1}{\left(\frac{1}{x} + \frac{b_1 \sigma_{s,d}^2}{\sin^2 \theta_1} \right) \left(\frac{1}{x} + \frac{b_1 \sigma_{s,l_1}^2}{\sin^2 \theta_2} \right)} d\theta_1 d\theta_2 \\
&= \frac{1}{\pi^2} \int_0^{(M-1)\pi/M} \int_0^{(M-1)\pi/M} \lim_{x \rightarrow \infty} \frac{1}{\left(\frac{1}{x} + \frac{b_1 \sigma_{s,d}^2}{\sin^2 \theta_1} \right) \left(\frac{1}{x} + \frac{b_1 \sigma_{s,l_1}^2}{\sin^2 \theta_2} \right)} d\theta_1 d\theta_2 \\
&= \frac{1}{\pi^2} \int_0^{(M-1)\pi/M} \int_0^{(M-1)\pi/M} \frac{\sin^2 \theta_1 \sin^2 \theta_2}{b_1^2 \sigma_{s,d}^2 \sigma_{s,l_1}^2} d\theta_1 d\theta_2 = \frac{g_1(1)^2}{b_1^2 \sigma_{s,d}^2 \sigma_{s,l_1}^2}.
\end{aligned}$$

Next,

$$\begin{aligned}
\lim_{x,y \rightarrow \infty} xy I_2(x,y) &= \lim_{x,y \rightarrow \infty} F_1 \left(\left(\frac{1}{x} + \frac{b_1 \sigma_{s,d}^2}{\sin^2 \theta} \right) \left(\frac{1}{y} + \frac{b_1 \sigma_{r_1,d}^2}{\sin^2 \theta} \right) \right) \left[1 - F_1 \left(1 + \frac{x b_1 \sigma_{s,l_1}^2}{\sin^2 \theta} \right) \right] \\
&= \frac{1}{\pi} \int_0^{(M-1)\pi/M} \lim_{x,y \rightarrow \infty} \frac{1}{\left(\frac{1}{x} + \frac{b_1 \sigma_{s,d}^2}{\sin^2 \theta} \right) \left(\frac{1}{y} + \frac{b_1 \sigma_{r_1,d}^2}{\sin^2 \theta} \right)} d\theta \\
&\quad - \frac{1}{\pi^2} \int_0^{(M-1)\pi/M} \int_0^{(M-1)\pi/M} \lim_{x,y \rightarrow \infty} \frac{1}{\left(\frac{1}{x} + \frac{b_1 \sigma_{s,d}^2}{\sin^2 \theta_1} \right) \left(\frac{1}{y} + \frac{b_1 \sigma_{r_1,d}^2}{\sin^2 \theta_1} \right) \left(1 + \frac{x b_1 \sigma_{s,l_1}^2}{\sin^2 \theta_2} \right)} d\theta_1 d\theta_2 \\
&= \frac{1}{\pi} \int_0^{(M-1)\pi/M} \frac{\sin^4 \theta}{b_1^2 \sigma_{s,d}^2 \sigma_{r_1,d}^2} d\theta = \frac{g_1(2)}{b_1^2 \sigma_{s,d}^2 \sigma_{r_1,d}^2},
\end{aligned}$$

where the function $g_1(\cdot)$ is defined in (2.27). The approximate expression for the SER for the single relay scenario then follows as provided in(2.53).

As can be seen, the proof for tightness depends on simple evaluation of some limit functions as the SNR tends to infinity, and the proofs for the multinode case and M-QAM follow in the same footsteps.

Chapter 3

Relay Assignment Protocols for Coverage Expansion

In the previous chapter, the relay nodes were assumed to be already assigned to the source node. For practical implementation of cooperative communications in wireless networks, we need to develop protocols by which nodes are assigned to cooperate with each other. In most of the previous works on cooperation [18,22,23,43], the cooperating relays are assumed to exist and are already coupled with the source nodes in the network. These works also assumed a deterministic network topology, i.e., deterministic channel gain variances between different nodes in the network. If the random users' spatial distribution, and the associated propagation path losses between different nodes in the network are taken into consideration, then these assumptions, in general, are no longer valid.

Moreover, it is of great importance for service providers to improve the coverage area in wireless networks without cost of more infrastructure and under the same quality of service requirements. This poses challenges for deployment of wireless networks because of the difficult and unpredictable nature of wireless channels.

In this chapter, we address the relay-assignment problem for implementing cooperative diversity protocols to extend coverage area in wireless networks. We study the problem under the knowledge of the users' spatial distribution which determines the channel statistics, as the variance of the channel gain between any two nodes is a function of the distance between these two nodes. We consider an uplink scenario where a set of users are trying to communicate to a base-station (BS) or access point (AP) and propose practical algorithms for relay assignment. To better assess the performance of the proposed protocols, we derive a lower bound on the outage probability of any practical relay-assignment protocol. The lower bound is derived by assuming a Genie-aided protocol.

Related work for relay assignment assumes the availability of a list of candidate relays and develop relay-selection algorithms from among the list [56, 57]. In [57], two approaches for selecting a best relay are provided: Best-Select in the Neighbor Set and Best-Select in the Decoded Set. The Best-Select in the Neighbor Set algorithm is based on the average received SNRs, or equivalently the distance, while the latter is based on the instantaneous channel fading realization. Our proposed protocols do not assume a given candidate list to search for the best relay, instead we assume a random node distribution across the network and take this into consideration when analyzing the performance. We further develop and analyze distributed relay assignment protocols and benchmark their performance by the derived lower bounds.

3.1 System Model

We consider a wireless network with a circular cell of radius ρ . The BS/AP is located at the center of the cell, and N users are uniformly distributed within the

cell. The probability density function of the user's distance r from the BS/AP is thus given by

$$q(r) = \frac{2r}{\rho^2}, \quad 0 \leq r \leq \rho, \quad (3.1)$$

and the user's angle is uniformly distributed between $[0, 2\pi)$. Two communications schemes are going to be examined in the sequel. Non-cooperative transmission, or direct transmission, where users transmit their information directly to the BS/AP, and cooperative communications where users can employ a relay to forward their data.

In the direct transmission scheme, which is employed in current wireless networks, the signal received at the destination d (BS/AP) from source user s , can be modeled as

$$y_{sd} = \sqrt{PKr_{sd}^{-\alpha}}h_{sd}x + n_{sd}; \quad (3.2)$$

where P is the transmitted signal power, x is the transmitted data with unit power, h_{sd} is the channel fading gain between the two terminals. The channel fade of any link is modeled as a zero mean circularly symmetric complex Gaussian random variable with unit variance. In (3.2), K is a constant that depends on the antennas design, α is the path loss exponent, and r_{sd} is the distance between the two terminals. K, α , and P are assumed to be the same for all users. The term n_{sd} in (3.2) denotes additive noise. All the noise components are modeled as white Gaussian noise (AWGN) with variance N_o . From (3.2), the received signal-to-noise ratio is

$$\text{SNR}(r_{sd}) = \frac{|h_{sd}|^2 K r_{sd}^{-\alpha} P}{N_o}. \quad (3.3)$$

We characterize the system performance in terms of outage probability. Outage is defined as the event that the received SNR falls below a certain threshold γ_{nc} , where the subscript nc denotes non-cooperative transmission. The probability of

outage \mathcal{P}_{nc} for non-cooperative transmission is defined as,

$$\mathcal{P}_{nc} = \mathcal{P}(\text{SNR}(r) \leq \gamma_{nc}). \quad (3.4)$$

The SNR threshold γ_{nc} is determined according to the application and the transmitter/receiver structure. If the received SNR is higher than the threshold γ_{nc} , the receiver is assumed to be able to decode the received message with negligible probability of error. If an outage occurs, the packet is considered lost.

For the cooperation protocol, a hybrid version of the incremental and selection relaying proposed in [22] is employed. In this hybrid protocol, if a user's packet is lost, the BS/AP broadcasts negative acknowledgement (NACK) so that the relay assigned to this user can re-transmit this packet again. This introduces spatial diversity because the source message can be transmitted via two independent channels as depicted in Fig. 3.1. The relay will only transmit the packet if it is capable of capturing the packet, i.e., if the received SNR at the relay is above the threshold. In practice, this can be implemented by utilizing a cyclic redundancy check (CRC) code in the transmitted packet. The signal received from the source to the destination d and the relay l in the first stage can be modeled as,

$$y_{sd} = \sqrt{PKr_{sd}^{-\alpha}}h_{sd}x + n_{sd}, \quad y_{sl} = \sqrt{PKr_{sl}^{-\alpha}}h_{sl}x + n_{sl}. \quad (3.5)$$

If the SNR of the signal received at the destination from the source falls below the cooperation SNR threshold γ_c , the destination requests a second copy from the relay. Then if the relay was able to receive the packet from the source correctly, it forwards it to the destination

$$y_{ld} = \sqrt{PKr_{ld}^{-\alpha}}h_{ld}x + n_{ld}, \quad (3.6)$$

The destination will then combine the two copies of the message x as follows,

$$y_d = a_{sd}y_{sd} + a_{ld}y_{ld}; \quad (3.7)$$

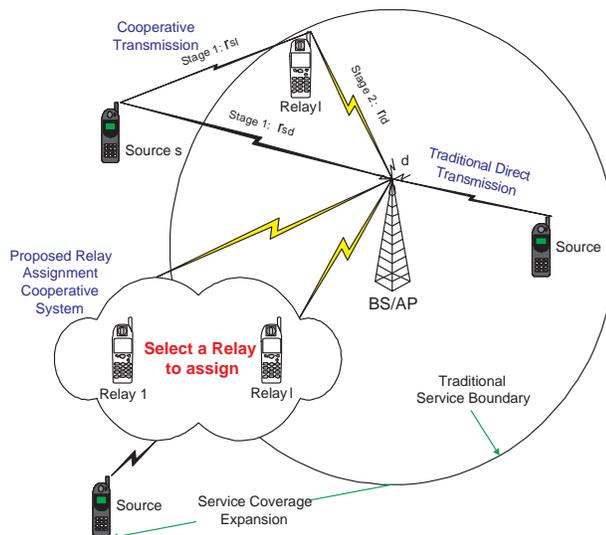


Figure 3.1: Illustrating the difference between the direct and cooperative transmission schemes, and the coverage extension prospected by cooperative transmission.

where $a_{sd} = \sqrt{PKl_{sd}^{-\alpha}h_{sd}^*}$, and $a_{rd} = \sqrt{PKl_{rd}^{-\alpha}h_{rd}^*}$. The formulation in (3.7) allows us to consider two scenarios at the destination: if $I = 1$ then the combining at the destination is a maximal-ratio-combiner (MRC), on the other hand if $I = 0$, then the destination only uses the relay message for decoding. The later scenario might be useful in case where the destination can not store an analogue copy of the source's message from the previous transmission.

3.2 Relay Assignment: Protocols and Analysis

In this section, we start with driving the average outage for direct transmission. Then we calculate the conditional outage probability for cooperative transmission and try to use the formulas to deduce the best relay location.

3.2.1 Direct Transmission

As discussed before, the outage is defined as the event that the received SNR is lower than a predefined threshold which we denote by γ_{nc} . The outage probability for the direct transmission mode \mathcal{P}_{OD} conditioned on the user's distance can be calculated as

$$\mathcal{P}_{OD}(r_{sd}) = \mathcal{P}(\text{SNR}(r_{sd}) \leq \gamma_{nc}) = 1 - \exp\left(-\frac{N_o \gamma_{nc} r_{sd}^\alpha}{KP}\right) \simeq \frac{N_o \gamma_{nc} r_{sd}^\alpha}{KP}, \quad (3.8)$$

where the above follows because $|h_{sd}|^2$, the magnitude square of the channel fade, has an exponential distribution with unit mean. The approximation in (3.8) is at high SNR.

To find the average outage probability over the cell, we need to average over the user distribution in (3.1). The average outage probability is thus given by

$$\begin{aligned} \mathcal{P}_{OD} &= \int_0^\rho \mathcal{P}_{OD}(r_{sd}) q(r_{sd}) dr_{sd} = \int_0^\rho \frac{2r_{sd}}{\rho^2} \left(1 - \exp\left(-\frac{N_o \gamma_{nc} r_{sd}^\alpha}{KP}\right)\right) dr_{sd} \\ &= 1 - \frac{2}{\alpha \rho^2} \left(\frac{KP}{N_o \gamma_{nc}}\right)^{\frac{2}{\alpha}} \Gamma\left(\frac{2}{\alpha}, \frac{N_o \gamma_{nc} \rho^\alpha}{KP}\right) \simeq \frac{2 \gamma_{nc} \rho^\alpha N_o}{KP(\alpha + 2)}, \end{aligned} \quad (3.9)$$

where $\Gamma(\cdot, \cdot)$ is the incomplete Gamma function, and it is defined as [58]

$$\Gamma(a, x) = \int_0^x \exp^{-t} t^{a-1} dt. \quad (3.10)$$

3.2.2 Cooperative Transmission: Conditional Outage Probability

Consider a source-destination pair that are r_{sd} units distance apart. Let us compute the conditional outage probability for given locations of the user and the helping relay. Using (3.5), the SNR received at the BS/AP d and the relay l from the source s is given by

$$\text{SNR}(r_{sd}) = \frac{|h_{sd}|^2 K r_{sd}^{-\alpha} P}{N_o}, \quad \text{SNR}(r_{sl}) = \frac{|h_{sl}|^2 K r_{sl}^{-\alpha} P}{N_o}. \quad (3.11)$$

While from (3.7), the SNR of the combined signal received at the BS/AP is given by

$$\text{SNR}_d = \mathbb{I} \frac{|h_{sd}|^2 K r_{sd}^{-\alpha} P}{N_o} + \frac{|h_{ld}|^2 K r_{ld}^{-\alpha} P}{N_o}. \quad (3.12)$$

The terms $|h_{sd}|^2$, $|h_{sl}|^2$, and $|h_{ld}|^2$ are mutually independent exponential random variables with unit mean. The outage probability of the cooperative transmission P_{OC} conditioned on the fixed topology of the user s and the relay l can be calculated as follows. Using the law of total probability we have

$$\mathcal{P}_{OC} = \Pr(\text{Outage} | \text{SNR}_{sd} \leq \gamma_c) \Pr(\text{SNR}_{sd} \leq \text{SNR}_{sd}) \quad (3.13)$$

where the probability of outage is zero if $\text{SNR}_{sd} > \gamma_c$. The outage probability conditioned on the event that the source-destination link is in outage is given by

$$\Pr(\text{Outage} | \text{SNR}_{sd} \leq \gamma_c) = \Pr(\text{SNR}_{sl} \leq \gamma_c) + \Pr(\text{SNR}_{sl} > \gamma_c) \Pr(\text{SNR}_d \leq \gamma_c | \text{SNR}_{sd} \leq \gamma_c), \quad (3.14)$$

where the addition of the above probabilities because they are disjoint events, and the multiplication is because the source-relay link is assumed to fade independently from the other links. The conditioning was removed for the same reason.

For the case where MRC is allowed at the destination, then the conditional outage probability at the destination is given by

$$\Pr(\text{SNR}_d \leq \gamma_c | \text{SNR}_{sd} \leq \gamma_c) = \frac{\Pr(\text{SNR}_d \leq \gamma_c)}{\Pr(\text{SNR}_{sd} \leq \gamma_c)}. \quad (3.15)$$

Using (3.15) and (3.14) in (3.13), the conditional outage probability for cooperative communications with MRC can be calculated as

$$\begin{aligned} \mathcal{P}_{OC}(r_{sd}, r_{sl}, r_{ld}) &= (1 - f(\gamma_c, r_{sd})) (1 - f(\gamma_c, r_{sl})) \\ &+ f(\gamma_c, r_{sl}) \left[1 - \frac{r_{sd}^{-\alpha}}{r_{sd}^{-\alpha} - r_{ld}^{-\alpha}} f(\gamma_c, r_{sd}) - \frac{r_{ld}^{-\alpha}}{r_{ld}^{-\alpha} - r_{sd}^{-\alpha}} f(\gamma_c, r_{ld}) \right] \end{aligned} \quad (3.16)$$

where $f(x, y) = \exp(-\frac{N_oxy^\alpha}{KP})$. The above expression can be simplified as follows

$$\mathcal{P}_{OC}(r_{sd}, r_{sl}, r_{ld}) = (1 - f(\gamma_c, r_{sd})) - \frac{r_{ld}^{-\alpha}}{r_{ld}^{-\alpha} - r_{sd}^{-\alpha}} f(\gamma_c, r_{sl}) (f(\gamma_c, r_{ld}) - f(\gamma_c, r_{sd})). \quad (3.17)$$

For the $I = 1$ case, or when MRC is used at the destination, then using the approximation $\exp(-x) \simeq 1 - x + \frac{x^2}{2}$ for small x , the above outage expression can be approximated at high SNR to

$$\begin{aligned} \mathcal{P}_{OC}(r_{sd}, r_{sl}, r_{ld}) &\simeq \frac{N_o}{KP} r_{sd}^\alpha - \frac{N_o^2}{2K^2P^2} r_{sd}^{2\alpha} - \frac{r_{sd}^\alpha}{r_{sd}^\alpha - r_{ld}^\alpha} \\ &\times \left[\frac{N_o}{KP} (r_{sd}^\alpha - r_{ld}^\alpha) + \frac{N_o^2}{2K^2P^2} ((r_{ld}^\alpha - r_{sd}^\alpha)(2r_{sl}^\alpha + r_{ld}^\alpha + r_{sd}^\alpha)) \right] \end{aligned} \quad (3.18)$$

Simplifying the above expression, we get

$$\mathcal{P}_{OC}(r_{sd}, r_{sl}, r_{ld}) \simeq \frac{N_o^2}{2K^2P^2} r_{sd}^{2\alpha} \left[2 \frac{r_{sl}^\alpha}{r_{sd}^\alpha} + \frac{r_{ld}^\alpha}{r_{sd}^\alpha} \right] \quad (3.19)$$

For the $I = 0$ case, or when no-MRC is used at the destination, then the conditional outage expression in (3.15) simplifies to

$$\Pr(SNR_d \leq \gamma_c | SNR_{sd} \leq \gamma_c) = \Pr(SNR_d \leq \gamma_c). \quad (3.20)$$

This is because the SNR received at the destination in this case is just due to the signal received from the relay-destination path. The conditional outage expression in this case can be shown to be given by

$$\mathcal{P}_{OC}(r_{sd}, r_{sl}, r_{ld}) = (1 - f(\gamma_c, r_{sd})) [1 - f(\gamma_c, r_{ld}) f(\gamma_c, r_{sl})]. \quad (3.21)$$

3.2.3 Optimal Relay Position

To find the optimal relay position, we need to find the pair (r_{sl}, r_{ld}) that minimizes the conditional outage probability expression in (3.17). First we consider the $I = 1$ scenario, where MRC is utilized at the receiver.

MRC Case:

In the following, we are going to prove that the optimal relay position, under fairly general conditions, is towards the source and on the line connecting the source and destination. Examining the conditional outage expression in (3.17), it is clear that for any value of r_{ld} , the optimal value for r_{sl} that minimizes the outage expression is the minimum value for r_{sl} . And since for any value of r_{ld} the minimum r_{sl} lies on the straight line connecting the source and destination, we get the first intuitive result that the optimal relay position is on this straight line.

Now, we prove that the optimal relay position is towards the source. Normalizing with respect to r_{sd} by substituting $x = \frac{r_{ld}}{r_{sd}}$ in (3.17) and $1 - x = \frac{r_{sl}}{r_{sd}}$, we have

$$\mathcal{P}_{OC}(x) = \frac{N_o^2}{2K^2P^2} r_{sd}^{2\alpha} [2(1-x)^\alpha + x^\alpha] \quad (3.22)$$

taking the derivative with respect to x we get

$$\frac{\partial \mathcal{P}_{OC}(x)}{\partial x} = \frac{N_o^2}{2K^2P^2} r_{sd}^{2\alpha} [-2\alpha(1-x)^{\alpha-1} + \alpha x^{\alpha-1}]. \quad (3.23)$$

Equating the above derivative to zero we get the unique solution

$$x^* = \frac{1}{1 + \left(\frac{1}{2}\right)^{\frac{1}{\alpha-1}}}. \quad (3.24)$$

Checking for the second order conditions, we get, that $\mathcal{P}_{OC}''(x) \geq 0$, which shows that the problem is convex, and x^* specified in (3.24) is indeed the optimal relay position.

Note from the optimal relay position in (3.24), that for propagation path loss $\alpha \geq 2$, we have that $x^* > 0.5$, which means that the optimal relay position is closer to the source node. In Fig. 3.2, we plot the conditional outage probability expression for different source-destination separation distances, and different values

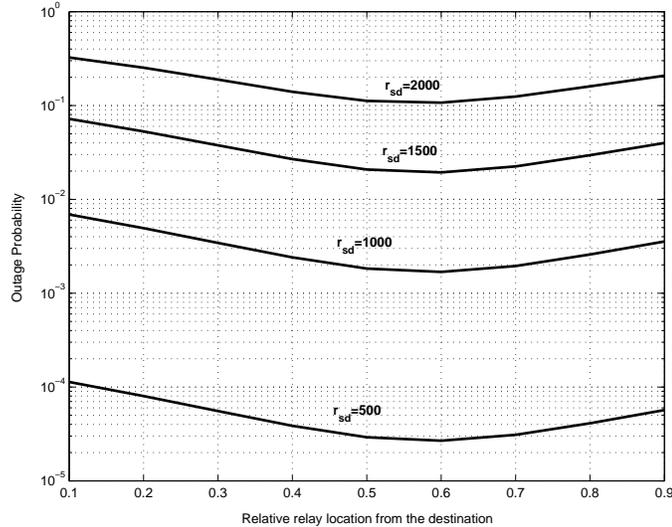


Figure 3.2: The effect of the relay location on the outage probability.

for the relay location x . It is clear from the figure that the optimal relay position is, for a lot of cases, around $x = 0.6$ to $x = 0.75$. This is the motivation for proposing the nearest-neighbor relay selection protocol in the next Section.

No-MRC case:

Next, we determine the optimal relay location for the $I = 0$ case. From the conditional outage expression in (3.21), it can be seen that if we have the freedom to put the relay anywhere in the two-dimensional plane of the source-destination pair, then the optimal relay position should be on the line joining the source and the destination- this is because of the fact that if the relay is located at any position in the two-dimensional plane, then its distances to both the source and the destination are always larger than their corresponding projections on the straight line joining the source-destination pair. In this case, we can substitute for r_{ld} by $r_{sd} - r_{sl}$. The optimal relay position can be found via solving the following

optimization problem

$$r_{sl}^* = \arg \min_{r_{sl}} \mathcal{P}_{OC}(r_{sd}, r_{sl}), \quad \text{subject to } 0 \leq r_{sl} \leq r_{sd}. \quad (3.25)$$

Since the minimization of the expression in (3.17) with respect to r_{sl} is equivalent to minimizing the exponent in the second bracket, solving the optimization problem in (3.25) is equivalent to solving

$$r_{sl}^* = \arg \min_{r_{sl}} r_{sl}^\eta + (r_{sd} - r_{sl})^\eta, \quad \text{subject to } 0 \leq r_{sl} \leq r_{sd}. \quad (3.26)$$

The above optimization problem can be simply analytically solved, and the optimal relay position can be shown to be equal to $r_{sl}^* = \frac{r_{sd}}{2}$ for $\eta > 1$. Therefore, the optimal relay position is exactly in the middle between the source and destination when no MRC is used at the destination. For this case, we are able to drive a lower bound on the performance of any relay assignment protocol as will be discussed later.

3.3 Relay Assignment Algorithms

In this section, we propose two distributed relay assignment algorithms. The first is a user-user cooperation protocol in which the nearest neighbor is assigned as a relay. The second considers the scenario where fixed relays are deployed in the network to help the users.

3.3.1 Nearest-Neighbor Protocol

In this subsection, we propose the Nearest-Neighbor protocol for relay assignment that is both distributed and simple to implement. In this protocol, The relay assigned to help is the nearest neighbor to the source as demonstrated in Fig. 3.3.

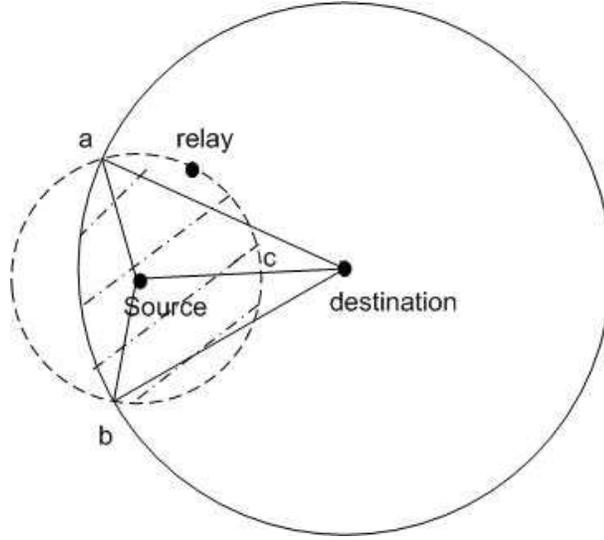


Figure 3.3: Illustrating cooperation under nearest neighbor protocol: The nearest neighbor is at a distance r_{sl} from the source. Therefore, the shaded area should be empty from any users.

The source sends a “Hello” message to his neighbors and selects the signal received with the largest SNR, or the shortest arrival time, to be its closest neighbor.

The outage probability expression, which we refer to as \mathcal{P}_{ONN} , for given source-relay-destination locations is still given by (3.17). To find the total probability, we need to average over all possible locations of the user and the relay. The user’s location distribution with respect to the BS/AP is still given as in the direct transmission case (3.1). The relay’s location distribution, however, is not uniform. In the sequel we calculate the probability density function of the relay’s location. According to our protocol, the relay is chosen to be the nearest neighbor to the user. The probability that the nearest neighbor is at distance r_{sl} from the source is equivalent to calculating the probability that the shaded area in Fig. 3.3 is empty.

Denote this area, which is the intersection of the two circles with centers s and d , by $A(r_{sd}, r_{sl})$. For $0 < r_{sl} \leq \rho - r_{sd}$, the area of intersection is a circle with radius r_{sl} and center s . The probability density function of r_{sl} , $p_{r_{sl}}(x)$, can be calculated as

$$\begin{aligned} p_{r_{sl}}(x) &= \frac{\partial}{\partial x} (1 - \mathcal{P}(r_{sl} > x)) \\ &= \frac{\partial}{\partial x} \left(1 - \left(1 - \frac{x^2}{\rho^2} \right)^{N-1} \right) \\ &= \frac{2(N-1)x}{\rho^2} \left(1 - \frac{x^2}{\rho^2} \right)^{N-2}, \quad 0 < r_{sl} \leq \rho - r_{sd}. \end{aligned} \quad (3.27)$$

For $\rho - r_{sd} < r_{sl} \leq \rho + r_{sd}$, the intersection between the two circles can be divided into three areas: 1) the area of the sector acb in circle s ; 2) area of the triangle asb ; 3) area enclosed by the chord ab in circle d . Hence the intersection area, denoted by $A(r_{sd}, r_{sl})$ can be written as

$$A(r_{sd}, r_{sl}) = r_{sl}^2 \theta + \frac{1}{2} r_{sl}^2 \sin(2\theta) + \left(\rho^2 \phi - \frac{1}{2} \rho^2 \sin(2\phi) \right) \quad (3.28)$$

where $\theta = \cos^{-1} \left(\frac{\rho^2 - r_{sl}^2 - r_{sd}^2}{2r_{sl}r_{sd}} \right)$, and $\phi = \cos^{-1} \left(\frac{r_{sl}^2 - \rho^2 - r_{sd}^2}{2\rho r_{sd}} \right)$. The probability density function for r_{sl} for this range is given by

$$p_{r_{sl}}(x) = \frac{\partial}{\partial x} \left(1 - \left(1 - \frac{A(r_{sd}, r_{sl})}{\pi \rho^2} \right)^{N-1} \right), \quad \rho - r_{sd} < r_{sl} \leq \rho + r_{sd}. \quad (3.29)$$

This completely defines the probability density function for the nearest neighbor and the average can be found numerically as the integrations are extremely complex. In the sequel, we derive an approximate expression for the outage probability under the following two assumptions. Since the relay is chosen to be the nearest neighbor to the source, the SNR received at the relay from the source is rarely below the threshold γ_c , hence, we assume that the event of the relay being in outage is negligible. The second assumption is that the nearest neighbor always

lies on the intersection of the two circles, as points a or b in Fig. 3.3. This second assumption is a kind of worst case scenario, because a relay at distance r_{sl} from the source can be anywhere on the arc \widehat{acb} , and a worst case scenario is to be at points a or b . This simplifies the outage calculation as the conditional outage probability (3.17) is now only a function of the source distance r_{sd} as follows

$$\mathcal{P}_{ONN}(r_{sd}) \simeq 1 - f(\gamma_c, r_{sd}) - \frac{N_o\gamma_c}{KP} r_{sd}^\gamma f(\gamma_c, r_{sd}). \quad (3.30)$$

Averaging (3.30) over the user distribution (3.1) and using the definition of the incomplete Gamma function in (3.10), we get

$$\mathcal{P}_{ONN} \simeq 1 - \frac{2}{\alpha\rho^2} \left(\frac{KP}{N_o\gamma_c}\right)^{\frac{2}{\alpha}} \Gamma\left(\frac{2}{\alpha}, \frac{N_o\gamma_c\rho^\alpha}{KP}\right) - \frac{2}{\alpha\rho^2} \left(\frac{KP}{N_o\gamma_c}\right)^{\frac{2}{\alpha}} \Gamma\left(\frac{2}{\alpha} + 1, \frac{N_o\gamma_c\rho^\alpha}{KP}\right) \quad (3.31)$$

Using the same approximation as above, the conditional outage probability for the no-MRC case is given by

$$\mathcal{P}_{ONN}(r_{sd}) = (1 - f(\gamma_c, r_{sd}))^2. \quad (3.32)$$

Averaging the above expression over r_{sd} we have

$$\mathcal{P}_{ONN} \simeq 1 - \frac{4}{\alpha\rho^2} \left(\frac{KP}{N_o\gamma_c}\right)^{\frac{2}{\alpha}} \Gamma\left(\frac{2}{\alpha}, \frac{N_o\gamma_c\rho^\alpha}{KP}\right) + \frac{2}{\alpha\rho^2} \left(\frac{KP}{2N_o\gamma_c}\right)^{\frac{2}{\alpha}} \Gamma\left(\frac{2}{\alpha}, \frac{2N_o\gamma_c\rho^\alpha}{KP}\right). \quad (3.33)$$

3.3.2 Fixed Relays Strategy

In some networks, it might be easier to deploy fixed nodes in the cell to act as relays. This will reduce the overhead of communications between users to pair for cooperation. Furthermore, in wireless networks users who belong to different authorities might act selfishly to maximize their own gains, i.e., selfish nodes. For such scenarios protocols for enforcing cooperation or to introduce incentives for

the users to cooperate need to be implemented. In the subsection, we propose deploying nodes in the network that act as relays and do not have their own information. Each user will be associated with one relay to help in forwarding the dropped packets. The user can select the closets relay, which can be implemented using the exchange of "Hello" messages and selecting the signal with shortest arrival time, for example.

Continuing with our circular model for the cell, with uniform users distribution, the relays are deployed uniformly by dividing the cell into a finite number m of equal sectors, equal to the number of fixed relays to be deployed. Fig. 3.4 depicts a network example for $m = 3$. The relays are deployed at a distance r_{ld} from the destination. This distance should be designed to minimize the average outage probability as follows

$$r_{ld}^* = \arg \min \mathcal{P}_{OC}(r_{ld}), \quad \text{s.t. } 0 < r_{ld} < \rho \quad (3.34)$$

where the average outage probability $\mathcal{P}_{OC}(r_{ld})$ is defined as

$$\mathcal{P}_{OC} = \int_0^\rho \frac{2l_{sd}}{\rho^2} \int_{-\frac{\pi}{m}}^{\frac{\pi}{m}} \mathcal{P}_{OC}(r_{sd}, r_{sl}(\theta), r_{ld}) \frac{m}{2\pi} d\theta dl_{sd} \quad (3.35)$$

where $\mathcal{P}_{OC}(l_{sd}, l_{sr}, l_{rd})$ is defined in (3.17), and the distance from the source to the fixed relay is given by

$$r_{sl}(\theta) = \sqrt{r_{sd}^2 + r_{ld}^2 - 2r_{sd}r_{ld} \cos(\theta)}, \quad (3.36)$$

where θ is uniformly distributed between $[-\frac{\pi}{m}, \frac{\pi}{m}]$. Solving the above optimization problem is very difficult, hence, we are going to consider the following heuristic. In the Nearest-Neighbor protocol, the relay was selected to be the nearest neighbor to the user. Here, we are going to calculate the relay position that minimizes the mean square distance between the users in the sector and the relay. Without loss of

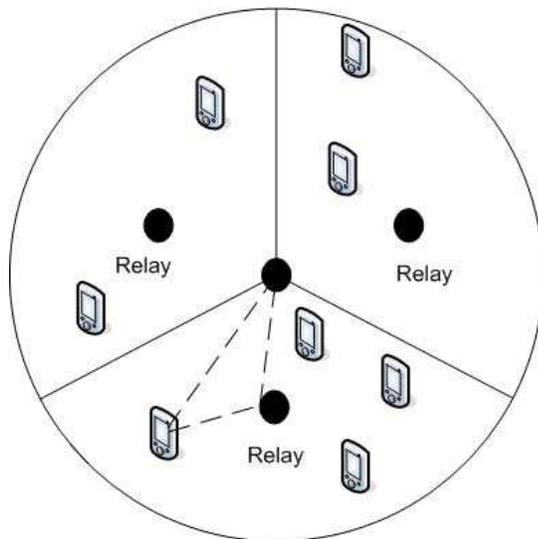


Figure 3.4: Illustrating cooperation under nearest neighbor protocol: The nearest neighbor is at a distance r_{sl} from the source. Therefore, the shaded area should be empty from any users.

generality, assuming the line dividing the sector to be the x-axis, the mean square distance between a user at distance r and angle θ from the center of the cell and the relay is given by the following function

$$q(r_{ld}) = \text{E} (\|r e^{j\theta} - r_{ld}\|^2), \quad (3.37)$$

where $j = \sqrt{-1}$, and E denotes the joint statistical expectation over the random variables r and θ . Solving for the optimal r_{ld} that minimizes $q(r_{ld})$

$$r_{ld}^* = \arg \min \text{E} (\|r e^{j\theta} - r_{ld}\|^2) \quad (3.38)$$

we get

$$r_{ld}^* = \frac{2m}{3\pi} \sin\left(\frac{\pi}{m}\right) \rho. \quad (3.39)$$

Fig. 3.5 depicts the average outage probability versus the number of relays deployed in the network for different cell sizes. The numerical results are for the

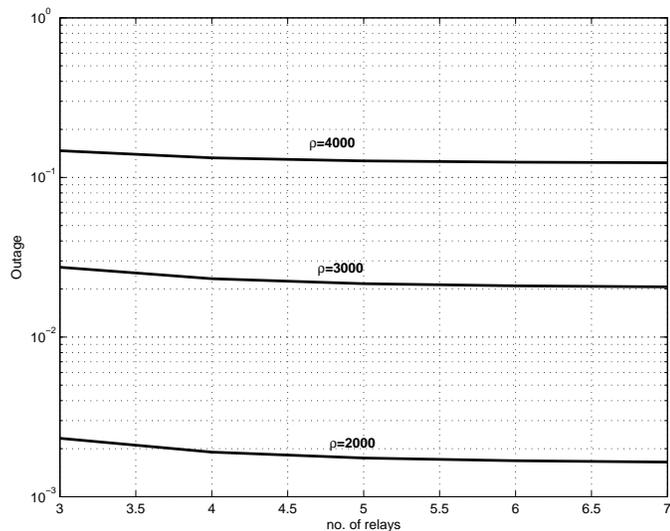


Figure 3.5: Average outage probability versus the number of relays in fixed relaying.

following parameters: $K = 1$, $\alpha = 3$, $P = 0.05$, $R = 1$, and $N_o = 10^{-12}$. We can see from the results that the performance saturates at approximately $m = 6$ relays, which suggests that dividing the cell into 6 sectors with a relay deployed in each sector can provide good enough performance.

3.3.3 Lower Bound on the Performance: The Genie-aided algorithm:

For both the MRC and no-MRC case, we determined the optimal relay location. For the MRC case, the optimal relay position is towards the source and can be determined according to (3.24). For the no-MRC case, we showed that the optimal relay position is in the mid-point between the source and the destination. We will drive a lower bound on the outage probability for any relay assignment protocol based on a Genie aided approach. This bound serves as a benchmark for the perfor-

mance of the Nearest-Neighbor protocol, and the fixed-relaying scheme proposed in the chapter. The Genie-Aided protocol works as follows. For any source node in the network, a Genie is going to put a relay at the optimal position on the line joining this source node and the destination (BS/AP).

Next we analyze the average outage performance of the Genie-Aided protocol. For the MRC case, substituting the optimal relay position in (3.24) in the conditional outage expression in (3.17) we get

$$\mathcal{P}_{OC}(r_{sd}) = 1 - f(\gamma_c, r_{sd}) - \frac{1}{1 - x^*} f(\gamma_c, (1 - x^*)r_{sd}) (f(\gamma_c, x^*r_{sd}) - f(\gamma_c, r_{sd})). \quad (3.40)$$

Averaging the above expression over the user distribution, the average outage probability for the Genie-aided lower bound for the MRC case $I = 1$ is given by

$$\begin{aligned} \mathcal{P}_{OG,1} = & 1 - \frac{2}{\alpha\rho^2} \left(\frac{KP}{N_o\gamma_c} \right)^{\frac{2}{\alpha}} \Gamma \left(\frac{2}{\alpha}, \frac{N_o\gamma_c\rho^\alpha}{KP} \right) \\ & - \frac{1}{1 - (x^*)^\alpha} \frac{2}{\alpha\rho^2} \left(\frac{KP}{N_o\gamma_c((1 - x^*)^\alpha + (x^*)^\alpha)} \right)^{\frac{2}{\alpha}} \Gamma \left(\frac{2}{\alpha}, \frac{((1 - x^*)^\alpha + (x^*)^\alpha) N_o\gamma_c\rho^\alpha}{KP} \right) \\ & + \frac{1}{1 - (x^*)^\alpha} \frac{2}{\alpha\rho^2} \left(\frac{KP}{N_o\gamma_c((1 - x^*)^\alpha + 1)} \right)^{\frac{2}{\alpha}} \Gamma \left(\frac{2}{\alpha}, \frac{((1 - x^*)^\alpha + 1) N_o\gamma_c\rho^\alpha}{KP} \right) \end{aligned} \quad (3.41)$$

We will denote the average probability of outage for the no-MRC case by $\mathcal{P}_{OG,2}$. Substituting the optimal relay position r_{sl}^* in the conditional outage expression (3.17), we get

$$\mathcal{P}_{OG}(r_{sd}) = \left(1 - \exp\left(-\frac{N_o\gamma_c r_{sd}^\alpha}{KP}\right) \right) \left(1 - \exp\left(-\frac{2N_o\gamma_c \left(\frac{r_{sd}}{2}\right)^\alpha}{KP}\right) \right). \quad (3.42)$$

Averaging the above expression over all possible users' locations,

$$\begin{aligned} \mathcal{P}_{OG,2} = & 1 + \frac{2}{\alpha\rho^2} \left(\frac{kP}{N_o\gamma_c(1+2^{1-\alpha})} \right)^{\frac{2}{\alpha}} \Gamma\left(\frac{2}{\alpha}, \frac{N_o\gamma_c(1+2^{1-\alpha})\rho^\alpha}{kP}\right) \\ & - \frac{2}{\alpha\rho^2} \left(\frac{kP}{N_o\gamma_c} \right)^{\frac{2}{\alpha}} \Gamma\left(\frac{2}{\alpha}, \frac{N_o\gamma_c\rho^\alpha}{kP}\right) - \frac{2}{\alpha\rho^2} \left(\frac{kP}{N_o\gamma_c2^{1-\alpha}} \right)^{\frac{2}{\alpha}} \Gamma\left(\frac{2}{\alpha}, \frac{N_o\gamma_c2^{1-\alpha}\rho^\alpha}{kP}\right). \end{aligned} \quad (3.43)$$

3.4 Numerical Results

We performed some computer simulations to compare the performance of the proposed relay-assignment protocols and validate the theoretical results we derived in this chapter. In all of our simulations, we compared the outage performance of three different transmission schemes: Direct transmission, Nearest-Neighbor protocol, and fixed-relaying. In all of the simulations, the channel fading between any two nodes (either a user and the BS/AP or two users) is modeled as a random Rayleigh fading channel with unit variance.

For fairness in comparison between the proposed cooperative schemes and the direct transmission scheme, the spectral efficiency is kept fixed in both cases and this is done as follows. Since a packet is either transmitted once or twice in the cooperative protocol, the average rate in the cooperative case can be calculated as

$$E(R_c) = R_c\mathcal{P}_{OD,\gamma_c}(r_{sd}) + \frac{R_c}{2}\mathcal{P}_{OD,\gamma_c}(r_{sd}). \quad (3.44)$$

where R_c is the spectral efficiency in b/s/Hz for cooperative transmission, and $\mathcal{P}_{OD,\gamma_c}(r_{sd})$ denotes the outage probability for the direct link at rate R_c . In (3.44), note that one time slot is utilized if the direct link is not in outage, and two time slots are utilized if it is in outage. Note that the later scenario is true even if the relay does not transmit because the time slot is wasted anyway. Averaging over

the source destination separation, the average rate is given by

$$\bar{R}_c = \frac{R_c}{2} \left(1 + \frac{2}{\alpha \rho^2} \left(\frac{KP}{\gamma_c N_o} \right)^{\frac{2}{\alpha}} \Gamma \left(\frac{2}{\alpha}, \frac{N_o \gamma_c \rho^\alpha}{KP} \right) \right) \quad (3.45)$$

We need to calculate the SNR threshold γ_c corresponding to transmitting at rate R_c . The resulting SNR threshold γ_c should generally be larger than γ_{nc} required for non-cooperative transmission. It is in general very difficult to find an explicit relation between the SNR threshold γ_c and the transmission rate R_c , and thus we render to a special case to capture the insights of this scenario. Let the outage be defined as the event that the mutual information I between two terminals is less than some specific rate R [31]. If the transmitted signals are Gaussian, then according to our channel model, the mutual information is given by $I = \log(1 + \text{SNR}_{sd})$. The outage event for this case is defined as

$$O_I \triangleq \{h_{sd} : I < R\} = \{h_{sd} : \text{SNR}_{sd} < 2^R - 1\}. \quad (3.46)$$

The above equation implies that if the outage is defined in terms of the mutual information and the transmitted signals are Gaussian, then the SNR threshold γ_c and the spectral efficiency R are related as $\gamma_c = 2^R - 1$, i.e., they exhibit an exponential relation.¹ For the sake of comparison \bar{R}_c should be equal to R , the spectral efficiency of direct transmission. Thus for a given R one should solve for R_c . This can lead to many solutions for R_c , and we are going to choose the minimum R_c [22]

¹Intuitively, under a fixed modulation scheme and fixed average power constraint, one can think of the SNR threshold as being proportional to the minimum distance between the constellation points, which in turn depends on the number of constellation points for fixed average power, and the later has an exponential relation to the number of bits per symbol that determines the spectral efficiency R .

In the following simulation comparisons, we study the outage probability performance when varying three basic quantities in our communication setup: the transmission rate, the transmit power, and the cell radius. In all the scenarios, we consider direct transmission, nearest-neighbor, fixed relaying with 6 relays deployed in the network, and the Genie-aided lower bound. For all the cooperative transmission cases, both MRC and no-MRC is examined.

Fig. 3.6 depicts the outage probability versus the transmit power in dBW. It is clear from the slopes of the curves that cooperation yields more steeper curves due to the diversity gain. Fixed relaying with MRC has the best performance, and it is very close to the Genie lower bound with no MRC. Fixed relay has generally better performance than nearest-neighbor protocols. Cooperation yields around 7dbW savings in the transmit power with respect to direct transmission

Fig. 3.7 depicts the outage probability curves versus the cell radius. Fixed relaying also has the best performance. There is a 200% increase in the cell radius at an 0.01 outage. We can see that the gap between direct transmission and cooperation decrease with increasing the cell size. The rationale here is that with increasing the cell size, the probability of packets in outage increases, and hence, the probability that the relay will forward the source's packet increases. This reduces the bandwidth efficiency of the system, and hence increases the overall outage probability. This tradeoff between the spectral efficiency and the diversity gain of cooperation makes direct transmission good enough for larger cell sizes.

Similar conclusions can be drawn out of Fig. 3.8, which plots the outage probability versus the spectral efficiency.

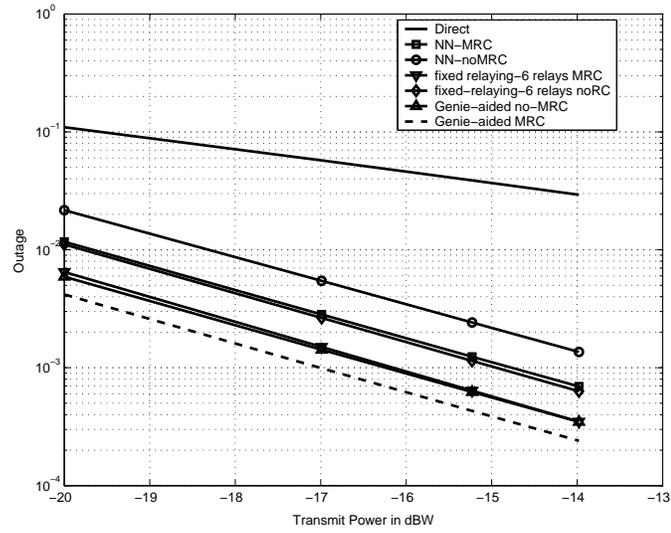


Figure 3.6: Average outage probability versus the transmit power.

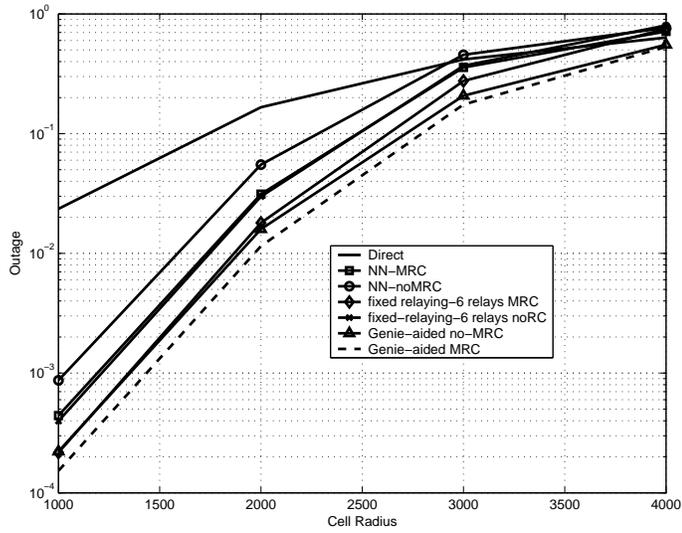


Figure 3.7: Average outage probability versus the cell radius.

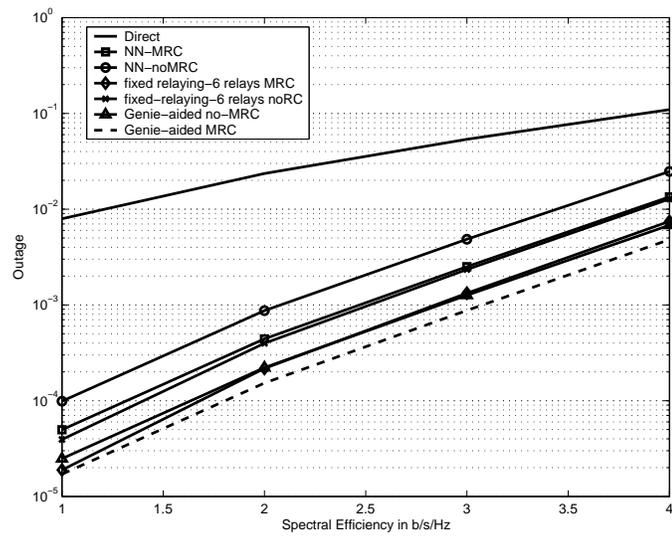


Figure 3.8: Average outage probability versus the spectral efficiency.

Chapter 4

Cognitive Cooperative

Multiple-Access

Despite the promised gains of cooperative communication demonstrated in the previous chapters, the impact of cooperation on higher network levels is not completely understood yet. Most of the previous work on cooperation assume the user has always a packet to transmit which is not generally true in a wireless network. For example in a network, most of the sources are bursty in nature which leads to periods of silence in which the users may have no data to transmit. Such a phenomenon may affect important system parameters that are relevant to higher network layers, for example, buffer stability and packet delivery delay.

We focus on the multiple-access layer in this chapter. One can ask many important questions now. Can we design cooperation protocols taking these higher layer network features into account? Can the gains promised by cooperation at the physical layer be leveraged to the multiple-access layer? More specifically, what is the impact of cooperation on important multiple-access performance metrics such as stable throughput region and packet delivery delay? We try to address all of

these important questions to demonstrate the possible gains of cooperation at the multiple-access layer. A slotted time division multiple-access (TDMA) framework in which each time slot is assigned only to one terminal, i.e., orthogonal multiple-access is considered. If a user does not have a packet to transmit in his time slot, then this time slot is not utilized. These unutilized time slots are wasted channel resources that could be used to enhance the system performance. Recently, the concept of cognitive radio has been introduced to allow the utilization of unused channel resources by enabling the operation of a secondary system overlapping with the original system (see [59] and references therein).

We propose a novel cognitive multiple-access strategy with the concept of cooperation. In the proposed protocol the cognitive relay tries to “*smartly*” utilize the periods of source silence to cooperate with other terminals in the network, i.e., to increase the reliability of communications against random channel fades. In particular, when the relay senses the channel for empty time slots, the slots are then used to help other users in the network by forwarding their packets lost in some previous transmissions. Thus this new protocol has cooperative cognitive aspects in the sense that unused channel resources are being utilized by the relay to cooperate with other users in the network. It should be pointed out that the proposed cooperative protocol does not result in any bandwidth loss because there are no channel resources reserved for the relay to cooperate. We demonstrate later that this important feature of the proposed protocol can lead to significant gains especially in high spectral efficiency regimes.

We develop two protocols to implement this new cooperative cognitive multiple-access (CCMA) strategy. The first protocol is CCMA within a single frame (CCMA-S), where the relay keeps a lost packet no more than one time frame

and then drops the packet if it was not able to deliver it successfully to the destination. The dropped packet then has to be retransmitted by the originating user. It turns out that in this protocol the relay's queue is always bounded, and that the terminals queues are interacting. To analyze the stability of the system's queues we resort to a stochastic dominance approach. Analyzing the stability of interacting queues is a difficult problem that has been addressed for ALOHA systems initially in [60]. Later in [61], the dominant system approach was explicitly introduced and employed to find bounds on the stable throughput region of ALOHA with collision channel model. Many other works followed that to study the stability of ALOHA. In [62], necessary and sufficient conditions for the stability of a finite number of queues were provided, however, the stable throughput region was only explicitly characterized for a 3-terminals system. In [63], the authors provided tighter bounds on the stable throughput region for the ALOHA system using the concept of *stability ranks*, which was also introduced in the same paper. The stability of ALOHA systems under a multi-packet reception model (MPR) was considered in [64] and [65]. Characterizing the stable throughput region for interacting queues with $M > 3$ terminals is still an open problem. In CCMA-S, the interaction between the queues arise due to the role of the relay in enabling cooperation, which is different from the intrinsic cause of interaction in the ALOHA system. To analyze the stability of CCMA-S, we introduce a new dominant system to resolve this interaction.

The second protocol that we propose, named CCMA-Multiple-enhanced (CCMA-Me), differs in the way the relay handles the lost packets. In CCMA-Me, if a packet is captured by the relay and not by the destination, then this packet is removed from the corresponding user's queue and it becomes the relay's responsibility to

deliver it to the destination. The term enhanced refers to the design of the protocol, as the relay only helps the users with inferior channel gains reflected in the distances from the terminals to the destination. Different from CCMA-S, the size of the relay's queue can possibly be unbounded and so its stability must be taken into consideration.

In addition to characterizing the stable throughput region of our proposed protocols, we also analyze the queueing delay performance. Delay is an important performance measure and network parameter that may affect the tradeoff between rate and reliability of communication. Delay analysis for interacting queues is a notoriously hard problem that has been investigated in [66] and [67] for ALOHA. We consider a symmetric 2-users scenario when analyzing the delay performance of the proposed protocols.

We summarize the major contributions in this chapter in the following.

- Different from the main thrust of work on cooperative communications that focus on the physical layer, we investigate the impact of cooperation on higher network levels, specifically, the multiple-access layer. Moreover, the approach we are taking has a cross-layer nature as we consider physical layer parameters in our framework.
- We consider the intrinsic role of source burstiness that results in periods of silence and wasted channel resources in the design of our proposed multiple-access protocol. In particular, we propose a cognitive multiple-access strategy that enables cooperation during the unused channel resources. Thus the proposed cooperation protocol causes no bandwidth loss.
- We analyze the performance of our proposed protocol in terms of important network measures. In particular, we characterize the maximum stable

throughput region and the delay performance of the proposed protocols.

- We demonstrate that our proposed protocols have significantly better performance over TDMA without relaying, ALOHA, selection, and incremental decode-and-forward especially at high spectral efficiency regimes.
- This work demonstrates a very important fact that the design and development of cooperation protocols for wireless networks should have a broader view of different network layers as this cross-layer view can lead to significant performance gains.

Related work that study the impact of cooperation on the multiple-access layers are few. Cooperation in random access networks has been considered in [68], [69], [70]. In [69], the authors proposed a distributed version of network diversity multiple-access (NDMA) [71] protocol and they provided pairwise error probability analysis to demonstrate the diversity gain. In [68] and [70], the authors presented the notion of utilizing the spatial separation between users in the network to assign cooperating pairs (also groups) to each other. In [70], spread spectrum random access protocols were considered in which nearby inactive users are utilized to gain diversity advantage via cooperation assuming a symmetrical setup where all terminals are statistically identical. However, the previously cited works still focus on physical layer parameters as the diversity gains achieved and the outage probability. The work in this chapter presents, to the best of our knowledge, the first true investigation of the impact of cooperation on the multiple-access layer.

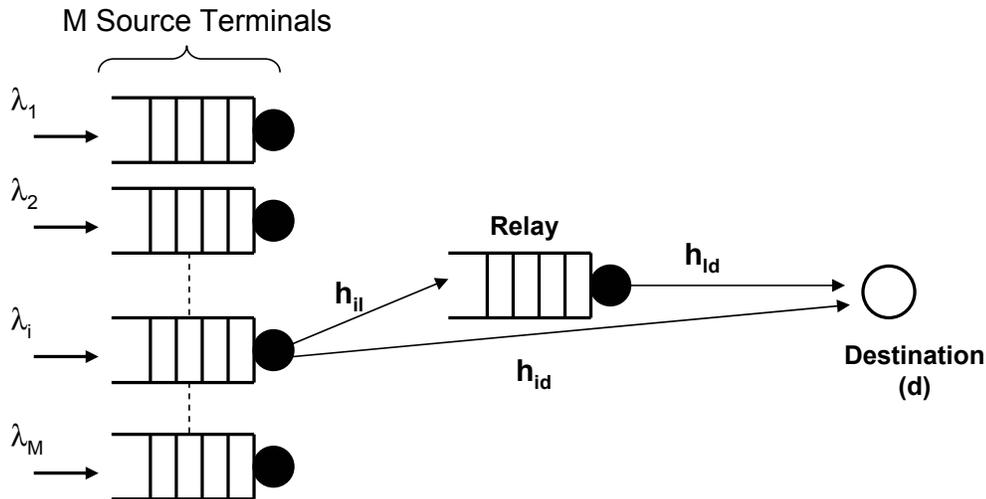


Figure 4.1: Network and channel Model.

4.1 System Model

We consider the uplink of a TDMA system. The network consists of a finite number $M < \infty$ of source terminals numbered $1, 2, \dots, M$, a relay node l ¹, and a destination node d , see Fig. 4.1. Let $\mathcal{T} = \{\mathcal{M}, l\}$ denote the set of transmitting nodes, where $\mathcal{M} = \{1, 2, \dots, M\}$ is the set of source terminals, and $\mathcal{D} = \{l, d\}$ denotes the set of receiving nodes or possible destinations. For simplicity of presentation, in the following we use terminal to refer to a source terminal.

First, we describe the queueing model for the multiple-access channel. Each of the M terminals and the relay l has an infinite buffer for storing fixed length packets. The channel is slotted, and a slot duration is equal to a packet duration. The arrival process at any terminal's queue is independent identically distributed (i.i.d.) from one slot to another, and the arrival processes are independent from one terminal to another. The arrival process at the i -th queue ($i \in \{1, 2, \dots, M\}$)

¹We use l to denote the relay not to confuse with r that denotes distance

is assumed stationary with mean λ_i . Terminals access the channel by dividing the channel resources, time in this case, among them, hence, each terminal is allocated a fraction of the time. Let $\Omega = [\omega_1, \omega_2, \dots, \omega_M]$ denote a resource-sharing vector, where $\omega_i \geq 0$ is the fraction of the time allocated to terminal $i \in \mathcal{M}$, or it can represent the probability that terminal i is allocated the whole time slot [72]. The later notation is used as it allows to consider fixed duration time slots with continuous values of the resource-sharing vector. A time frame is defined as M consecutive time slots. The set of all feasible resource-sharing vectors is specified as follows

$$F \triangleq \left\{ \Omega = [\omega_1, \omega_2, \dots, \omega_M] \in \mathfrak{R}_+^M : \sum_{i \in \mathcal{M}} \omega_i \leq 1 \right\}. \quad (4.1)$$

A fundamental performance measure of a communication network is the stability of its queues. Stability can be loosely defined as having a certain quantity of interest kept bounded. In our case, we are interested in the queue size and the packet delivery delay to be bounded. More rigourously, stability can be defined as follows. Denote the queue sizes of the transmitting nodes at any time t by the vector $\mathbf{Q}^t = [Q_i^t, i \in \mathcal{T}]$. We adopt the following definition of stability used in [62]

Definition 1 *Queue $i \in \mathcal{T}$ of the system is stable, if*

$$\lim_{t \rightarrow \infty} Pr [Q_i^t < x] = F(x) \quad \text{and} \quad \lim_{x \rightarrow \infty} F(x) = 1. \quad (4.2)$$

If $\lim_{x \rightarrow \infty} \liminf_{t \rightarrow \infty} Pr [Q_i^t < x] = 1$, the queue is called substable.

From the definition, if a queue is stable then it is also substable. If a queue is not substable, then it is unstable. An arrival rate vector $[\lambda_1, \lambda_2, \dots, \lambda_M]$ is said to be stable if there exists a resource-sharing vector $\Omega \in F$ such that all the queues in $\mathcal{T} = \{\mathcal{M}, l\}$ are stable. The multidimensional stochastic process \mathbf{Q}^t can be easily shown to be an irreducible and aperiodic discrete-time Markov process with

countable number of states and state space $\in \mathbb{Z}_+^{M+1}$. For such a Markov chain, the process is stable if and only if there exists a positive probability for every queue being empty [63], i.e.,

$$\lim_{t \rightarrow \infty} \Pr [Q_i(t) = 0] > 0, \quad i \in \mathcal{T}. \quad (4.3)$$

If the arrival and service processes of a queueing system are strictly stationary, then one can apply Loynes's theorem to check for stability conditions [73]. This theorem states that if the arrival process and the service process of a queueing system are strictly stationary, and the average arrival rate is less than the average service rate, then the queue is stable; if the average arrival rate is greater than the average service rate then the queue is unstable.

Next, we describe the physical channel model. The wireless channel between any two nodes in the network is modeled as a Rayleigh narrowband flat fading channel with additive Gaussian noise. The transmitted signal also suffers from propagation path loss that causes the signal power to attenuate with distance. The signal received at a receiving node $j \in \mathcal{D}$ from a transmitting node $i \in \mathcal{T}$ at time t can be modeled as

$$y_{ij}^t = \sqrt{Gr_{ij}^{-\gamma}} h_{ij}^t x_i^t + n_{ij}^t, \quad i \in \mathcal{T}, j \in \mathcal{D}, i \neq j, \quad (4.4)$$

where G is the transmitting power, assumed to be the same for all transmitting terminals, r_{ij} denotes the distance between the two nodes i, j , γ is the path loss exponent, and h_{ij}^t captures the channel fading coefficient at time t and is modeled as i.i.d. zero-mean, circularly-symmetric complex Gaussian random process with unit variance. The term x_i^t denotes the transmitted packet with average unit power at time t , and n_{ij}^t denotes i.i.d. additive white Gaussian noise with zero mean and variance N_o . Since the arrival, the channel gains, and the additive noise processes

are assumed stationary, we can drop the index t without loss of generality. We consider the scenario in which the fading coefficients are known to the appropriate receivers, but are not known at the transmitters.

In this chapter, we characterize the success and failure of packet reception by outage events and outage probability, which is defined as follows. For a target signal-to-noise (SNR) ratio β , if the received SNR as a function of the fading realization h is given by $\text{SNR}(h)$, then the outage event O is the event that $\text{SNR}(h) < \beta$, and $\Pr[\text{SNR}(h) < \beta]$ denotes the outage probability. This definition is equivalent to the *capture model* in [74], [75]. The SNR threshold β is a function of different parameters in the communication system; it is a function of the application, the data rate, the signal-processing applied at encoder/decoder sides, error-correction codes, and other factors. For example, varying the data rate and fixing all other parameters, the required SNR threshold β to achieve certain system performance is a monotonically increasing function of the data rate. Also, increasing the signal-processing and encode/decoder complexity in the physical layer reduces the required SNR threshold β for a required system performance.

For the channel model in (4.4), the received SNR of a signal transmitted between two terminals i and j can be specified as follows

$$\text{SNR}_{ij} = \frac{|h_{ij}|^2 r_{ij}^{-\gamma} G}{N_o}, \quad (4.5)$$

where $|h_{ij}|^2$ is the magnitude channel gain square and has an exponential distribution with unit mean. The outage event for a SNR threshold β is equivalent to

$$O_{ij} = \{h_{i,j} : \text{SNR}_{ij} < \beta\} = \{h_{i,j} : |h_{ij}|^2 < \frac{\beta N_o r_{i,j}^\gamma}{G}\} \quad (4.6)$$

Accordingly, the probability of outage is given by,

$$\Pr [O_{ij}] = \Pr \left[|h_{ij}|^2 < \frac{\beta N_o r_{i,j}^\gamma}{G} \right] = 1 - \exp \left(-\frac{\beta N_o r_{i,j}^\gamma}{G} \right), \quad (4.7)$$

where the above follows from the exponential distribution of the received SNR. Since we will use the above expression frequently in our subsequent analysis, and for compactness of representation, we will use the following notation to denote the success probability (no outage) at SNR threshold β

$$f_{ij} = \exp \left(-\frac{\beta N_o r_{i,j}^\gamma}{G} \right). \quad (4.8)$$

4.2 Cooperative Cognitive Multiple Access (CCMA) Protocols

In a TDMA system without relays, if a terminal does not have a packet to transmit, its time slot remains idle, i.e., wasted channel resources. We investigate the possibility of utilizing these wasted channel resources by employing a relay. In this section, we introduce our proposed cognitive multiple-access strategy based on employing relays in the wireless network. Furthermore, we develop two protocols to implement this new approach. We assume that the relay can sense the communication channel to detect empty time slots and we assume that the errors and delay in packet acknowledgement feedback is negligible.

First, we describe the new multiple-access strategy. Due to the broadcast nature of the wireless medium, the relay can listen to the packets transmitted by the terminals to the destination. If the packet is not received correctly by the destination, the relay stores this packet in its queue, given that it was able to decode this packet correctly. Thus, the relay's queue contains packets that have not been

transmitted successfully by the terminals. At the beginning of each time slot, the relay listens to the channel to check whether the time slot is empty (not utilized for packet transmission) or not. If the time slot is empty the relay will retransmit the packet at the head of its queue, hence utilizing this channel resource that was previously wasted in a TDMA system without a relay. Moreover, this introduces spatial diversity in the network as the channel fades between different nodes in the network are independent. In the following we develop two different protocols to implement the proposed cognitive multiple-access approach . The proposed protocol is cognitive in the sense that it introduces a relay in the network that tries detecting unutilized channel resources and use them to help other terminals by forwarding packets lost in previous transmissions.

4.2.1 CCMA-Single frame (CCMA-S)

The first protocol that we propose is cooperative cognitive multiple-access within a single frame duration or (CCMA-S). The characteristic feature of CCMA-S is that any terminal keeps its lost packet in its queue until it is captured successfully at the destination. CCMA-S operates according to the following rules.

- Each terminal transmits the packet at the head of its queue in its assigned slot, if the terminal's queue is empty the slot is free.
- If destination receives a packet successfully, it sends an ACK which can be heard by both the terminal and the relay. If the destination does not succeed in receiving the packet correctly but the relay does, then the relay stores this packet at the end of its queue. The corresponding terminal still keeps the lost packet at the head of its queue

- The relay senses the channel, and at each empty time slot the relay transmits the packet at the head of its queue, if its queue is nonempty. If the transmitted packet is received correctly by the destination it sends an ACK and the corresponding terminal removes this packet from its queue.
- If the relay does not succeed in delivering a packet to the destination during a time frame starting from the time it received this packet, then the relay drops the packet from its queue. In this case the corresponding terminal becomes responsible for delivering the packet to the destination.

Following are some important remarks on the above protocol. According to the above description of CCMA-S, the relay's queue has always a finite number of packets (at most has M backlogged packets). This follows because according to the protocol, the relay can have at most one packet from each terminal. Thus the stability of the system is only determined by the stability of the terminals' queues. Secondly, successful service of a packet in a frame depends on whether the other terminals have idle time slots or not. Therefore individual terminals' queues are *interacting*.

4.2.2 CCMA-Multiple frames (CCMA-M)

In this section, we describe the implementation of protocol CCMA-M. The main difference between protocols CCMA-S and CCMA-M is in the role of the relay and the behavior of the terminals' regarding their backlogged packets. More specifically, a terminal removes a packet from its queue if it is received successfully by either the destination or the relay. CCMA-M operates according to the following rules.

- Each terminal transmits the packet at the head of its queue in its assigned time slot. If the queue is empty the time slot is free.

- If a packet is received successfully by either the destination or the relay, the packet is removed from the terminal's queue (the relay needs to send an ACK if one is not heard by the destination in this case).
- If a packet is not received successfully by both the relay and the destination, the corresponding terminal retransmits this packet in its next assigned time slot.
- At each sensed empty time slot, the relay retransmits the packet at the head of its queue.

One can now point out the differences between the queues in system CCMA-S and CCMA-M: i) The size of the relay's queue can possibly grow in CCMA-M as it can have more than one packet from each terminal, however, it can not exceed size M in CCMA-S; ii) The terminal's queues in CCMA-M are not interacting as in CCMA-S. This is because the terminal removes the packets which were received correctly by the relay or the destination. In other words, servicing the queue of any terminal depends only on the channel conditions from that terminal to the destination and relay, and does not depend on the status of the other terminals' queues; iii) The stable throughput region of CCMA-M requires studying the stability of both the terminals' queues and the relay's queue.

4.3 Stability Analysis

The aim of this section is to characterize the stable throughput region of the proposed cooperation protocols. Furthermore, we compare our results against the stable throughput regions of TDMA without relaying, ALOHA, selection decode-and-forward, and incremental decode-and-forward.

4.3.1 Stability Analysis of CCMA-S

The 2-Terminal Case

In CCMA-S, we observed in the previous section that the relay's queue size is always finite, hence it is always stable. For the 2-terminal case, the queues evolve as a two dimensional Markov-chain in the first quadrant. From the protocol description in the previous Section, one can observe that the system of queues in CCMA-S are interacting. In other words, the transition probabilities differ according to whether the size of the queues are empty or not. For example, if one of the two terminals' queues was empty for a long time, then the relay serves the lost packets from the other terminal more often. On the other hand, if one of the two terminals queues never empties, then the other terminal will never get served by the relay. Studying stability conditions for interacting queues is a difficult problem that has been addressed for ALOHA systems [61], [63]. The concept of dominant systems was introduced and employed in [61] to help finding bounds on the stability region of ALOHA with collision channel. The dominant system in [61] was defined by allowing a set of terminals with no packets to transmit to continue transmitting dummy packets.

To analyze the stability of CCMA-S, we develop a dominant system to decouple the interaction of the terminals originating from the role of the relay in cooperation. Using our developed dominant system we are able to characterize the stability region of CCMA-S for a fixed resource sharing vector, and hence the whole stability region. The following Lemma states the stability region of CCMA-S for a fixed resource sharing vector $[\omega_1, \omega_2]$.

Lemma 1 *The stability region of CCMA-S for a fixed resource sharing vector*

$[\omega_1, \omega_2]$ is given by $\mathcal{R}(S_1) \cup \mathcal{R}(S_2)$ where

$$\mathcal{R}(S_1) = \{[\lambda_1, \lambda_2] \in R_+^2 : \lambda_2 < h(\lambda_1; w_1, w_2, f_{1d}, f_{2d}, f_{2l}, f_{1d}), \text{ for } \lambda_1 < \omega_1 f_{1d}\} \quad (4.9)$$

and

$$\mathcal{R}(S_2) = \{[\lambda_1, \lambda_2] \in R_+^2 : \lambda_1 < h(\lambda_2; w_2, w_1, f_{2d}, f_{1d}, f_{1l}, f_{1d}), \text{ for } \lambda_2 < \omega_2 f_{2d}\} \quad (4.10)$$

where $h(x; \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6) = \alpha_2 \alpha_4 + \alpha_1 \alpha_2 \left(1 - \frac{x}{\alpha_1 \alpha_3}\right) (1 - \alpha_3) \alpha_5 \alpha_6$.

Proof Lemma 1 The proof depends on constructing a dominant system that decouples the interaction between the queues and thus renders the analysis tractable. By dominance, we mean that the queues in the dominant system stochastically dominate the queues in the original CCMA-S system, i.e., with the same initial conditions for queue sizes in both the original and dominant systems, the queue sizes in the dominant system are not smaller than those in the original system [61].

We define the dominant system for CCMA-S as follows. For $j \in \{1, 2\}$, define S_j as

- Arrivals at queue $i \in \mathcal{M}$ in S_j is the same as CCMA-S,
- The channel realizations h_{kl} , where $k \in \mathcal{T}$ and $l \in \mathcal{D}$, for both S_j and CCMA-S are identical,
- Time slots assigned to user $i \in \mathcal{M}$ are identical for both S_j and CCMA-S,
- The noise generated at receiving ends of both systems are identical,
- The packets successfully transmitted by the relay for user j are not removed from user's j queue in S_j .

The above definition of the dominant system implies that queue j evolves exactly as in a TDMA system without a relay. If both the dominant system and the original CCMA-S started with the same initial queue sizes, then the queues in system S_j are always not shorter than those in CCMA-S. This follows because a packet successfully transmitted for queue j in S_j is always successfully transmitted from the corresponding queue in CCMA-S. However, the relay can succeed in forwarding some packets from queue j in CCMA-S in the empty time slots of the other terminal. This implies that queue j empties more frequently in CCMA-S and therefore the other terminal is better served in CCMA-S compared to S_j . Consequently, stability conditions for the dominant system S_j ($j \in \{1, 2\}$) are sufficient for the stability of the original CCMA-S system. In the following, we first derive the sufficient conditions for stability of CCMA-S.

Consider system S_1 in which the relay only helps terminal 2 and terminal 1 acts exactly as in a TDMA system. In order to apply Loynes' theorem, we require the arrival and service processes for each queue to be stationary. The queue size for terminal $i \in \{1, 2\}$ in system S_1 at time t , denoted by $Q_i^t(S_1)$, evolves as follows

$$Q_i^{t+1}(S_1) = (Q_i^t(S_1) - Y_i^t(S_1))^+ + X_i^t(S_1), \quad (4.11)$$

where $X_i^t(S_1)$ represents the number of arrivals in slot t and is a stationary process by assumption with finite mean $E[X_i^t(S_1)] = \lambda_i$. The function $()^+$ is defined as $(x)^+ = \max(x, 0)$. $Y_i^t(S_1)$ denotes the possible (*virtual*) departures from queue i at time t ; by virtual we mean that $Y_i^t(S_1)$ can be equal to 1 even if $Q_i^t(S_1) = 0$. We assume that departures occur before arrivals, and the queue size is measured at the beginning of the slot [61]. For terminal $i = 1$, the service process can be modeled as

$$Y_1^t(S_1) = \mathbf{1} \left[A_1^t \cap \overline{O_{1,d}^t} \right], \quad (4.12)$$

where $\mathbf{1}[\cdot]$ is the indicator function, A_1^t denotes the event that slot t is assigned to terminal 1, and $\overline{O_{1,d}^t}$ denotes the complement of the outage event between terminal 1 and the destination d at time t .² Due to the stationarity assumption of the channel gain process $\{h_{i,d}^t\}$, and using the outage expression in (4.7), the probability of this event is given by $\Pr[\overline{O_{1,d}^t}] = f_{1d}$. From the above, it is clear that the service process $Y_1^t(S_1)$ is stationary and has a finite mean given by $E[Y_1^t(S_1)] = \omega_1 f_{1d}$, where $E[\cdot]$ denotes statistical expectation. According to Loynes, stability of queue 1 in the dominant system S_1 is achieved if the following condition holds

$$\lambda_1 < \omega_1 f_{1d}. \quad (4.13)$$

Consider now queue 2 in system S_1 . The difference between the evolution of this queue and queue 1 is in the definition of the service process $Y_2^t(S_1)$. A packet from queue 2 can be served in a time slot in either one of the two following events: 1) If the time slot belongs to queue 2 and the associated channel $h_{2,d}^t$ is not in outage; or 2) the time slot belongs to queue 1, queue 1 is empty, in the previous time slot there was a successful "maybe virtual" reception of a packet at the relay from terminal 2, and the relay-destination channel is not in outage. This can be modeled as

$$Y_2^t(S_1) = \mathbf{1}\left[A_2^t \cap \overline{O_{2,d}^t}\right] + \mathbf{1}\left[A_1^t \cap \{Q_1^t(S_1) = 0\} \cap A_2^{t-1} \cap \overline{O_{2,l}^{t-1}} \cap O_{2,d}^{t-1} \cap \overline{O_{l,d}^t}\right] \quad (4.14)$$

where $\{Q_1^t(S_1) = 0\}$ denotes the event that terminal's 1 queue is empty in time slot t . The two indicator functions in the right hand side of equation (4.14) are

² $\overline{(\cdot)}$ denotes the complement of the event.

mutually exclusive, hence, the average rate of the service process is given by

$$E[Y_2^t(S_1^1)] = \omega_2 f_{2d} + \omega_1 \Pr[\{Q_1^t(S_1^1) = 0\}] \omega_2 (1 - \Pr[O_{2l}]) \Pr[O_{2d}] (1 - \Pr[O_{ld}]), \quad (4.15)$$

where $\Pr[O_{ij}]$ is the probability of outage between nodes i and j . Using Little's theorem [76] and (4.13), the probability of queue 1 empty is given by

$$\Pr[\{Q_1^t(S_1^1) = 0\}] = 1 - \frac{\lambda_1}{\omega_1 f_{1d}}. \quad (4.16)$$

Using the expression of the outage probability in (4.7) and Loynes conditions for stability [73], the stability condition for queue 2 in the dominant system S_1 is given by

$$\lambda_2 < \omega_2 f_{2d} + \omega_1 \omega_2 \left(1 - \frac{\lambda_1}{\omega_1 f_{1d}}\right) (1 - f_{2d}) f_{2l} f_{ld}. \quad (4.17)$$

Both conditions (4.13) and (4.17) represent the stability region for system S_1 for a specific resource-sharing vector (ω_1, ω_2) pair. Call this region $\mathcal{R}(S_1)$. Using parallel arguments for the dominant system S_2 , we can characterize the stability region $\mathcal{R}(S_2)$ for this system by the following pair of inequalities

$$\lambda_2 < \omega_2 f_{2d}, \quad \lambda_1 < \omega_1 f_{1d} + \omega_2 \omega_1 \left(1 - \frac{\lambda_2}{\omega_2 f_{2d}}\right) (1 - f_{1d}) f_{1l} f_{ld}. \quad (4.18)$$

Since stability conditions for a dominant system is sufficient for the stability of CCMA-S, any point inside the regions $\mathcal{R}(S_1)$ and $\mathcal{R}(S_2)$ can be achieved by the original system CCMA-S, hence $\mathcal{R}(S_1) \cup \mathcal{R}(S_2)$ is a subset from the stability region of CCMA-S for a fixed resource sharing pair (ω_1, ω_2) . This region is depicted in Fig. 4.2.

Up to this point we only proved the sufficient conditions for the stability of CCMA-S in the Lemma. To prove the necessary conditions, we follow a similar argument that was used by [61] and [64] for ALOHA systems to prove the

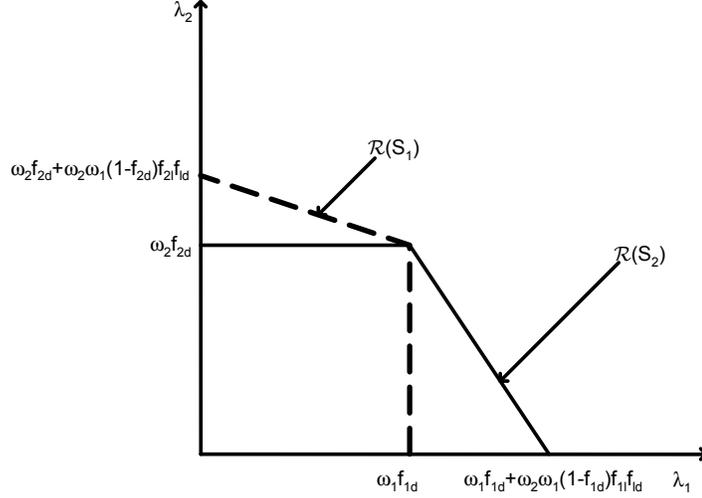


Figure 4.2: Stable throughput region for system CCMA-S for a fixed resource-sharing vector (ω_1, ω_2) given by $\mathcal{R}(S_1) \cup \mathcal{R}(S_2)$.

indistinguishability of the dominant and original systems at saturation. The argument is as follows. Consider the dominant system S_1 whose stability region is characterized by the pair of inequalities (4.13), (4.17). Note that if queue 2 does not empty, packets of user 1 are always dropped by the relay in both S_1 and CCMA-S, and both systems become identical. In S_1 , for $\lambda_1 < \omega_1 f_{1d}$, if $\lambda_2 > \omega_2 f_{2d} + \omega_1 \omega_2 \left(1 - \frac{\lambda_1}{\omega_1 f_{1d}}\right) (1 - f_{2d}) f_{2d} f_{1d}$ then using same argument as before Q_1^t is stable and Q_2^t is unstable by Loynes, i.e., $\lim_{t \rightarrow \infty} Q_2^t \rightarrow \infty$ almost surely. If Q_2^t tends to infinity almost surely, i.e., does not empty, then S_1 and CCMA-S are identical, and if both systems are started from the same initial conditions, then on a set of sample paths of positive probability Q_2^t in CCMA-S never returns to zero for $t \geq 0$. Hence Q_2^t in CCMA-S tends to infinity with positive probability, i.e., CCMA-S is also unstable. This means that the boundary for the stability region of the dominant system is also a boundary for the stability region of the original CCMA-S system. Thus, conditions for stability of the dominant system

is sufficient and necessary for stability of the original system. This completes the proof of Lemma 1 ■

The whole stability region for system CCMA-S can be determined by taking the union over all feasible resource-sharing vectors as follows.

$$\mathcal{R}(\text{CCMA-S}) = \bigcup_{\Omega \in \mathcal{F}} \left\{ \mathcal{R}_1(S^1) \cup \mathcal{R}_2(S^1) \right\}. \quad (4.19)$$

We give a complete characterization of the stability region of CCMA-S in the following Theorem.

Theorem 4 *The stability region for a 2-user CCMA-S system is given by*

$$\mathcal{R}(\text{CCMA-S}) = \{[\lambda_1, \lambda_2] \in R_+^2 : \lambda_2 < \max[g_1(\lambda_1), g_2(\lambda_1)]\} \quad (4.20)$$

where the functions $g_1(\cdot)$ and $g_2(\cdot)$ are defined as follows

$$g_1(\lambda_1) = \begin{cases} K_2 \left(\frac{\lambda_1 + f_{1d}}{2f_{1d}} - \frac{f_{2d}}{2K_2} \right)^2 - \frac{K_2 \lambda_1}{f_{1d}} + f_{2d}, & 0 \leq \lambda_1 \leq f_{1d} - \frac{f_{1d}f_{2d}}{K_2}, \\ f_{2d} - \frac{f_{2d}}{f_{1d}} \lambda_1, & f_{1d} - \frac{f_{1d}f_{2d}}{K_2} < \lambda_1 \leq f_{1d}. \end{cases} \quad (4.21)$$

And the function $g_2(\cdot)$ is specified as

$$g_2(\lambda_1) = \begin{cases} f_{2d} - \frac{f_{2d}}{f_{1d}} \lambda_1, & 0 \leq \lambda_1 < \frac{f_{1d}^2}{(1-f_{1d})f_{1l}f_{ld}}, \\ \frac{f_{1d}f_{2d}}{K_1} + f_{2d} - 2f_{2d} \sqrt{\frac{\lambda_1}{K_1}}, & \frac{f_{1d}^2}{(1-f_{1d})f_{1l}f_{ld}} \leq \lambda_1 \leq \lambda_1^*. \end{cases} \quad (4.22)$$

where

$$\lambda_1^* = \begin{cases} f_{1d}, & (1-f_{1d})f_{1l}f_{ld} < f_{1d}, \\ \frac{1}{4K_1}(f_{1d} + K_1)^2, & f_{1d} \leq (1-f_{1d})f_{1l}f_{ld}. \end{cases} \quad (4.23)$$

and $K_i = (1-f_{id})f_{il}f_{ld}$, $i \in \{1, 2\}$.

Proof Theorem 4 See Appendix A. ■

An interesting observation that we make from the above Theorem is that both functions $g_1(\lambda_1)$ and $g_2(\lambda_1)$ are linear on some part of their domain and strictly convex on the other part. It is not obvious however whether both functions are strictly convex over their domain of definition, and hence, whether the boundary of the stability region of CCMA-S given by $\max\{g_1, g_2\}$ is convex or not. In the following Lemma, we prove this property.

Lemma 2 *The boundary of the stability region of system CCMA-S given by $\max\{g_1, g_2\}$ is convex.*

Proof Lemma 2 See Appendix B. ■

The above Lemma will prove useful in characterizing the relation among the stability regions of the different multiple-access protocols considered in this chapter. The first relation that we state is that between the stability regions of TDMA and CCMA-S.

Lemma 3 *The stability region of TDMA is contained inside that of CCMA-S. In other words*

$$\mathcal{R}(TDMA) \subseteq \mathcal{R}(CCMA-S). \quad (4.24)$$

The two regions are identical if the following two conditions are satisfied simultaneously

$$(1 - f_{2d})f_{2l}f_{ld} < f_{2d}, \quad (1 - f_{1d})f_{1l}f_{ld} < f_{1d}. \quad (4.25)$$

Proof Lemma 3 We use Theorem 4 and Lemma 2 in the proof of the this Lemma. The stability region of TDMA is determined according to the following parametric inequalities

$$\lambda_1 \leq \omega_1 f_{1d}, \quad \lambda_2 \leq \omega_2 f_{2d} \quad (4.26)$$

or equivalently

$$\frac{\lambda_1}{f_{1d}} + \frac{\lambda_2}{f_{2d}} \leq 1. \quad (4.27)$$

From Lemma 2, both functions g_1 and g_2 that determines the boundary of the stability region of CCMA-S are convex. From the proof of the convexity, we note that the straight line $\lambda_2 = f_{2d} - \frac{f_{2d}}{f_{1d}}\lambda_1$ is a tangent for both functions, hence, it lies below both functions. Since this straight line is itself the boundary for the TDMA stability region, then the stability region of TDMA is a subset of that of CCMA-S. To prove the second part of the Lemma we use the definitions of the functions g_1, g_2 in Theorem 4. From (4.93), observe that if $(1 - f_{1d})f_{1l}f_{ld} < f_{1d}$ then the maximum λ_1 is determined by f_{1d} . If simultaneously $(1 - f_{2d})f_{2l}f_{2d} < f_{2d}$, then by substituting both conditions in the domain definitions of the functions g_1, g_2 , it can be seen that both functions reduce to

$$g_1(\lambda_1) = g_2(\lambda_1) = f_{2d} - \frac{f_{2d}}{f_{1d}}\lambda_1, \quad 0 \leq \lambda_1 \leq f_{1d}, \quad (4.28)$$

which is the boundary for the stability region of TDMA. Hence if both conditions in (4.25) are satisfied, CCMA-S and TDMA have the same stable throughput regions. This completes the proof of the Lemma. ■

The Symmetric M -Terminal Case

Stability analysis for the general M -terminal case is very complicated. For the ALOHA case, only bounds on the stability region have been derived [63], [64]. In this chapter, we only focus on the symmetric scenario. We define the dominant system for M -terminal CCMA-S as follows. For $1 \leq j \leq N$, define S_j^M as

- Arrivals at queue i in S_j^M is the same as CCMA,

- The channel realizations h_{kl} , where $k \in \mathcal{T}$ and $l \in \mathcal{R}$, for both S_j^M and CCMA are identical,
- User $i \in \mathcal{M}$ is assigned the same time slots in both systems,
- The noise generated at receiving ends of both systems are identical,
- The packets served by the relay for the first j terminals are not removed from these terminals' queues.

The last rule implies that the first j queues act as in a TDMA system without a relay. The relay, however, can help the other users $j + 1 \leq k \leq N$ in the empty slots of the TDMA frame.

Now consider system S_M^M in which the relay does not help any of the users. It is clear that the queue sizes in this system are never smaller than those in the original system CCMA-S. For S_M^M , the success probability of transmitting a packet is equal for all terminals and is given by

$$P_s(S_M^M) = \Pr[\text{SNR} \geq \beta] = f_{1d}. \quad (4.29)$$

The service rate per terminal is thus given by $\mu(S_M^M) = \frac{f_{1d}}{M}$, due to the symmetry of the problem. Since system S_M^M acts as a TDMA system without a relay, the queues are decoupled and hence the arrival process and departure process of each of them is strictly stationary. Applying Loynes theorem, the stability condition for S_M^M is given by

$$\lambda < f_{1d}, \quad (4.30)$$

where λ is the aggregate arrival rate for the M terminals.

Next, let us consider the stability of symmetric CCMA-S. Since S_M^M as described before dominates CCMA-S, if S_M^M is stable then CCMA-S is also stable. Therefore,

for $\lambda < f_{1d}$, system CCMA-S is stable. On the other hand, if all the queues in S_M^M are unstable, then none of these queues ever empty, hence, the relay loses its role and both systems CCMA-S and S_M^M are *indistinguishable* if both started with the same initial conditions. Therefore, if we have $\lambda > f_{1d}$ then all the queues in S_M^M are unstable and accordingly system CCMA-S is unstable as well. Therefore, the maximum stable throughput for system CCMA-S can be summarized in the following theorem.

Theorem 5 *The maximum stable throughput $\lambda_{MST}(CCMA-S)$ for system CCMA-S is equal to that of a TDMA system without a relay and is given by*

$$\lambda_{MST}(CCMA-S) = f_{1d}. \quad (4.31)$$

However, we conjecture that for the general asymmetric M -terminals scenario, the whole stability region of TDMA will be contained inside that of CCMA-S. Another important issue to point out is that although CCMA-S and TDMA have the same maximum stable throughput for the symmetrical case, the two systems do not have the same delay performance as will be discussed later.

4.3.2 Stability Analysis of CCMA with multiple frames (CCMA-M)

In CCMA-M the relay's queue can possibly grow and hence should be taken into account when studying the system stability. This means that for stability we require both the M terminals' queues and the relay's queue to be stable. The stability region of the whole system is the intersection of the stability regions of the M terminals and that of the relay. First, we consider the $M = 2$ -terminal case. According to the operation of system CCMA-M, a terminal succeeds in

transmitting a packet if either the destination or the relay receives this packet correctly. The success probability of terminal i in CCMA-M can thus be calculated as

$$\Pr[\text{Success of terminal } i] = \Pr[\overline{O_{i,l}} \cup \overline{O_{i,d}}], \quad (4.32)$$

where $\overline{O_{i,l}}$ denotes the event that the relay received the packet successfully, and $\overline{O_{i,d}}$ denotes the event that the destination received the packet successfully. The success probability of terminal $i \in \{1, 2\}$ in CCMA-M can thus be specified as follows

$$P_i = f_{id} + f_{il} - f_{id}f_{il}. \quad (4.33)$$

We first consider the stability region for the system determined just by the terminals' queues. Since for each queue $i \in \mathcal{M}$, the queue behaves exactly as in a TDMA system with the success probability determined by (4.33), the stability region $\mathcal{R}_{\mathcal{M}}(\text{CCMA-M})$ for the set of queues in \mathcal{M} is given by

$$\mathcal{R}_{\mathcal{M}}(\text{CCMA-M}) = \{[\lambda_1, \lambda_2, \dots, \lambda_M] \in R_+^M : \lambda_i < \omega_i P_i, \forall i \in \mathcal{M}, \text{ and } [\omega_1, \omega_2, \dots, \omega_M] \in F\}. \quad (4.34)$$

Next we study the stability of the relay's queue l . The evolution of the relay's queue can be modeled as

$$Q_l^t = (Q_l^t - Y_l^t)^+ + X_l^t, \quad (4.35)$$

where X_l^t denotes the number of arrivals at time slot t and Y_l^t denotes the possibility of serving a packet at this time slot from the relay's queue ($Y_l^t(G)$ takes values in $\{0, 1\}$). Now we establish the stationarity of the arrival and service processes of the relay. If the terminals' queues are stable, then by definition the departure processes from both terminals are stationary. A packet departing from a terminal queue is stored in the relay's queue (i.e., counted as an arrival) if simultaneously

the following two events happen: the terminal-destination channel is in outage and the terminal-relay channel is not in outage. Hence, the arrival process to the queue can be modeled as follows

$$X_l^t = \sum_{i \in \mathcal{M}} \mathbf{1} \left[A_i^t \cap \{Q_i^t \neq 0\} \cap O_{id}^t \cap \overline{O_{il}^t} \right]. \quad (4.36)$$

In (4.36), $\{Q_i^t \neq 0\}$ denotes the event that terminal's i queue is not empty, i.e., the terminal has a packet to transmit, and according to Little's theorem it has probability $\lambda_i/(\omega_i P_i)$, where P_i is terminal's i success probability and is defined in (4.33). The random processes involved in the above expressions are all stationary, hence, the arrival process to the relay is stationary. The expected value of the arrival process can be computed as follows

$$\lambda_l = \sum_{i \in \mathcal{M}} \lambda_i \frac{(1 - f_{id})f_{il}}{P_i}. \quad (4.37)$$

Similarly, we establish the stationarity of the service process from the relay's queue. The service process of the relay's queue depends by definition on the empty slots available from the terminals, and the channel from the relay to the destination being not in outage. By assuming the terminals' queues to be stable, they offer stationary empty slots (stationary service process) to the relay. Also the channel statistics is stationary, hence, the relay's service process is stationary. The service process of the relay's queue can be modeled as

$$Y_l^t = \sum_{i \in \mathcal{M}} \mathbf{1} \left[A_i^t \cap \{Q_i^t = 0\} \cap \overline{O_{id}^t} \right], \quad (4.38)$$

and the average service rate of the relay can be determined from the following equation

$$E[Y_l^t] = \sum_{i \in \mathcal{M}} \omega_i \left(1 - \frac{\lambda_i}{\omega_i P_i}\right) f_{ld}. \quad (4.39)$$

Using Loynes and equations (4.37) and (4.39), the stability region for the relay $\mathcal{R}_l(\text{CCMA-M})$ is determined by the condition $E[X_l^t] < E[Y_l^t]$. The total stability region for system CCMA-M is given by the intersection of two regions $\mathcal{R}_{\mathcal{M}}(\text{CCMA-M}) \cap \mathcal{R}_l(\text{CCMA-M})$ which is easily shown to be equal to $\mathcal{R}_l(\text{CCMA-M})$. The stability region for CCMA-M with 2 terminals is thus characterized as follows

$$\mathcal{R}(S^2) = \left\{ [\lambda_1, \lambda_2] \in R^{+2} : \frac{\lambda_1}{P_1} ((1 - f_{1d})f_{1l} + f_{ld}) + \frac{\lambda_2}{P_2} ((1 - f_{2d})f_{2l} + f_{ld}) < f_{ld} \right\} \quad (4.40)$$

For $M = 2$, this reveals that the stability region of CCMA-M is bounded by a straight line. Since the stability region for TDMA is also determined by a straight line, when comparing both stability regions it is enough to compare the intersection of these lines with the axes. These intersections for CCMA-M are equal to

$$\lambda_1^*(\text{CCMA-M}) = \frac{f_{ld}P_1}{f_{ld} + (1 - f_{1d})f_{1l}}, \quad \lambda_2^*(\text{CCMA-M}) = \frac{f_{ld}P_2}{f_{ld} + (1 - f_{2d})f_{2l}}, \quad (4.41)$$

while the corresponding values for TDMA are given by

$$\lambda_1^*(\text{TDMA}) = f_{1d}, \quad \lambda_2^*(\text{TDMA}) = f_{2d}. \quad (4.42)$$

It is clear that the stability region for TDMA is completely contained inside the stability region of CCMA-M if $\lambda_1^*(\text{CCMA-M}) > \lambda_1^*(\text{TDMA})$ and $\lambda_2^*(\text{CCMA-M}) > \lambda_2^*(\text{TDMA})$. Using (4.41) and (4.42), these two conditions are equivalent to

$$f_{ld} > f_{1d}, \quad f_{ld} > f_{2d}. \quad (4.43)$$

These conditions have the following intuitive explanation. If the channel between the relay and destination has higher success probability than the channel between the terminal and destination, then it is better to have the relay help the terminal transmit its packets. Note that (4.43) implies that TDMA can offer better performance for the terminal whose success probability does not satisfy (4.43). This

possible degradation in the performance does not appear in CCMA-S because the design of CCMA-S does not allow the relay to store the packets it received for ever. Hence, the performance of protocol CCMA-S can not be less than that of TDMA. In protocol CCMA-M, however, the relay becomes responsible for all of the packets it receives, and if the relay-destination channel has a higher probability of outage than the terminal-destination channel then the system encounters a loss in the performance. This calls for the development of an enhanced version of protocol CCMA-M that takes this into account.

Enhanced Protocol CCMA-Me

The previous discussion motivates the design of an *enhanced* version of CCMA-M, which we refer to as CCMA-Me. In this enhanced strategy, the relay only helps the terminals which are in worst channel condition than the relay itself. In other words, the relay helps the terminal whose outage probability to the destination satisfy $f_{id} > f_{rd}$ for $i \in \mathcal{M}$. Other terminals that do not satisfy this inequality operate as in TDMA, i.e., the relay does not help them.

Next we calculate the stability region for the enhanced system and consider $M = 2$ terminals for illustration. Assume that the relay only helps terminal 1. Similar to our calculations for the arrival and service processes for the relay in CCMA-M, we can show that the average arrival rate to the relay in CCMA-Me is given by

$$E[X_i^t(\text{CCMA-Me})] = \frac{\lambda_1}{P_1}(1 - f_{1d})f_{1l}, \quad (4.44)$$

and the average service rate to the relay is given by

$$E[Y_i^t(\text{CCMA-Me})] = \left(\omega_1 \left(1 - \frac{\lambda_1}{\omega_1 P_1}\right) + \omega_2 \left(1 - \frac{\lambda_2}{\omega_2 P_2}\right) \right) f_{1d}. \quad (4.45)$$

Using Loynes theorem [73] and equations (4.44) and (4.45), the stability region $\mathcal{R}(\text{CCMA-Me})$ is given by

$$\mathcal{R}(\text{CCMA-Me}) = \left\{ [\lambda_1, \lambda_2] \in R_+^2 : \frac{\lambda_1}{P_1} ((1 - f_{1d})f_{1l} + f_{1d}) + \lambda_2 \frac{f_{1d}}{f_{2d}} < f_{1d} \right\}. \quad (4.46)$$

The stability region for the enhanced protocol CCMA-Me is no less than the stability region of TDMA $\mathcal{R}(\text{TDMA}) \subseteq \mathcal{R}^{2,e}$, and the proof simply follows from the construction of the enhanced protocol CCMA-Me.

For a general M -terminal case, the analysis is the same and the stability region for CCMA-Me can be fully characterized as follows.

Theorem 6 *The stability region for M -terminals CCMA-Me is specified as*

$$\mathcal{R}(\text{CCMA-Me}) = \left\{ [\lambda_1, \lambda_2, \dots, \lambda_M] \in R_+^M : \sum_{i \in \mathcal{M}_1} \frac{\lambda_i}{P_i} ((1 - f_{id})f_{il} + f_{id}) + \sum_{j \in \mathcal{M}_2} \lambda_j \frac{f_{jd}}{f_{jd}} < f_{id} \right\}. \quad (4.47)$$

where $\mathcal{M}_1 = \{i \in \mathcal{M} : f_{id} > f_{id}\}$, or the set of terminals that the relay helps, and $\mathcal{M}_2 = \{i \in \mathcal{M} : f_{id} < f_{id}\}$ is the complement set.

We observe that the stability region of CCMA-Me is still bounded by a straight line. This follows because the stability of the system of queues in CCMA-Me is determined by the stability of a single queue, which is the relay's queue. It remains to specify the relation between the stability regions of CCMA-S and CCMA-Me, which is characterized in the following Theorem.

Theorem 7 *The stability region of CCMA-Me contains that of CCMA-S. In other words,*

$$\mathcal{R}(\text{TDMA}) \subseteq \mathcal{R}(\text{CCMA-S}) \subseteq \mathcal{R}(\text{CCMA-Me}). \quad (4.48)$$

Proof of Theorem 7: See Appendix C. ■

4.3.3 Existing Cooperation Protocols: Stability Analysis

In this subsection we discuss stability results for some existing decode-and-forward cooperation protocols. In particular, we consider the family of adaptive relaying proposed in [22], which comprises selection and incremental relaying. In the following, we discuss stability results for these two protocols and compare them to our proposed CCMA protocol.

Stability Region for Selection Decode-and-Forward

In this subsection, we characterize the stability region of selection decode-and-forward (SDF) described as follows. The cooperation is done in two phases. In the first phase, the source transmits and both the relay and the destination listen. In the second phase, if the relay is able to decode the signal correctly, then it is going to forward the received packet to the destination, otherwise the source retransmits the packet. Accordingly, there is always a specified channel resource dedicated for the relay to help the source, which is different from the opportunistic nature of cooperation in our proposed algorithms. Different from [22], we do not allow the destination to store analog signals in order to do maximum ratio combining (MRC). This is to have a fair comparison with all the protocols presented in this chapter which does not utilize MRC. Note that all of the results obtained in this chapter can be extended to the scenario where the destination saves copies of received signals and applies MRC, however, it would not add new insights to the results. For the sake of the analysis of SDF, we also assume that the channel fade changes independently from one time slot to another, which is different from the channel model in [22]. Note also that this assumption is in favor of SDF, and in general if the channel is correlated from one time slot to another, the performance

of SDF will degrade because of diversity loss.

In the following we analyze the outage probability for SDF under two scenarios. In the first scenario, the structure of the packets arrivals at the terminal is not allowed to be altered. Thus, each packet is transmitted in two consecutive time slots with the original spectral efficiency (for example using the same modulation scheme). This, however, results in SDF having half the bandwidth efficiency of TDMA and CCMA because each packet requires two time slots for transmission. In the second scenario, the bandwidth efficiency is preserved among all protocols. This can be done by allowing the terminal and the relay to change the structure of the incoming packets so that each of them transmit at twice the incoming rate (twice the spectral efficiency). Hence each packet is now transmitted in one time slot again, and the average spectral efficiency for SDF under this scenario is equal to that of TDMA and CCMA [22].

For the first scenario, an outage occurs if both the source-destination link and the source-relay-destination link are in outage. This can be specified as follows

$$\Pr_{SDF}(O) = \Pr \left[\left(\{\text{SNR}_{i,d} < 2\beta\} \cap \{\text{SNR}_{i,l} < 2\beta\} \cap \{\text{SNR}_{i,d} < 2\beta\} \right) \cup \left(\{\text{SNR}_{i,d} < 2\beta\} \cap \{\text{SNR}_{i,l} > 2\beta\} \cap \{\text{SNR}_{l,d} < 2\beta\} \right) \right] \quad i \in \mathcal{M} \quad (4.49)$$

The factor 2 in front of the SNR threshold β is to account to the fact that the transmitted power is divided by 2 in the first scenario to have the same energy per bit (Note that the bandwidth efficiency of SDF in the first scenario is half of that of CCMA, hence, we need to reduce the transmit power by half). The first term in the right-hand side of (4.49) corresponds to the event that both the source-destination and the source-relay link were in outage in the first time slot, and the source-destination link remained in outage in the second time slot. The

second term in (4.49) corresponds to the event that the source-destination link was in outage and the source-relay link was not in outage in the first time slot, but the relay-destination link was in outage in the second time slot. The probability in (4.49) can be expressed as

$$\Pr_{i,SDF}(O) = (1 - f_{id}^2)^2 (1 - f_{il}^2) + (1 - f_{id}^2) f_{il}^2 (1 - f_{ld}^2), \quad (4.50)$$

where f_{ij} is defined in (4.8). Since a single packet is transmitted in two time slots, one can think of this protocol as a modified TDMA system with the cooperation time slot has twice the length of the time slot in TDMA. The average arrival rate per cooperation time-slot is $2\lambda_i$ for $i \in \mathcal{M}$. Loynes condition for stability is given by

$$\frac{\lambda_1}{1 - \Pr_{1,SDF}(O)} + \frac{\lambda_2}{1 - \Pr_{2,SDF}(O)} < \frac{1}{2}, \quad (4.51)$$

where $1 - \Pr_{SDF}(O)$ is the success probability for terminal i .

For the second scenario, we need to calculate the SNR threshold β' corresponding to transmitting at twice the rate. The resulting SNR threshold β' should generally be larger than β required for transmission at the original rate. It is in general very difficult to find an explicit relation between the SNR threshold β and the transmission rate, and thus we render to a special case to capture the insights of this scenario. Let the outage be defined as the event that the mutual information I between two terminals is less than some specific rate R [31]. If the transmitted signals are Gaussian, then according to our channel model, the mutual information between terminal $i \in \mathcal{T}$ and terminal $j \in \mathcal{D}$ is given by $I = \log(1 + \text{SNR}_{ij})$. The outage event for this case is defined as

$$O_I \triangleq \{h_{ij} : I < R\}. \quad (4.52)$$

The above equation implies that if the outage is defined in terms of the mutual

information and the transmitted signals are Gaussian, then the SNR threshold β and the spectral efficiency R are related as $\beta = 2^R - 1$, i.e., they exhibit an exponential relation. Hence for protocol SDF when transmitting at twice the rate the corresponding SNR threshold β' is given by $\beta' = 2^{2R} - 1$, and given β one can find β' through the previous equation. Note that we do not reduce the power in this second scenario because both SDF and CCMA will have the same spectral efficiency.³

Stability Region for Incremental Decode-and-Forward

The second relaying strategy that we are considering in our comparison is incremental relaying. In such a strategy, feedback from the destination in the form of ACK or NACK is utilized at the relay node to decide whether to transmit or not. Amplify-and-forward incremental relaying was proposed in [22] in which the source transmits in the first phase, and if the destination was not able to receive correctly it sends a NACK that can be received by the relay. The relay then amplifies and forwards the signal it received from the source in the first phase. It can be readily seen that such a strategy is more bandwidth efficient than SDF because the relay only transmits if necessary.

In our comparison, we consider a modified version of the incremental relaying strategy proposed in [22]. In particular, we consider a decode-and-forward incremental relaying with selection capability at the relay (SIDF). In SIDF, the first

³Intuitively, under a fixed modulation scheme and fixed average power constraint, one can think of the SNR threshold as being proportional to the minimum distance between the constellation points, which in turn depends on the number of constellation points for fixed average power, and the latter has an exponential relation to the number of bits per symbol that determines the spectral efficiency R .

phase is exactly as amplify-and-forward incremental relaying. In the second phase, if the destination does not receive correctly then the relay, if it was able to decode the source signal correctly, forwards the re-encoded signal to the destination, otherwise the source retransmits again. One can think of this protocol as combining the benefits of selection and incremental relaying.

Next we analyze the outage probability of SIDF. As we did when studying SDF, we are also going to consider two scenarios for SIDF, namely, when the packet structure is not allowed to be changed and the scenario of equal spectral efficiency. First we consider the first scenario where the packet structure is not allowed to be changed. The spectral efficiency of SIDF in this case is less than TDMA or CCMA because the relay is occasionally allocated some channel resources for transmission with positive probability. Since both SDF and SIDF have the same mechanism for the outage event, it is readily seen that the outage event for SIDF is also given by (4.49) with the difference that we only use β in this case without the term 2 because SIDF will use the same transmit power⁴. The outage event is thus given by

$$\Pr_{SIDF}(O) = (1 - f_{id})^2 (1 - f_{il}) + (1 - f_{id}) f_{il} (1 - f_{id}). \quad (4.53)$$

The above expression represents the success probability of transmitting a packet in one or two consecutive time slots. Terminal i uses one time slot with probability f_{id} and two time slots with probability $1 - f_{id}$. The average number of time slots used by terminal i during a frame in SIDF is thus given by $2 - f_{id}$. The set of

⁴This is in favor of SIDF because the transmit power should be reduced to account for the reduction in the average spectral efficiency.

queues are not interacting in this case and the stability region is simply given by

$$\sum_{i \in \mathcal{M}} \frac{\lambda_i (2 - f_{id})}{1 - \Pr_{SIDF}(O)} < 1. \quad (4.54)$$

Next, we consider the second scenario of SIDF where the spectral efficiency is preserved for SIDF as for TDMA or CCMA. In this scenario, both the terminals and the relay will be transmitting at a higher rate \tilde{R} such that the average spectral efficiency is equal to the spectral efficiency R of TDMA or CCMA. The average spectral efficiency $R(SIDF)$ of SIDF when transmitting at a spectral efficiency \tilde{R} is given by

$$R(SIDF) = \tilde{R} \tilde{f}_{id} + \frac{\tilde{R}}{2} (1 - \tilde{f}_{id}) \triangleq v(\tilde{R}), \quad (4.55)$$

where \tilde{f}_{id} is the success probability for terminal's i -destination link when operating at spectral efficiency \tilde{R} , and we denote the whole function in the above expression by $v(\cdot)$. For the sake of comparison $R(SIDF) = v(\tilde{R})$ should be equal to R . Thus for a given R one should solve for $\tilde{R} = v^{-1}(R)$. This function can lead to many solutions for \tilde{R} , and we are going to choose the minimum \tilde{R} [22]. The stability region is thus given by

$$\sum_{i \in \mathcal{M}} \frac{\lambda_i}{1 - \Pr_{SIDF}(\tilde{O})} < 1. \quad (4.56)$$

where $\Pr_{SIDF}(\tilde{O})$ has the same form as $\Pr_{SIDF}(O)$ but evaluated at spectral efficiency \tilde{R} .

4.3.4 Numerical Results

We compare the stability regions of $M = 2$ -users TDMA, CCMA-S, CCMA-Me, the two forms of adaptive relaying proposed in [22] (selection and incremental relaying), and ALOHA as an example of random access.

In ALOHA, a terminal transmits a packet with some positive probability p if it has a packet to transmit. This means that there can be collisions among different terminals due to simultaneous transmissions in a time slot. The stability region for a general multipacket reception (MPR) ALOHA system was characterized in [64]. To make this chapter self-contained, we state the results from [64] here. We first introduce the notations used in [64]. Let $q_{i,i}$ denote the probability of successfully decoding a packet transmitted by user $i = 1, 2$ given that user i only transmitted, and $q_{i,\{1,2\}}$ denote the probability that user i 's packet is successfully decoded given that both users transmit. If $q_{i,i} \geq q_{i,\{1,2\}}$ then [64] characterizes the stability region of slotted ALOHA as $\mathcal{R}_1 \cap \mathcal{R}_2$ where

$$\mathcal{R}_1 \triangleq \{[\lambda_1, \lambda_2] \in R_+^2 : [\lambda_1, \lambda_2] \text{ lies below the curve } \lambda_2 = z(\lambda_1; q_{1,1}, q_{1,\{1,2\}}, Q_1, Q_2)\} \quad (4.57)$$

and

$$\mathcal{R}_2 \triangleq \{[\lambda_1, \lambda_2] \in R_+^2 : [\lambda_1, \lambda_2] \text{ lies below the curve } \lambda_1 = z(\lambda_2; q_{2,2}, q_{2,\{1,2\}}, Q_2, Q_1)\} \quad (4.58)$$

where $Q_1 = q_{1,1} - q_{1,\{1,2\}}$ and $Q_2 = q_{2,2} - q_{2,\{1,2\}}$. The function z is defined as follows [64]

$$z(\lambda; \alpha, \theta, \varepsilon, \delta) = \begin{cases} \theta - \frac{\lambda\delta}{\alpha - \varepsilon}, & \lambda \in \mathcal{I}_1, \\ \frac{(\sqrt{\alpha\theta}\sqrt{\lambda\delta})^2}{\varepsilon}, & \lambda \in \mathcal{I}_2 \end{cases} \quad (4.59)$$

where $\mathcal{I}_1 = \left[0, \frac{\theta(\alpha - \varepsilon)^2}{\alpha\delta}\right]$ and $\mathcal{I}_2 = \left[\frac{\theta(\alpha - \varepsilon)^2}{\alpha\delta}, \frac{\alpha\theta}{\delta}\right]$.

We are going to specify the parameters $q_{i,i}$ and $q_{i,\{1,2\}}$ for the capture channel [64] equivalent to the outage event defined in our work. It is clear that the success probability $q_{i,i}$ is equivalent to f_{id} in our notation. It remains to compute $q_{i,\{1,2\}}$,

$i = 1, 2$. The received signal model for $i = 1$ is

$$y_{1d}^t = \sqrt{Gr_{1d}^{-\gamma}}h_{1d}^t x_1^t + \sqrt{Gr_{2d}^{-\gamma}}h_{2d}^t x_2^t + n_d^t, \quad (4.60)$$

According to the capture model, a packet is captured if the received signal-to-interference-and-noise ratio (SINR) exceeds the threshold β . The SINR for terminal 1 is given by

$$\text{SINR}_1 = \frac{Gr_{1d}^{-\gamma} |h_{1d}|^2}{N_o + Gr_{2d}^{-\gamma} |h_{2d}|^2}. \quad (4.61)$$

The probability of user's 1 packet captured is given by

$$q_{1,\{1,2\}} = \Pr [\text{SINR}_1 > \beta] = \Pr \left[|h_{1d}|^2 > \frac{N_o r_{1d}^\gamma \beta}{G} + \beta \left(\frac{r_{1d}}{r_{2d}} \right)^\gamma |h_{2d}|^2 \right]. \quad (4.62)$$

The above equation can be easily computed and the success probability is given by

$$q_{1,\{1,2\}} = \frac{1}{1 + \beta \left(\frac{r_{1d}}{r_{2d}} \right)^\gamma} q_{1,\{1\}}. \quad (4.63)$$

Similar formula can be derived for $q_{2,\{1,2\}}$.

In Figs 4.3 and 4.4 we plot the stability regions for TDMA, CCMA-S, CCMA-Me, selection decode-and-forward (SDF), incremental decode-and-forward (SIDF), and ALOHA for a SNR threshold of $\beta = 35$, and $\beta = 64$, respectively. For SDF and SIDF we use the first scenario in which the packet structure is not changed. The parameters used to depict these results are as follows. The distances in meters between different terminals are given by $r_{1,d} = 120$, $r_{2,d} = 110$, $r_{l,d} = 40$, $r_{1,l} = 85$, $r_{2,l} = 80$. The propagation path-loss is given by $\gamma = 3.6$, the transmit power $G = 0.01$ watt, and $N_o = 10^{-11}$. In both Figs 4.3 and 4.4, CCMA-Me has the largest stable throughput region. In Fig. 4.3, CCMA-S and TDMA have identical stable throughput region, and it can be checked that conditions (4.25) are satisfied for $\beta = 35$. In Fig. 4.4, TDMA is contained inside CCMA-S. Both CCMA-S

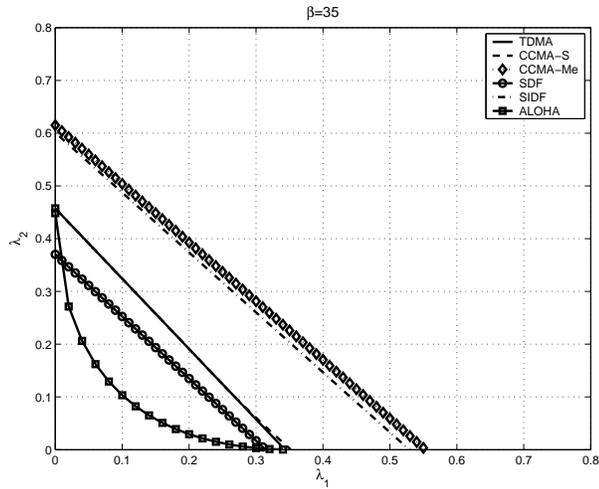


Figure 4.3: Stability regions for the different considered protocols at a SNR threshold equal to $\beta = 35$. For this value of β CCMA-S is equivalent to TDMA as depicted. CCMA-Me has the largest throughput region.

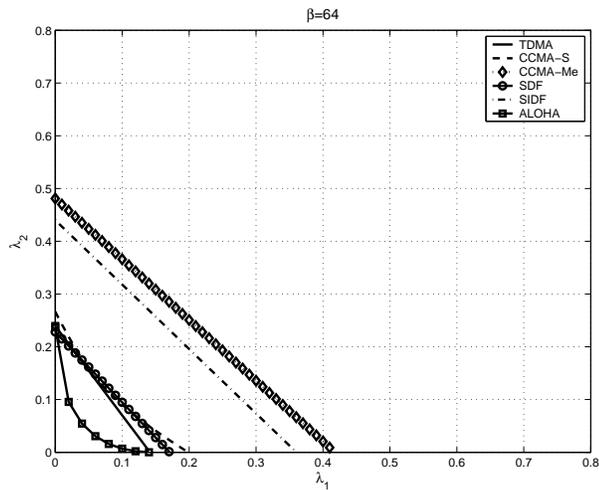


Figure 4.4: Stability regions for the different considered protocols at a SNR threshold equal to $\beta = 64$. TDMA is contained in CCMA-S, and the gap between SIDF and CCMA-Me increases in this case.

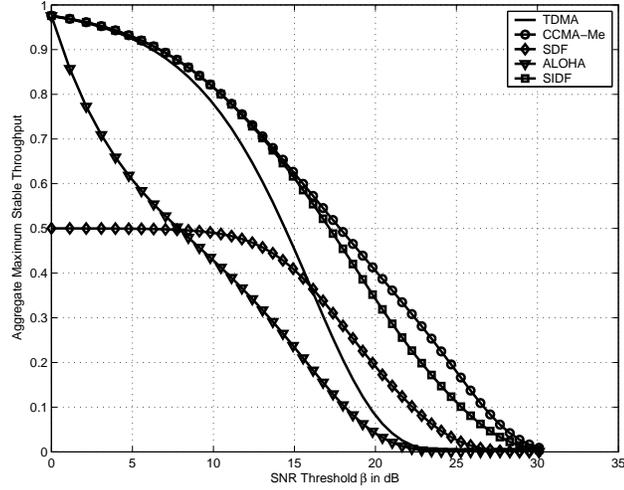


Figure 4.5: Aggregate maximum stable throughput versus SNR threshold β in dB. The propagation path loss is set to $\gamma = 3.5$. First scenario is used for SDF and SIDF. CCMA-Me has the best tradeoff curve among all the other protocols.

and CCMA-Me provide larger stable throughput region over SDF and ALOHA. This is because of the lost bandwidth efficiency in SDF and the interference in ALOHA due to collisions. SIDF is very close to CCMA-Me for smaller values of β , and the gap between them increases with increasing β as depicted in Fig. 4.4. This is because the bandwidth efficiency of SIDF reduces with increasing β which increases the probability of using a second time slot by the relay.

Next we demonstrate the tradeoff between the maximum stable throughput (MST) versus the SNR threshold β and the transmission rate R . For SDF and SIDF, we consider the two scenarios described in section 4.3.3, where in the first scenario the incoming packet structure is not changed and hence the two protocols have less bandwidth efficiency compared to TDMA and CCMA. While in the second strategy, the packet structure is changed to preserve the bandwidth efficiency. The maximum stable throughput results for the two scenarios are depicted in Figs

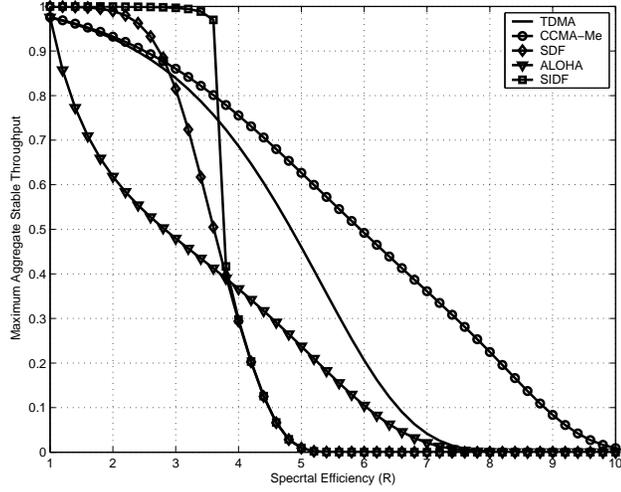


Figure 4.6: Aggregate maximum stable throughput versus spectral efficiency R in b/s/Hz. The propagation path loss is set to $\gamma = 3.5$. Second scenario is used for SDF and SIDF. SIDF has the best performance for low spectral efficiency but it suffers from a catastrophic degradation when increasing the spectral efficiency R . CCMA-Me has a graceful degradation due to its bandwidth efficiency, and it has the best tradeoff for medium to high spectral efficiency.

4.5 and 4.6, respectively. In both figures, the relative distance between terminals are, $r_{1,d} = r_{2,d} = 130$, $r_{l,d} = 50$, and $r_{1,l} = r_{2,l} = 80$. In Fig. 4.5, the MST is plotted against the SNR threshold β and the propagation path loss is set to $\gamma = 3.5$. For SDF and SIDF, we use the first scenario. CCMA-Me has the best tradeoff for the whole range. TDMA and CCMA-S have identical performance as proven before for the symmetric case. The maximum attained MST for SDF is 0.5 as one can expect because of the time slot repetition. ALOHA has better performance over SDF for low SNR threshold, but for medium and high values of β , SDF has better performance. SIDF has close performance to CCMA-Me for low values of β , and CCMA-Me outperforms it for the rest of the SNR threshold range.

In Fig. 4.6 the MST is plotted against the transmission rate R , and we use the second scenario for SDF and SIDF. For this case the MST of SDF starts from 1 for low rates R but decays exponentially after that. SIDF and SDF have the best performance for low spectral efficiency regimes where sacrificing the bandwidth by transmitting at higher rate is less significant than the gains achieved by diversity. SIDF performs better than SDF because it is more bandwidth efficient. For higher spectral efficiency regimes, the proposed CCMA-Me provides significantly higher stable throughput compared to SDF or SIDF. An important point to observe from Fig. 4.6 is the graceful degradation in the performance of CCMA-Me, while the sudden catastrophic performance loss in SDF and SIDF. The rationale here is that the cognitive feature of the proposed CCMA-Me results in no bandwidth loss because cooperation is done in the idle time slots, while both SDF and SIDF suffer from a bandwidth loss which increases for SIDF with increasing R . Our results reveal a very interest observation that utilizing empty time slots to increase system reliability via cooperation is a very promising technique in designing cooperative

relaying strategies for wireless networks.

4.4 Throughput Region

In the characterization of the stable throughput region in the previous section, the source burstiness is taken into consideration. Consider now the scenario under which all terminals queues are saturated, i.e., each terminal has infinite number of packets waiting transmission. The maximum throughput supported by any terminal can be defined under such a scenario by the average maximum number of packets that can be transmitted successfully by that terminal [65]. The set of all such saturated throughput for different resource-sharing vectors defines the throughput region.

Since in both CCMA-S and CCMA-Me the relay role depends on having empty time slots to enable cooperation, There is no surprise that under such saturated queues scenario the relay loses its role and CCMA-S and CCMA-Me reduce to TDMA without relaying. We state this in the following Corollary.

Corollary 1 *The throughput regions of TMDA, CCMA-S, and CCMA-Me are equivalent.*

$$\mathcal{C}(TDMA) \equiv \mathcal{C}(CCMA-S) \equiv \mathcal{C}(CCMA-Me). \quad (4.64)$$

From the above corollary, we conclude that the saturated throughput region is a subset of the stability region for both CCMA-S and CCMA-Me. This is an important observation because for ALOHA systems it is conjectured in [65] that the maximum stable throughput region is identical to the throughput region. It is of interest then to point out that CCMA-S and CCMA-Me are examples of multiple-access protocols where the stable throughput region is different from the

throughput region.

4.5 Delay Analysis

In this section we characterize the delay performance of the proposed cognitive cooperative multiple-access protocols, CCMA-S and CCMA-Me.

4.5.1 Delay Performance for CCMA-S

In CCMA-S, a packet does not depart a terminal's queue until it is successfully transmitted to the destination. Therefore, the delay encountered by a packet is the one encountered in the terminal's queue. Delay analysis for interacting queues in ALOHA has been studied in [66], [67] and more recently for ALOHA with MPR channels in [64], and it turns out to be a notoriously hard problem. Most of the known results are only for the 2-users ALOHA case. In this section we consider a symmetric 2-users CCMA-S scenario and characterize its delay performance.

Define the moment generating function of the joint queues' sizes processes (Q_1^t, Q_2^t) as follows

$$G(u, v) = \lim_{t \rightarrow \infty} E \left[u^{Q_1^t} v^{Q_2^t} \right]. \quad (4.65)$$

From the queue evolution equations in (4.11), we have

$$E \left[u^{Q_1^{t+1}} v^{Q_2^{t+1}} \right] = E \left[u^{X_1^t} v^{X_2^t} \right] E \left[u^{(Q_1^t - Y_1^t)^+} v^{(Q_2^t - Y_2^t)^+} \right], \quad (4.66)$$

where the above equation follows from the independence assumption of the future arrival processes from the past departure and arrival processes. Since the arrival processes are assumed to follow Bernoulli random process, the moment generating function of the joint arrival processes is given by

$$A(u, v) \triangleq \lim_{t \rightarrow \infty} E \left[u^{X_1^t} v^{X_2^t} \right] = (u\lambda + 1 - \lambda)(v\lambda + 1 - \lambda), \quad (4.67)$$

where due to the symmetry of the two terminals, each has an arrival rate λ . From the definition of the service process of CCMA-S (4.14), it follows that

$$\begin{aligned} E \left[u^{(Q_1^t - Y_1^t)^+} v^{(Q_2^t - Y_2^t)^+} \right] &= E \left[\mathbf{1}(Q_1^t = 0, Q_2^t = 0) \right] + B(u) E \left[\mathbf{1}(Q_1^t > 0, Q_2^t = 0) u^{Q_1^t} \right] \\ &+ B(v) E \left[\mathbf{1}(Q_1^t = 0, Q_2^t > 0) v^{Q_2^t} \right] + D(u, v) E \left[\mathbf{1}(Q_1^t > 0, Q_2^t > 0) u^{Q_1^t} v^{Q_2^t} \right], \end{aligned} \quad (4.68)$$

where

$$\begin{aligned} B(z) &= \frac{w f_{1d} + w^2 f_{1l} f_{ld} (1 - f_{1d})}{z} + 1 - (w f_{1d} + w^2 f_{1l} f_{ld} (1 - f_{1d})), \\ D(u, v) &= w f_{1d} \left(\frac{1}{u} + \frac{1}{v} \right) + 2w(1 - f_{1d}), \end{aligned} \quad (4.69)$$

in which we use w to denote symmetric resource-sharing portion of each terminal.

Substituting (4.67), (4.68), into (4.66) and taking the limits we get

$$\begin{aligned} G(u, v) &= A(u, v) (G(0, 0) + B(u) [G(u, 0) - G(0, 0)] + B(v) [G(0, v) \\ &- G(0, 0)] + D(u, v) [G(u, v) + G(0, 0) - G(u, 0) - G(0, v)]). \end{aligned} \quad (4.70)$$

We can rewrite the above equation as follows $G(u, v) = \frac{H(u, v)}{F(u, v)}$, where

$$\begin{aligned} H(u, v) &= G(0, 0) + B(u) [G(u, 0) - G(0, 0)] + B(v) \times [G(0, v) - G(0, 0)] \\ &+ D(u, v)(G(0, 0) - G(u, 0) - G(0, v)), \end{aligned} \quad (4.71)$$

and $F(u, v) = 1 - A(u, v)D(u, v)$.

Define $G_1(u, v) \triangleq \frac{\partial G(u, v)}{\partial u}$. Due to symmetry, the average queue size is given by $G_1(1, 1)$. Hence to find the average queuing delay, we need to compute $G_1(1, 1)$.

First we find a relation between $G(0, 0)$ and $G(1, 0)$ using the following two properties which follow from the symmetry of the problem: $G(1, 1) = 1$ and $G(1, 0) = G(0, 1)$. Applying these two properties to (4.70) along with a simple application of L'Hopital limit theorem we get

$$(w^2 f_{1l} f_{ld} (1 - f_{1d})) G(0, 0) + (w f_{1d} - w^2 f_{1l} f_{ld} (1 - f_{1d})) G(1, 0) = w f_{1d} - \lambda. \quad (4.72)$$

Taking the derivative of (4.70) with respect to u , applying L'Hopital twice, and using the relation in (4.72) we get

$$G_1(1, 1) = \frac{\lambda(1 - \lambda)}{wf_{1d} - \lambda} - \frac{w^2 f_{1l} f_{1d} (1 - f_{1d})}{wf_{1d} - \lambda} G_1(1, 0). \quad (4.73)$$

To find another equation relating $G_1(1, 1)$ and $G_1(1, 0)$, we compute $\frac{\partial G(u, u)}{\partial u}$ at $u = 1$. After some tedious but straightforward calculations, we get

$$\frac{\partial G(u, u)}{\partial u} \Big|_{u=1} = 2\lambda - 1 + \frac{wf_{1d} - w^2 f_{1l} f_{1d} (1 - f_{1d})}{wf_{1d} - \lambda} G_1(1, 0) - \frac{4wf_{1d}\lambda - 2wf_{1d} - \lambda^2}{2(wf_{1d} - \lambda)}. \quad (4.74)$$

Due to the symmetry of the problem, we have the following property

$$\frac{\partial G(u, u)}{\partial u} \Big|_{u=1} = 2G_1(1, 1). \quad (4.75)$$

Using the above equation, and solving (4.73) and (4.74) we get

$$G_1(1, 1) = \frac{-(2wf_{1d} + w^2 f_{1l} f_{1d} (1 - f_{1d})) \lambda^2 + 2wf_{1d}\lambda}{2(wf_{1d} + w^2 f_{1l} f_{1d} (1 - f_{1d}))(wf_{1d} - \lambda)}. \quad (4.76)$$

The queueing delay for system CCMA-S can thus be determined as in the following Theorem.

Theorem 8 *The average queueing delay for a symmetrical two terminals CCMA-S system is given by*

$$D(CCMA-S) = \frac{G_1(1, 1)}{\lambda} = \frac{-(2wf_{1d} + w^2 f_{1l} f_{1d} (1 - f_{1d})) \lambda + 2wf_{1d}}{2(wf_{1d} + w^2 f_{1l} f_{1d} (1 - f_{1d}))(wf_{1d} - \lambda)}. \quad (4.77)$$

From Theorem 8, it can be observed that at $\lambda = \omega f_{1d}$ the delay of the system becomes unbounded, i.e., the system becomes saturated. This confirms our previous results in Corollary 1 that for a symmetrical system, both TDMA and CCMA-S have the same maximum stable throughput of $\lambda = \omega f_{1d}$.

4.5.2 Delay Performance of CCMA-Me

Due to the symmetrical scenario considered in analyzing the delay performance, if the relay helps one terminal then it helps all terminals, in which case both CCMA-M, CCMA-Me become equivalent. In CCMA-M, a packet can encounter two queuing delays; the first in the terminal's queue and the second in the relay's queue. If a packet successfully transmitted by a terminal directly goes to the destination, then this packet is not stored in the relay's buffer. Denote this event by ξ . The total delay encountered by a packet in CCMA-M can thus be modeled as

$$T(\text{CCMA-M}) = \begin{cases} T_t, & \xi, \\ T_t + T_l, & \bar{\xi}, \end{cases} \quad (4.78)$$

where T_t is the queuing delay in the terminal's queue, and T_l is the queuing delay at the relay's queue. We can elaborate more on (4.78) as follows. For a given packet in the terminal's queue, if the first successful transmission for this packet is to the destination, then the delay encountered by this packet is only the queuing delay in the terminal's queue. On the other hand, if the first successful transmission for this packet is not to the destination, then the packet will encounter the following delays: queuing delay in the terminal's queue in addition to the queuing delay in the relay's queue.

First, we find the queuing delay either in the terminal's or relay's queue, as both queues have similar evolution equations, with the difference being in the average arrival and departure rates. Using the same machinery utilized in the analysis of the queuing delay in CCMA-S to analyze the delay performance of CCMA-M, the average queue size can be found as

$$E[N] = \frac{\lambda(1 - \lambda)}{\mu - \lambda}, \quad (4.79)$$

where λ denotes the average arrival rate and μ denotes the average departure rate. We now compute the average delay in (4.78). The probability that, for any packet, the first successful transmission from the terminal's queue is to the destination is given by

$$\Pr[\xi] = \frac{f_{1d}}{f_{1d} + f_{1l} - f_{1d}f_{1l}} = \frac{f_{1d}}{P_1}. \quad (4.80)$$

From (4.78), (4.79), and (4.80), the average delay for system CCMA-M is thus given by

$$D(\text{CCMA-M}) = \frac{f_{1d}}{P_1} \frac{1 - \lambda}{wP_1 - \lambda} + \frac{f_{1l}(1 - f_{1d})}{P_1} \left(\frac{1 - \lambda}{wP_1 - \lambda} + \frac{1 - \lambda_l}{\mu_l - \lambda_l} \right), \quad (4.81)$$

where λ_l and μ_l are the average arrival and departure rates, respectively, for the relay's queue defined in (4.36) and (4.38). After simplifying the above equation, the average queueing delay for system CCMA-M can be summarized in the following Theorem

Theorem 9 *The average queueing delay for a packet in a symmetrical 2-terminal CCMA-M system is given by*

$$D(\text{CCMA-M}) = \frac{1 - \lambda}{wP_1 - \lambda} + \frac{f_{1l}(1 - f_{1d})}{P_1} \left(\frac{1 - \lambda_l}{\mu_l - \lambda_l} \right). \quad (4.82)$$

4.5.3 Numerical Results for Delay Performance

We illustrate the delay performance of the proposed multiple-access schemes with varying SNR threshold β through some numerical examples. As in the case for the stability region, we include the delay performance of ALOHA, TDMA without relaying, SDF, and SIDF in our results.

Delay performance results for the 2-terminals symmetric ALOHA with capture has been derived in [64]. We provide the results here for reference. Using the same

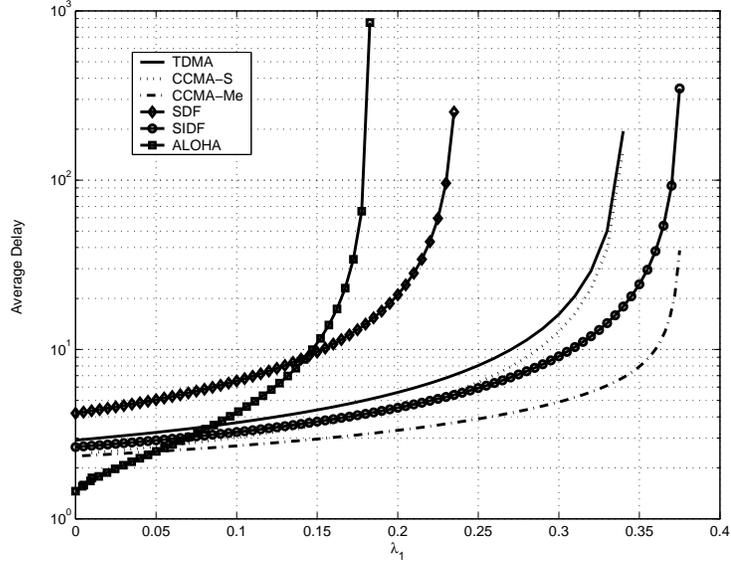


Figure 4.7: Average queueing delay per terminal versus the arrival rate for a SNR threshold of $\beta = 15$.

notation in [64], the average delay performance was proved to be

$$D(\text{ALOHA}) = \frac{1}{q_{1|1}} \left[\frac{q_{1|1}(1 - \lambda) + p(q_{1|\{1,2\}} - q_{1|1})(1 - \lambda/2)}{pq_{1|1} + p^2(q_{1|\{1,2\}} - q_{1|1}) - \lambda} \right], \quad (4.83)$$

where $q_{1|1}$ and $q_{1|\{1,2\}} = q_{2|\{1,2\}}$ are defined as before. In (4.83), p denotes the transmission probability for both users and it can be optimized to minimize the average system delay, and the results can be found in [64].

To compare the delay performance of the different multiple-access protocols considered in this chapter, we plotted the analytical expressions we got for the queueing delay. The system parameters are the same used to generate the MST plots in Figs. 4.5 and 4.6. Figs. 4.7 and 4.8 depict the delay results for SNR thresholds $\beta = 15$ and $\beta = 64$, respectively. From Fig. 4.7, at very low arrival rates, ALOHA has the best delay performance. Increasing the arrival rate λ , both CCMA-S and SDF outperforms other strategies. For higher values of λ , CCMA-Me

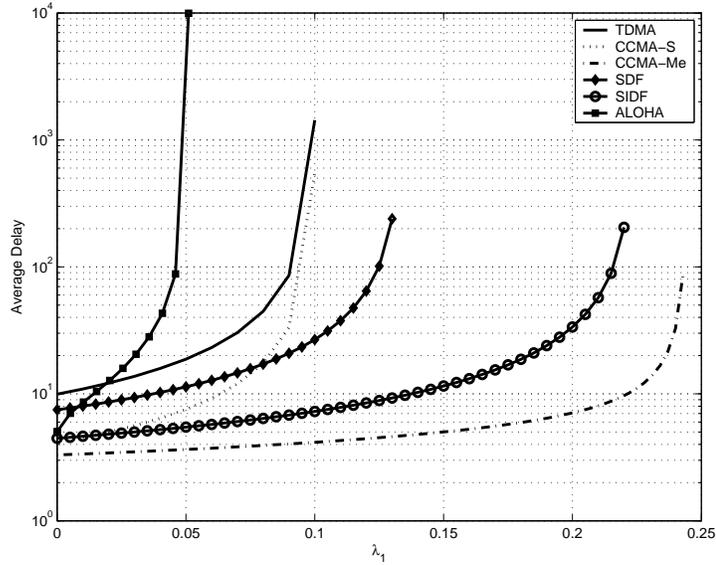


Figure 4.8: Average queuing delay per terminal versus the arrival rate for a SNR threshold of $\beta = 64$.

has the best performance

The situation changes in Fig. 4.8 for $\beta = 64$ as both CCMA-S and SDF outperform ALOHA even for very small arrival rates. The intuition behind this is the more stringent system requirements reflected by the higher SNR threshold $\beta = 64$, which makes the interference in ALOHA more severe. This makes our cognitive multiple-access protocol CCMA-S and CCMA-M perform better than ALOHA because of its high bandwidth efficiency and the gains of cooperation. Another important remark is that although CCMA-S and TDMA has the same MST for the symmetric case, as proven before and as clear from Figs. 4.7 and 4.8 where both protocols saturate at the same arrival rate, CCMA-S has always better delay performance than TDMA.

Appendices

Appendix A: Proof of Theorem 4

The stable throughput region of CCMA-S for a fixed resource-sharing vector (ω_1, ω_2) is specified in Lemma 1. In order to find the whole stability region of the protocol, we need to take the union over all possible values of (ω_1, ω_2) in F . One method to characterize this union is to solve a constrained optimization problem to find the maximum feasible λ_2 corresponding to each feasible λ_1 . For a fixed λ_1 , the maximum stable arrival rate for queue 2 is given by solving the following optimization problem

$$\begin{aligned} \max_{w_1, w_2} \lambda_2 &= w_2 f_{2d} + w_1 w_2 K_2 - \frac{\lambda_1 w_2 K_2}{f_{1d}}, \\ \text{s.t. } w_1 + w_2 &\leq 1, \quad \lambda_1 \leq w_1 f_{1d}, \end{aligned} \quad (4.84)$$

where $K_i = (1 - f_{2d})f_{2i}f_{1d}$, and $i \in \{1, 2\}$. To put this problem in a standard form, we can equivalently write it as the minimization of $-\lambda_2$. The Lagrangian of this optimization problem is given by

$$\begin{bmatrix} -w_2 K_2 \\ -w_1 K_2 + \frac{\lambda_1 K_2}{f_{1d}} - f_{2d} \end{bmatrix} + u_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + u_2 \begin{bmatrix} -f_{1d} \\ 0 \end{bmatrix} = 0, \quad (4.85)$$

where u_1, u_2 are the complementary slackness variables that are non-negative. Solving for the complementary slackness variables we get the following relation between u_1 and u_2

$$u_1 = w_2 K_2 + u_2 f_{1d}. \quad (4.86)$$

This shows that $u_1 > 0$, i.e., the first constraint ($w_1 + w_2 = 1$) in (4.84) is met with equality.

Substituting $w_2 = 1 - w_1$ in the cost function in (4.84), we get

$$\lambda_2 = (1 - w_1) f_{2d} + w_1 (1 - w_1) K_2 - (1 - w_1) \frac{\lambda_1 K_2}{f_{1d}}. \quad (4.87)$$

Taking the first derivative of the above equation with respect to w_1 , we get

$$\frac{\partial \lambda_2}{\partial w_1} = -f_{2d} + K_2 - 2w_1 K_2 + \frac{\lambda_1 K_2}{f_{1d}}. \quad (4.88)$$

Since K_2 is nonnegative, the second derivative is negative, hence, the cost function in (4.87) is concave in w_1 and the necessary conditions for optimality KKT [77] are also sufficient. Equating the first derivative in (4.88) to zero, the solution w_1^* is given by

$$w_1^* = \frac{1}{2K_2} \left(K_2 - f_{2d} + \frac{\lambda_1 K_2}{f_{1d}} \right). \quad (4.89)$$

From the first constraint in (4.84), the minimum value for w_1 that guarantees the stability of queue 1 in system S_1 is given by $w_{1,min} = \frac{\lambda_1}{f_{1d}}$. Hence, if $w_1^* > w_{1,min}$, and given concavity of the cost function, the optimal solution is just w_1^* , otherwise it is given by $w_{1,min}$. Characterizing this in terms of the channel parameters, the optimal solution for the optimization problem in (4.84) is given by

$$w_1^* = \begin{cases} \frac{1}{2K_2} \left(K_2 - f_{2d} + \frac{\lambda_1 K_2}{f_{1d}} \right), & \lambda_1 \leq f_{1d} - \frac{f_{1d} f_{2d}}{K_2} \\ \frac{\lambda_1}{f_{1d}}, & f_{1d} - \frac{f_{1d} f_{2d}}{K_2} < \lambda_1 < f_{1d}. \end{cases} \quad (4.90)$$

If $\lambda_1 > f_{1d}$, then the first queue can never be stable by construction of S_1 .

Now we solve for the other branch in the stability region given by the dominant system S_2 . The equations for this branch are given by (4.17). Similar to the first stability region branch, solving for the Lagrangian of this branch gives the necessary condition $w_1 + w_2 = 1$. First, we find the maximum achievable stable rate for the first queue, which is achieved when $\lambda_2 = 0$. Substituting in (4.17), the arrival rate λ_1 can be written as

$$\lambda_1 = w_1 f_{1d} + w_1 (1 - w_1) K_1. \quad (4.91)$$

The above equation is obviously concave. Taking the first derivative and equating

to zero we get

$$w_1^*|_{\lambda_2=0} = \frac{1}{2K_1}(f_{1d} + K_1). \quad (4.92)$$

Since $w_1 \leq 1$, then if $w_1^*|_{\lambda_2=0} > 1$, i.e., $f_{1d} > K_1$, and given the concavity of the cost function, we let $w_1^*|_{\lambda_2=0} = 1$. Substituting $w_1^*|_{\lambda_2=0}$ in (4.91), the maximum achievable rate λ_1 can be given by

$$\lambda_1^* = \begin{cases} f_{1d}, & f_{1d} > K_1, \\ \frac{1}{4K_1}(f_{1d} + K_1)^2, & 0 \leq f_{1d} \leq K_1. \end{cases} \quad (4.93)$$

Next for a fixed λ_1 , we solve for the optimal λ_2 that can be achieved from the second branch. We can write (4.17) in terms of λ_2 as follows

$$\lambda_2 = \frac{f_{1d}f_{2d}}{K_1} + (1 - w_1)f_{2d} - \lambda_1 \frac{f_{2d}}{w_1 K_1}. \quad (4.94)$$

Taking the first derivative with respect to w_1 we get

$$\frac{\partial \lambda_2}{\partial w_1} = -f_{2d} + \lambda_1 \frac{f_{2d}}{K_1 w_1^2}. \quad (4.95)$$

The second derivative is negative, which renders the whole function concave. Equating the first derivative to zero we get

$$w_1^* = \sqrt{\frac{\lambda_1}{(1 - f_{1d})f_{1d}f_{2d}}}. \quad (4.96)$$

From (4.17), we have the following constraint for the stability of the second queue, $\lambda_2 \leq w_2 f_{2d}$, which can be written in terms of w_1 as follows

$$w_1 \leq 1 - \frac{\lambda_2}{f_{2d}}. \quad (4.97)$$

Substituting λ_2 from (4.97) into (4.94), we get that the maximum value for w_1 in terms of λ_1 is given by $w_1 \leq \frac{\lambda_1}{f_{1d}}$. Therefore, the optimal value for w_1 can be written as

$$w_1^* = \min \left(\sqrt{\frac{\lambda_1}{(1 - f_{1d})f_{1d}f_{2d}}}, \frac{\lambda_1}{f_{1d}} \right). \quad (4.98)$$

The value of the expression in (4.98) can be shown never to exceed 1 by substituting the maximum value of λ_1 given by (4.93) in the above equation. After some manipulations, the optimum w_1^* in (4.98) can be further simplified as follows

$$w_1^* = \begin{cases} \frac{\lambda_1}{f_{1d}}, & \text{if } \lambda_1 \leq \frac{f_{1d}^2}{(1-f_{1d})f_{1l}f_{1d}}, \\ \sqrt{\frac{\lambda_1}{(1-f_{1d})f_{1l}f_{1d}}}, & \text{otherwise.} \end{cases} \quad (4.99)$$

We summarize equations (4.87,4.90,4.94,4.99) describing the envelopes of the first and second branches as follows. For the first branch given by equations (4.87, 4.90), substituting (4.90) in (4.87) we get

$$g_1(\lambda_1) = \begin{cases} K_2 \left(\frac{\lambda_1}{2f_{1d}} + 0.5 - \frac{f_{2d}}{2K_2} \right)^2 - \frac{K_2\lambda_1}{f_{1d}} + f_{2d}, & 0 \leq \lambda_1 \leq f_{1d} - \frac{f_{1d}f_{2d}}{K_2}, \\ f_{2d} - \frac{f_{2d}}{f_{1d}}\lambda_1, & f_{1d} - \frac{f_{1d}f_{2d}}{K_2} < \lambda_1 \leq f_{1d}. \end{cases} \quad (4.100)$$

Similarly, by substituting (4.99) in (4.94), we get for the second branch

$$g_2(\lambda_1) = \begin{cases} f_{2d} - \frac{f_{2d}}{f_{1d}}, & \lambda_1 < \frac{f_{1d}^2}{(1-f_{1d})f_{1l}f_{1d}}, \\ \frac{f_{1d}f_{2d}}{K_1} + f_{2d} - 2f_{2d}\sqrt{\frac{\lambda_1}{K_1}}, & \frac{f_{1d}^2}{(1-f_{1d})f_{1l}f_{1d}} \leq \lambda_1 \leq \lambda_1^*. \end{cases} \quad (4.101)$$

The value for λ_1^* is given by (4.93). Since $g_1(\lambda_1)$ and $g_2(\lambda_2)$ are achieved by the dominant systems S_1 and S_2 respectively, then they are both achieved by CCMA-S.

This proves Theorem 1. ■

Appendix B: Proof of Lemma 2

We prove the convexity of the envelope of region $\mathcal{R}(\text{CCMA-S})$. First we consider the envelope $g_1(\lambda_1)$ for the first branch. As shown in Fig. 4.9(a), $g_1(\lambda_1)$ is defined over two regions: For $\lambda_1 \in [0, f_{1d} - \frac{f_{1d}f_{2d}}{K_2}]$, $g_1(\lambda_1) \triangleq g_{11}(\lambda_1) = K_2 \left(\frac{\lambda_1}{2f_{1d}} + 0.5 - \frac{f_{2d}}{2K_2} \right)^2 - \frac{K_2\lambda_1}{f_{1d}} + f_{2d}$, while for $\lambda_1 \in (f_{1d} - \frac{f_{1d}f_{2d}}{K_2}, f_{1d}]$, $g_1(\lambda_1) \triangleq g_{12}(\lambda_1)$ is a straight line given

by $f_{2d} - \frac{f_{2d}}{f_{1d}}$. It can be readily seen that $g_1(\lambda_1)$ is convex over both regions because its second derivative is nonnegative. Thus to prove that $g_1(\lambda_1)$ is convex over the whole region, we check the continuity and the first derivative at the intersection point $f_{1d} - \frac{f_{1d}f_{2d}}{K^2}$. It is simple to show that $g_1(\lambda_1)$ is continuous at this point. Now, it remains to check the slope of the tangent at the intersection point. Taking the first derivative of g_{11} and g_{12} , we can show that

$$\frac{\partial g_{11}(\lambda_1)}{\partial \lambda_1} = \frac{\partial g_{12}(\lambda_1)}{\partial \lambda_1} \Big|_{\lambda_1 = f_{1d} - \frac{f_{1d}f_{2d}}{K^2}}. \quad (4.102)$$

Therefore, $g_1(\lambda_1)$ is also differentiable, which proves that $g_1(\lambda_1)$ is convex over its domain of definition. Similar arguments apply for $g_2(\lambda_1)$ depicted in Fig. 4.9(b) to prove its convexity.

The envelope of the stability region of $\mathcal{R}(\text{CCMA-S})$ is given by $\max [g_1(\lambda_1), g_2(\lambda_2)]$. The maximum of two convex functions can be shown to be convex as follows. Let $0 < a < 1$, and $\lambda_{11}, \lambda_{12}$ belong to the feasible region of λ_1 , then we have

$$\begin{aligned} & \max [g_1(a\lambda_{11} + (1-a)\lambda_{12}), g_2(a\lambda_{11} + (1-a)\lambda_{12})] \\ & \leq \max [ag_1(\lambda_{11}) + (1-a)g_1(\lambda_{12}), ag_2(\lambda_{11}) + (1-a)g_2(\lambda_{12})], \quad (1) \quad (4.103) \\ & \leq \max [ag_1(\lambda_{11}), ag_2(\lambda_{11})] + \max [(1-a)g_1(\lambda_{12}), (1-a)g_2(\lambda_{12})] \quad (2) \end{aligned}$$

where (1) follows by the convexity of $g_1(\cdot)$ and $g_2(\cdot)$, and (2) follows by the properties of the max function. This proves that the envelope of the stability region for system CCMA-S is convex. ■

Appendix C: Proof of Theorem 7

In this appendix, we prove that the stability region of CCMA-S is a subset of that of CCMA-Me. In the proof of this Theorem, we use two facts: the envelope of $\mathcal{R}(\text{CCMA-S})$ is convex as proved in Lemma 2, and the envelope of $\mathcal{R}(\text{CCMA-Me})$ is

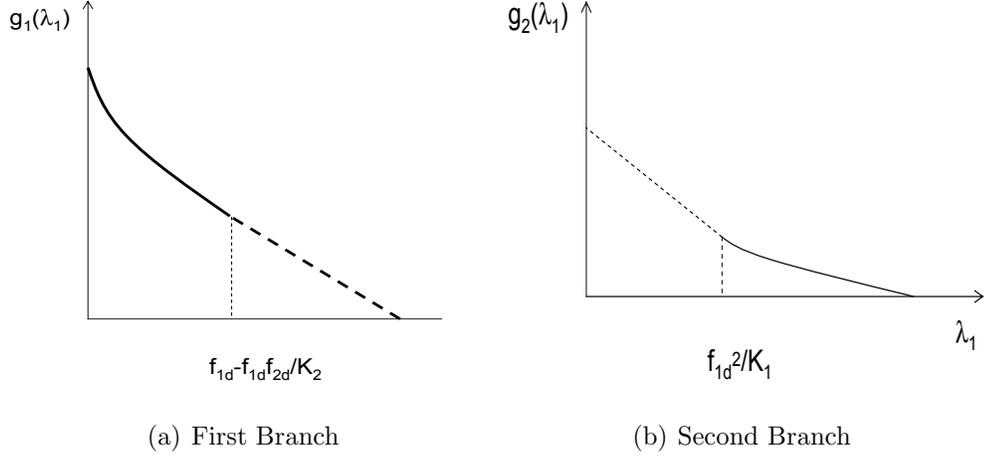


Figure 4.9: Envelopes for the stability region of CCMA-S.

a straight line. Hence, to prove Theorem 3, it suffices to show that the intersections of the envelope of region $\mathcal{R}(\text{CCMA-S})$ with the λ axes are not greater than those for region $\mathcal{R}(\text{CCMA-Me})$. We consider the scenario when both CCMA-S and CCMA-Me have larger stability regions than TDMA, or equivalently $f_{id}f_{iu}(1 - f_{id}) \geq f_{id}$, $i = 1, 2$, because if this condition is not satisfied the stability region of CCMA-S becomes identical to that of TDMA, and CCMA-Me was shown to outperform TDMA, and hence our Theorem is true.

We will consider only the intersections with the λ_1 axis, and similar arguments follow for the λ_2 axis. The intersection of $\mathcal{R}(\text{CCMA-S})$ with the λ_1 axis, or the maximum stable arrival rate for the first queue is given by (4.93), while that for $\mathcal{R}(\text{CCMA-Me})$ is given by (4.41). Denote the difference between these two quantities by δ given as follows

$$\delta = \frac{f_{id}(f_{1d} + f_{1l} - f_{1d}f_{1l})}{f_{id} + (1 - f_{1d})f_{1l}} - \frac{[f_{1d} + (1 - f_{1d})f_{1l}f_{id}]^2}{4(1 - f_{1d})f_{1l}f_{id}}. \quad (4.104)$$

If we prove that δ is nonnegative, then we are done. For an arbitrary fixed value for the pair (f_{1d}, f_{id}) , we consider the range of f_{1l} that satisfies the constraint

$f_{1d}f_{1l}(1 - f_{1d}) \geq f_{1d}$. The first derivative of δ with respect to f_{1l} is given by

$$\frac{\partial \delta}{\partial f_{1l}} = \frac{f_{1d}(1 - f_{1d})(f_{1d} - f_{1d})}{(f_{1d} + (1 - f_{1d})f_{1l})^2} - \frac{1}{4}(1 - f_{1d})f_{1d} + \frac{f_{1d}^2}{4(1 - f_{1d})f_{1l}^2 f_{1d}}, \quad (4.105)$$

Since $f_{1d} > f_{1d}$, the second derivative of the above function is easily seen to be negative, hence, δ is concave in f_{1l} in the region of interest.

f_{1l} takes values in the range $[\frac{f_{1d}}{(1-f_{1d})f_{1d}}, 1]$. The function δ evaluated at the minimum value of f_{1l} is

$$\begin{aligned} \delta|_{f_{1l, \min}} &= \frac{f_{1d}f_{1d} + f_{1d}}{f_{1d} + (1 - f_{1d})f_{1l}} - f_{1d} \\ &= \frac{f_{1d} - f_{1d}(1 - f_{1d})f_{1l}}{f_{1d} + (1 - f_{1d})f_{1l}} > 0. \end{aligned} \quad (4.106)$$

Hence, δ is positive at the left most boundary of the feasible region of f_{1l} . For the maximum feasible value of f_{1l} which is equal to 1, the function δ is given by

$$\delta = \frac{f_{1d}}{1 - f_{1d} + f_{1d}} - \frac{(f_{1d} + (1 - f_{1d})f_{1d})^2}{4(1 - f_{1d})f_{1d}}. \quad (4.107)$$

Fig. 4.10 depicts δ for $f_{1l} = 1$ over the feasible region of the pair (f_{1d}, f_{1d}) . As shown in the figure, the value of δ is always positive.

Hence, for an arbitrary feasible value of the pair (f_{1d}, f_{1d}) , δ is concave in f_{1l} and positive at the extreme end points of the feasible region of f_{1l} , therefore, δ is positive for the whole range of f_{1l} for an arbitrary vale of the pair (f_{1d}, f_{1d}) . This proves that δ is always positive over the feasible region of success probabilities, and therefore, $\mathcal{R}(\text{CCMA-S})$ is a subset from $\mathcal{R}(\text{CCMA-Me})$. This proves Theorem 6. ⁵■

⁵The above Theorem has another proof that does not need numerical evaluation, however, it is complicated and will not add new insights to the results.

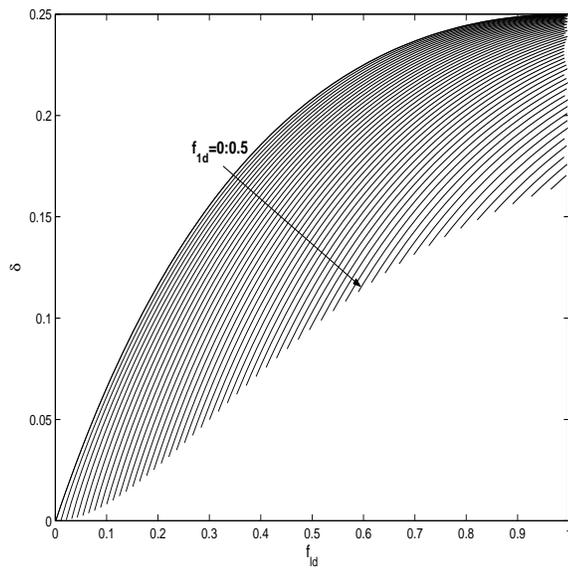


Figure 4.10: Numerical evaluation of δ for $f_{1l} = 1$.

Chapter 5

On the Energy Efficiency of Cooperative Communications

In the previous chapters, the gains of cooperative communications is studied under the ideal model of negligible listening and computing power. In some types of wireless networks, as in sensor networks, and depending on the type of nodes used, the power consumed in receiving and processing may constitute a significant portion of the total consumed power. Cooperative diversity can provide gains in terms of savings in the required transmit power in order to achieve a certain performance requirement because of the spatial diversity it adds to the system. However, if one takes into account the extra processing and receiving power consumption at the relay and destination nodes required for cooperation, then there is obviously a tradeoff between the gains in the transmit power and the losses due to the receive and processing powers when applying cooperation. Hence such a tradeoff between the gains promised by cooperation and this extra overhead in terms of the energy efficiency of the system should be taken into consideration in the network design.

In this chapter we investigate such a tradeoff and characterize the gains of co-

operation under such extra overhead. Moreover, we also consider some practical system parameters as the power amplifier loss, the quality of service (QoS) required, the relay location, and the optimal number of relays. We compare between two communications architectures, direct transmission and cooperative transmission. Our performance metric for comparison between the two architectures is the energy efficiency of the communication scheme. More specifically, for both architectures we compute the optimal total power consumption to achieve certain QoS requirements and we calculate the cooperation gain defined as the ratio between the power required for direct transmission and cooperation. When this ratio is smaller than one, this indicates that direct transmission is more energy efficient, and that the extra overhead induced by cooperation outweighs its gains in the transmit power. Moreover, we compare between optimal power allocation at the source and relay nodes and equal power allocation. The results reveal that under some scenarios, equal power allocation is almost equivalent to optimal power allocation. We also investigate the effect of relay location on the performance to provide guidelines for relay assignment algorithms. Finally, we generalize the above results to the case of multiple relays trying to answer the important question of how many relays should be used for cooperation given some communication setup.

Related work to studying energy efficiency of cooperative transmission in sensor networks can be found in [78]. In [78], distributed space-time codes are utilized to perform cooperation between clusters of nodes. It is assumed that the intermediate hops between the source and the destination can decode correctly without errors. In our work, we take into consideration the possibility of the wireless channel being in outage between any two nodes in the network. We also consider an incremental relaying cooperation strategy which is more bandwidth efficient com-

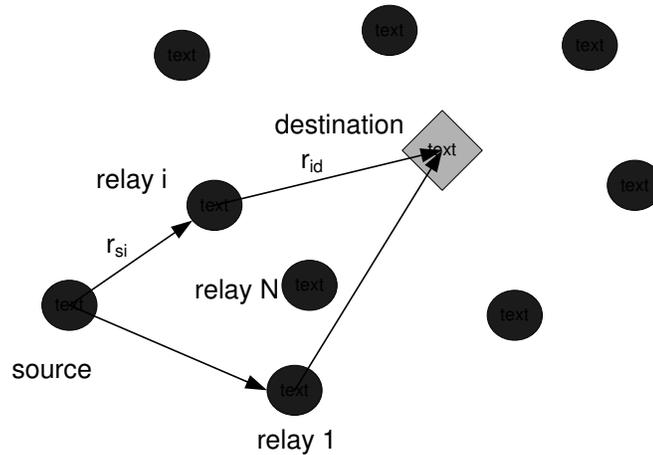


Figure 5.1: System Model

pared to distributed space-time codes. Moreover, it is easier to implement than distributed space-time codes, as the later requires synchronization between the spatially separated relays performing the distributed space-time code.

5.1 System Model

We consider a single source-destination pair separated by distance r_{sd} . The number of potential relays available to help the source is N . This is illustrated in Fig. 5.1, where the distances between source and relay i , and relay i and destination are r_{si} and r_{id} , respectively, and $i \in \{1, 2, \dots, N\}$. First we analyze the performance of the single relay scenario, and later we extend the results for arbitrary finite N .

We compare the performance of two communication scenarios. In the first scenario only direct transmission between the source and destination nodes is allowed, and this accounts for conventional direct transmission. In direct transmission, if the channel link between the source and destination encounters a deep fade or strong shadowing for example, then the communication between these two nodes

fails. Moreover, if the channel is slowly varying, which is the case in sensor networks due to the stationarity or limited mobility of the nodes, then the channel might remain in the deep fade state for long time (strong time correlation), hence conventional automatic repeat request (ARQ) might not help in this case.

In the second communication scenario, we consider a two phase cooperation protocol. In the first phase, the source transmits a signal to the destination, and due to the broadcast nature of the wireless medium the relay can overhear this signal. If the destination receives the packet from this phase correctly, then it sends back an acknowledgement (ACK) and the relay just idles. On the other hand, if the destination can not decode the received packet correctly, then it sends back a negative acknowledgement (NACK). In this case, if the relay was able to receive the packet correctly in the first phase, then it forwards it to the destination. So the idea behind this cooperation protocol is to introduce a new ARQ in another domain, which is the spatial domain, as the links between different pairs of nodes in the network fade independently. The assumptions of high temporal correlation and independence in the spatial domain will be verified through experiments as discussed in Section 5.4.

Next the wireless channel and system models are described. We consider a sensor network in which the link between any two nodes in the network is subject to narrowband Rayleigh fading, propagation path-loss, and additive white Gaussian noise (AWGN). The channel fades for different links are assumed to be statistically mutually independent. This is a reasonable assumption as the nodes are usually spatially well separated. For medium access, the nodes are assumed to transmit over orthogonal channels, thus no mutual interference is considered in the signal model. All nodes in the network are assumed to be equipped with single-element

antennas, and transmission at all nodes is constrained to the half-duplex mode, i.e., any terminal cannot transmit and receive simultaneously.

The power consumed in a transmitting or receiving stage is described as follows. If a node transmits with power P , only $P(1 - \alpha)$ is actually utilized for RF transmission, where $(1 - \alpha)$ accounts for the efficiency of the RF power amplifier which generally has a non-linear gain function. The processing power consumed by a transmitting node is denoted by P_c . Any receiving node consumes P_r power units to receive the data. The values of the parameters α, P_r, P_c are assumed the same for all nodes in the network and are specified by the manufacturer. Following, we describe the received signal model for both direct and cooperative transmissions.

First, we describe the received signal model for the direct transmission mode. In the direct transmission scheme, which is employed in current wireless networks, each user transmits his signal directly to the next node in the route which we denote as the destination d here. The signal received at the destination d from source user s , can be modeled as

$$y_{sd} = \sqrt{P_s^D(1 - \alpha)r_{sd}^{-\gamma}}h_{sd}x + n_{sd}, \quad (5.1)$$

where P_s^D is the transmission power from the source in the direct communication scenario, x is the transmitted data with unit power, h_{sd} is the channel fading gain between the two terminals s and d . The channel fade of any link is modeled as a zero mean circularly symmetric complex Gaussian random variable [4] with unit variance. In (5.1), γ is the path loss exponent, and r_{sd} is the distance between the two terminals. The term n_{sd} in (5.1) denotes additive noise; the noise components are modeled as white Gaussian noise (AWGN) with variance N_o .

Second, we describe the signal model for cooperative transmission. The cooperative transmission scenario comprises two phases as illustrated before. The

signals received from the source at the destination d and relay 1 in the first stage can be modeled respectively as,

$$y_{sd} = \sqrt{P_s^c(1 - \alpha)r_{sd}^{-\gamma}}h_{sd}x + n_{sd}, \quad y_{s1} = \sqrt{P_s^c(1 - \alpha)r_{s1}^{-\gamma}}h_{s1}x + n_{s1}, \quad (5.2)$$

where P_s^c is the transmission power from the source in the cooperative scenario. The channel gains h_{sd} and h_{s1} between the source-destination and source-relay are modeled as zero-mean circular symmetric complex Gaussian random variables with zero mean. If the SNR of the signal received at the destination from the source falls below the threshold β , the destination broadcasts a NACK. In this case, if the relay was able to receive the packet from the source correctly in the first phase, it forwards the packet to the destination with power P_1

$$y_{1d} = \sqrt{P_1(1 - \alpha)r_{1d}^{-\gamma}}h_{1d}x + n_{1d}. \quad (5.3)$$

Cooperation results in additional spatial diversity by introducing this artificial multipath through the relay link. This can enhance the transmission reliability against wireless channel impairments as fading, but will also result in extra receiving and processing power. In the next Section, we discuss this in more details.

5.2 Performance Analysis and Optimum Power Allocation

We characterize the system performance in terms of outage probability. Outage is defined as the event that the received SNR falls below a certain threshold β , hence, the probability of outage P_O is defined as,

$$\mathcal{P}_O = \mathcal{P}(\text{SNR} \leq \beta). \quad (5.4)$$

If the received SNR is higher than the threshold β , the receiver is assumed to be able to decode the received message with negligible probability of error. If an outage occurs, the packet is considered lost. The SNR threshold β is determined according to the application and the transmitter/receiver structure. For example, larger values of β is required for applications with higher quality of service (QoS) requirements. Also increasing the complexity of transmitter and/or receiver structure, for example applying strong error coding schemes, can reduce the value of β for the same QoS requirements.

Based on the derived outage probability expressions, we formulate a constrained optimization problem to minimize the total consumed power subject to a given outage performance. We then compare the total consumed power for the direct and cooperative scenarios to quantify the energy savings, if any, gained by applying cooperative transmission.

5.2.1 Direct Transmission

As discussed before, the outage is defined as the event that the received SNR falls below a predefined threshold which we denoted by β . From the received signal model in (5.1), the received SNR from a user at a distance r_{sd} from the destination is given by

$$\text{SNR}(r_{sd}) = \frac{|h_{sd}|^2 r_{sd}^{-\gamma} P_s^D (1 - \alpha)}{N_o}, \quad (5.5)$$

where $|h_{sd}|^2$ is the magnitude square of the channel fade and follows an exponential distribution with unit mean; this follows because of the Gaussian zero mean distribution of h_{sd} . Hence, the outage probability for the direct transmission mode P_{OD} can be calculated as

$$\mathcal{P}_{OD} = \mathcal{P}(\text{SNR}(r_{sd}) \leq \beta) = 1 - \exp\left(-\frac{N_o \gamma r_{sd}^\gamma}{(1 - \alpha) P_s^D}\right). \quad (5.6)$$

The total transmitted power P_{tot}^D for the direct transmission mode is given by

$$P_{tot}^D = P_s^D + P_c + P_r, \quad (5.7)$$

where P_s^D is the power consumed at the RF stage of the source node, P_c is the processing power at the source node, and P_r is the receiving power at the destination. The requirement is to minimize this total transmitted power subject to the constraint that we meet a certain QoS requirement that the outage probability is less than a given outage requirement, which we denote by \mathcal{P}_{out}^* . Since both the processing and receiving powers are fixed, the only variable of interest is the transmitting power P_s^D .

The optimization problem can be formulated as follows

$$\min_{P_s^D} P_{tot}^D, \quad \text{s.t. } \mathcal{P}_{OD} \leq \mathcal{P}_{out}^*. \quad (5.8)$$

The outage probability \mathcal{P}_{OD} is a decreasing function in the power P_s^D . Substituting \mathcal{P}_{out}^* in the outage expression in (5.6), we get after some simple arithmetics that the optimal transmitting power is given by

$$P_s^{D*} = -\frac{\beta N_o r_{sd}^\gamma}{(1 - \alpha) \ln(1 - \mathcal{P}_{out}^*)}. \quad (5.9)$$

The minimum total power required for direct transmission in order to achieve the required QoS requirement is therefore given by

$$P_{tot}^* = P_c + P_r - \frac{\beta N_o r_{sd}^\gamma}{(1 - \alpha) \ln(1 - \mathcal{P}_{out}^*)}. \quad (5.10)$$

In the next subsection we formulate the optimal power allocation problem for the cooperative communication scenario.

5.2.2 Cooperative Transmission

For the optimal power allocation problem in cooperative transmission, we consider two possible scenarios. In the first scenario, the relay is allowed to transmit

with different power than the source and hence the optimization space is two-dimensional: source and relay power allocations. The solution for this setting provides the minimum possible total consumed power. However, the drawback of this setting is that the solution for the optimization problem is complex and might not be feasible to implement in sensor nodes. The second setting that we consider is constraining the source and relay nodes to transmit with equal powers. This is much easier to implement as the optimization space is one dimensional in this case, moreover, a relaxed version of the optimization problem can render a closed form solution. Clearly the solution of the equal power allocation problem provides a suboptimal solution to the general case in which we allow different power allocations at the source and the relay. It is interesting then to investigate the conditions under which these two power allocation strategies have close performance.

First, we characterize the optimal power allocations at the source and relay nodes. Consider a source-destination pair that are r_{sd} units distance. Let us compute the conditional outage probability for given locations of the source and the helping relay. As discussed before, cooperative transmission encompasses two phases. Using (5.2), the SNR received at the destination d and relay 1 from the source s in the first phase are given by

$$\text{SNR}_{sd} = \frac{|h_{sd}|^2 r_{sd}^{-\gamma} P_s^C (1 - \alpha)}{N_o}, \quad \text{SNR}_{s1} = \frac{|h_{s1}|^2 r_{s1}^{-\gamma} P_s^C (1 - \alpha)}{N_o}. \quad (5.11)$$

While from (5.3), the SNR received at the destination from the relay in the second phase is given by

$$\text{SNR}_{1d} = \frac{|h_{1d}|^2 r_{1d}^{-\gamma} P_1 (1 - \alpha)}{N_o}. \quad (5.12)$$

Note that the second phase of transmission is only initiated if the packet received at the destination from the first transmission phase is not correctly received. The

terms $|h_{sd}|^2$, $|h_{s1}|^2$, and $|h_{1d}|^2$ are mutually independent exponential random variables with unit mean.

The outage probability of the cooperative transmission \mathcal{P}_{OC} can be calculated as follows

$$\begin{aligned}\mathcal{P}_{OC} &= \mathcal{P}((\text{SNR}_{sd} \leq \beta) \cap (\text{SNR}_{sl} \leq \beta)) \\ &\quad + \mathcal{P}((\text{SNR}_{sd} \leq \beta) \cap (\text{SNR}_{ld} \leq \beta) \cap (\text{SNR}_{sl} > \beta)) \\ &= (1 - f(r_{sd}, P_s^C)) (1 - f(r_{sl}, P_s^C)) + (1 - f(r_{sd}, P_s^C)) (1 - f(r_{ld}, P_l)) f(r_{sl}, P_s^C),\end{aligned}\tag{5.13}$$

where $f(x, y) = \exp(-\frac{N_o \beta x^\gamma}{y(1-\alpha)})$. The first term in the above expression corresponds to the event that both the source-destination and the source-relay channels are in outage, and the second term corresponds to the event that both the the source-destination and the relay-destination channels are in outage while the source-relay channel is not. The above expression can be simplified as follows

$$\mathcal{P}_{OC} = (1 - f(r_{sd}, P_s^C)) (1 - f(r_{ld}, P_l) f(r_{sl}, P_s^C)).\tag{5.14}$$

The total average consumed power for cooperative transmission to transmit a packet is given by

$$\begin{aligned}E[P_{tot}^C] &= (P_s^C + P_c + 2P_r) \mathcal{P}(\text{SNR}_{sd} \geq \beta) \\ &\quad + (P_s^C + P_c + 2P_r) \mathcal{P}(\text{SNR}_{sd} < \beta) \mathcal{P}(\text{SNR}_{s1} < \beta) \\ &\quad + (P_s^C + P_1 + 2P_c + 3P_r) \mathcal{P}(\text{SNR}_{sd} < \beta) \mathcal{P}(\text{SNR}_{s1} > \beta),\end{aligned}\tag{5.15}$$

where the first term in the right hand side corresponds to the event that the direct link in the first phase is not in outage, therefore, the total consumed power is only given by that of the source node, and the 2 in front of the received power term P_r is to account for the relay receiving power. The second term in the summation corresponds to the event that both the direct and the source-relay links are in

outage, hence the total consumed power is still given as in the first term. The last term in the total summation accounts for the event that the source-destination link is in outage while the source-relay link is not, and hence we need to account for the relay transmitting and processing powers, and the extra receiving power at the destination. Using the Rayleigh fading channel model, the average total consumed power can be given as follows

$$\begin{aligned}
P_{tot}^C &= (P_s^C + P_c + 2P_r) f(r_{sd}, P_s^C) \\
&+ (P_s^C + P_c + 2P_r) (1 - f(r_{sd}, P_s^C)) (1 - f(r_{sl}, P_s^C)) \\
&+ (P_s^C + P_1 + 2P_c + 3P_r) (1 - f(r_{sd}, P_s^C)) \times f(r_{sl}, P_s^C).
\end{aligned} \tag{5.16}$$

We can formulate the power minimization problem in a similar way to (5.8) with the difference that there are two optimization variables in the cooperative transmission mode, namely, the transmit powers P_s^C and P_1 at the source and relay nodes respectively. The optimization problem can be stated as follows

$$\min_{P_s^C, P_1} P_{tot}^C(P_s^C, P_1), \quad \text{s.t. } \mathcal{P}_{OC}(P_s^C, P_1) \leq \mathcal{P}_{out}^*. \tag{5.17}$$

This optimization problem is nonlinear and does not admit a closed form solution. Therefore we resort to numerical optimization techniques in order to solve for this power allocation problem at the relay and source nodes, and the results are shown in the simulations section.

In the above formulation we considered optimal power allocation at the source and relay node in order to meet the outage probability requirement. The performance attained by such an optimization problem provides a benchmark for the cooperative transmission scheme. However, in a practical setting, it might be difficult to implement such a complex optimization problem at the sensor nodes. A more practical scenario would be that all the nodes in the network utilize the same

power for transmission. Denote the equal transmission power in this case by P_{CE} ; the optimization problem in this case can be formulated as

$$\min_{P_{CE}} P_{tot}^C(P_{CE}), \quad \text{s.t. } \mathcal{P}_{OC}(P_{CE}) \leq \mathcal{P}_{out}^*. \quad (5.18)$$

Beside being a one-dimensional optimization problem that can be easily solved, the problem can be relaxed to render a closed form solution. Note that at enough high SNR the following approximation holds $\exp(-x) \simeq (1 - x)$; where x here is proportional to $1/SNR$.

Using the above approximation in (5.16), and after some mathematical manipulation, the total consumed power can be approximated as follows

$$P_{tot}^C \simeq P_{CE} + P_c + 2P_r + (P_{CE} + P_c + P_r) \frac{k_1}{P_{CE}} - (P_{CE} + P_c + P_r) \frac{k_1 k_2}{P_{CE}^2}. \quad (5.19)$$

Similarly, the outage probability can be written as follows

$$\mathcal{P}_{OC} \simeq \frac{k_1 k_2}{P_{CE}^2} + \frac{k_1 k_3}{P_{CE}^2} - \frac{k_1 k_2 k_3}{P_{CE}^3}, \quad (5.20)$$

where $k_1 = \frac{\beta N_o r_{sd}^\gamma}{1-\alpha}$, $k_2 = \frac{\beta N_o r_{sl}^\gamma}{1-\alpha}$, and $k_3 = \frac{\beta N_o r_{ld}^\gamma}{1-\alpha}$. This is a constrained optimization problem in one variable and its Lagrangian is given by

$$\frac{\partial P_{tot}^C}{\partial P_{CE}} + \lambda \frac{\partial \mathcal{P}_{OC}}{\partial P_{CE}} = 0, \quad (5.21)$$

where the derivatives of the total power consumption P_{tot}^C and the outage probability \mathcal{P}_{OC} with respect to the transmit power P_{CE} are given by

$$\begin{aligned} \frac{\partial P_{tot}^C}{\partial P_{CE}} &= 1 + \frac{k_1 k_2 - (P_c + P_r) k_1}{P_{CE}^2} + \frac{2k_1 k_2 (P_c + P_r)}{P_{CE}^3}; \\ \frac{\partial \mathcal{P}_{OC}}{\partial P_{CE}} &= \frac{-2(k_1 k_2 + k_1 k_3)}{P_{CE}^3} + \frac{3k_1 k_2 k_3}{P_{CE}^4}, \end{aligned} \quad (5.22)$$

respectively. Substituting the derivatives in (5.22) into the Lagrangian in (5.21), and doing simple change of variables $1/P_{CE} = x$, the Lagrangian can be written

in the following simple polynomial form

$$1 + (k_1k_2 - (P_c + P_r)k_1)x^2 + 2(k_1k_2(P_c + P_r) - \lambda(k_1k_2 + k_1k_3))x^3 + 3\lambda k_1k_2k_3x^4 = 0, \quad (5.23)$$

under the outage constraint

$$(k_1k_2 + k_1k_3)x^2 - k_1k_2k_3x^3 = \mathcal{P}_{out}^*. \quad (5.24)$$

The constraint equation above is only a polynomial of order three, so it can be easily solved and we can find the root that minimizes the cost function.

5.3 Multi-Relay Scenario

In this section, we extend the protocol described in Section 5.1 to the case when there is more than one potential relay. Let N be the number of relays assigned to help a given source. The cooperation protocol then works as an N -stage ARQ protocol as follows. The source node transmits its packets to the destination and the relays try to decode this packet. If the destination does not decode the packet correctly, it sends a NACK that can be heard by the relays. If the first relay is able to decode the packet correctly, it forwards the packet with power P_1 to the destination. If the destination does not receive correctly again, then it sends a NACK and the second candidate relay, if it received the packet correctly, forwards the source's packet to the destination with power P_2 . This is repeated until the destination gets the packet correctly or the N trials corresponding to the N relays are exhausted.

We model the status of any relay by 1 or 0, corresponding to whether the relay received the source's packet correctly or not, respectively. Writing the status of all the relays in a column vector results in a $N \times 1$ vector whose entries are

either 0 or 1. Hence, the decimal number representing this $N \times 1$ vector can take any integer value between 0 and $2^N - 1$. Denote this vector by S_k where $k \in \{0, 1, 2, \dots, 2^N - 1\}$.

For a given status of the N relays, an outage occurs if and only if the links between the relays that decoded correctly and the destination are all in outage. Denote the set of the relays that received correctly by $\chi(S_k) = \{i : S_k(i) = 1, 1 \leq i \leq N\}$, and $\chi^c(S_k)$ as the set of relays that have not received correctly, i.e., $\chi^c(S_k) = \{i : S_k(i) = 0, 1 \leq i \leq N\}$. The conditional probability of outage given the relays status S_k is thus given by

$$\mathcal{P}_{OC|S_k} = \mathcal{P} \left(SNR_{sd} \leq \beta \bigcap_{j \in \chi(S_k)} (SNR_{jd} \leq \beta) \right), \quad (5.25)$$

The total outage probability is thus given by

$$\mathcal{P}_{OC} = \sum_{k=0}^{2^N-1} \mathcal{P}(S_k) \mathcal{P}_{OC|S_k}. \quad (5.26)$$

We then need to calculate the probability of the set S_k , which can then be written as

$$\mathcal{P}(S_k) = \mathcal{P} \left(\bigcap_{i \in \chi(S_k)} (SNR_{si} \geq \beta) \bigcap_{j \in \chi^c(S_k)} (SNR_{sj} \leq \beta) \right). \quad (5.27)$$

The average outage probability expression can thus be given by

$$\mathcal{P}_{OC} = \sum_{k=0}^{2^N-1} (1 - f(r_{sd}, P_s^c)) \prod_{j \in \chi(S_k)} (1 - f(r_{jd}, P_j)) f(r_{sj}, P_s^c) \prod_{j \in \chi^c(S_k)} (1 - f(r_{sj}, P_s^c)). \quad (5.28)$$

where P_j , $j \in \{1, 2, \dots, N\}$, is the power allocated to the j -th relay.

Next we compute the average total consumed power for the N -relays scenario.

First we condition on some relays' status vector $\chi(S_k)$

$$E [P_{tot}^c] = E [E [P_{tot}^c | \chi(S_k)]] = \sum_{k=0}^{2^N-1} P(\chi(S_k)) E [P_{tot}^c | \chi(S_k)]; \quad (5.29)$$

For a given $\chi(S_k)$, we can further condition on whether the source get the packet through from the first trial or not. This event happens with probability $f(r_{s,d}, P_s^c)$, and the consumed power in this case is given by

$$P_{tot}^{c,1} = P_s^c + (N + 1)P_r + P_c; \quad (5.30)$$

The complementary event that the source failed to transmit its packet from the direct transmission phase happens with probability $1 - f(r_{s,d}, P_s^c)$, and this event can be further divided into two mutually exclusive events. The first is when the first $|\chi(S_k)| - 1$ relays from the set $\chi(S_k)$ fails to forward the packet and this happens with probability $\prod_{i=1}^{|\chi(S_k)|-1} (1 - f(r_{i,d}, P_i))$ and the corresponding consumed power is given by

$$P_{tot}^{c,2} = P_s^c + (N + 1 + |\chi(S_k)|) P_r + (|\chi(S_k)| + 1) P_c + \sum_{n=1}^{|\chi(S_k)|} P_{\chi(S_k)(n)}; \quad (5.31)$$

And the second is when one of the intermediate relays in the set $\chi(S_k)$ successfully forwards the packet and this happens with probability $\prod_{m=1}^{j-1} (1 - f(r_{m,d}, P_m)) f(r_{j,d}, P_j)$ if this intermediate relay was relay number j , and the corresponding power is given by

$$P_{tot}^{c,3,j} = P_s^c + (N + 1 + j)P_r + (1 + j)P_c + \sum_{i=1}^j P_{\chi(S_k)(i)}. \quad (5.32)$$

From (5.29), (5.30), (5.31), and (5.32), the average total consumed power can be given by

$$\begin{aligned} E [P_{tot}^c] = & \sum_{k=0}^{2^N-1} P(\chi(S_k)) \left\{ f(r_{s,d}, P_s^c) P_{tot}^{c,1} \right. \\ & + (1 - f(r_{s,d}, P_s^c)) \left[\prod_{i=1}^{|\chi(S_k)|-1} (1 - f(r_{i,d}, P_i)) P_{tot}^{c,2} \right. \\ & \left. \left. + \sum_{j=1}^{|\chi(S_k)|-1} \prod_{m=1}^{j-1} (1 - f(r_{m,d}, P_m)) f(r_{j,d}, P_j) P_{tot}^{c,3,j} \right] \right\}. \end{aligned} \quad (5.33)$$

The optimization problem can then be written as

$$\min_{\mathbf{P}} P_{tot}^C(\mathbf{P}), \quad \text{s.t. } \mathcal{P}_{OC}(\mathbf{P}) \leq \mathcal{P}_{out}^*. \quad (5.34)$$

where $\mathbf{P} = [P_s^c, P_1, P_2, \dots, P_N]^T$.

5.4 Experimental and Simulation Results

5.4.1 Experimental Results

In our system model we have assumed the channel independence between the following links: the source-relay link, the source-destination link, and the relay-destination link. Moreover, a strong motivation for applying cooperative transmission instead of ARQ in the time domain, is the assumption of high temporal correlation which results in delay and requires performing interleaving at the transmitter side. In this section, we have conducted a set of experiments to justify these two fundamental assumptions.

The experiments are set up as follows. We have three wireless nodes in the experiments, one of them acts as the sender and the other two act as receivers. Each wireless node is computer equipped with a IEEE 802.11g wireless card, specifically, we utilized three LINKSYS wireless-G USB network adaptors. The sender's role is to broadcast data packets with a constant rate, while the two receivers' role is to decode the packets and record which packet is erroneous. The traffic rate is 100 packets per second, and the size of each packet is 554 bytes (including packet headers). The two receivers are placed together, with the distance between them being 20cm. The distance between the transmitter and the receiver is around 5 meters. The experiments have been mainly conducted in office environments. The experiments results, which are illustrated next, have revealed two important

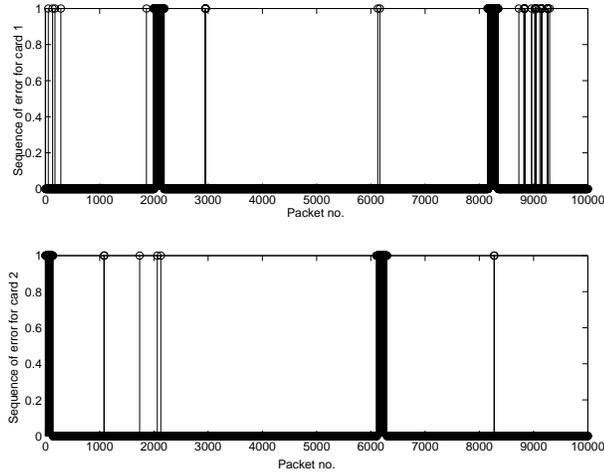


Figure 5.2: Sequence of packet errors at the two utilized wireless cards.

observations: the channels exhibit strong time correlation for each receiver, while there is negligible dependence between the two receivers. Fig. 5.2 illustrates one instantiation of the experiments. The first figure illustrates the results obtained at the first receiver and the second figure is for the second receiver.

For each figure, the horizontal axis denotes the sequence number of the first 100000 packets, and the vertical axis denotes whether a packet is erroneous or not. First, from these results we can see that packet errors exhibit strong correlation in time. For example, for the first receiver, most erroneous packets cluster at around 22^{nd} second and around 83^{rd} second. Similar observations also hold for the second receiver. If we take a further look at the results we can see that in this set of experiments the duration for the cluster is around 2 seconds. To help better understand the time correlation of erroneous packets, we have also used a two-state Markov chain to model the channel, as illustrated in Fig. 5.3. In this model “1” denotes that the packet is correct, and “0” denotes that the packet is erroneous. $P_{i|j}$ denotes the transition probability from state i to state j , that

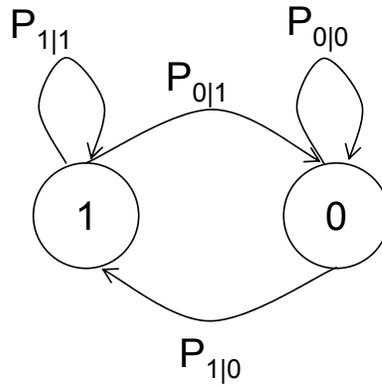


Figure 5.3: Modelling the channel by a two (on-off) state Markov chain to study the time correlation.

is, the probability to reach state j given the previous state is i . The following transition probabilities have been obtained after using the experimental results to train the model: $P_{1|0} = 0.03$, $P_{1|1} = 0.999$, $P_{0|0} = 0.97$, $P_{0|1} = 0.001$. These results also indicate strong time correlation. For example, given the current received packet is erroneous, the probability that the next packet is also erroneous is around $P_{0|0} = 0.97$.

Now we take a comparative look at the results obtained at the two receivers. From these results we can see that although there exists slight correlation in packet errors between the two receivers, it is almost negligible. To provide more concrete evidence of independence, we have estimated the correlation between the two receivers using the obtained experiment results. Specifically, we have measured the correlation coefficient between the received sequences at the two receivers and we found that the correlation coefficient is almost 0 which indicates a strong spatial independence between the two receivers.

5.4.2 Simulation Results

As discussed in the previous sections, there are different system parameters that can control whether we can gain from cooperation or not. Among which are the received power consumption, the processing power, the SNR threshold, the power amplifier loss, and the relative distances between the source, relay, and destination.

In order to understand the effect of each of these parameters, we are going to study the performance of cooperative and direct transmission when varying one of these parameters and fixing the rest. This is described in more details in the following. In all of the simulations, the aforementioned parameters take the following values when considered fixed: $\alpha = 0.3$, $\beta = 10$, $N_o = 10^{-3}$, $P_c = 10^{-4}$ Watt, $P_r = 5 \times 10^{-5}$, $QoS = \mathcal{P}_{out}^* = 10^{-4}$. We define the cooperation gain as the ratio between the total power required for direct transmission to achieve a certain QoS, and the total power required by cooperation to achieve the same QoS.

First, we study the effect of varying the receive power P_r as depicted in Fig. 5.4. We plot the cooperation gain versus the distance between the source and the destination for different values of receive power $P_r = 10^{-4}, 5 \times 10^{-5}, 10^{-5}$ Watt. At source-destination distances below 20m, the results reveal that direct transmission is more energy efficient than cooperation, i.e., the overhead in receive and processing power due to cooperation outweighs its gains in saving the transmit power. For $r_{sd} > 20m$, the cooperation gain starts increasing as the transmit power starts constituting a significant portion of the total consumed power. This ratio increases until the transmit power is the dominant part of the total consumed power and hence the cooperation gain starts to saturate.

In the plotted curves, the solid lines denote the cooperation gain when utilizing optimal power allocation at the source and the relay, while the dotted curves

denote the gain for equal power allocation. For $r_{sd} \leq 100\text{m}$, both optimal power allocation and equal power allocation almost yield the same cooperation gain. For larger distances, however, a gap starts to appear between optimal and equal power allocation. The rationale behind these observations is that at small distances the transmit power is a small percentage of the total consumed power and hence optimal and equal power allocation almost have the same behavior, while at larger distances, transmit power plays a more important role and hence a gap starts to appear.

In Fig. 5.5 we study the effect of changing the SNR threshold β . The distance between source and destination r_{sd} is fixed to 100m. It is clear that the cooperation gain increases with increasing β , and that for the considered values of the system parameters, equal power allocation provides almost the same gains as optimal power allocation. In Fig. 5.6 we study the effect of the power amplifier loss α . In this case, we plot the total consumed power for cooperation and direct transmission versus distance for different values of α . Again below 20m separation between the source and the destination, direct transmission provides better performance over cooperation. It can also be seen from the plotted curves that the required power for direct transmission is more sensitive to variations in α than the power required for cooperation. The reason is that the transmit power constitutes a larger portion in the total consumed power in direct transmission than in cooperation, and hence the effect of α is more significant. The QoS, measured by the required outage probability, has similar behavior and the results are depicted in Fig. (5.7).

Next we study the effect of varying the relay location. We consider three different positions for the relay, close to the source, in the middle between the source and the destination, and close to the destination. In particular, the relay

position is taken equal to $(r_{sl} = 0.2r_{sd}, r_{ld} = 0.8r_{sd})$, $(r_{sl} = 0.5r_{sd}, r_{ld} = 0.5r_{sd})$, and $(r_{sl} = 0.8r_{sd}, r_{ld} = 0.2r_{sd})$.

Figs. 5.8 and 5.9 depict the power required for cooperation and direct transmission versus r_{sd} for equal power and optimal power allocation, respectively. In the equal power allocation scenario, the relay in the middle gives the best results, and the other two scenarios, relay close to source and relay close to destination provide the same performance. This can be expected because for the equal power allocation scenario the problem becomes symmetric in the source-relay and relay-destination distances. For the optimal power allocation scenario depicted in Fig. 5.9, the problem is no more symmetric because different power allocation is allowed at the source and relay. In this case, numerical results show that the closer the relay to the source the better the performance. The intuition behind this is that when the relay is closer to the source, the source-relay channel is very good and almost error-free.

From both figures, it is also clear that for small source-destination separation r_{sd} , equal and optimal power allocation almost provide the same cooperation gain while for larger r_{sd} optimal power allocation provides more gain. Another important observation is that at small distances below 100m, the location of the relay does not affect the performance much. This makes the algorithms required to select a relay in cooperative communications simpler to implement for source-destination separations in this range. Finally, the threshold behavior below 20m still appears where direct transmission becomes more energy efficient.

Fig. 5.10 depicts the multiple relays scenario for different values of outage probability \mathcal{P}_{out}^* . The results are depicted for a source-destination distance of 100m, and for $N = 0, 1, 2, 3$ relays, where $N = 0$ refers to direct transmission. As

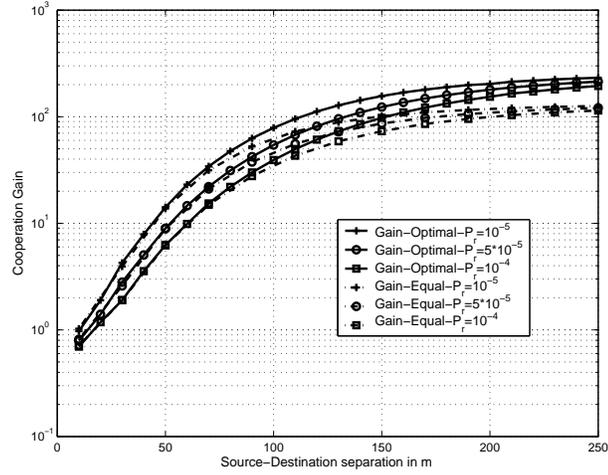


Figure 5.4: Cooperation gain versus the source-destination distance for different values of received power consumption .

shown in Fig. 5.10, for small values of required outage probability, one relay is more energy efficient than two or three relays. As we increase the required QoS, reflected by \mathcal{P}_{out}^* , the optimal number of relays increases. Hence, our analytical framework can also provide guidelines to determining the optimal number of relays under any given scenario.

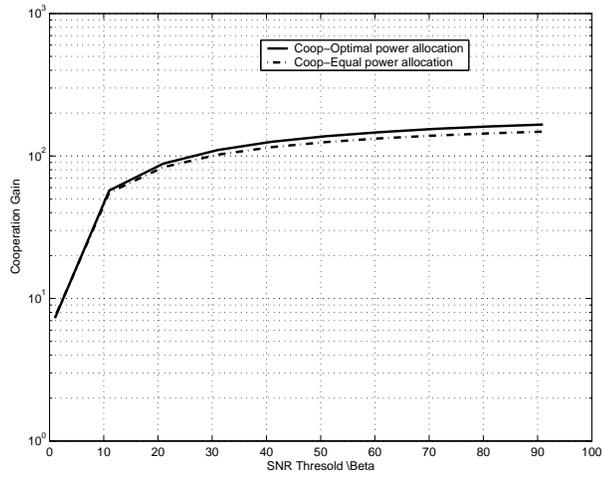


Figure 5.5: Cooperation gain versus the SNR threshold β .

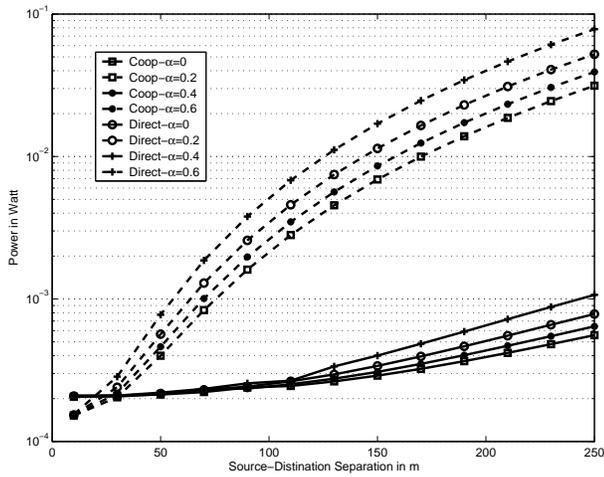


Figure 5.6: Optimal power consumption for both cooperation and direct transmission scenarios for different values of power amplifier loss α .

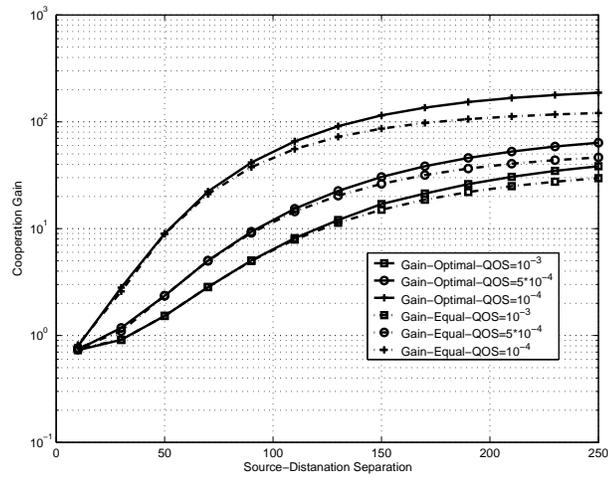


Figure 5.7: Cooperation gain versus the source-destination distance for different values of QoS.

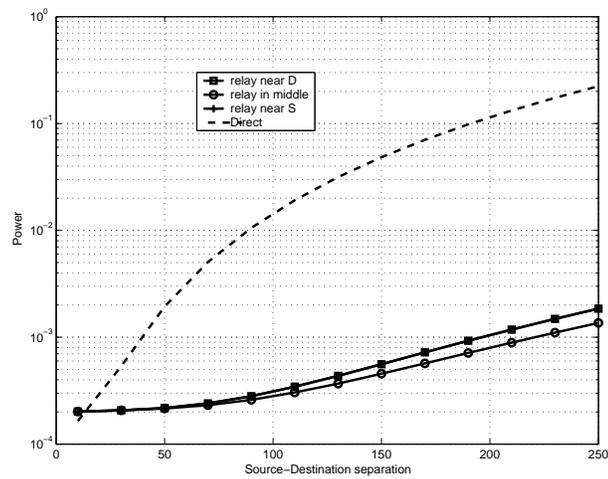


Figure 5.8: Optimal Consumed Power versus distance for different relay locations for equal power allocation at source and relay.

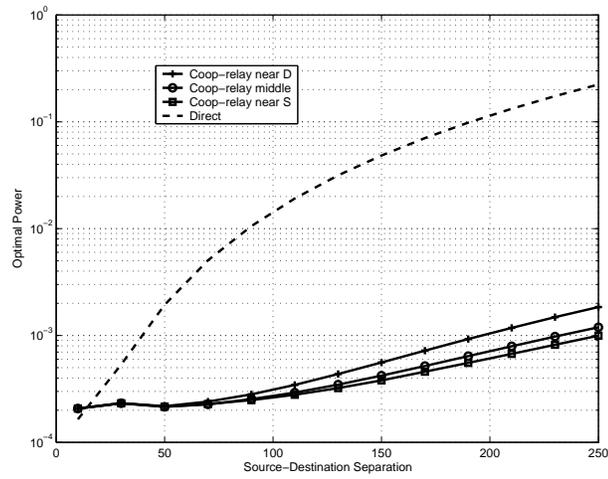


Figure 5.9: Optimal Consumed Power versus distance for different relay locations for optimal power allocation at source and relay.

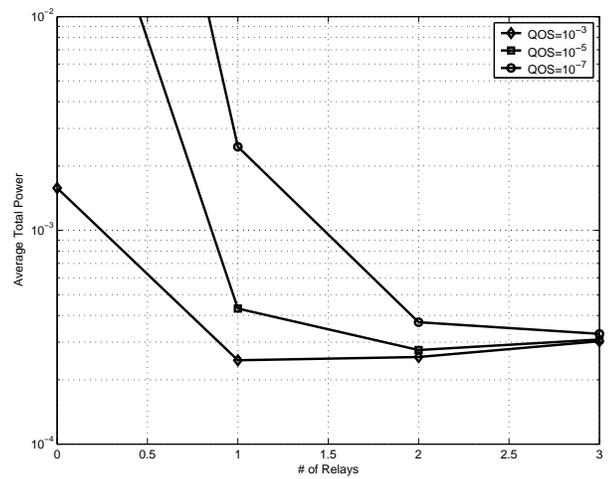


Figure 5.10: Optimal Consumed Power versus number of relays for different values of required outage probability.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

In this thesis we have developed and analyzed cooperative communications protocols for wireless networks that can achieve significantly higher energy and bandwidth efficiency compared to non-cooperative schemes. Our results indicate that cooperative communications indeed offers a paradigm shift to the design of wireless networks that can yield to achieving the increasing demands of future wireless applications. More specifically, we have addressed the following problems.

In Chapter 2, we proposed a class of cooperative diversity protocols for multi-node wireless networks employing decode-and forward relaying. This class of protocols consists of schemes in which each relay can combine the signals arriving from an arbitrary but fixed number of previous relays along with that received from the source. We derived exact expressions for the SER of a general cooperation scheme for both MPSK and MQAM modulation. Also, we provided approximations for the SER which are shown to be tight at high enough SNR. Our theoretical analysis reveals a very interesting result: this class of cooperative protocols shares the

same asymptotic performance at high enough SNR. Thus the performance of a simple cooperation scenario in which each relay combines the signals arriving from the previous relay and the source is asymptotically exactly the same as that for the most complicated scenario in which each relay combines the signals arriving from all the previous relays and the source. The analysis also reveals that the proposed protocols achieve full diversity gain in the number of cooperating terminals. Moreover, we formulated the optimal power allocation problem, and show that the optimum power allocated at the nodes for an arbitrary network follow a certain ordering. We found that the optimal power allocation scheme does not depend on the quality of the direct link between the source and the destination. Furthermore, we provided closed form solutions for the optimal power allocation for some network topologies of practical interest, and we showed through numerical examples that our theoretical results match with the simulation results.

In Chapter 3, we addressed the relay-assignment problem for coverage extension in cooperative transmission over wireless networks based on the knowledge of the channel statistics governed by the users' spatial distribution. We proposed two distributed relay-assignment protocols. The Nearest-neighbor is a simple algorithm in which the relay is selected to be the nearest neighbor to the user. We also considered the scenario where fixed relays are deployed in the network to help the existing users. Outage performance of the proposed protocols was analyzed. We further developed lower bounds on the performance of any relay assignment protocol via a Genie aided method. Our numerical results indicate significant gains in the system performance. In particular, fixing the average transmit power, significant increase in the coverage area of the network can be achieved by our simple distributed protocols. Similarly, fixing the cell radius, the average power

required to achieve a certain outage probability is significantly reduced.

In Chapter 4, we studied the impact of cooperative communications at the multiple-access layer. We introduced a new cognitive multiple-access protocol in the presence of a relay in the network. The relay senses the channel for idle channel resources and exploits them to cooperate with the terminals in forwarding their packets. We developed two protocols to implement the proposed multiple-access strategy, namely, CCMA-S and CCMA-Me. We characterized the maximum stable throughput region of the proposed protocols and compared them to some existing adaptive relaying strategies, non-cooperative TDMA, and random-access ALOHA. Moreover, we studied the delay performance of the proposed protocols. Our analysis reveals significant performance gains of the proposed protocols over their non-cognitive counterparts. This is because the proposed multiple-access strategies do not result in any bandwidth loss, as cooperation is enabled only in idle “*unused*” channel resources, which results in a graceful degradation of the maximum stable throughput when increasing the communication rate. On the other hand, the maximum stable throughput of non-cognitive relaying strategies as selection and incremental relaying suffer from catastrophic degradation with increasing the communication rate, because of their inherent bandwidth inefficiency. Furthermore, we showed that the throughput region of the proposed protocols is a subset from their maximum stable throughput region, which is different from the case in ALOHA where it is conjectured that both regions are identical.

In Chapter 5, we consider a practical framework for analyzing the performance of cooperative transmission in sensor networks by considering the extra overhead induced by enabling cooperation. This extra overhead appears in the extra processing and receiving powers at the relays and destination. Our analytical and

numerical results reveal an interesting threshold behavior that separates regions where direct transmission is better from regions where cooperation prevails. We also show that under certain scenarios, equal power allocation has very close performance to optimal power allocation. Moreover, we show that for small distances between the source and the destination, the performance is not sensitive to relay location, which leads to simpler relay assignment algorithms. Our results also reveal that cooperative communication is more robust to poor power amplifier designs compared to direct transmission. Our analytical framework can also be utilized to determine the optimal number of relays for any given scenario. In summary, we provide important guidelines for wireless sensor network designers to decide when and how to apply the cooperative communication paradigm, and when is direct transmission more energy efficient.

6.2 Future Work

There are several problems that can lead to fruitful research in cooperative communications and networking, among which we believe the following are of high relevance.

6.2.1 Cooperation among Correlated Sources

This problem is of high relevance to sensor networks as the measurements taken at different sensors may have some correlation structure. For the three node model, we can think of the relay as having a correlated version of the signal transmitted by the source. New relaying strategies that take into account such correlation structure need to be developed. In decode-and-forward schemes, for example, the relay

can utilize this correlation as side information for decoding the source's message. Another possibility is that the relay can communicate its information simultaneously to the base-station using some orthogonal codes or distributed space-time codes saving the bandwidth loss in conventional cooperation. Distributed source-coding can be done jointly with cooperation (distributed space-coding) to optimize some distortion function depending on the problem.

6.2.2 Cognitive Cooperative Communications

It is well established now that the scarce spectrum is under-utilized (spectral holes). One way to reuse it is by introducing a secondary system that opportunistically shares the spectrum of the primary users through the utilization of cognitive radio. In a much broader sense, one can think that these spectral holes can also be used to enhance the performance of the primary system. An interesting and important problem to investigate is how to divide the new resources “*spectral holes*” between cooperation (to enhance the primary system performance), and spectrum sharing by secondary users, and what is the fundamental tradeoff between them. This depends on the type of data and required quality of service for both systems. Another important problem that is relevant to the two scenarios above, is the issue of spectrum sensing. This should be done in a distributed cooperative manner to be robust against the effects of channel fading, shadowing, and the uncertainty of the interference signal.

BIBLIOGRAPHY

- [1] J. G. Proakis, *Digital Communications*, McGraw-Hill Inc., 1994.
- [2] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.
- [3] E. Telatar, “Capacity of multi-antenna gaussian channels,” *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.
- [4] G. J. Foschini and M. Gans, “On the limits of wireless communication in a fading environment when using multiple antennas,” *Wireless Personal Communication*, vol. 6, pp. 311–335, Mar 1998.
- [5] S. M. Alamouti, “A simple transmit diversity technique for wireless communications,” *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.
- [6] V. Tarokh, N. Seshadri, and A. R. Calderbank, “Space-time codes for high data rate wireless communication: Performance criterion and code construction,” *IEEE Trans. Information Theory*, vol. 44, no. 2, pp. 744–765, March 1998.
- [7] V. Tarokh, H. Jafarkhani, and A.R. Calderbank, “Space–time block coding for wireless communications: Performance results,” *IEEE Journal on Select Areas in Communications*, vol. 17, no. 3, pp. 451–460, Mar. 1999.
- [8] L. J. Cimini, “Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing,” *IEEE Transactions on Communications*, vol. 33, pp. 665–675, July 1985.
- [9] H. Bölcskei and A. J. Paulraj, “Space-frequency codes for broadband fading channels,” *Proc. ISIT’2001*, June 2001.
- [10] W. Su, Z. Safar, and K. J. R. Liu, “Full-rate full-diversity space-frequency codes with optimum coding advantage,” *IEEE Trans. Information Theory*, vol. 51, no. 1, pp. 229–249, Jan. 2005.

- [11] W. Su, Z. Safar, M. Olfat, and K. J. R. Liu, "Obtaining full-diversity space-frequency codes from space-time codes via mapping," *IEEE Trans. on Signal Processing (Special Issue on MIMO Wireless Communications)*, vol. 51, pp. 2905–2916, Nov 2003.
- [12] W. Su, Z. Safar, and K. J. R. Liu, "Towards maximum achievable diversity in space, time and frequency: performance analysis and code design," *IEEE Trans. on Wireless Communications*, vol. 4, pp. 1847, July 2005.
- [13] A. K. Sadek, W. Su, and K. J. R. Liu, "Diversity analysis for frequency-selective mimo-ofdm systems with general spatial and temporal correlation model," *IEEE Transactions on Communications*, vol. 54, pp. 878–888, May 2006.
- [14] A. K. Sadek, W. Su, and K. J. R. Liu, "Maximum achievable diversity for arbitrary spatially correlated mimo-ofdm systems," *Proc. IEEE Globecom*, Dec 2004.
- [15] A. K. Sadek, W. Su, and K. J. R. Liu, "Eigen-selection approach for joint beam-forming and space-frequency coding in mimo-ofdm systems with spatial correlation feedback," *Proceedings IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2005.
- [16] A. K. Sadek, W. Su, and K. J. R. Liu, "Transmit beamforming design for space-frequency coded mimo-ofdm systems with spatial correlation feedback," *IEEE Transactions on Communications, revised*, 2006.
- [17] A. Sendonaris, E. Erkip, and B. Aazhang, "Increasing uplink capacity via user cooperation diversity," *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Aug 1998.
- [18] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity, part I: System description," *IEEE Trans. Communications*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.
- [19] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity, part II: Implementation aspects and performance analysis," *IEEE Trans. Communications*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.
- [20] J. N. Laneman and G. W. Wornell, "Exploiting distributed spatial diversity in wireless networks," in *Proc. Allerton Conf. Communications, Control, and Computing*, Oct 2000.
- [21] J. N. Laneman, G. W. Wornell, and D. N. C. Tse, "An efficient protocol for realizing cooperative diversity in wireless networks," *Proc. IEEE Int. Symp. Information Theory (ISIT)*, June 2001.

- [22] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, pp. 3062–3080, Dec 2004.
- [23] J. N. Laneman and G. W. Wornell, "Distributed space-time coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2415–2525, Oct. 2003.
- [24] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Probab.*, 1971.
- [25] T. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. on Information Theory*, vol. 25, no. 5, pp. 572–584, Sept 1979.
- [26] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley Inc., 1991.
- [27] G. Kramer, M. Gatspar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Information Theory*, vol. 51, no. 9, pp. 3037–3063, Sept. 2005.
- [28] L. Sankaranarayanan, G. Kramer, and N. B. Mandayam, "Capacity theorems for the multiple-access relay channel," *42nd Annual Allerton Conference on Communication, Control, and Computing, Allerton, IL*, Sept. 2004.
- [29] Chris T. K. Ng and Andrea J. Goldsmith, "Capacity gain from transmitter and receiver cooperation," *IEEE International Symposium on Information Theory (ISIT), Adelaide, Australia*, Sept. 2005.
- [30] T. E. Hunter and A. Nosratinia, "Performance analysis of coded cooperation diversity," *2003 Int. Conf. on Comm. ICC03*, vol. 4, pp. 2688–2692, 2003.
- [31] L. H. Ozarog, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, no. 5, pp. 359–378, May 1994.
- [32] P. A. Anghel, G. Leus, and M. Kaveh, "Multi-user space-time coding in cooperative networks," *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 6-10 2003.
- [33] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *the 3rd IEEE Sensor Array and Multi-Channel Signal Processing Workshop*, July 18-21 2004.
- [34] G. Scutari and S. Barbarossa, "Distributed space-time coding for regenerative relay networks," *IEEE Trans. on Wireless Communications*, vol. 4, no. 5, pp. 2387 – 2399, Sept 2005.

- [35] Huining Hu, Halim Yanikomeroglu, David D. Falconer, and Shalini Periyalwar, “Distributed space-time coding for regenerative relay networks,” *Proceedings IEEE Globecom*, Dec 2004.
- [36] K. G. Seddik, A. K. Sadek, and K. J. R. Liu, “Protocol-aware design criteria and performance analysis for distributed space-time coding,” in *Proc. IEEE GLOBECOM*, Dec. 2006.
- [37] W. Su, A. K. Sadek, and K. J. R. Liu, “Ser performance analysis and optimum power allocation for decode-and-forward cooperation protocol in wireless networks,” *Proceedings of WCNC*, pp. 984 – 989, Mar 2005.
- [38] W. Su, A. K. Sadek, and K. J. R. Liu, “Cooperative communications in wireless networks: Performance analysis and optimum power allocation,” *submitted to Wireless Personal Communications, in revision*.
- [39] P. A. Anghel and M. Kaveh, “Exact symbol error probability of a cooperative network in a rayleigh fading environment,” *IEEE Transactions on Wireless Communications*, vol. 3, pp. 1416–1421, Sep. 2004.
- [40] J. Boyer, D. D. Falconer, and H. Yanikomeroglu, “Multihop diversity in wireless relaying channels,” *IEEE Trans. on Communications*, vol. 52, no. 10, October 2004.
- [41] A. K. Sadek, W. Su, and K. J. R. Liu, “Performance analysis for multi-node decode-and-forward relaying in collaborative wireless networks,” *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, March 19-23 2005.
- [42] A. K. Sadek, W. Su, and K. J. R. Liu, “A class of cooperative communication protocols for multi-node wireless networks,” *Proc. IEEE International Workshop on Signal Processing Advances in Wireless Communications SPAWC*, pp. 560 – 564, June 2005.
- [43] A. K. Sadek, W. Su, and K. J. R. Liu, “Multi-node cooperative communications in wireless networks,” *IEEE Transactions on Signal Processing*, vol. 55, pp. 341–355, Jan 2007.
- [44] A. K. Sadek, Z. Han, and K.J.R. Liu, “Multi-node cooperative resource allocation to improve coverage area in wireless networks,” *Proc. IEEE Globecom*, Nov 2005.
- [45] A. K. Sadek, Z. Han, and K. J. R. Liu, “A distributed relay-assignment algorithm for cooperative communications in wireless networks,” *Proc. IEEE International Conference on Communications (ICC), Istanbul.*, June 2006.

- [46] A. K. Sadek, Z. Han, and K. J. R. Liu, "Relay-assignment protocols for coverage extension in cooperative communications over wireless networks," *IEEE Transactions on Wireless Communications*, submitted, 2007.
- [47] A. K. Sadek, K. J. R. Liu, and A. Ephremides, "Collaborative multiple-access for wireless networks: Protocols design and stability analysis," *Proc. IEEE CISS, Princeton, NJ*, March 2006.
- [48] A. K. Sadek, K. J. R. Liu, and A. Ephremides, "Collaborative multiple-access protocols for wireless networks," *Proc IEEE International Conference on Communications (ICC)*, June 2006.
- [49] A. K. Sadek, K. J. R. Liu, and A. Ephremides, "Cognitive multiple-access via cooperation: Protocol design and performance analysis," *IEEE Transactions on Information Theory*, revised, 2007.
- [50] A. K. Sadek, W. Yu, and K. J. R. Liu, "Does cooperation have better performance in sensor networks?," *Proc. IEEE SECON*, 2006.
- [51] A. K. Sadek, W. Yu, and K. J. R. Liu, "On the energy efficiency of cooperative communications in wireless sensor networks," *ACM Transactions on Sensor Networks*, submitted, 2006.
- [52] K. G. Seddik, A. K. Sadek, Weifeng Su, and K. J. R. Liu, "Outage analysis and optimal power allocation for multi-node relay networks," *accepted for publication IEEE Signal Processing Letters*.
- [53] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. Information Theory*, vol. 49, pp. 1073–1096, May 2003.
- [54] M. K. Simon and M. S. Alouini, "A unified approach to the performance analysis of digital communication over generalized fading channels," *Proc. IEEE*, vol. 86, no. 9, pp. 1860–1877, September 1998.
- [55] H. S. Wang and P.-C. Chang, "On verifying the first-order markovian assumption for a rayleigh fading channel model," *IEEE Trans. Vehicular Technology*, vol. 45, no. 2, pp. 353–357, 1996.
- [56] Z. Lin and E. Erkip, "Relay search algorithms for coded cooperative systems," in *Proc. of 2005 GLOBECOM Communication Theory Symposium*, Dec 2005.
- [57] J. Luo, R. S. Blum, L. J. Greenstein, L. J. Cimini, and A. M. Haimovich, "New approaches for cooperative use of multiple antennas in ad-hoc wireless networks," *Proc. IEEE Vehicular Technology Conference (VTC)-Fall*, Sept 2004.

- [58] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 4-th edition, Academic Press, 1980.
- [59] H. Zheng and C. Peng, "Collaboration and fairness in opportunistic spectrum access," *IEEE International Conference on Communications (ICC)*, May 2005.
- [60] B. Tsybakov and W. Mikhailov, "Ergodicity of slotted aloha system," *Probl. Peredachii Info.*, vol. 15, pp. 73–87, 1979.
- [61] R. Rao and A. Ephremides, "On the stability of interacting queues in a multi-access system," *IEEE Trans. Inform. Theory*, vol. 34, pp. 918–930, Sept. 1988.
- [62] W. Szpankowski, "Stability conditions for some multiqueue distributed systems: Buffered random access systems," *Adv. Appl. Probab.*, vol. 26, pp. 498–515, 1994.
- [63] W. Luo and A. Ephremides, "Stability of n interacting queues in random-access systems," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1579–1587, July 1999.
- [64] V. Naware, G. Mergen, and L. Tong, "Stability and delay of finite-user slotted aloha with multipacket reception," *IEEE Trans. Info. Theory*, vol. 51, no. 7, pp. 2636–2656, July 2005.
- [65] J. Luo and A. Ephremides, "On the throughput, capacity and stability regions of random multiple access," *accepted in IEEE Trans. on Information Theory*.
- [66] P. Nain, "Analysis of a two node aloha network with infinite capacity buffers," in *Proc. Int. Sem. Computer Networking, Performance Evaluation*, p. 2.2.12.2.16, 1985.
- [67] M. Sidi and A. Segall, "Two interfering queues in packet-radio networks," *IEEE Trans. Commun.*, vol. COM-31, no. 1, pp. 123129, Jan. 1983.
- [68] G. Scutari, S. Barbarossa, and D. Ludovici, "Impact of multiple transmit antennas in a queued sdma/tdma downlink," *Proc. of the IEEE Int. Workshop on Signal Processing Advances for Wireless Communications*, pp. 170–174, June 2003.
- [69] R. Lin and A. P. Petropulu, "Cooperative transmission for random access wireless networks," in *Proc. of Intl. Conf. on Acoust. Speech and Sig. Proc.*, March 2005.

- [70] A. Ribeiro, N.D. Sidiropoulos, and G.B. Giannakis, “Achieving wireline random access throughput in wireless networking via user cooperation,” in *Proc. of IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC '05)*, pp. 5–8, June 2005.
- [71] M. K. Tsatsanis, R. Zhang, and S. Banerjee, “Network-assisted diversity for random access wireless networks,” *IEEE Trans. on Signal Processing*, vol. 48, pp. 702–711, March 2000.
- [72] M. Kobayashi, G. Caire, and D. Gesbert, “Impact of multiple transmit antennas in a queued sdma/tdma downlink,” in *Proc. SPAWC 2005, 6th IEEE Workshop on Signal Processing Advances in Wireless Communications*, June 2005.
- [73] R. M. Loynes, “The stability of a queue with non-independent inter-arrival and service times,” *Proc. Cambridge Philos. Soc.*, pp. 497–520, 1962.
- [74] B. Hajek, A. Krishna, and R. O. LaMaire, “On the capture probability for a large number of stations,” *IEEE Trans. on Communications*, vol. 45, no. 2, pp. 254–260, Feb. 1997.
- [75] M. Zorzi and R. Rao, “Capture and retransmission control in mobile radio,” *IEEE JSAC*, vol. 12, pp. 1289–1298, Oct. 1994.
- [76] D. Bertsekas and R. Gallager, *Data Networks, 2nd edition.*, Prentice Hall, 1991.
- [77] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge Univ. Press, 2004.
- [78] S. Cui, A. J. Goldsmith, and A. Bahai, “Energy-efficiency of mimo and cooperative mimo in sensor networks,” *IEEE Journal on Selected Areas of Communications*, vol. 22, no. 6, pp. 1089–1098, August 2004.