COOPERATIVE STRATEGIES FOR WIRELESS RELAY NETWORKS

BY IVANA MARIĆ

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Electrical and Computer Engineering

Written under the direction of

Professor Roy D. Yates

and approved by

New Brunswick, New Jersey

October, 2006

ABSTRACT OF THE DISSERTATION

Cooperative Strategies for Wireless Relay Networks

by Ivana Marić

Dissertation Director: Professor Roy D. Yates

The goal of this thesis is to understand the mechanisms and potential benefits of relaying and node cooperation in wireless networks. We analyze specific cooperative schemes, present practical cooperative protocols for large networks, and derive capacity results for systems with limited transmitter cooperation.

Motivated by sensor applications, the first part of the thesis considers cooperation in large, energy-constrained networks that have to deliver multicast data. We use the insights offered by network information theory to propose an *accumulative broadcast* strategy that allows nodes to collect energy of unreliably received overheard signals. As a message is forwarded through the network, nodes will have multiple opportunities to reliably decode the message by collecting energy during each retransmission. We analyze two problems concerned with energy-efficient multicast/broadcast. First, we formulate the minimum-energy accumulative broadcast problem. We show that the problem is NP-complete and propose an energy-efficient heuristic algorithm. We then address the maximum lifetime multicast problem and present the *Maximum Lifetime Accumulative Broadcast (MLAB)* algorithm that finds the optimum solution. The resulting transmit power levels ensure that the lifetimes of the active relays are the same, causing them to fail simultaneously.

The proposed broadcast scheme employs a decode-and-forward (DF) relay strategy. In general, however, the optimal relay strategy is unknown. In the second part of the thesis, we evaluate the performance of amplify-and-forward (AF) strategies for energy-constrained networks. For a single source-destination pair, we characterize the optimum AF bandwidth and present the optimum power allocation among the AF relays which can be viewed as a form

of maximum ratio combining. Motivated by large bandwidth resources we further consider orthogonal signaling at the nodes. While the result for the optimum bandwidth still holds, the relay power solution in this case has the form of water-filling. In contrast, in a network with unconstrained bandwidth, the DF strategy operates in the wideband regime and requires a different choice of relays. Thus, in a large scale network, the choice of a cooperative strategy goes beyond determining a coding scheme at a node; it also determines the operating bandwidth and the best distribution of the relay power.

In the third part of the thesis, we introduce limited transmitter cooperation to the interference channel with two independent sources and two receivers. Transmitter cooperation enabled by side-channel links with finite capacities allows for a partial message exchange between encoders. After cooperation, each encoder will know a *common* message partially describing the original messages, and its own *private* message. We first determine the capacity region of the compound multiple-access channel (MAC) in which both common and private messages are decoded at both receivers. We then relax the decoding constraint and consider the *interference channel with common information* in which each private message is decoded only by one receiver. We determine the strong interference conditions under which the capacity region of this channel is found to coincide with the compound MAC capacity region. Finally, we determine the strong interference conditions for the *interference channel with unidirectional cooperation* in which messages sent at one encoder are known to the other encoder, but not vice versa.

Acknowledgements

I find this page to be the most impossible page in my thesis and the hardest to write. No matter what, it will fall short in acknowledging the true contribution of many people, especially that of my advisor, Professor Roy D. Yates. This thanks is against his advice. It is here to demonstrate my incapability to express the extent to which I admire him as a scientist, and as a person. I am grateful for the thoughtfulness he put into helping me develop my research skills, for teaching me high standards of research and everything that I know about networks operating in the low SNR regime. I am thankful for his patience with my slow learning at the beginning and pickiness in finding a job at the end. I am most grateful for all these years he let me step into his office whenever I needed to. His views and wisdom about both research and life have been and will always be a true inspiration for me.

I want to thank Dr. Larry Greenstein and Professor Christopher Rose for being on both my proposal and thesis committee, and for their suggestions that improved this thesis. I want to thank Professor Andrea Goldsmith and Professor Shlomo Shamai (Shitz) for being a part of the committee all the way from Stanford and Technion. I am very grateful to Professor Goldsmith for her kindness and support for my first ever ISIT presentation, and for her keen interest in my scientific future. Special thanks to Professor Shamai for all the advice and encouragement in the several months preceding this defense. I am indebted to him for all the long research discussions during which he shared with me some of his remarkable knowledge in the area on Information Theory.

Special thanks to Dr. Gerhard Kramer. His methodical approach and insight into Multiuser Information Theory is what I hope to acquire in years to come. My collaboration with him, Professor Shamai and Professor Goldsmith was an exceptional experience for me.

I am grateful to Professor Sergio Verdú for his interest in my work, a beautiful course on Information Theory that he allowed me to take, and for inviting me to present part of this thesis at Princeton.

I am also grateful to Professor Tsachy Weissman for giving me an opportunity to present my work at the Information Systems Lab at Stanford, and to Professor Giuseppe Caire for my presentation at Eurocom.

I wish to thank Ruoheng Liu for the exciting collaboration we had in the last several months.

WINLAB has been a wonderful place to work, and I thank Ivan Seskar for making me aware of it and handing me the application forms. I am grateful to Professor David Goodman and Professor Dipankar Raychaudhury for their support. Thank you, Kevin for putting up with all the nights of uncooked dinners, empty refrigerator and long discussions about my career. Finally, special thanks to my parents and my brother who, by life example showed me the beauty of an academic career and unselfishly pushed me to pursue it far from my home.

Table of Contents

List of Figures

Chapter 1

Background and Thesis Overview

The performance limits of wireless networks are still unknown. In the point-to-point channel, the optimum performance is characterized by its capacity, whereas in the multiple source-destination scenario occuring in wireless networks, the best possible performance is determined by the *capacity region.* For a given channel and a given set of power constraints at the transmitters, the capacity region is the closure of all the achievable rates for all the communicating pairs. The capacity region defining the best set of rates achievable in the network is known only for specific scenarios. One such scenario is the discrete memoryless multiaccess channel (MAC) [1, 2] in which many nodes wish to communicate independent messages to a single receiver. The MAC capacity region was determined by Ahlswede [1] and Liao [2], and for the case in which the transmitters share a common information in addition to their private messages, by Slepian and Wolf in [3]. A more recent success was the full understanding of the multiple-input multiple-output Gaussian (MIMO) broadcast channel with perfect channel state information (CSI) at both the transmitter and the receivers [4–7].

Under a simpler model that assumes only one source-destination pair, the problem becomes finding the capacity of the active link. The wireless network becomes the *multiple-relay channel* [8, 9] with all the nodes except the source and the sink acting as pure relays. However, the simpler assumption does not allow for a simple solution: the capacity of the relay channel is still unknown even for the case of a single relay node. The difficulty of these problems comes from their generality: a relay is allowed to use any causal coding strategy under its power constraint and the challenge is to find the optimal one. This problem is thus inherent for any multihop system due to the presence of the relays. While the answer to the ultimate question of wireless network capacity remains elusive to researchers, existing cooperative schemes have shown potential gains in wireless systems. As relaying is an unavoidable component of multi-terminal systems, building practical communication networks with good performance will require a choice of suitable cooperative schemes. The goal of this thesis is to understand the mechanisms and potential benefits of relaying and node cooperation in specific wireless networks.

Several coding strategies for relay networks have been proposed that led to capacity results for certain relay channels. For the special case of the physically degraded single-relay channel, the capacity was found in [10]. The physically degraded model however, assumes that the received signal at the destination is a noisy copy of the signal received at the relay and thus has no value given the observation made at the relay. It does not, therefore, model well the wireless situation where two receivers experience independent noise. Thus, the observations at both relay and destination are valuable in offering a diversity gain. Nevertheless, the coding strategies proposed in [10] were shown to be crucial also for more general network models [11–14]. The capacity-achieving coding strategy in [10] required block Markov superposition encoding, random partitioning of the message set and successive decoding. Two modifications that somewhat simplify the above scheme were proposed in [15–17]. To avoid the random partitioning, *backward decoding* [15] and *windowed decoding* [16] were used instead. The approach of [16] was extended to the general relay channel model and referred to as *multi-hopping* in [14]. When relays are close to the source, this strategy achieves the capacity for some wireless relay network models [13]. All of these strategies require the relay to reliably decode the source message before forwarding and are thus referred to as *decode-and-forward* (DF) [18].

A different paradigm in which a relay does not decode the message, but sends the compressed received values to the destination was proposed in [10] and extended to the multiple-relay channel in [12]. When the relays are close to the destination, this strategy, referred to as *quantize-and-forward* (QF) achieves the *antenna-clustering capacity* [12].

In another strategy that does not require reliable decoding at the relays, called *amplify-and-forward* (AF) [18], a relay forwards the scaled version of the received noisy copy of the source signal. Hence, the data is sent through only two hops with no cooperation among relays. Under the assumption of uncoded transmission at the source, it was shown that the two-hop AF strategy achieves the asymptotic capacity in a relay network in the limit as the number of relays becomes large [19,20]. It was further shown that in a random network, power efficiency of such strategy increases with the number of relays [21]. Even though the relay power in [19–22] was allocated suboptimally, the favorable scaling was achieved due to the coherent combining of the relay signals that increases the received SNR at the destination node.

Hybrid schemes based on DF, QF and AF strategies have been proposed in [23]. For a fading relay channel and co-located receivers these schemes were shown to have a superior performance, without requiring the relay to have the destination channel state information.

As the capacity-achieving examples above illustrate, the performance and hence the benefit of a relay coding strategy vary with network topology. In networks with a large number of nodes, a relay position might then determine the most suitable coding strategy for that relay, or alternatively, the best subset of relays may be chosen to employ a particular coding strategy. Furthermore, in the scenarios where the power allocation among relays is possible, we expect that there exists an optimal relay power allocation. For certain network and traffic scenarios, we will show such an optimum power allocation. Even beyond, we will argue that the choice of the coding strategy determines the optimum bandwidth allocation in a network.

Motivated by sensor network applications, the focus of the first part of this thesis will be on the energy-constrained networks in which energy, rather than bandwidth, is the limiting resource. We start by considering energy-efficient multicast and broadcast in Chapter 2. Cooperative strategies for networks with a single source-destination pair are considered in Chapter 3. In Chapter 4, we analyze *the interference channel with limited cooperation* - a network with two cooperating sources and two destinations. In the reminder of this chapter, we present a thesis overview.

1.1 Energy-Efficient Multicast and Broadcast

The capacity result in [24] revealed the limiting performance and the behavior of wireless networks. Result showed that the maximum throughput per user decreases to zero as the number of users grows. In their model, the authors assumed a restricted point-to-point coding scheme in which, at any given time, a receiver decodes a signal from one sender while considering the other signals purely as noise. On the other hand, the problem of network capacity assumes arbitrarily complex coding and decoding in the network. The result in [16] demonstrates that the more general network coding can indeed change the scaling behavior of the network capacity. The crucial thing is then for the destination to keep all the received *overheard* signals instead of treating them as interference. Even in the case of a one-relay physically degraded channel, utilizing the unreliable overheard information was essential for achieving capacity. Yet, the previous work on energy-efficient broadcast in wireless networks does not incorporate the overheard signals at the receivers. The wireless formulation of the conventional broadcast problem assumes that a node can benefit from a certain transmission only if the received power is above a threshold required for reliable communication. The solution to the problem is then specified by a broadcast tree. The arcs in the broadcast tree uniquely determine the power levels for each transmission; a relay that is the parent of a group of siblings in the broadcast tree transmits with the power needed to reliably reach the most disadvantaged sibling in the group. In Chapter 2, we propose the cooperative broadcast strategy that allows nodes to accumulate the energy of unreliable receptions. We refer to this cooperative strategy as *accumulative broadcast.*

In Section 2.4, we address the minimum-energy accumulative broadcast problem with the objective to broadcast data reliably to all network nodes at a fixed rate with minimum transmit power. The minimum-energy broadcast problem was formulated as a minimum-energy broadcast tree problem in [25, 26]. In a wired network, the minimum-cost broadcast tree can be found in $O(n^2)$ operations [27]. However, in the wireless network, this problem was shown in [28] to be NP-hard and later on, in [29–31] to be NP-complete. The greater difficulty of the wireless broadcast tree problem stems from the *wireless multicast advantage* [26], the fact that a wireless transmission can be received by all nodes in the transmission range. Authors in [26] proposed the Broadcast Incremental Power (BIP) algorithm, a greedy heuristic that uses the principle of Prim's algorithm [32] while assigning costs to the nodes in a way that exploits the wireless multicast advantage. The analytical results for performance of BIP are given in [33]. Several other heuristics for constructing energy-efficient broadcast trees have been proposed in the literature and evaluated through simulations (see [28–31, 34, 35] and references therein).

While the minimum-energy broadcast problem results in energy-efficient solutions in terms of the total network power, the different transmitter power levels generally assigned to the relays cause higher drain relays to fail first. Distributing the traffic more evenly throughout the network requires a different performance objective. The problem of maximizing the network lifetime where the *network lifetime* is the time duration until the first node battery is fully drained [36] addresses this issue. In Section 2.5, we address the network lifetime problem under the accumulative broadcast cooperative strategy. Finding a broadcast tree that maximizes network lifetime was considered in [37–39]. The similar problem of maximizing the network lifetime during a multicast was addressed in [40, 41]. Because the energies of the nodes in a tree are drained unevenly, the optimal tree changes in time and therefore the authors [37, 39, 41] distinguished between the *static* and *dynamic* maximum lifetime problem. In a static problem, a single tree is used throughout the broadcast session whereas the dynamic problem allows a sequence of trees to be used. Since the latter approach balances the traffic more evenly among the nodes, it generally performs better. For the static problem, an algorithm was proposed that finds the optimum tree [37]. For the special case of identical initial battery energy at the nodes, the optimum tree was shown to be the minimum spanning tree. In a dynamic problem, a series of trees were used that were periodically updated [37] or used with assigned duty cycles [39].

In Chapter 2, we re-examine these two broadcast problems under the assumption that nodes exploit the energy of unreliable receptions and demonstrate the benefit of accumulative broadcast compared to the conventional broadcast. In Section 2.4, we present the minimum-energy accumulative broadcast problem and the approach to its solution [42–44]. We show that finding the best solution to the minimum-energy accumulative broadcast problem is NP-complete. On the other hand, we show in Section 2.5 that the maximum lifetime multicast problem has a simple optimal solution and propose the *Maximum Lifetime Accumulative Broadcast (MLAB)* algorithm that finds it [45, 46]. The simplicity of the solution is due to the accumulative broadcast that facilitates load balancing by relaxing the constraint that a relay has to transmit with power sufficient to reach its most disadvantaged child. The power levels given by the solution ensure that the lifetimes of relay nodes are the same and thus, their batteries die simultaneously.

1.2 Two-Hop Relay Networks

The proposed accumulative broadcast scheme employs a decode-and-forward relay strategy. As pointed out earlier however, not only is the optimal relay strategy unknown, but also certain network scenarios have been identified for which unreliable forwarding is optimal [19–21]. This motivates the performance evaluation of amplify-and-forward strategies for energy-constrained networks. In Chapter 3, we investigate the impact of the choice of the relay strategy on the network operation. For a single source-destination pair, we characterize the optimum AF bandwidth and show that transmitting in the optimum bandwidth allows the network to operate in the linear regime where the achieved rate increases linearly with the available network power [47]. We then present the optimum power allocation among the AF relays [48]. The solution, which can be viewed as a form of maximum ratio combining, indicates the favorable relay positions in the network. Motivated by the large bandwidth resources we further consider a network that uses orthogonal transmissions at the nodes. While the above result for the optimum bandwidth still holds, a different set of relays should optimally be employed. In this case, the relay power solution can be viewed as a form of water-filling. In contrast, in a network with unconstrained bandwidth, the DF strategy will operate in the wideband regime and will require a different choice of relays. The two-hop DF strategies are analyzed in Section 3.5.

1.3 The Interference Channel with Limited Transmitter Cooperation

The work presented in Chapters 2-3 focuses on single-source networks. However, the wireless ad hoc and sensor networks generally consist of multiple source-destination pairs. Multiple data streams cause the interference at the receivers if they are transmitted at the same time and the transmitter of one data stream is sufficiently close to the intended receiver of another data stream. Furthermore, if sources produce correlated data, as often is the case in sensor networks, exploiting the correlation leads to more efficient data transmission through the network [49]. Hence, the multiple source problem adds a new dimension to the cooperation problem.

The channel model that incorporates the communication situation in which two separate transmitters wish to communicate their independent messages to two corresponding receivers is the interference channel [50, 51]. The capacity region of the interference channel is an open problem. The best known achievable region consisting of seven MAC-like bounds per receiver was proposed by Han and Kobayashi [52]. Recently, a new achievable region was presented in [53] based on the observation that one of the H-K bounds can be omitted. For the Gaussian interference channel, the new outer bounds were recently presented by Kramer in [54]. The bounds improve the best known outer bounds derived by Sato [51] and Carleial [55]. The full capacity region of the interference channel is known in the case of the *strong interference* channel [56] satisfying

$$I(X_1;Y_1|X_2) \leq I(X_1;Y_2|X_2) \tag{1.1}$$

$$I(X_2;Y_2|X_1) \leq I(X_2;Y_1|X_1) \tag{1.2}$$

for all product distributions on the inputs X_1 and X_2. The capacity region in this case coincides with the capacity region of the two-sender, two-receiver channel in which both messages are decoded at both receivers, as determined by Ahlswede [57].

In Chapter 4, we introduce limited transmitter cooperation into the interference channel model. The broadcast nature of the wireless channel allows a signal transmitted from one node to be overheard at the other and could be used as a base for transmitter cooperation. Under full transmitter and receiver cooperation with no cost for cooperation in terms of power and bandwidth, this channel becomes a 2×2 MIMO channel. Thus MIMO channel capacity [58] is the upper bound to the sum-rate of the interference channel with cooperation. Under full transmitter cooperation which again incurs no cost in terms of power and bandwidth, the channel reduces to the MIMO broadcast channnel. In the Gaussian case, the capacity region is known and is achieved by *dirty-paper coding* [4, 5, 7, 59]. For the interference channel with cooperation, it has been demonstrated that dirty paper coding brings significant gains as long as the sources are clustered together [60, 61]; for the destination cluster, gains are obtained employing receiver cooperation through amplify-and-forward [60]. As in a single source-destination Gaussian network, the performance and the suitability of a particular cooperative scheme depends on the network topology, that is, relative positions of sources and receivers [61–64]. This motivates the analysis of discrete cooperative channel models that incorporate different geometries.

A problem in which encoders partially cooperate in a discrete memoryless channel was proposed by Willems for a multiaccess channel (MAC) [65]. To model transmitter cooperation, two side-channel communication links with finite capacities are introduced between the two encoders. The amount of information exchanged between the two transmitters is bounded by the capacities of the communication links. The proposed discrete channel model enables investigation of transmitter cooperation gains. In the discrete memoryless MAC with partially cooperating encoders [65], the outcome of the cooperation is referred to as a *conference*. Willems determined the capacity region of this channel and thus specified the optimum conference. In this work, we are interested in the channel model with two receivers. In Section 4.4, we extend Willems' result to a compound MAC in which two decoders wish to decode messages sent from the encoders [66]. We show that the same form of conference as in [65] remains optimal and determine the capacity region [66].

When cooperating over links with finite capacities, encoders obtain partial information about each other's messages. This partial information may consist of a fraction of bits sent by the encoder, where the fraction is determined by the capacity of the cooperating link. This information becomes a *common* message as it is known to both encoders after the conference. When the capacity of a cooperating link is large enough to carry the code rate, full information about the source message is obtained by the other encoder. In general, however, each encoder will still have a *private* message, independent information that remains unknown to the other encoder. Both common and private messages are decoded at a single receiver in the case of the MAC [65], or at both receivers in the case of a compound MAC [66].

In Section 4.5, we consider the interference channel with both private and common messages at the encoders [67, 68]. We relax the requirement that both private messages need to be decoded at both decoders. Instead, a private message at an encoder is intended for a corresponding decoder whereas the common message is to be received at both decoders. We derive conditions under which the capacity region of this channel coincides with the capacity region of the channel in which both private messages are required at both receivers. We show that the obtained conditions are equivalent to the strong interference conditions determined by Costa and El Gamal for the interference channel with independent messages.

Finally, in Section 4.6, we analyze the interference channel with a different form of transmitter cooperation, that we refer to as *unidirectional cooperation*. We assume that messages sent at one encoder are known to the other encoder, but not vice versa. The encoder that knows both messages can exploit that information to improve the achievable rates. We derive conditions under which the capacity region of this channel coincides with the capacity region of the channel in which both messages are decoded at both receivers. We compare the obtained conditions with the strong interference conditions in the interference channel with independent messages as well as in the interference channel with both private and common messages [69, 70].

Chapter 2

Accumulative Broadcast

2.1 Introduction

In this chapter, we consider the problem of energy-efficient multicast and broadcast in a wireless network. In the multicast problem, a message from a source node is to be delivered efficently to a set of destination nodes. When the set of destination nodes includes all the network nodes (except the source), the multicast problem reduces to the broadcast problem. We present a cooperative broadcast and multicast strategy that allows nodes to exploit the unreliably received signals. We consider the multicast/broadcast problems with two different objectives. In Section 2.4, we present the *minimum-energy accumulative broadcast problem* and the approach to its solution [42–44]. Because it allows for more radiated broadcast energy to be captured, it is straightforward to show [43] that accumulative broadcast increases the energy-efficiency of conventional broadcast. We show that finding the best solution to the minimum-energy accumulative broadcast problem is NP-complete. This motivates a heuristic algorithm that finds energy-efficient solutions. In Section 2.4.1, we propose a centralized algorithm that requires global knowledge of the channel gains. In Section 2.4.2, we then present a distributed version of the accumulative broadcast heuristic algorithm that uses only local information at the nodes and is thus better suited for networks with a large number of nodes.

We then consider the *maximum network lifetime multicast problem.* We show in Section 2.5 that this problem has a simple optimal solution and we propose a *Maximum Lifetime Accumulative Broadcast (MLAB)* algorithm that finds it [45, 46]. The simplicity of the solution is due to the accumulative broadcast that facilitates load balancing by relaxing the constraint that a relay has to transmit with power sufficient to reach its most disadvantaged child. The power levels given by the solution ensure that the lifetimes of relay nodes are the same and thus, their batteries die simultaneously. In Section 2.5.4 we present a distributed MLAB algorithm that determines the transmit power levels locally at the nodes.

2.2 System Model

In this work, we seek to employ overheard broadcast information in a large scale network. We focus on techniques that can be implemented as distributed network layer algorithms in which nodes use local information and coarse timing and synchronization. In particular, we make the following assumptions:

- *Loose Synchronization:* Nodes cannot synchronize transmissions for coherent signal combining at a receiver.
- *Reliable Forwarding:* A node can forward a message only after reliably decoding that message.

The advantages of coherent signaling and unreliable forwarding have been recognized for networks in which one or more relay nodes forward to a destination node, [10, 18–20]. However, coherent signal combining is challenging to implement as it requires the precise sychronization of transmitting nodes and exact knowledge of radio path delays needed for coherent combining at a single receiver. By contrast, unreliable forwarding is practically implementable and has been shown to be superior to reliable forwarding in certain scenarios [18]. Nevertheless, we will see that reliable forwarding can simplify the system architecture and the optimization of retransmission strategies while still allowing us to benefit from unreliable overheard information.

As stated earlier, motivated by applications for wireless sensor networks, the focus of our work will be on power-constrained networks. Such networks operate in the wideband regime [71], as the data rate is very small compared to the bandwidth, resulting in a low spectral efficiency. In the sensor networks where the energy-efficiency is the primary goal [72], operating in the wideband regime seems like the right choice: at the expense of using the large number of degrees of freedom per transmitted bit, the transmit energy per bit can be minimized [73]. However, finding the minimum energy per bit in networks with relays is still an open problem.

We will show that for a network operating in the wideband regime, a single codebook can be used by all forwarding nodes. While there is a benefit of using more general codes with incremental redundancy [74] in a general wireless network, this benefit diminishes when broadcasting in a network operating in the wideband regime [75].

We consider a stationary wireless network of N nodes such that from each transmitting node k to each receiving node m, there exists an AWGN channel of bandwidth W characterized by a frequency non-selective link gain h_{mk}. Each channel is assumed to be time-invariant with a constant link gain representing the signal path loss. We further assume large enough bandwidth resources to enable each transmission to occur in an orthogonal channel, thus causing no interference to other transmissions. Each node has both a transmitter and a receiver capable of operating over all channels.

A receiver node j is said to be in the transmission range of transmitter i if the received power at j is above a threshold that ensures the capacity of the channel from i to j is above the code rate of node i. We assume that each node can use different power levels, which will determine its transmission range. The nodes beyond the transmission range will receive an unreliable copy of a transmitted signal. Those nodes can exploit the fact that a message is sent through multiple hops on its way to other nodes. Repeated transmissions act as a repetition code for all nodes beyond the transmission range.

The assumption of an AWGN channel oversimpifies the reality, since fading is one of the salient characteristics of wireless channels. There are a few scenarios in which this assumption actually holds. For example, when terminals are in the proximity of each other, multipath may not be significant and fading can be neglected. In wideband CDMA with a lot of resolvable multipath, the signal at the output of the RAKE receiver can be viewed as the signal at the output of the AWGN channel.

In general however, the power response of the channel will depend both on frequency and time due to multipath. Consider a single-user flat fading channel

$$y_i = h_i x_i + z_i \tag{2.1}$$

where h_i are the samples of the fading process with power $v_i = |h_i|^2$, $E[v] = 1$, $E[|x_i|^2] \leq P$ and $E[|z_i|^2] = \sigma^2$. As pointed out in [76], it follows from the Jensen's inequality that the capacity of this channel

$$C = E_v \left[\frac{1}{2} \log \left(1 + \frac{Pv}{\sigma^2} \right) \right] \tag{2.2}$$

is upper bounded by the capacity of an AWGN channel with the same average power.

The same conclusion holds for the more general case of the frequency-selective channel with the frequency response at time t denoted $C(t, f)$. The capacity of this channel is [77]

$$C = E_v \left[\int_{-\infty}^{\infty} \frac{1}{2} \log \left(1 + \frac{P|C(t,f)|^2}{\sigma^2} \right) df \right], \tag{2.3}$$

where the expectation is taken with respect to the statistics of the random process $C(t, f)$. Under the ergodic assumption, the statistics are independent of both t and f and the capacity (2.3) is thus again smaller than the AWGN channel capacity. Hence, assuming no channel state information at the transmitter, the rate achievable in the AWGN channel gives an upper bound on the performance achievable in the fading channel.

Gains from cooperation in fading channels can also be examined from the perspective of *cooperative diversity* [18, 78]. The multicast problems presented in Chapter 2 and the two-hop relay problem presented in Chapter 3 could be re-considered from this perspective as well. Since we started this work, there has been a lot of progress demonstrating diversity gains due to cooperation, initiated with the work by Laneman et. al. [18]. More recently, various cooperation schemes were investigated from the multiplexing-diversity tradeoff perspective [79].

We note that algorithms proposed in Chapters 2 and 3 can readily be applied in the fading environment. It is the analysis that would have to be re-done to evaluate the performance gains. In fact, the cooperative strategy shown to maximize the network lifetime in Chapter 2.5 was independently proposed to improve the performance of collocated users in a fading channel [23]. More recently, the same strategy, refered to as *Dynamic Decode-and-Forward*, was shown to achieve the optimum diversity-multiplexing tradeoff [80].

A word on notation. In this thesis, power labeled in capital letters represents power per dimension [Watts/(sHz)]; power denoted in small letters is used to mean the total power in units of [Watts]; N_0 denotes single-sided noise power spectral density in [Watts/Hz]. For convenience only and with no loss of generality, for numerical results we assume unit double-side noise power spectral density. Equivalently, the power per dimension when multiplied by a path loss G between the transmitter and the receiver will yield the received SNR. The path loss is calculated as $G = r^{-n}$ where r denotes a transmitter-receiver distance and n is the propagation exponent again, for convenience only. Consequently, G is of the order $1 - 10^{-4}$, rather than $10^{-8} - 10^{-12}$, as on an actual radio path.

We view each orthogonal channel as a discrete-time Gaussian channel by representing a waveform of duration T as a vector in the $n = 2WT$ dimensional space [81]. Then, during the ith slot, a source node, labeled node 1, transmits a codeword (vector of length n) $\mathbf{x}(i)$ from a $(2^{nR}, n)$ Gaussian code that is generated according to the distribution $p(\mathbf{x}) = \prod_{l=1}^{n} p(x_l)$ where $p(x) \sim N(0,1)$. Under the reliable forwarding constraint, a node j is permitted to retransmit (forward) codeword $\mathbf{x}(i)$ only after reliably decoding $\mathbf{x}(i)$. With an appropriate set of retransmissions at appropriate power levels, eventually every node will have reliably decoded $\mathbf{x}(i)$. Henceforth, we drop the index i and focus on the broadcast of a single codeword $\mathbf{x}$. We will say a node is *reliable* once it has reliably decoded $\mathbf{x}$. During the multicast, the message is repeatedly transmitted until the set of destination nodes $\mathcal{D}$ becomes reliable.

The constraint of reliable forwarding imposes an ordering on the network nodes. In particular, a node m will decode $\mathbf{x}$ from the transmissions of a specific set of transmitting nodes that became reliable prior to node m. Starting with node 1, the source, as the first reliable node, a solution to the accumulative multicast/broadcast problem will be characterized by a *reliability schedule*, which specifies the order in which the nodes become reliable. Since during the multicast, the broadcast stops after the message has been delivered to D destination nodes, a reliability schedule will not necessarily contain all the network nodes. In general, a multicast reliability schedule is an ordered subsequence of the list of nodes of length M, $D < M \leq N$, that starts with node 1, and contains all destination nodes and a subset of network nodes that relay the message. In the broadcast case, a reliability schedule $[n_1, n_2, n_2, \ldots, n_N]$ is simply a permutation of $[1, 2, \ldots, N]$ that always starts with the source node $n_1 = 1$.

For a given reliability schedule, we refer to the ith node in the schedule as simply node i. After each node $k \in \{1, \ldots, m-1\}$ transmits codeword $\mathbf{x}$ with average energy per symbol P_k, the received signal at node m for each symbol x in the codeword is

$$\mathbf{y}_m = \mathbf{h}_m x + \mathbf{z}, \tag{2.4}$$

where $\mathbf{h}_m = [\sqrt{h_{m1}P_1}, \ldots, \sqrt{h_{mm-1}P_{m-1}}]^T$ has kth element $\sqrt{h_{mk}P_k}$ equal to the received energy corresponding to the transmission of node k and $\mathbf{z}$ is is a random noise vector with covariance matrix $\mathbf{K_z} = \sigma^2 \mathbf{I}_K$. The mutual information is given by

$$I(x; \mathbf{y}_m) = \frac{1}{2} \log_2 \left(1 + \frac{\sum_{k=1}^{m-1} h_{mk} P_k}{\sigma^2} \right) \tag{2.5}$$

as in a multi-antenna system with $m-1$ transmitting antennas and one receiving antenna [58]. It follows from (2.5) that the maximal number of bits per second that can be transmitted in the system given by (2.4) is

$$r_m = W \log_2 \left(1 + \frac{\sum_{k=1}^{m-1} h_{mk} p_k}{N_0 W} \right) \qquad \text{bits/s}, \tag{2.6}$$

where p_k is the transmit power at node k and N_0 is the one-sided power spectral density of the noise.

Let the required data rate for broadcasting $\overline{r}$ be given by

$$\overline{r} = W \log_2 \left(1 + \frac{\overline{P}}{N_0 W} \right) \qquad \text{bits/s}. \tag{2.7}$$

Rate $\overline{r}$ has to be achieved at every reliable node m. From (2.6) and (2.7), achieving $r_m = \overline{r}$ implies that the total received power at node m is above the threshold $\overline{P}$, that is,

$$\sum_{k=1}^{m-1} h_{mk} p_k \geq \overline{P}. \tag{2.8}$$

After the data has been successfully delivered to the destination nodes, all those nodes are reliable and the feasibility constraint (2.8) is satisfied at every destination node m.

In the system where the nodes are power limited and the data rate $\overline{r}$ is small relative to the channel bandwidth W, the system operates in the wideband regime [71]. The increase in rate with power is linear

$$\overline{r}_\infty = \lim_{W\to\infty} \overline{r} = \lim_{W\to\infty} W \log_2\left(1 + \frac{\overline{P}}{WN_0}\right) = \frac{\overline{P}}{N_0 \log 2} \qquad \text{bits/s.} \tag{2.9}$$

We emphasize that the system operates at a low spectral efficiency due to the low transmit powers and does not imply the large operating bandwidth W. From (2.9), when communicating at rate $\overline{r}_\infty$, the required signal energy per bit has the minimum value $E_b = \overline{P}/\overline{r}_\infty = N_0 \log 2$ Joules/bit. Thus, the system uses the energy in the most economical way possible to communicate reliably [73] because the system uses a large number of degrees of freedom per information bit. This energy can be collected at a node m during one transmission interval $[0, T]$ when a transmitter j is signaling with power $p_j = \overline{P}/h_{mj}$, as commonly assumed in wireless broadcast problems [26,29,30,37–39,41]. However, during the accumulative broadcast in the system (2.4), the required energy E_b is collected in $m-1$ repeated transmissions. In the wideband regime, the maximum achievable rate at node m given by (2.6) becomes

$$\lim_{W\to\infty} r_m = \frac{1}{N_0 \log 2} \sum_{k=1}^{m-1} h_{mk} p_k. \tag{2.10}$$

In [82], it was shown that TDMA is first-order optimal in the wideband regime as it achieves the minimum energy per information bit of a multiaccess channel. Using (2.10), it is straightforward to conclude that the first-order optimality is preserved even if the repetition code described above is employed. We formally state this conclusion in the next theorem.

Theorem 1 *For the wideband regime, with fixed transmitted powers* $\{p_1, \ldots, p_N\}$ *and a reliability schedule* $[1, 2, \ldots, N]$*, the maximum rate achievable from the source to node* m *is given by Equation (2.10) and is achieved by repetition coding.*

The proof is given in Appendix A.1.

2.3 Approach

As pointed out in Chapter 1, in the conventional broadcast problem, the broadcast tree uniquely determines the transmission levels; a relay that is the parent of a group of siblings in the broadcast tree transmits with the power needed to reliably reach the most disadvantaged sibling in the group.

In the accumulative broadcast, however, there is no a clear parent-child relationship between nodes because nodes collect energy from the transmissions of many nodes. Furthermore, the optimum solution may require that a relay transmits with a power level different from the level precisely needed to reach a group of nodes reliably; the nodes may collect the rest of the needed energy from the future transmissions of other nodes. In fact, the optimum solution often favors such situations because all nodes beyond the range of a certain transmission are collecting energy while they are unreliable; the more such nodes, the more efficiently the transmitted energy is being used.

The differences from the conventional broadcast problem dictate a new approach. The optimum solution must specify the reliability schedule as well as the transmit power levels used at each node. A schedule is an ordered subsequence of M nodes from a network of N nodes,

$$\mathbf{x} = [x_1, \ldots, x_M], \tag{2.11}$$

with $x_1 = 1$. We say that the *length* of the subsequence $\mathbf{x}$ in (2.11) is $\|\mathbf{x}\| = M$. In the broadcast case, $M = N$. Let

$$\{\mathbf{x}\} = \{x_1, \ldots, x_{\|\mathbf{x}\|}\} \tag{2.12}$$

denote the set of nodes in a schedule $\mathbf{x}$ and let Π_N denote the set of all variable-length ordered subsequences of $\{1, \ldots, N\}$. It follows that the family of all possible schedules is

$$\mathcal{X}_N(\mathcal{D}) = \{\mathbf{x} \in \Pi_N | \mathcal{D} \in \{\mathbf{x}\}, x_1 = 1\} \tag{2.13}$$

Given a schedule $\mathbf{x}$, we define a gain matrix $\mathbf{G}(\mathbf{x})$

$$[\mathbf{G}(\mathbf{x})]_{ij} = \begin{cases} h_{x_i x_j} & i > j \\ 0 & \text{otherwise} \end{cases}$$

for $1 \leq i, j \leq M$. Thus, channel gains corresponding to any node j that is not in schedule $\mathbf{x}$ are not included in $\mathbf{G}(\mathbf{x})$; since a node j does not participate in the retransmission of the message it can be omitted from the problem formulation.

Given a schedule $\mathbf{x}$, we can define the accumulative multicast and broadcast problem with specific performance objective as a linear program (LP) that will find the optimum solution for that schedule. Such a solution will identify the nodes that should transmit and their transmission power levels. To define the LP for a certain schedule $\mathbf{x}$, we use the observation that every node selected to transmit by the optimum solution, needs to transmit only once. This fact is given by the next theorem.

Theorem 2 *For the wideband regime, given a solution to the accumulative broadcast problem consisting of a sequence of transmissions where a node j is assigned to transmit K times with power levels $P_j^1, \ldots P_j^K$ then there is a feasible optimum solution in which node j transmits once with power level $\sum_{k=1}^{K} P_j^k$.*

The proof is given in Appendix A.2.

2.4 Minimum-Energy Cooperative Broadcast

We next consider the problem of minimum-energy broadcast, with the objective being to broadcast data reliably to all network nodes at a fixed rate with minimum transmitted power.

As explained in the previous section, this problem can be divided into two subproblems. The crucial step is finding the best schedule. We can define the minimum-energy cooperative broadcast problem for schedule $\mathbf{x}$ in terms of the vector $\mathbf{p}$ of transmitted powers as the LP for schedule $\mathbf{x}$ is

$$\rho(\mathbf{x}) = \min \ \mathbf{1}^T \mathbf{p} \tag{2.14}$$

$$\text{subject to } \mathbf{G}(\mathbf{x})\mathbf{p} \geq \mathbf{1}\overline{P}, \tag{2.14a}$$

$$\mathbf{p} \geq \mathbf{0}. \tag{2.14b}$$

The inequality (2.14a) contains $N-1$ constraints as in (2.8) requiring that the received power at all the nodes but the source is above the required threshold $\overline{P}$. Given a schedule

$\mathbf{x}$, we will use $\mathbf{p}^*(\mathbf{x})$ to denote a power vector $\mathbf{p}$ that achieves total transmitted power $\rho(\mathbf{x})$.

In a schedule, all N nodes are given a chance to transmit (since p_j can be greater than 0 for every node) and only the order of transmission is different for different schedules. Since the source always transmits first, there are $(N-1)!$ schedules corresponding to the number of permutations of $N-1$ elements. Thus, out of $N^{(N-2)}$ broadcast trees, we consider a subset of $(N-1)!$ schedules. If the best solution is that only a subset of nodes should be transmitting, the LP for the best schedule will find that solution by setting appropriate powers to zero. In general, however, the problem of finding a best schedule for minimum-energy cooperative broadcast is intractable.

Theorem 3 *The existence of a schedule* $\mathbf{x}$ *such that* $\rho(\mathbf{x}) \leq B$ *is an NP-complete problem.*

Proof:

Let Π_i denote the set of all vectors $\pi = [\pi_0, \ldots, \pi_i]$ that are permutations of $[0, 1, \ldots, i]$. A formal statement of the ACCUMULATIVE BROADCAST (AB) problem is

AB Given a nonnegative matrix specified by $\{h_{j,k} | 1 \leq j \leq m, 0 \leq k \leq m\}$, and a constant c, does there exist a permutation $\pi \in \Pi_k$ with $\pi_0 = 0$ and a non-negative vector $\mathbf{p} = [p_0, p_1, \ldots, p_m]$ such that $\sum_{k=0}^{m} p_k \leq c$ and $\sum_{k=0}^{j-1} h_{\pi_j, \pi_k} p_{\pi_k} \geq 1, \quad j = 1, \ldots, m$.

Thus an instance of AB is specified by the pair $(\{h_{j,k}\}, c)$. Note that we set the reliability threshold to unit power since any scaling can be specified by the constant c. We observe that AB is in NP since given a permutation π and vector $\mathbf{p}$, it is easy to check whether the AB constraints are met.

We will show that the ACCUMULATIVE BROADCAST problem is NP complete by a polynomial time reduction of the DIRECTED HAMILTON PATH (DHP) [27] problem. Formally the DHP problem is

DHP Given a directed graph $G = (V, A)$ with nodes $V = \{0, \ldots, n\}$, does there exist a permutation $\pi \in \Pi_n$ such that $\pi_0 = 0$ and $(\pi_i, \pi_{i+1}) \in A$ for $i = 0, \ldots, n-1$.

We now describe the transformation of DHP into an instance of AB. Without loss of generality, we assume that the instance of DHP is such that node 0 has a single outgoing arc $(0, 1)$ and that node n is a sink node reachable by an arc (i, n) from each node $i \in \{1, \ldots, n-1\}$. Note that if this condition does not hold, we can add such source and sink nodes and solve an equivalent DHP. Thus, for each such graph, the Hamilton path, if it exists, will start at node 0 and terminate at node n.

Given $G = (V, A)$ for DHP, we construct a set of nodes G' and matrix $\{h_{j,k}\}$ for an instance of AB. In particular, for each node $k \in G$, we construct a cluster of nodes $C_k \subset G'$. In particular, the cluster C_k includes a node $i_{j,k}$ for each incident arc $(j, k) \in A$ and a node $o_{k,l}$ for each outgoing arc $(k, l) \in A$. That is, in terms of each arc $(j, k) \in A$, we have created an *incident node* $i_{j,k} \in C_k$ and an *outgoing node* $o_{j,k} \in C_j$. Note that cluster C_0 contains only the single node $o_{0,1}$ and that the sink node n has the cluster $C_n = \{i_{j,n} | 1 \le j < n\}$ of only incoming nodes.

To avoid an explicit enumeration of the nodes in G', we describe the matrix $\{h_{j,k}\}$ in terms of a function $h(a, b)$ that gives the channel gain from node b to node a. Similarly, we will use the notation $p(a)$ to denote the transmitted power of the node a. Corresponding to each arc $(j, k) \in A$, we have $h(i_{j,k}, o_{j,k}) = 1$. Within each cluster C_k, we have that for any pair of incident nodes $i_{j,k}$ and $i_{j',k}$, $h(i_{j,k}, i_{j',k}) = 1$. In addition, for each outgoing node $o_{k,l} \in C_k$, and each incoming node $i_{j,k} \in C_k$, $h(o_{k,l}, i_{j,k}) = 1$. For all other pairs of nodes $a, b \in G'$, we set $h(a, b) = 0$. Keep in mind that if $h(a, b) = 1$, then $p(b) = 1$ yields received power $h(a, b)p(b) = 1$ at node a. We will see in our AB construction, each node a will use power $p(a) \in \{0, 1\}$.

To prove that AB is NP-complete, we show that the graph G has a Hamilton path if and only if the resulting instance $(h(\cdot, \cdot), c = 2n)$ of AB is feasible. Consider a Hamilton path that starts at node zero and proceeds through all nodes to node n. Suppose the Hamilton path uses arc (j, k), then for the AB problem, we set $p(o_{j,k}) = 1$, $p(i_{j,k}) = 1$, $p(o_{j,k'}) = 0$ for all $k' \neq k$, and $p(i_{j',k}) = 0$ for all $j' \neq j$. In the context of AB, node $o_{j,k}$ transmits to make node $i_{j,k}$ reliable and then node $i_{j,k}$ transmits to make all nodes in cluster C_k reliable. If the next arc in the Hamilton path is (k, l), then in the AB, $o_{k,l}$, which has already been made reliable by the transmission of $i_{j,k}$, will transmit to make $i_{k,l}$ reliable. We call the event that an incoming node $i_{j,k}$ is made reliable a *visit* to cluster C_k. The sequence of nodes in the Hamilton path corresponds exactly to the sequence of cluster visits. To calculate the total transmitted power, note that in cluster C_0, node $o_{0,1}$ will transmit. In clusters $1, \ldots, n-1$, one incoming node and one outgoing node will transmit. Lastly, in cluster C_n, one incoming node will transmit to make the other incoming nodes in C_n reliable. The total transmitted power will be exactly $2n$. We note that the node ordering required by the formal statement of AB will not be uniquely specified. If cluster C_k is visited before cluster C_l, then all nodes in C_k must be ordered ahead of nodes in C_l. In a cluster C_k, if incoming node $i_{j,k}$ is made reliable then $i_{j,k}$ must be first in the cluster but other nodes in the cluster can be ordered arbitrarily.

To complete the proof, suppose we have a solution to the AB problem. This AB solution must make every node in the graph G' reliable. For each cluster C_k, $1 \leq k \leq n$, at least one incoming node $i_{j,k}$ must be made reliable by the transmission of the corresponding outgoing node $o_{j,k}$. However, since this transmission of $o_{j,k}$ makes only $i_{j,k}$ reliable, one such transmission is needed for each cluster C_k. Over all clusters C_k, $1 \leq k \leq n$, we require n such transmissions. Further, within each cluster, the outgoing nodes can be made reliable only by the transmission of an incoming node in the cluster. Thus for each cluster C_k, $1 \leq k \leq n$, at least one incoming node $i_{j,k}$ must transmit to make all other nodes in the cluster reliable; this requires n additional transmissions. Thus $2n$ is a lower bound to the number of transmissions for the AB problem. Moreover, if the solution to AB achieves the minimum $2n$, then each outgoing node transmission must be to an incoming node in a cluster that has had no other incoming nodes receive a transmission from its corresponding outgoing node. That is, each cluster can be visited only once for the $2n$ lower bound to be met. Starting with node 0 and cluster C_0, node $o_{0,1}$ will transmit to make node $i_{0,1}$ reliable. Node $i_{0,1}$ must then transmit to make all other nodes in cluster C_1 reliable. An outgoing node $o_{1,k}$ will then transmit to make a node $i_{1,k}$ reliable, constituting a visit to cluster k. To achieve the $2n$ lower bound, each cluster will be visited precisely once, with termination at cluster C_n. Since moving from cluster C_j to visit C_k can occur only if (j, k) is an arc in G, the AB solution corresponds to a Hamilton path in the graph G. □

Because of the complexity of the problem of finding the best schedule, we now propose a heuristic algorithm that finds a good schedule.

2.4.1 Scheduling Heuristic

Once the schedule is determined using the algorithm, the LP for that schedule is solved to find the optimum power levels. We evaluate the performance of the algorithm through simulation and compare its power efficiency to the optimum solution as well as to the performance of BIP.

We observe that we can restrict ourselves to scheduling nodes in an order in which they can become reliable one at a time. When a node j is scheduled to be the next node in a schedule after a set of nodes S, then a transmission from that set has to make node j reliable. If the power that is needed to reach node j is enough to reach another unreliable node i as well, then we could have done better by assigning node i for transmission before node j. This is because i cannot benefit from a transmission from node j (since it is made reliable before j) but j might benefit from a transmission from i. If, in fact the optimal solution is to simultaneously make the two nodes i and j reliable by a transmission from the same set S, then those two nodes do not need to overhear each other's transmission. Thus, all the schedules in which nodes i and j are scheduled one right after the other, in any order, will have the same performance. This reasoning will be used in the proposed heuristic algorithm.

The algorithm pseudocode is given in Figure 2.1.

The algorithm starts with a partial reliability schedule $\mathbf{s} = [1]$ that contains only the source. Given a partial schedule $\mathbf{s}$, a step of the *Greedy filling algorithm* does the following:

1. We find the reliable node k that maximizes the *fill rate* of the unreliable set, where

$\mathbf{s} = [1]$; $\mathbf{p} = 0$
`while` $(|S| < N)$ `do`
$k = \arg\max_{i \in S} \sum_{j \in U} h_{ji}$;
$j = \arg\min_{m \in U} (\overline{P} - \sum_{i \in S} h_{mi} p_i)/h_{mk}$;
$p_k \leftarrow p_k + (\overline{P} - \sum_{i \in S} h_{ji} p_i)/h_{jk}$;
$\mathbf{s} \leftarrow [\mathbf{s}, j]$
`end`
Given a partial schedule $\mathbf{s}$, S *is the unordered set of nodes in* $\mathbf{s}$.
Its complement U *is the set of unreliable nodes. The cardinality of* S *is given by* $|S|$.

Figure 2.1: Greedy Filling Algorithm for minimum-energy cooperative broadcast

the *fill rate*

$$R_k = \sum_{j \in U} h_{jk} \tag{2.15}$$

is the sum of the link gains from node k to the set U of all unreliable nodes.

2. We increase p_k such that the transmission by k adds one more node, node j, to the reliable set.

3. We append node j to the partial schedule.

Once the schedule is complete, the LP is solved to find the optimum power levels for that schedule.

We evaluated the performance of the algorithm and compared it to the optimal solution as well as to the performance of the Broadcast Incremental Power (BIP) algorithm for networks with a small number $(5 - 10)$ of randomly positioned nodes. The BIP algorithm, proposed in [26], is greedy heuristic that uses the principle of Prim's algorithm [32] while assigning costs to the nodes in a way that exploits the wireless multicast advantage. We also compared the performance of two heuristics for more dense networks with a maximum of 150 nodes. Nodes were uniformly distributed in an area of size 10×10. The transmitted power was attenuated as d_{jk}^{α} for three different values of propagation exponent, $\alpha = 2, 3, 4$. The received power threshold was chosen to be $\overline{P} = 1$. Results were based on the performance of 100 randomly chosen networks. In small networks, the performance metric used was the normalized total transmit power in the network. In each simulation run, the power used when a heuristic algorithm was employed was normalized by the power used in the optimum solution. Results

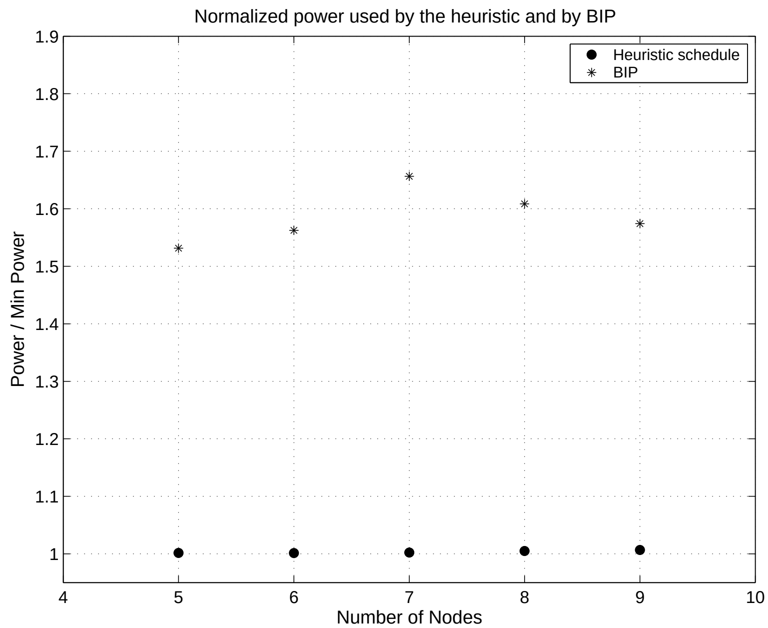

Figure 2.2: Normalized power used for broadcast. Normalization is with respect to the optimum power for minimum-energy cooperative broadcast.

are shown in Figure 2.2 as a function of the number of network nodes for $\alpha = 2$. Results show the heuristic algorithm performance very close to the optimum. This is a desirable and important characteristic, given the complexity of finding the optimum solution. The simulation results also show 2 dB savings in average power of accumulative broadcast over the minimum-energy broadcast tree approach with the broadcast tree found by BIP.

For the networks with a larger number of nodes, performance comparison of the greedy filling algorithm and BIP are shown in Figure 2.3. The metric used was the average total power used for broadcasting. We observe that total power decreases with the number of nodes due to the increased number of shorter hops. The decrease in the case of the accumulative broadcast is steeper since the increased number of transmissions allows for more energy to be collected. Hence the relative improvement over BIP increases with the node density of the network. For smaller values of propagation exponent α, the smaller path loss allows for the higher gains from the accumulative broadcast and we observe up to 5 dB savings per node for $\alpha = 2$. Results also show

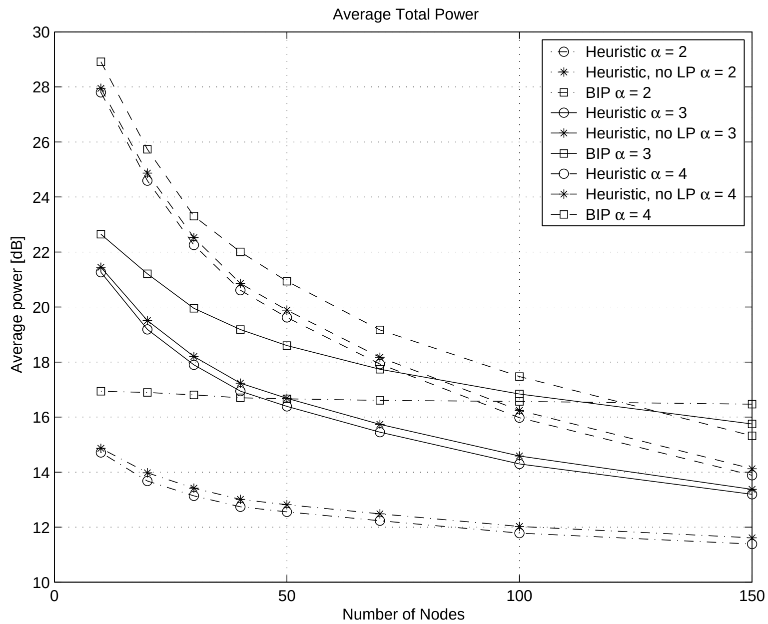

Figure 2.3: Average total power used for broadcast.

that, for a larger number of nodes the total power required is smaller for larger values of α. This counterintuitive result is due to the fact that, as the networks becomes more dense, most of the distances d_{jk} become less than 1 so that $1/d_{jk}^4 > 1/d_{jk}^2$.

Figure 2.3 also shows the loss in the performance when the LP is not employed to determine the optimum power levels for a greedy filling schedule. Instead, the power levels $\{p_k | k = 1, \ldots, N\}$ found by the greedy filling algorithm are used for broadcasting. We observe only a small loss in the performance. Thus, finding the optimum power levels is not as crucial as finding a good schedule. We will use this observation to formulate the distributed version of the greedy filling algorithm next.

2.4.2 Distributed Algorithm

In the greedy filling algorithm, we assumed full knowledge of the link gains when forming a schedule. In particular, we assumed that the fill rates of the reliable nodes are known in every step, so that the transmitting node k could be chosen. Also, we assumed that the transmit power needed to make one more node reliable could be determined. In this section, we propose a distributed version of the greedy filling algorithm that

```
At each node i do:
initialize S_i = ∅;  E_i = 0
while (E_i < P̄T) do
    when  ACK_j is reliably received with power P:
        calculate h_ij = P_ACK/P
        transmit LG_i control packet with power P̄_{N_R}/h_ij
        S_i ← S_i ∪ {j}
    when data received from node k with power h_ik p:
        if k ∈ N_R(i):  LastTransmitNode=k
        E_i ← E_i + h_ik pT
        if ( E_i ≥ P̄T )
            send ACK_i with power P_ACK
            S_i ← S_i ∪ {i}
            wait for LG control packets from j ∈ N_R(i) \ S_i
            initialize U_i,  R_i = Σ_{j∈U_i} h_ji,  P_i = 0
            transmit R_i with power P_ACK
        end %if
end %while
while( |U_i| > 0 )
    when  ACK_j received with power P ≥ P̄_{N_R}:
        if transmitting:  stop
        S_i ← S_i ∪ {j},  U_i ← U_i \ {j},  R_i ← R_i − h_ji
        transmit R_i with power P_ACK
        if ( i = arg max_{j∈S_i}{R_j} )
            transmit data with power p
            when an ACK received after T':  stop
            P_i ← P_i + pT'/T
            if (i never transmitted):  s_i = [LastTransmitNode, i]
        end %if
    when data received from node k ∈ N_R(i):
        if (i never transmitted):  LastTransmitNode=k
end %while
```

For the control packets sent at a common rate, P_{ACK} is the transmit power and $\overline{P}_{N_R}$ is the received power threshold required in a neighborhood.
For a node i, E_i denotes energy collected from data packets and R_i denotes its fill rate.

Figure 2.4: Distributed Greedy Filling Algorithm.

assumes only local information at the nodes. The distributed algorithm is based on the observation that the greatest contribution to a fill rate of a reliable node i will be made by the fill rates of its neighbors that, together with node i, define a *neighborhood* $N_R(i)$ of node i. As we specify later, the transmit broadcast energy of node i will be determined by acknowledgments (ACKs) sent by unreliable neighbors as they become reliable. In addition, ACK packets will allow node i to determine the neighborhood $N_R(i)$. Specifically, an ACK is sent with a fixed power level P_{ACK} chosen such that it guarantees the network connectivity [83]. Distributed algorithms for determining such a power level have been proposed (see [35]). The neighborhood $N_R(i)$ is defined as the set of nodes that receive the ACK from node i (ACK_i), with received power above a threshold $\overline{P}_{N_R}$ that assures reliable reception. The formed links are bidirectional which is a desirable property in a wireless network [34, 84]. All control packets used in the algorithm will be received reliably within each neighborhood. We assume that ACKs are always received correctly. In practice however, the ACK packets are often lost. Such conditions can lead to a power efficiency loss as transmitting nodes rely on ACKs to decide on when to transmit and with what power level.

We now give a detailed description of the distributed algorithm. The algorithm pseudocode is given in Figure 2.4.

We let $S_i = S \cap N_R(i)$ and $U_i = U \cap N_R(i)$ denote respectively reliable and unreliable neighbors of node i. Initially, each node i sets $S_i = \emptyset$. While node i is unreliable, node i will collect the energy of overheard transmissions including those from nodes outside its neighborhood. In addition, it will listen for any ACK_j that is received with power above the threshold $\overline{P}_{N_R}$. When it receives ACK_j, node i will identify that $j \in N_R(i)$ and will respond by sending a link gain (LG) control packet containing h_{ij} reliably to node j. This also informs node j that $i \in N_R(j)$. Node i will then update S_i by adding node j to S_i.

Once it becomes reliable, node i will itself send an ACK_i. An ACK_i will prompt every node j in U_i to send an LG packet to node i, enabling node i to calculate its fill rate R_i and broadcast it in $N_R(i)$.

After a reliable node i receives an ACK_j, it will do the following: if it was transmitting data at the time it received an ACK_j, it will stop transmitting. It will update sets S_i and U_i by moving node j from U_i to S_i, update its fill rate R_i and notify its neighbors of its new fill rate. The reliable node that has the maximum fill rate in S_i will decide to transmit.

At all the times prior to its first transmission, node i will keep track of the identity of the reliable neighbor (a node in S_i) from which it received the last data packet. Thus, in case node i decides to transmit next, it will know the node whose transmission preceeded its own and will use that information for future data broadcasting. The algorithm will stop at a node i when U_i is empty.

When transmitting, node i will repeatedly send the same data packet of duration T at the rate $\overline{r}_\infty$ given by (2.9) and at the fixed low power level p until it hears an ACK after some time T'. Each of these data packets will contribute $h_{ji}pT$ to the energy collected at the unreliable node j. Once the total collected energy at node E_j becomes $E_j = \overline{P}T$, node j will send an ACK_j. By that time, node i has transmitted T'/T data packets and it can determine that the actual transmit power level needed to make j reliable was pT'/T. Power level p is assumed to be chosen small enough so that the negligible excessive power is received at an unreliable node before it has a chance to transmit an ACK. At the end of the algorithm, a node i will know the total broadcast energy it used E_i^t, and thus the power level $P_i = E_i^t/T$ that it will use for time T, to transmit new data that will arrive. Node i will also know the identity of the neighbor after which it should transmit the new data. We assume that P_i is small enough to allow the network to operate in the wideband regime.

In the algorithm, the action of nodes are triggered by receptions of the ACK messages and we let each *step* of the algorithm start with the transmission of an ACK. The algorithm will terminate in $N-1$ steps. It is easy to see that deadlocks cannot occur: Since the network is connected, every unreliable node will eventually have a reliable neighbor, causing the fill rate of that neighbor to be nonzero. At each step, one of the fill rates must be the maximum and, therefore, at least one reliable node will decide to transmit. Once all the nodes are reliable, all the fill rates will be zero and the algorithm

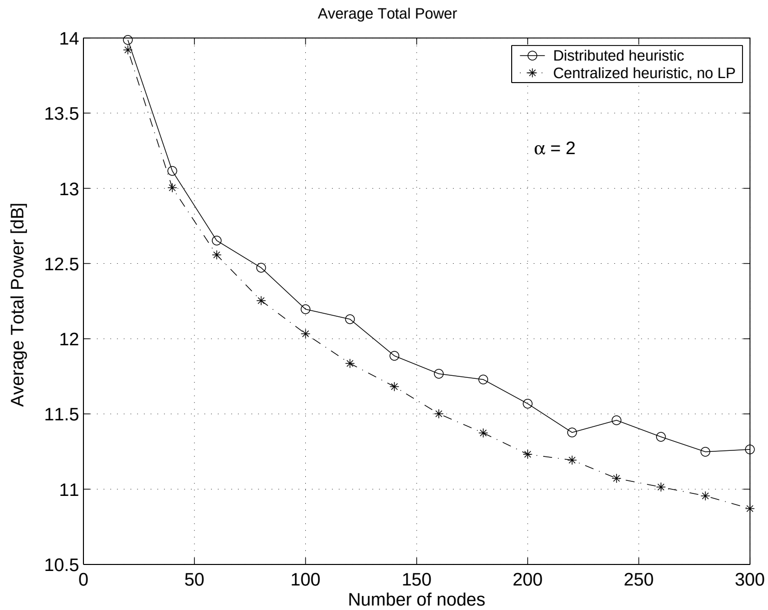

Figure 2.5: Performance comparison between distributed and centralized versions of the Greedy Filling Algorithm.

will stop.

The benefit of the overheard information will be highest in the neighborhood of a transmitting node. For that reason, the distributed greedy heuristic, unlike its centralized counterpart, allows simultaneous transmissions from nodes that are not in the same neighborhood. We examine the impact of the limited knowledge at the nodes to the performance of the heuristic. Performance of the algorithm depends on the choice of range R. For large enough R, the performance of the distributed algorithm approaches the performance of its centralized version. For the smallest R that provided the network connectivity, Figure 2.5 shows the performance comparison of the distributed algorithm with its centralized counterpart as a function of node density. In this case, the actual power levels found by the greedy filling algorithm were used instead of the optimum power levels. Comparison with the centralized algorithm using an LP as well as with BIP, is shown in Figure 2.6. We observe that the performance of the distributed algorithm is close to the performance of its centralized version.

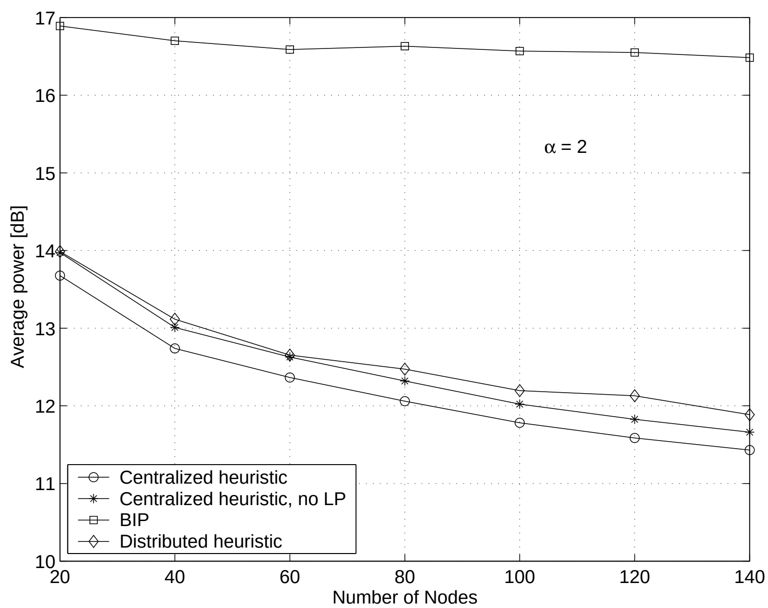

Figure 2.6: Performance comparison between distributed and centralized versions of the Greedy Filling Algorithm and BIP.

2.5 Cooperative Multicast for Maximum Network Lifetime

We next formulate the maximum lifetime problem for cooperative multicast and broadcast and present a *Maximum Lifetime Accumulative Broadcast (MLAB)* algorithm that finds the optimal solution.

2.5.1 Problem Formulation

A *lifetime of a node* i transmitting with power p_i is given by $T_i(p_i) = e_i/p_i$ where e_i is initial battery energy at node i. The *network lifetime* is the time until the first node failure

$$T_{net}(\mathbf{p}) = \min_i T_i(p_i) \tag{2.16}$$

where $\mathbf{p}$ is a vector of transmitted node powers. The problem is to maximize the network lifetime under the constraint that all destination nodes become reliable. For the multicast problem, broadcasting until the subset of destination nodes becomes reliable will solve the problem.

We can define the problem of maximizing the network lifetime for schedule $\mathbf{x}$ in terms of the vector $\mathbf{p}$ of transmitted powers as

$$\min \max_{\mathbf{p}} \frac{p_i}{e_i} \tag{2.17}$$

$$\text{subject to } \mathbf{G}(\mathbf{x})\mathbf{p} \geq \mathbf{1}\overline{P} \tag{2.17a}$$

$$\mathbf{p} \geq \mathbf{0}. \tag{2.17b}$$

The inequality (2.17a) contains $M-1$ constraints as in (2.8), requiring that the accumulated received power at all nodes in the schedule (except the source) is above the threshold $\overline{P}$. It should be apparent that the ith power in $\mathbf{p}$ corresponds to the transmit power of node x_i for every node x_i in the schedule $\mathbf{x}$. Alternatively, we can define the problem in terms of *normalized* node powers $\overline{p}_i = p_i e_1/e_i$ that account for different battery capacities at the nodes; the lifetime at every node i in terms of the normalized power is as if all the batteries were the same: $T_i = e_i/p_i = e_1/\overline{p}_i$. In terms of normalized node powers, Problem (2.17) can be defined as

$$\min \max_{\overline{\mathbf{p}}} \overline{p}_i \tag{2.18}$$

$$\text{subject to } \overline{\mathbf{G}}(\mathbf{x})\overline{\mathbf{p}} \geq \mathbf{1}\overline{P}$$

$$\overline{\mathbf{p}} \geq \mathbf{0}$$

where each column $\overline{\mathbf{g}}_i$ of the normalized gain matrix $\overline{\mathbf{G}}(\mathbf{x})$ is obtained from the corresponding column $\mathbf{g}_i$ of matrix $\mathbf{G}(\mathbf{x})$ as $\overline{\mathbf{g}}_\mathbf{i} = \mathbf{g}_\mathbf{i} e_i/e_1$.

For any schedule $\mathbf{x}$, we can formulate Problem (2.18) as an LP in terms of transmit power levels $\overline{\mathbf{p}}$,

$$\hat{p}^*(\mathbf{x}) = \min_{\overline{\mathbf{p}}} \ \hat{p} \tag{2.19}$$

$$\text{subject to } \overline{\mathbf{G}}(\mathbf{x})\overline{\mathbf{p}} \geq \mathbf{1}\overline{P} \tag{2.19a}$$

$$\overline{\mathbf{p}} \leq \mathbf{1}\hat{p} \tag{2.19b}$$

$$\overline{\mathbf{p}} \geq \mathbf{0}. \tag{2.19c}$$

If $\hat{p} = \hat{p}^*(\mathbf{x})$, then there exists a power vector $\overline{\mathbf{p}}$ such that (2.19a)-(2.19c) are satisfied. It follows that for any $p > \hat{p}$, $\overline{\mathbf{p}} \leq \mathbf{1}p$. Thus, for any power $\hat{p} \geq \hat{p}^*(\mathbf{x})$, we say that power $\hat{p}$ is *feasible* for schedule $\mathbf{x}$. Over all possible schedules, the optimum power is

$$p^* = \min_{\mathbf{x} \in \mathcal{X}_N(\mathcal{D})} \hat{p}^*(\mathbf{x}). \tag{2.20}$$

Equation (2.20) is a formal statement of the problem from which finding the best schedule corresponding to p^* is not apparent. We will see that the power p^*, may, in fact, be the solution to (2.19) for a set of schedules, $\mathcal{X}^*$. From now on, we will consider only normalized powers and we therefore drop the overline notation; $\mathbf{H}$ will denote the ordinary gain matrix, $\mathbf{G}(\mathbf{x})$ will denote the gain matrix permuted for schedule $\mathbf{x}$, and the power vector will be simply $\mathbf{p}$, with p_i representing either the power of node i or node x_i, as appropriate for the context.

Rather than identifying $\mathcal{X}^*$, we employ a simple procedure that, for any power p, determines a collection of schedules for which power p is feasible. In particular, to distribute a broadcast message, we let each node retransmit with power p *as soon as possible*, namely as soon as it becomes reliable. We refer to such a distribution as the *ASAP*(p) *distribution*. During the ASAP(p) distribution, the message will be resent in a sequence of retransmission stages from sets of nodes $Z_1(p), Z_2(p), \ldots$ with power p where in each stage i, a set Z_i that became reliable during stage $i-1$, transmits and makes Z_{i+1} reliable.

Let $S_i(p)$ and $U_i(p)$ denote the reliable nodes and unreliable nodes at the start of stage i. $U_{Di}(p) \subset U_i(p)$ is the set of unreliable destination nodes at the start of stage i. Then, $Z_1(p) = 1$ and $S_i(p) = Z_1(p) \cup \ldots \cup Z_i(p)$. The set $Z_{i+1}(p)$ is given by

$$Z_{i+1}(p) = \{z \in U_i(p) : p \sum_{k \in S_i(p)} h_{zk} \geq \overline{P}\}. \tag{2.21}$$

Note that if power p is too small, the ASAP(p) distribution can *stall* at stage i with $S_{i+1}(p) = S_i(p)$ and $U_{Di}(p) \neq \emptyset$. In this case, ASAP(p) fails to distribute the message to all destination nodes. When $U_{Di}(p) = \emptyset$ at any i, the ASAP(p) distribution *terminates successfully*. We will say that the ASAP(p) distribution is a *feasible multicast* if it terminates successfully.

The partial node ordering, $Z_1(p), Z_2(p), \ldots$, specifies the sequence in which nodes became reliable during the ASAP(p) distribution. In particular, any schedule $\mathbf{x}$ that is consistent with this partial ordering is a feasible schedule for power p. Nodes that become reliable during the same stage of ASAP(p) can be scheduled in an arbitrary order among themselves since these nodes do not contribute to each other's received power. The following theorem verifies that in terms of maximizing the network lifetime it is sufficient to consider only schedules consistent with the ASAP(p) distribution.

Theorem 4 *If $\tilde{p}$ is a feasible power for a schedule $\tilde{\mathbf{x}}$, then the ASAP($\tilde{p}$) distribution is a feasible multicast.*

In particular, Theorem 4 implies that for optimum power p^*, the ASAP(p^*) distribution is feasible. The proof of this theorem is given in Appendix A.3.

We next present the *Maximum Lifetime Accumulative Broadcast (MLAB)* algorithm that determines the optimum power p^*. Once the power p^* is given, broadcasting with ASAP(p^*) will maximize the network lifetime.

2.5.2 The MLAB Algorithm

We label node 1 as the source node and 2 as its closest neighbor (more precisely, the node with the highest link gain to the source). The MLAB algorithm finds the optimum power p^* through a series of ASAP(p) distributions, starting with the smallest possible *candidate broadcast power*, $p = \overline{P}/h_{21}$. Whether ASAP(p) stalls or terminates successfully, we define $\tau(p)$ as the terminating stage. When $p = p^*$, the ASAP(p^*) distribution will terminate in $\tau^* = \tau(p^*)$ stages. When the ASAP(p) distribution stalls at stage $\tau(p)$, we determine the minimum power increase δ for which ASAP($p + \delta$) will not stall at stage $\tau(p)$, in the following way. The increase in broadcast power δ_j needed to make a node $j \in U_{\tau(p)}(p)$ reliable must satisfy

$$\overline{P} = (p + \delta_j) \sum_{k \in S_{\tau(p)}(p)} h_{jk}. \tag{2.22}$$

We choose $\delta = \min_{j \in U_{\tau(p)}(p)} \delta_j$. We then increase p to $p + \delta$ and *restart* the MLAB algorithm. The algorithm stops when an ASAP(p) distribution terminates successfully.

The pseudocode of the algorithm is given in Figure 2.7. The MLAB algorithm ends after at most $N - 1$ restarts. There exists a set of feasible schedules that are consistent

```
Initialize:  p = \overline{P}/h_{21}
Start:  Set S_1(p) = {1}; U_1(p) = S^c
          apply the ASAP(p) distribution;
If ASAP(p) stalls at stage τ(p):
    for all j ∈ U_{τ(p)}(p) calculate:
        δ_j = \overline{P}/ Σ_{k∈S_{τ(p)}(p)} h_{jk} − p;
    Set:  δ = min_{j∈U_{τ(p)}(p)} δ_j;  p ← p + δ;
    go to Start;
end
```

The cardinality of S is given by $|S|$. S^c denotes the complement.

Figure 2.7: MLAB Algorithm.

with the partial ordering given by the ASAP(p) distribution. The normalized transmit power at all nodes in $S_{\tau(p)}(p)$ is p. Note that the last transmitting set $Z_{\tau(p)}$ could in fact, transmit with power less than p if it is enough for the last set of unreliable destination nodes, $U_{D,\tau(p)}(p)$, to become reliable. Thus, choosing the power level at all nodes to be p is not necessarily a unique solution. While this won't change the network lifetime, the latter solution will reduce the total transmit power in the network. Next we show that the power found by MLAB is in fact the optimum power, that is, $p = p^*$.

Theorem 5 *The MLAB algorithm finds the optimum power p^* such that the ASAP(p^*) distribution maximizes the network lifetime.*

The proof is given in Appendix A.4.

Finally, we note that the full restarts of the MLAB algorithm are used primarily to simplify the proof of Theorem 5. In fact, when MLAB stalls, it is sufficient for the reliable nodes to offer incremental retransmissions at power Δ^*. This observation will be the basis of distributed algorithm proposed in Section 2.5.4.

2.5.3 Performance

We now evaluate the benefit of accumulative broadcast to the network lifetime and compare it to the conventional network broadcast that discards overheard data in a network. In particular, networks with randomly positioned nodes in a 10 x 10 square region were generated. The transmitted power was attenuated with distance d as d^α for different values of propagation exponent $\alpha = 2, 3, 4$. The received power threshold was

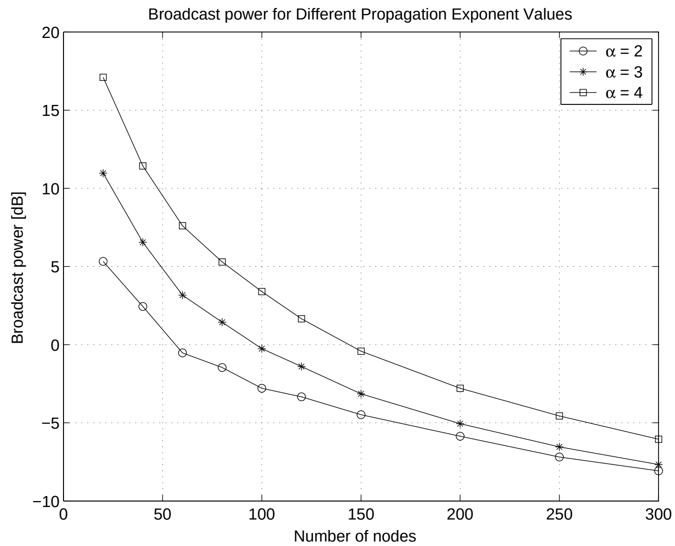

Figure 2.8: Broadcast power for maximum network lifetime for different propagation exponent values. Powers are determined by MLAB Algorithm.

chosen to be $\overline{P} = 1$. Results were based on the performance of 100 randomly chosen networks. Figure 2.8 shows the power p for different values of propagation exponent in networks with different node densities. The observed power decrease is due to shorter hops between nodes in denser networks. For equal battery capacities at the nodes, the corresponding network lifetime, i.e. the time until the first node failure, is shown in Figure 2.9.

Figures 2.10 and 2.11 show the benefit of accumulative broadcast as compared to conventional broadcast in terms of network lifetime. For conventional broadcast, the authors in [37,38] proposed two algorithms, MSNL and MST, that maximize the static network lifetime as well as WMSTSW, a greedy algorithm that increases the dynamic lifetime. We compare the performance of these algorithms for three different battery energy distribution, as given in [37, 38], to the network lifetime found by the MLAB algorithm. We first assume that all the nodes have identical batteries. Then, we consider two different node battery scenarios in which the initial battery energies at the nodes are independent uniform $(0, 1000)$ or uniform $(500, 1000)$ random variables. Several other algorithms to increase the dynamic network lifetime were evaluated in [38] with similar performance to WMSTSW. As expected, we see that solution found by

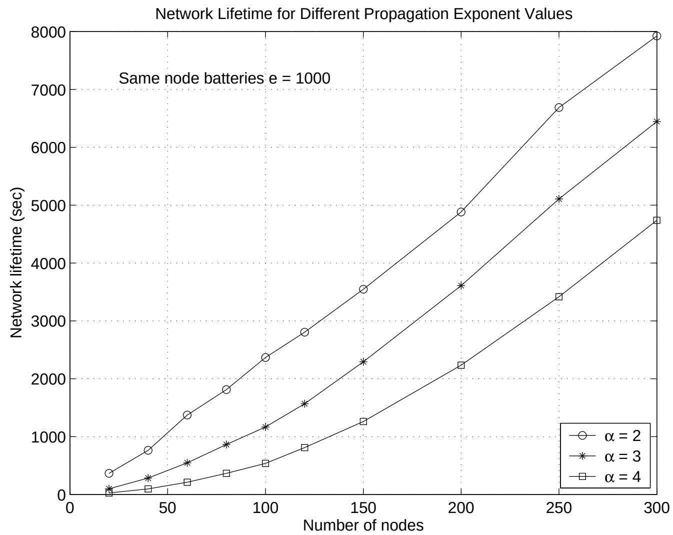

Figure 2.9: Network lifetime for different propagation exponent values. Powers are determined by MLAB Algorithm.

MLAB considerably increases network lifetime. Typically, MLAB increased the network lifetime by a factor of 2 or more. The reason is twofold. First, in MLAB the broadcast uses the energy of overheard information which enables for more radiated energy to be captured. Second, the accumulative broadcast enables MLAB to distribute the load more evenly among the nodes than does the dynamic load balancing in conventional broadcast.

2.5.4 Distributed MLAB Algorithm

We next describe a distributed MLAB algorithm for accumulative broadcast that determines broadcast power locally at each node. Nodes are assumed to have no knowledge of link gains (distances) to other nodes at the beginning of the algorithm. The distributed algorithm will be run at the beginning of a broadcast session during the broadcast of the first message. Let q denote the broadcast power determined by distributed MLAB. Once the power q is determined, data will be broadcasted through the ASAP(q) distribution. In a static network where the same power q is used throughout a long broadcast session, the initial overhead to determine q will be small compared to the amount of broadcast data.

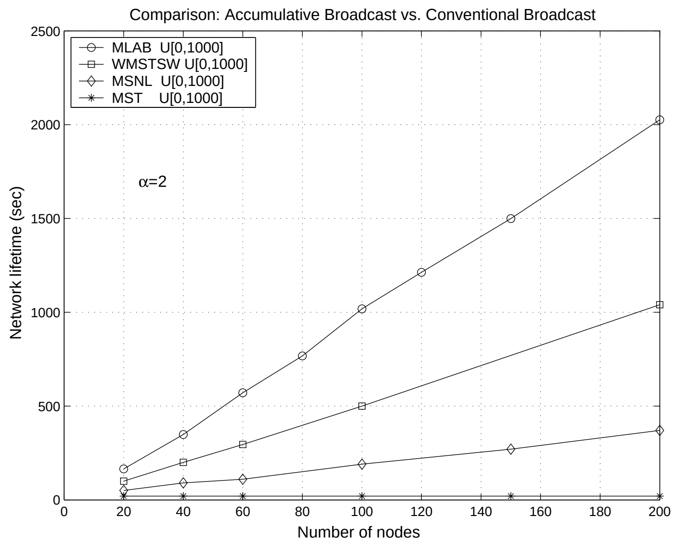

Figure 2.10: Network lifetime of accumulative broadcast and conventional broadcast.

The distributed implementation of the MLAB algorithm has to resolve the following:

1. *When should a reliable node decide to increase the broadcast power?*

2. *How much should a reliable node increase the broadcast power?*

When the ASAP(p) distribution stalls, determining the necessary power increase δ requires global knowledge of network gains and cannot be computed locally at a node. In the distributed MLAB algorithm, the broadcast power will be increased in steps of size Δ, for some small fixed power Δ. Further, during the initial broadcast phase while the algorithm is run to determine q, we let Δ be the transmit power of every transmission. A reliable node intending to transmit with power $n\Delta$ for some $n > 1$ will instead repeatedly transmit for n times, each time with power Δ. A transmission from a node i with power Δ will be overheard by a number of nodes that define a Δ-*neighborhood* $N_i(\Delta)$ of node i. Nodes will belong to $N_\Delta(i)$ if they can detect the presence of a signal sent at node i, although their received power may not be sufficient for reliable decoding. Overhearing a broadcast from a node k will enable node i to determine the link gain h_{ik} and identify node k as its reliable neighbor. During the algorithm, node i will keep track of its set of reliable neighbors, $R_i \subset N_i(\Delta)$. From the number of repeated transmissions at node k, node i will also be able to determine the current

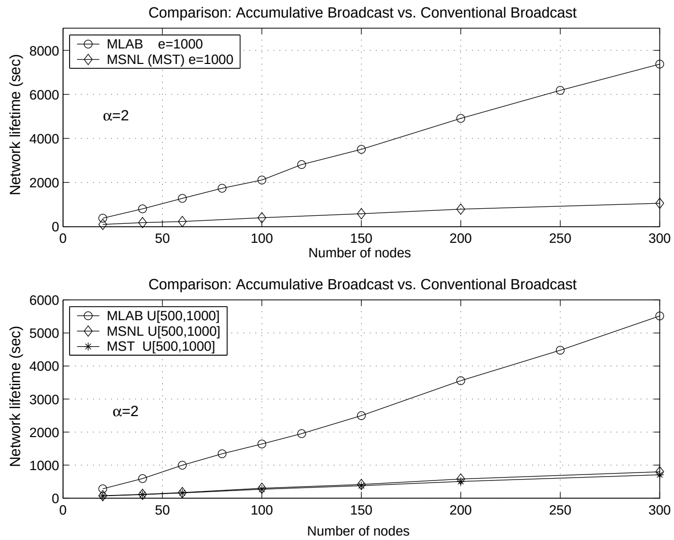

Figure 2.11: Network lifetime of accumulative broadcast and conventional broadcast.

transmit power at node k. Because the transmit power will not necessarily be the same at all nodes all the time, node i will keep track of transmit power $p_i(k) = n_i(k)\Delta$ for every $k \in R_i$, where $n_i(k)$ is a number of repeated broadcasts by node k. In addition, once reliable, node i will keep track of its unreliable neighbors, U_i. An unreliable node j will send NACK$_j$ control messages to identify itself. As node i becomes reliable, it will broadcast with maximum power among its reliable neighbors in R_i,

$$p(R_i) = \Delta \max_{k \in R_i}\{n_i(k)\}.$$

While reliable, whenever it overhears a transmission that increases the power $p(R_i)$, node i will repeat the broadcast to meet it. In that way, the current maximum transmit power in the network $\overline{q}$ will *propagate* until all reliable nodes have transmitted with that power. A reliable node that overhears no transmissions for time $T_o > \tau(\Delta)$ and has unreliable neighbors will decide to increase its transmit power. At the end of the algorithm, power $\overline{q}$ will determine the broadcast power q. A detailed description of the algorithm is given in pseudocode in Figure 2.12.

Constraining the power of each transmission to Δ defines Δ-neighborhoods and allows nodes to determine the link gains within their Δ-neighborhoods. Therefore, power Δ defines the network topology and has to be high enough to guarantee network

```
At each node i do:
initialize R_i = ∅, p_R = 0;
while (p_R < P̄) do:
    when data received with P from k:
        collect data; p_R ← p_R + P;
        if k ∉ R_i:
            h_ik = P/Δ, n_i(k) = 0, R_i ← R_i ∪ {k};
            send NACK_i reliably to k;
        end %if
        n_i(k) = n_i(k) + 1;
end % while
as (p_R ≥ P̄) do once:
    decode the message;
    set n_i(i) = max_{k∈R_i}{n_i(k), 1};
    broadcast the decoded message once;
    U_i = {j : j that responded with NACK_j};
    broadcast n_i(i) − 1 times;
while (p_R ≥ P̄) do:
    when data received from node k:
        update n_i(k) = n_i(k) + 1;
        if n_i(k) > n_i(i):
            n_i(i) ← n_i(i) + 1, broadcast;
        if k ∉ R_i:   R_i ← R_i ∪ {k}, U_i ← U_i \ {k};
    if no data received for T_o and U_i ≠ ∅:
        broadcast;
        n_i(i) ← n_i(i) + 1;
    end %if
end % while
```
Received power at a node is denoted p_R.

Figure 2.12: Distributed MLAB Algorithm.

connectivity [83]. In the distributed MLAB algorithm, it is sufficient that under power Δ, the network is connected in the overheard sense. That is, in the underlying graph, a link between two nodes exists if they can overhear each other. During MLAB, we assume that the network is connected under power Δ.

This assumption is not essential for the algorithm and can be relaxed by letting MLAB algorithm rely on preexisting network topology. Different distributed algorithms for determining network topology have been proposed (see [35], [84]) and typically employ short HELLO control packets exchanged at the nodes. Given the power P_c and rate r_c of control packets, HELLO packets define one-hop neighborhood $N_i(P_c)$ for node i as all nodes that can reliably receive a HELLO$_i$ packet sent at node i. A version of the distributed MLAB algorithm can then be run on the top of the topology defined by neighborhoods $N_i(P_c)$ instead of $N_i(\Delta)$.

Note that decreasing the rate r_c reduces the power P_c necessary for network connectivity by reducing the receiver power threshold needed for reliable communication. Connectivity in overheard sense, required for Δ-neighborhoods, reduces this threshold to its minimum value necessary to acquire a signal or decode a packet header and thus reduces necessary power for connectivity. Therefore, it may be reasonable to assume that under power Δ, the network is connected. The next theorem shows that the algorithm is correct and finishes in finite time.

Theorem 6 *The distributed MLAB algorithm makes every network node reliable in finite time.*

The proof is given in Appendix A.5.

The running time and performance of the algorithm are dependent on the values of the parameters Δ and T_o. In fact, we have the following theorem.

Theorem 7 *For large enough T_o, $T_o > \tau(\Delta)$, power q found by the distributed MLAB algorithm is within Δ of the optimum solution; that is, $q \in [p^*, p^* + \Delta)$.*

The proof is given in Appendix A.6.

Thus, by choosing a smaller Δ, the solution found by the distributed MLAB algorithm approaches the optimum, at the expense of longer running time due to the larger T_o and smaller step size Δ. When the distributed MLAB does not rely on a preexisting topology, there is a lower bound on Δ to guarantee network connectivity. At the other extreme, for Δ large enough to guarantee full connectivity (every node can overhear every transmission), T_o can be chosen to be 0. The optimal tuning of the algorithm parameters has yet to be determined.

2.6 Conclusion

In this chapter, we defined and solved two accumulative broadcast problems. We showed that finding the optimum solution for the minimum-energy accumulative broadcast problem is NP-complete, and proposed a simple energy-efficient heuristic algorithm. For the maximum lifetime problem we showed that there exists a simple optimal solution and proposed an MLAB algortihm that finds it.

While in accumulative broadcast nodes collect unreliably received signals, they forward the information only after reliably decoding a message. And yet, as pointed out earlier, the relay strategy where no decoding is performed at the relays is optimal for particular relay networks [12, 20]. For the networks of interest in this thesis, where power rather than bandwidth is the limiting resource, we wish to evaluate the energy-efficiency of the unreliable, amplify-and-forward strategy and compare it with the reliable, decode-and-forward strategy. We present such an analysis for the multiple-relay channel in the next chapter.

Chapter 3

Bandwidth and Power Allocation for Cooperative Strategies in Gaussian Relay Networks

3.1 Introduction

We next consider a network that consists of a single source-destination pair and M relays that dedicate their resources to relaying information for the source. Simple two-hop relaying is assumed, thus precluding communications among relays. The capacity of this network is not known for any finite M. The asymptotic capacity in this network, as the number of relays gets large, is achieved by the unreliable forwarding at the relays [20]. We examine the achievable rates with amplify-and-forward (AF) and decode-and-forward (DF) relaying strategies in this relay network. We show that the AF strategy does not necessarily benefit from the large available bandwidth. We characterize the optimum AF bandwidth and show that transmitting in the optimum bandwidth allows the network to operate in the linear regime where the achieved rate increases linearly with the available network power. We then present the optimum power allocation among the AF relays. Motivated by the large bandwidth resources we further consider a network that uses orthogonal transmissions at the nodes. While the above result for the optimum bandwidth still holds, we show that a different set of relays should optimally be employed.

The optimum AF bandwidth and the relay powers can be contrasted to the decode-and-forward solution. In a network with unconstrained bandwidth, the DF strategy will operate in the wideband regime to minimize the energy cost per information bit [71,85]. The wideband DF strategy requires again a different choice of relays; in the case of orthogonal signaling, we show that the data should be sent through only one DF relay. Thus, in general, in a large scale network, a choice of a coding strategy goes beyond determining a coding scheme at a node; it also determines the operating bandwidth as well as the best distribution of the relay power.

3.2 System Model

We consider a wireless Gaussian network with a single source, labeled node 0, the destination, labeled node $M+1$, and M relays. We consider two bandwidth allocations in the given network:

1. Shared bandwidth. All the relays transmit in a common bandwidth $W^{(c)}$.

2. Orthogonal channels. Every node is assigned an orthogonal channel of bandwidth $W^{(o)}$.

We adopt a discrete-time Gaussian channel model [81] and let the vector $\mathbf{X}[n] = [X_0[n], \ldots, X_M[n]]^T$ denote the channel inputs in time slot n. The channel input $X_0[n]$ depends on the source message and the channel input $X_m[n]$ at relay m depends on its past outputs $X_m[n] = f_m(Y_m[1], \ldots, Y_m[n-1])$. The output at relay m is denoted Y_{Rm} and at the destination $Y^{(c)}$ and $\mathbf{Y}^{(o)}$ for the shared bandwidth and orthogonal signaling, respectively.

In such a network, we consider *two-hop* forwarding strategies in which relays use only the information received from the source to choose their channel inputs and forward the messages to the destination. In the first hop, the source transmits. The channel output at relay m is

$$Y_{Rm}[n] = \sqrt{\alpha_m} X_0[n] + Z_{Rm}[n] \tag{3.1}$$

and at the destination,

$$Y^{(c)}[n] = \sqrt{\beta_0} X_0[n] + Z[n], \tag{3.2}$$

where $\sqrt{\alpha_m}$ and $\sqrt{\beta_0}$ are the source-relay m and source-destination channel gain, respectively, and $Z[n]$ is a zero-mean Gaussian random variable with variance $N_0/2$. In the second hop, relays transmit. In shared bandwidth, the channel output at the destination is

$$Y^{(c)}[n] = \sum_{m=1}^{M} \sqrt{\beta_m} X_m[n] + Z[n]. \tag{3.3}$$

However, when relays use orthogonal channels,

$$\mathbf{Y}^{(o)}[n] = \mathbf{B}\mathbf{X}[n] + \mathbf{Z}[n] \tag{3.4}$$

where $\mathbf{B} = \text{diag}\begin{pmatrix}\sqrt{\beta_0} & \dots & \sqrt{\beta_M}\end{pmatrix}$ and $\mathbf{Z}$ is a Gaussian noise vector with covariance matrix $\mathbf{K} = \sigma^2 \mathbf{I}_{M+1}$.

Using the Cut-Set Theorem [81, Thm. 14.10.1], it was shown in [19] that the capacity of this network is upper bounded by $\log M$, given that there is a *dead zone* around the source that contains no relays.

We consider two transmission strategies at the relays. As in [19], we consider the amplify-and-forward protocol at the relays, in which the noisy version of the source input X_0 received at relay m, $1 \leq m \leq M$, is amplified and forwarded with a unit delay. For amplification gain $b_m \geq 0$, in time slot n, relay m transmits

$$X_m[n] = \sqrt{b_m}\left(\sqrt{\alpha_m} X_0[n-1] + Z_{Rm}[n-1]\right). \tag{3.5}$$

We then consider the decode-and-forward strategy in which the source transmission is reliably received at a relay. The relay decodes, re-encodes using an independent codebook and transmits.

Rather than considering the power constraint imposed on each transmitter, we assume that the total power budget of p [Watts] is allocated to the network. The constraint is on the total power rather than on the power per dimension because DF and AF will not in general operate in the same bandwidth: as explained later on in this chapter, DF will operate in the wideband regime in order to improve the energy-efficiency. On the other hand, AF will operate in the smaller bandwidth in order to reduce the amount of amplified noise at the relays. Further, such a constraint allows a power allocation among the nodes such that $E[\mathbf{X}^T\mathbf{X}] \leq p/2W^{(i)}$, $i = 1, 2$.

3.3 Amplify-and-Forward: Optimum Bandwidth Allocation

We next consider the rates achievable with the AF strategy. Let p_m denote the power at node m and let $\mathbf{p} = \begin{bmatrix} p_0 & \dots & p_M \end{bmatrix}^T$ be the power allocation at all nodes. Vector $\mathbf{P} = \begin{bmatrix} P_0 & \dots & P_M \end{bmatrix}^T$ denotes nodes' powers per dimension and $P_m = p_m/2W^{(i)}$, $i =$

1, 2. The amplification gain b_m is chosen such that the transmit power at node m is p_m and is found from (3.5) to be

$$b_m = \frac{P_m}{\alpha_m P_0 + N_0/2}, \qquad m = 1, \dots M. \tag{3.6}$$

The achievable rates, given by the maximum mutual information between the channel input and the output can be found from the result in [58] to be

$$I_{AF}^{(i)}(\mathbf{P}) = \frac{1}{2}\log\left[1 + \frac{P_0}{N_0/2}\left(\beta_0 + G^{(i)}(\mathbf{P})\right)\right] \tag{3.7}$$

for $i = c, o$ where

- For the shared bandwidth,

$$G^{(c)}(\mathbf{P}) = \frac{\left(\sum_{m=1}^{M}\sqrt{\frac{\alpha_m \beta_m P_m}{\alpha_m P_0 + N_0/2}}\right)^2}{\left(\sum_{m=1}^{M}\frac{\beta_m P_m}{\alpha_m P_0 + N_0/2} + 1\right)}. \tag{3.8}$$

- For orthogonal channels,

$$G^{(o)}(\mathbf{P}) = \sum_{m=1}^{M}\frac{\alpha_m \beta_m P_m}{\alpha_m P_0 + \beta_m P_m + N_0/2}. \tag{3.9}$$

Rates (3.7) are normalized by the number of dimensions utilized by a node rather than by the total number of dimensions in the channel. For $G^{(i)}(\mathbf{P}) = 0$, (3.7) becomes the rate achieved in the single-user channel, by a direct source transmission at power P_0. Thus, we can view $G^{(i)}(\mathbf{P})$ as the gain obtained by employing the AF relays. The difference in the AF gains (3.8) and (3.9) comes from the coherent combining of the relay signals transmitted in a shared bandwidth, which is forfeited in the orthogonal channel system. The analysis presented in this section, however, applies to both cases and we therefore drop the (i) superscript. We next consider the total rate achieved by the AF strategy

$$r_{AF} = 2W I_{AF}(\mathbf{P}) \quad \text{bits/s}, \tag{3.10}$$

where I_{AF} is given by (3.7). As W becomes large, we observe that $G(\mathbf{p}/W)$ decreases to zero and therefore

$$\lim_{W\to\infty} r_{AF} = \lim_{W\to\infty} W \log\left(1 + \frac{\beta_0 p_0}{N_0 W}\right) = \frac{\beta_0 p_0}{N_0 \ln 2} \quad \text{bits/s}, \tag{3.11}$$

which is the rate achieved in the wideband regime by the source transmission. Therefore, there is no benefit from AF relays transmitting in the wideband regime. This behavior was previously observed in [86] in a parallel Gaussian network with two relays. Except for the very special case in which the source is in a favorable position compared to all the relays, the rate r_{AF} generally decreases for large W.

To characterize the optimum AF bandwidth, we formulate the AF power/bandwidth relay problem as

$$r^* = \max_{\mathbf{P},W} \quad 2WI_{AF}(\mathbf{P}) \tag{3.12}$$

$$\text{subject to } 2W\sum_{m=0}^{M} P_m \leq p, \tag{3.12a}$$

$$\mathbf{P} \geq 0, \tag{3.12b}$$

$$0 \leq W \leq W_{\max}. \tag{3.12c}$$

We assume that $W_{\max}$ is sufficiently large to allow the network to operate in the wideband regime. Let $(\mathbf{P}^*, W^*)$ denote the optimum power and bandwidth allocation that achieves r^* in (3.12). We first observe that, to achieve nonzero rate in (3.12), it has to hold that $W^* > 0$ and $P_0^* > 0$. Furthermore, constraint (3.12a) is always binding. Depending on the values of the channel gains, a solution to problem (3.12) may be a direct source transmission, that is, $P_m^* = 0$ for $m = 1, \ldots, M$, $W^* = W_{\max}$ and P_0^* given by (3.12a). Otherwise, there will be a set of $0 < K \leq M$ relays employing the AF strategy. Given $\mathbf{P}^*$, it will be convenient to relabel the nodes such that $m \in \{1, \ldots, K\}$ relays are the active transmitters with powers $P_m^* > 0$ while $P_m^* = 0$, for $m \in \{K+1, \ldots, M\}$. The Lagrangian in (3.12) is

$$\Lambda = 2WI_{AF}(\mathbf{P}) - \mu\left(2W\sum_{m=0}^{M} P_m - p\right) \tag{3.13}$$

The fact that rate r_{AF} is decreasing with W for large W implies that $W^* < W_{\max}$. Since in addition $W^* > 0$, the solution to (3.12) is never on the boundary (3.12c). By the Kuhn-Tucker conditions, this implies

$$\frac{\partial \Lambda}{\partial W} = 2I_{AF}(\mathbf{P}^*) - 2\mu\sum_{m=0}^{K} P_m^* = 0 \tag{3.14}$$

From (3.14), we obtain the Lagrange multiplier

$$\mu = \frac{I_{AF}(\mathbf{P}^*)}{\sum_{m=0}^{K} P_m^*}. \tag{3.15}$$

For nodes $k = 0, \ldots, K$ with non-zero transmitter powers, the Kuhn-Tucker conditions are

$$\frac{\partial \Lambda}{\partial P_k} = 2W^* \left[\frac{\partial I_{AF}(\mathbf{P}^*)}{\partial P_k} - \mu \right] = 0, \qquad k = 0, \ldots, K. \tag{3.16}$$

From (3.15) and (3.16),

$$\frac{\partial I_{AF}(\mathbf{P}^*)}{\partial P_k} = \mu = \frac{I_{AF}(\mathbf{P}^*)}{\sum_{m=0}^{K} P_m^*}, \qquad k = 0, \ldots, K. \tag{3.17}$$

The optimum power allocation $(P_0^*, \ldots, P_K^*)$ can then be determined from $K+1$ equations given by (3.17), and is independent of r and W^*. We present the solution for the optimum relay powers in the next section. The optimum bandwidth can be determined such that the solution lies on the feasibility region boundary (3.12a):

$$W^* = \frac{p}{2\sum_{m=0}^{K} P_m^*}. \tag{3.18}$$

From (3.12), (3.15) and (3.18),

$$r^* = 2W^* I_{AF}(\mathbf{P}^*) = \frac{I_{AF}(\mathbf{P}^*)}{\sum_{m=0}^{K} P_m^*} p = \mu p. \tag{3.19}$$

We thus proved the following:

Theorem 8 *The AF relay problem (3.12) has an optimum solution in which the optimum bandwidth W^*, the maximum rate r^* and the total power p have a linear relationship.*

We can view μ as a 'rate reward' (or the power efficiency); increasing the total available power in (3.12) by Δp increases the maximum achievable rate r^* by $\mu \Delta p$.

3.4 Amplify-and-Forward: Optimum Relay Power Allocation

We next consider a subproblem of (3.12) that determines the optimum relay powers per dimension, for any given source power P_0. We consider the shared bandwidth case first.

3.4.1 Shared Bandwidth

Given a source power P_0, we let

$$\gamma_m = \frac{\beta_m}{\alpha_m P_0 + N_0/2}. \tag{3.20}$$

To maximize the rate (3.7) over the relay powers $\hat{\mathbf{P}} = \begin{bmatrix} P_1 & \dots & P_M \end{bmatrix}^T$, we maximize the AF gain (3.8):

$$\max_{\hat{\mathbf{P}}} \quad \frac{\left(\sum_{m=1}^{M} \sqrt{\alpha_m \gamma_m P_m}\right)^2}{\sum_{m=1}^{M} \gamma_m P_m + 1} \tag{3.21}$$

$$\text{subject to } \sum_{m=1}^{M} P_m \leq P_R, \tag{3.21a}$$

$$\hat{\mathbf{P}} \geq 0 \tag{3.21b}$$

where $P_R = p/2W - P_0$ is the power allocated to the relays. To solve (3.21), we first argue that the solution is always on the boundary (3.21a). To see that, consider a feasible solution $\overline{\mathbf{P}}$ such that $\sum_{m=1}^{M} \overline{P}_m < P_R$. Then, there exist a constant $K > 1$ and a feasible solution $\mathbf{P}' = K\overline{\mathbf{P}}$ such that $\mathbf{P}'$ is on the boundary $K \sum_{m=1}^{M} \overline{P}_m = P_R$. Furthemore, it is easy to verify that $G(\mathbf{P}') > G(\overline{\mathbf{P}})$. We can thus let the constraint (3.21a) be satisfied with equality. The objective function (3.21) becomes

$$G(\hat{\mathbf{P}}) = \frac{\left(\sum_{m=1}^{M} \sqrt{\alpha_m \gamma_m P_m}\right)^2}{\sum_{m=1}^{M} \left(\gamma_m + 1/P_R\right) P_m}. \tag{3.22}$$

A solution to problem (3.21) can be found by representing the objective function (3.22) in the form of a Rayleigh quotient that would then be maximized [87]. A simpler approach, however, can be used by introducing variables $\mathbf{z} = [z_m]$

$$z_m = \sqrt{\gamma_m + 1/P_R}\sqrt{P_m}, \qquad m = 1, \dots, M \tag{3.23}$$

and a vector of coefficients $\mathbf{d} = [d_m]$ where

$$d_m = \sqrt{\frac{\alpha_m \gamma_m}{\gamma_m + 1/P_R}}, \qquad m = 1, \dots, M \tag{3.24}$$

Problem (3.21) can then be represented in a vector form

$$\max_{\mathbf{z}} \quad \frac{\left(\mathbf{d}^T \mathbf{z}\right)^2}{\mathbf{z}^T \mathbf{z}}. \tag{3.25}$$

Applying the Schwartz inequality, the solution to (3.25) is $\mathbf{z}^* = k\mathbf{d}$ where the constant k can be found from (3.21a) and (3.23). We get the optimum powers in the MRC form as

$$P_m^* = \frac{P_R \delta_m}{\sum_{k=1}^{M} \delta_k}. \tag{3.26}$$

where we define

$$\delta_m = \frac{\alpha_m \gamma_m}{(1 + \gamma_m P_R)^2} \tag{3.27}$$

The AF gain (3.8) becomes

$$G^{(1)}(P_0, P_R) = \sum_{m=1}^{M} \frac{\alpha_m \beta_m P_R}{\alpha_m P_0 + \beta_m P_R + N_0/2}. \tag{3.28}$$

The next lemma follows by comparing AF gains (3.28) and (3.9).

Lemma 1 *For any given source power P_0 and relay power P_R with the relays employing amplify-and-forward, signaling in shared bandwidth outperforms orthogonal signaling.*

Given the relay powers (3.26), the AF power/bandwidth problem (3.12) for any given W reduces to

$$\max_{P_0, P_R} \; I_{AF}^{(1)}(P_0, P_R) \tag{3.29}$$

$$\text{subject to } P_0 + P_R \le P_T = \frac{p}{2W}, \tag{3.29a}$$

$$P_0, P_R \ge 0. \tag{3.29b}$$

Lemma 2 *There exists a unique optimum solution $(P_0^*(W), P_R^*(W))$ to (3.29).*

The proof for the lemma follows from the observation that the optimum is on the boundary where $P_R = P_T - P_0$ and that $I_{AF}^{(1)}(P_0, P_T - P_0)$ is strictly concave in P_0.

Given $(P_0^*(W), P_R^*(W))$, the AF power/bandwidth problem (3.12) reduces to maximizing the rate with respect to bandwidth for $0 \le W \le W_{\max}$,

$$r^*(W) = \max_{W} 2W I_{AF}^{(1)}(P_0^*(W), P_R^*(W)). \tag{3.30}$$

3.4.2 Numerical Results

Numerical calculation of $r^*(W)$ is straightforward. The relay powers (3.26) are shown in Figures 3.1-3.4 for a scenario of $M = 2500$ relays positioned on a 100×100 square grid. The source and the destination are positioned on the two opposite sides of the grid. The propagation exponent $n = 2$ was chosen. As explained in the Introduction, for convenience we choose $N_0/2 = 1$ and path loss $G = r^{-n}$ where r denotes a transmitter-receiver distance. Using more realistic values for N_0, G and the distance units scale the results without impact on the power fraction allocated to a relay.

In the simulations, we consider a dense network with a large number of AF nodes transmitting signals that coherently combine at the destination receiver. It may therefore appear that such a setting may cause a physically impossible scenario in which the received power at the destination is higher than the transmitted power at the source. Calculating the received power for the actual values of the parameters reveals that the received power is still significantly smaller than the source transmit power. To illustrate this fact, assume that $M = 100$ relays are positioned on the line between the source and the destination at a distance of 1 meter from the destination and are transmitting with power $P_m = 1$ mW at a carrier frequency of 3 GHz. For the source power $P_0 = 1$ mW the resulting receiver power is 6×10^{-4} mW.

For large source power P_0, relay powers are shown in Figure 3.1. In this case, the received SNR at the relays is high and the network MAC side from the relays to the destination limits the performance. The relays that have a better channel to the destination are employed. Figure 3.2 shows the opposite case of a small power P_0 and a high power P_R. We observe a reversed relay power allocation compared to the previous case, as the network tries to improve the broadcast side performance by choosing the relays with high received SNR. Figure 3.3 shows the powers for larger values of P_0 and P_R. Finally, Figure 3.4 shows the relay powers when the network operates in a low SNR-regime due to small P_0 and P_R.

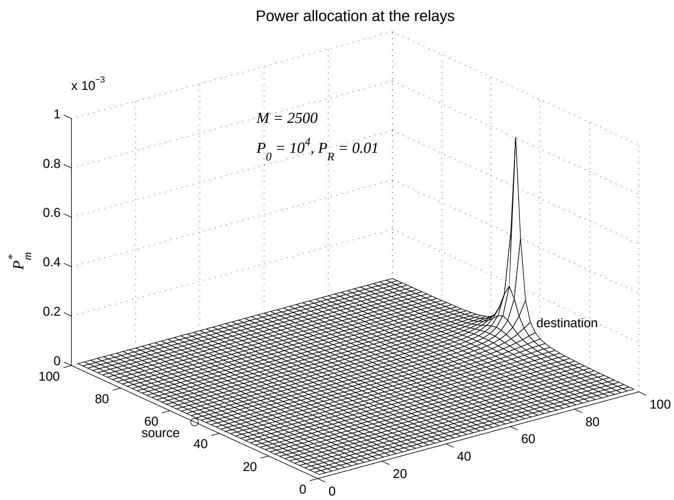

Figure 3.1: Relay powers for $P_0 = 10^4$, $P_R = 0.01$. Due to a small power available to the relays, the MAC side limits the network performance. The solution chooses relays that have a better relay-destination channel.

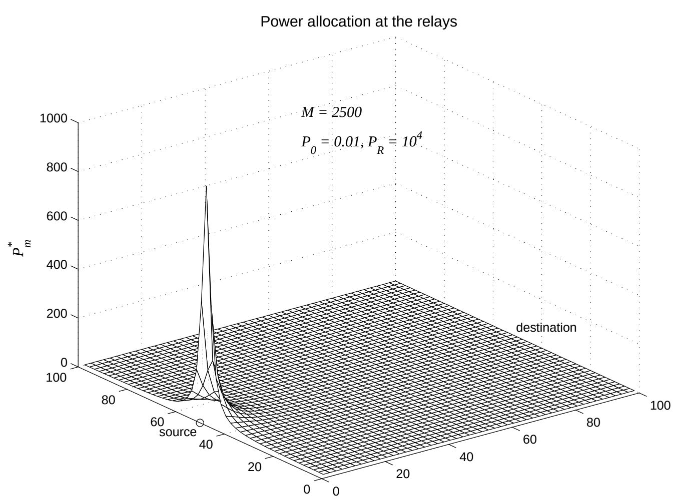

Figure 3.2: Relay powers for $P_0 = 0.01$, $P_R = 10^4$. Relays that are closer to the source are employed.

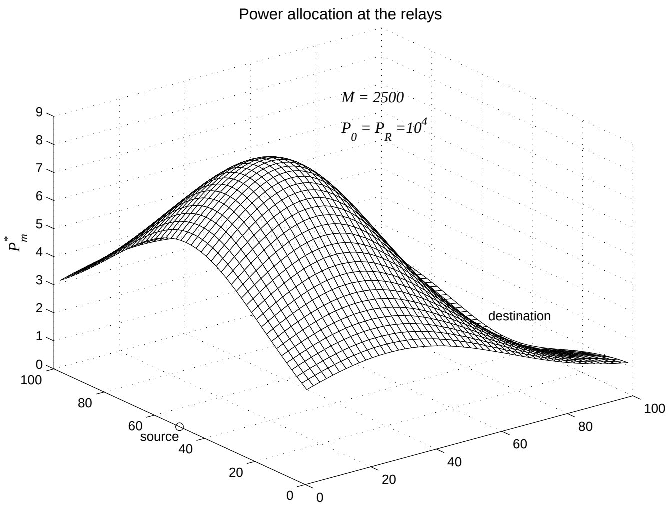

Figure 3.3: Relay powers for $P_0 = P_R = 10^4$ and $N_0/2 = 1$.

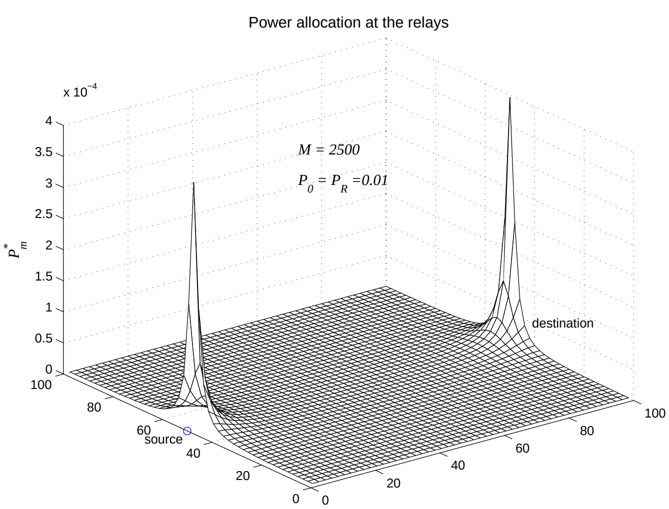

Figure 3.4: Relay powers for $P_0 = P_R = 0.01$ and $N_0/2 = 1$.

3.4.3 Orthogonal Channels

We next identify the best subset of AF relays and their powers for the case of orthogonal signaling. Given a source power P_0, we let

$$\gamma_m = \sqrt{\frac{\alpha_m \beta_m}{\alpha_m P_0 + N_0/2}}. \tag{3.31}$$

Again, to maximize the rate (3.7) over the relay powers $\hat{\mathbf{P}}$, we maximize the AF gain (3.9)

$$\max_{\hat{\mathbf{P}}} \sum_{m=1}^{M} \frac{\alpha_m \gamma_m^2 P_m}{\alpha_m + \gamma_m^2 P_m} \tag{3.32}$$

$$\text{subject to } \sum_{m=1}^{M} P_m \leq P_R, \tag{3.32a}$$

$$\hat{\mathbf{P}} \geq 0. \tag{3.32b}$$

From the Kuhn-Tucker conditions, the solution to (3.32) is in the form of water-filling

$$P_m^* = \frac{\alpha_m}{\gamma_m} \left[\frac{1}{\sqrt{\eta}} - \frac{1}{\gamma_m} \right]^+, \qquad m = 1, \ldots, M \tag{3.33}$$

where η is the Lagrange multiplier and is found such that constraint (3.32a) is satisfied with equality. Once again, the best choice of relays varies with the transmit source power. We observe that the AF relay network, depending on whether it operates in shared or orthogonal channels, will require two different relay power allocations, as given by (3.26) and (3.33).

3.4.4 Single-Relay Channel

A different AF paradigm can be used in a single-relay channel (or in a relay network with multiple relays that cannot hear each others' transmissions). Under the assumption that a relay can transmit and receive simultaneously, we can allow the source and the relay to transmit at the same time, in the shared bandwidth. As observed in [14], this strategy turns the relay channel into a unit-memory intersymbol interference channel, as the signal at the destination becomes

$$Y[n] = \sqrt{\beta_0} X[n] + \sqrt{\alpha_1 \beta_1 b_1} X[n-1] + \tilde{W}[n] \tag{3.34}$$

where

$$\tilde{W}[n] = \sqrt{\beta_1 b_1} Z[n] + W[n] \tag{3.35}$$

is the total noise at the destination. The amplification gain b_1 is given in (3.6), as before. Note that $\tilde{W}[n]$ has variance

$$\sigma^2_{\tilde{W}} = E[|\tilde{W}[n]|^2] = (\beta_1 b_1 + 1)N_0/2. \tag{3.36}$$

The capacity of this effective ISI channel is [81]

$$C = \int_{-1/2}^{1/2} \log_2 \left(\frac{v}{\sigma^2_{\tilde{W}}} |H(f)|^2 \right)^+ df \tag{3.37}$$

under the power constraint

$$\int_{-1/2}^{1/2} \left[v - \frac{\sigma^2_{\tilde{W}}}{|H(f)|^2} \right]^+ df = P_0 \tag{3.38}$$

where

$$|H(f)|^2 = \beta_0 + \alpha_1 \beta_1 b_1 + 2\sqrt{\alpha_1 \beta_0 \beta_1 b_1} \cos(2\pi f). \tag{3.39}$$

When using the large bandwidth W, the amplification gain b_1 goes to zero and the channel (3.34) becomes the point-to-point channel with no benefit from the relay. Thus, this AF strategy can again benefit from the bandwidth optimization. We illustrate that fact on the network example of [14] where a source, relay and destination are positioned on a line. The source-relay distance is denoted as d. For $P_0 = P_R = 10$, we repeat the performance comparison given in [14] in Figure 3.5. Figure 3.6 shows the comparison for $P_0 = P_R = 0.01$. Note that, when employing the DF strategy, a relay also transmits and receives simultaneously. We observe that the relative performance between DF and AF changes as a different power per dimension is used. Thus, in Figure 3.7, we compare the two strategies while allowing each of them to operate in its optimum bandwidth and thus optimum power per dimension, for the given power (Watts) at the nodes.

3.5 Decode-and-Forward

A multi-hopping strategy [14] in which the data sent at the source is successively decoded by the relays and finally by the destination was shown to achieve the rates [14,

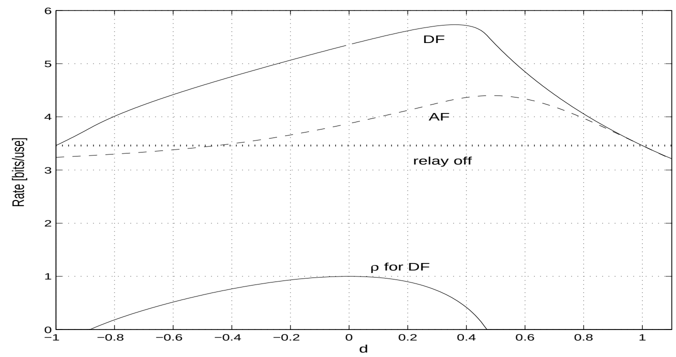

Figure 3.5: Achieved rate [bits/dim] in a single-relay channel for $P_0 = P_1 = 10$.

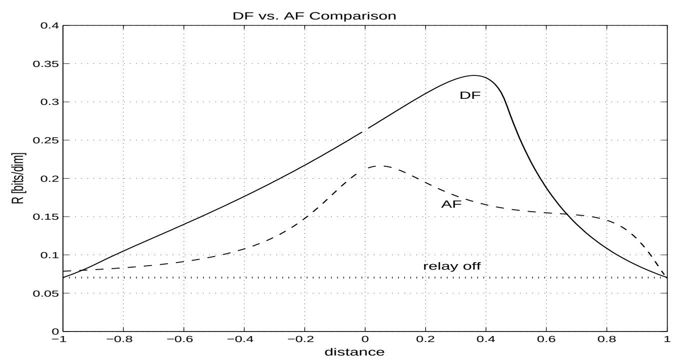

Figure 3.6: Achieved rate [bits/dim] in a single-relay channel for $P_0 = P_1 = 0.01$.

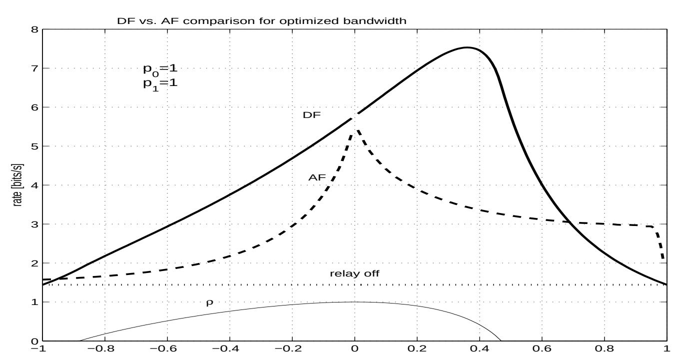

Figure 3.7: Achieved rate [bits/s] in a single-relay channel for $p_0 = p_1 = 1$.

Thm. 1]:

$$R_{DF} = \max_{\pi} \min_{0 \leq t \leq M} I(X_{\pi(0:t)}; Y_{\pi(t+1)} | X_{\pi(t+1:M)}) \tag{3.40}$$

where π is a permutation on the set of nodes and $\pi(0) = 0$, $\pi(M+1) = M+1$ and $\pi(i : j) = \{\pi(i), \ldots \pi(j)\}$. $\mathbf{X}_{i:j}$ denotes the channel inputs $\mathbf{X}_{i:j} = [X_i, X_{i+1} \ldots X_j]$. For a fixed covariance matrix $\mathbf{R} = E[\mathbf{X}\mathbf{X}^T]$, it follows from the conditional maximum entropy theorem [88, Lemma 1] that all the terms in (3.40) are maximized by choosing $\mathbf{X}$ as a zero-mean Gaussian vector.

The two-hop decode-and-forward is a more constrained case of multi-hopping and it imposes a constraint on the correlation between the inputs. The rate (3.40) then reduces to the minimum of the rates on each of the two hops; the first rate, achieved on the broadcast part of the network in the first hop:

$$R_{BC} = \frac{1}{2}\log(1 + \frac{\alpha P_0}{N_0/2}) \tag{3.41}$$

and the second rate, achieved on the MAC side in the hop from the relays to the destination:

$$R_{MAC}^{(1)} = \frac{1}{2}\log\left[1 + \frac{1}{N_0/2}\left(\beta_0 P_0 + (\sum_{j \in A(P_0)} \sqrt{\beta_j P_j})^2\right)\right]. \tag{3.42}$$

$A(P_0)$ denotes the subset of relays executing the DF strategy and $\alpha = \min_{j \in A(P_0)}\{\alpha_j\}$. The channel capacity from the source to any node in $A(P_0)$ is thus higher than the code rate R_{BC} and we say that the source makes a node in $A(P_0)$ reliable. In general, the source power is split into two parts: one part is used for the transmission to the relays, and the second part is used for helping them forward a message to the destination [10]. However, in (3.41) and (3.42), the source power is used for the first goal exclusively. The reason is that the MAC rate (3.42) increases with M and thus will become higher than the rate (3.41) for sufficiently large M.

In the case of the orthogonal signaling, the MAC rate is given by

$$R_{MAC}^{(2)} = \frac{1}{2}\log\left(1 + \frac{P_0\beta_0}{N_0/2}\right) + \sum_{j \in A(P_0)} \frac{1}{2}\log\left(1 + \frac{P_j\beta_j}{N_0/2}\right). \tag{3.43}$$

As in the AF case, the difference in the two bandwidth allocations is that the signaling in the common bandwidth allows for the coherent combining of the relay signals at the destination.

The achievable rate (3.40) reduces to

$$I_{DF}^{(i)} = \min\{R_{BC}, R_{MAC}^{(i)}\}, \quad i = 1, 2. \tag{3.44}$$

The achievable rate (3.44) is bounded by the worst source-relay link and by the MAC part of the relay network that, for any $A(P_0)$, is a Gaussian vector channel [58] with relays acting as a multiple-transmit antenna transmitter. For the given powers, the maximum rate in bits/s or equivalently, the minimum energy cost per information bit in both the point-to-point and Gaussian vector channel is achieved in the limit of large W [71]. Thus, the power-efficiency of the decode-and-forward strategy is maximized in the wideband regime. This behavior was recently analyzed in [85].

3.5.1 DF Orthogonal Signaling: Optimum Power Allocation

In this case, the MAC rate is given by $R_{MAC}^{(2)}$. For the given power P_0 and the rate r at the source, relay m will be able to execute the DF strategy only if the rate r can be communicated reliably from the source to relay m with power p_0. Thus it has to hold that

$$W \log\left(1 + \frac{\alpha_m P_0}{N_0/2}\right) \geq r. \tag{3.45}$$

When constraint (3.45) is met for node m, we say that the source makes node m reliable. To optimize the transmit powers, we have to find the best subset of nodes to be made reliable so that they can decode-and-forward the message. We use binary variables x_i to indicate which relays i will be in the active set $A(p_0)$ and formulate the maximization of $I_{DF}^{(2)}$ in the following way:

$$\max \quad r = W \log\left(1 + \frac{p_0 \beta_0}{W N_0}\right) + \sum_{i=1}^{M} x_i W \log\left(1 + \frac{p_i \beta_i}{W N_0}\right) \tag{3.46}$$

$$\text{subject to } W \log\left(1 + \frac{p_0 \alpha_i}{W N_0}\right) > x_i r \tag{3.46a}$$

$$\sum_{i=0}^{M} p_i \leq p, \tag{3.46b}$$

$$x_i \in \{0, 1\}, \tag{3.46c}$$

$$p_i \geq 0. \tag{3.46d}$$

Specifically, (3.46), sets $R_{MAC}^{(2)} = r$ while (3.46a) requires that rate r be achievable at each active relay. In the limit of large W, problem (3.46) simplifies to the orthogonal wideband DF relay problem:

$$\max \quad r = p_0\beta_0 + \sum_{i=1}^{M} x_i p_i \beta_i \tag{3.47}$$

$$\text{subject to } p_0\alpha_i > x_i r, \tag{3.47a}$$

$$\sum_{i=0}^{M} p_i \leq p, \tag{3.47b}$$

$$x_i \in \{0,1\}, \tag{3.47c}$$

$$p_i \geq 0. \tag{3.47d}$$

From (3.47) we observe that in terms of the set $A(p_0) = \{i | x_i = 1\}$ of active relays,

$$r = p_0\beta_0 + \sum_{i\in A(p_0)} p_i\beta_i \leq p_0\beta_0 + \left(\sum_{i\in A(p_0)} p_i\right) \max_{i\in A(p_0)} \beta_i. \tag{3.48}$$

Moreover, this upper bound is achievable by assigning the relay power budget $\sum_{i\in A(p_0)} p_i$ to a single relay k with $\beta_k = \max_{i\in A(p_0)} \beta_i$. This observation yields the following claim.

Theorem 9 *The orthogonal wideband DF relay problem (3.47) admits an optimal solution in which no more than one relay node transmits.*

The intuition of Theorem 9 is that the relays provide a set of parallel channels to the destination and under wideband operation, transmitted power per dimension is severely restricted. Thus waterfilling this power over the relay channels results in transmission only on the best channel to the destination.

By Theorem 9, it is sufficient to consider only policies that employ a single relay k. In this case, $x_k = 1$, and $x_i = 0$ for $i \neq k$. The problem (3.47) becomes the wideband single relay problem

$$\max \quad r_k = p_0\beta_0 + p_k\beta_k \tag{3.49}$$

$$\text{subject to } p_0\alpha_k \geq r_k, \tag{3.49a}$$

$$p_0 + p_k \leq p \tag{3.49b}$$

$$p_0, p_k \geq 0. \tag{3.49d}$$

In the problem (3.49), one can show that relay k is used with power $p_k > 0$ only if $\alpha_k > \beta_0$ and $\beta_k > \beta_0$. In this case, the transmit powers are

$$p_0^* = \frac{\beta_k}{\alpha_k + \beta_k - \beta_0} p, \qquad p_k^* = \frac{\alpha_k - \beta_0}{\alpha_k + \beta_k - \beta_0} p. \tag{3.50}$$

The achieved rate normalized by the noise variance is

$$r_k^* = \frac{\alpha_k \beta_k}{\alpha_k + \beta_k - \beta_0} p. \tag{3.51}$$

We emphasize that this is the optimal power assignment for using node k as long as node k is a useful relay, in the sense that k belongs to the set of useful relays

$$U = \{i | \alpha_i > \beta_0, \beta_i > \beta_0\}. \tag{3.52}$$

Finally, among all useful relays k, we choose that one which maximizes the rate r_k^*. We summarize our observations in the following theorem.

Theorem 10 *If the set U of useful relays is non-empty, the optimal solution to the orthogonal wideband DF relay problem (3.47) is for the source to employ relay*

$$k^* = \arg\min_{k \in U} \left[\frac{1}{\alpha_k} + \frac{1}{\beta_k} - \frac{\beta_0}{\alpha_k \beta_k} \right] \tag{3.53}$$

with power assignment given by (3.50); otherwise, if U is empty, then direct transmission from the source to the destination is optimal.

3.5.2 DF Coherent Combining: Optimum Power Allocation

When the decode-and-forward relays share the bandwidth, the maximum rate problem can be formulated as

$$\max \quad r = W \log \left(1 + \frac{p_0 \beta_0}{W N_0} + \frac{\left(\sum_{i=1}^{M} \sqrt{x_i \beta_i p_i} \right)^2}{W N_0} \right) \tag{3.54}$$

$$\text{subject to } W \log \left(1 + \frac{p_0 \alpha_i}{W N_0} \right) > x_i r \tag{3.54a}$$

$$\sum_{i=0}^{M} p_i \le p, \tag{3.54b}$$

$$x_i \in \{0, 1\}, \tag{3.54c}$$

$$p_i \ge 0. \tag{3.54d}$$

In the limit of large W, this problem simplifies to the wideband DF relay problem:

$$\max \quad r = p_0\beta_0 + \left(\sum_{i=1}^{M} \sqrt{x_i\beta_i p_i}\right)^2 \tag{3.55}$$

$$\text{subject to } p_0\alpha_i > x_i r, \tag{3.55a}$$

$$\sum_{i=0}^{M} p_i \leq p, \tag{3.55b}$$

$$x_i \in \{0,1\}, \tag{3.55c}$$

$$p_i \geq 0. \tag{3.55d}$$

Any choice of source power p_0 determines a reliable set of relays $A(p_0)$, for which (3.45) is satisfied. With total power $p_r = p - p_0$ allocated to relays, problem (3.55) simplifies to determining the optimum powers $\hat{\mathbf{p}}$ of relays within the set:

$$\max_{\hat{\mathbf{p}}} \left(\sum_{i\in A(p_0)} \sqrt{\beta_i p_i}\right)^2 \tag{3.56}$$

$$\text{subject to } \sum_{i=1}^{M} p_i \leq p_r, \tag{3.56b}$$

$$\hat{\mathbf{p}} \geq 0 \tag{3.56d}$$

It is straightforward to show that the solution is in the MRC form

$$\hat{p}_i^* = \frac{\beta_i p_r}{\sum_{k\in A(p_0)} \beta_k}. \tag{3.57}$$

Thus, unlike the orthogonal DF case, each reliable relay is employed in order to contribute to the coherent combining gain. The corresponding achievable rate (3.55) is

$$r(p_0) = \beta_0 p_0 + p_r \sum_{i\in A(p_0)} \beta_i. \tag{3.58}$$

Without loss of generality, we can assume that the relay nodes are labeled such that $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_M$. Thus, if $p_0\alpha_k \geq r$, then relay nodes 1 through k will be able to decode and forward. From (3.58), the achievable rate will become

$$r(p_0) = \beta_0 p_0 + p_r c_k \tag{3.59}$$

where

$$c_k = \sum_{i=1}^{k} \beta_k. \tag{3.60}$$

From (3.55) and (3.59), the wideband relay problem reduces to

$$\max \quad r_k = \beta_0 p_0 + p_k c_k \tag{3.61}$$

$$\text{subject to} \quad p_0 \alpha_k > r_k, \tag{3.61a}$$

$$p_0 + p_k \leq p, \tag{3.61b}$$

$$p_0, p_k \geq 0. \tag{3.61d}$$

We observe that the problem (3.61) is identical to problem (3.49) with c_k replacing β_k. In this case, however, node k will not be the only transmitting relay, but rather the transmitting relay with the kth smallest link gain to the source.

Using the same reasoning as in the case of (3.49), we conclude that a set of relays $\{1, \ldots, k\}$ is employed if $\alpha_k > \beta_0$ and $c_k > \beta_0$ for a given p_0. From (3.61), we obtain the optimum powers

$$p_0^* = \frac{c_k}{\alpha_k + c_k - \beta_0} p, \qquad\qquad p_k^* = \frac{\alpha_k - \beta_0}{\alpha_k + c_k - \beta_0} p. \tag{3.62}$$

The set of useful relays in this case is given by

$$U_c = \{k | \alpha_k > \beta_0, c_k > \beta_0\}. \tag{3.63}$$

We choose relay k such that

$$k^* = \arg \min_{k \in U_c} \left[\frac{1}{\alpha_k} + \frac{1}{c_k} - \frac{\beta_0}{\alpha_k c_k} \right]. \tag{3.64}$$

3.6 Conclusion

In this chapter, we characterized the optimum bandwidth of the amplify-and forward strategy employed in two-hop, multiple-relay networks. A similar strategy has recently been proposed in [89] and shown to achieve the outage capacity of a slow-fading single-relay channel operating in the low-SNR regime. For both AF and DF used in two-hop, multiple-relay networks, we also presented the optimum power allocation at the relays.

The work presented so far focuses on single networks. Wireless networks, however, generally consist of multiple source-destination pairs in which multiple data streams can cause interference at the receivers if they are transmitted at the same time and the transmitter of one data stream is sufficiently close to the intended receiver of another data stream. In the next chapter we consider the smallest such network i.e. the *interference channel* with two source-destination pairs. We introduce transmitter cooperation to this channel model and determine the capacity region under the conditions of strong interference.

Chapter 4

The Strong Intereference Channel with Limited Transmitter Cooperation

4.1 Introduction

While in previous chapters we considered single-source networks, in this chapter we focus on the communication situation in which two separate sources wish to communicate their independent messages to two corresponding receivers. Without cooperation, this communication situation is captured by the interference channel [50, 51]. In this work, we introduce limited transmitter cooperation in the interference channel.

We start by reviewing the capacity result on the *MAC channel with with common information*, by Slepian and Wolf [3] and Willems [15], in Section 4.2. We will use this result to derive capacity result in subsequent section. In Sections 4.3 and 4.4, we assume initially that both receivers wish to decode messages sent from both encoders, corresponding to a *compound multiple-access channel* (MAC). In Section 4.3, we consider the compound MAC with both private and common messages at the encoders and present the capacity region of this channel. In Section 4.4, we adopt the cooperation model proposed by Willems [65] that assumes encoders can cooperate over side-channel links with finite capacities. We determine the capacity region of this compound multiple-access channel. When cooperating over the links with finite capacities, encoders can exchange partial information about each other's messages. This information becomes a *common* message, as it is available to both encoders after message exchange. In addition, each encoder will still have a *private* message, independent information known to that encoder only.

In Section 4.5, we relax the constraint that both private messages need to be decoded at each receiver and consider the interference channel with private and common messages at the encoders. We assume that a private message at an encoder is intended for a corresponding decoder whereas the common message is to be received at both decoders. We show that, under the strong interference conditions determined by Costa and El Gamal [56], the capacity region of this channel coincides with the capacity region of the channel in which both private messages are required at both receivers.

Finally, in Section 4.6, we analyze the interference channel with a different form of transmitter cooperation, which we refer to as *unidirectional cooperation.* We assume that messages sent at one encoder are known to the other encoder, but not vice versa. We derive conditions under which the capacity region of this channel coincides with the capacity region of the corresponding compound MAC in which both messages are decoded at both receivers [69].

4.2 Review: The Discrete Memoryless Multiaccess Channel with Common Information

Limited transmitter cooperation allows encoders to exchange only a partial description of their messages. After such partial exchange, there will be common information about both messages known at both encoders, in addition to the private information. Consequently, the capacity region of the interference channel with limited transmitter cooperation is closely related to the capacity region of the multi-access channel in which private and common messages are transmittted to a single receiver, referred to as the *MAC with common information* [3]. We next review this channel and its the capacity region as determined in [3, 15]. The channel, shown in Figure 4.1, consists of finite sets $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}$ and a conditional probability distribution $p(y|x_1, x_2)$. Symbols $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$ are channel inputs and $y \in \mathcal{Y}$ are channel outputs. Each encoder t, $t = 1, 2$, wishes to send a private message $W_t \in \{1, \dots, M_t\}$ to the receiver in N channel uses. In addition, a common message $W_0 \in \{1, \dots, M_0\}$ needs to be communicated from the encoders to the receiver, as shown in Figure 4.1. The channel is memoryless and time-invariant in the sense that

$$p(y_n|\mathbf{x}_1^n, \mathbf{x}_2^n, \mathbf{y}^{n-1}, w_0, w_1, w_2) = p_{Y|X_1X_2}(y_n|x_{1,n}, x_{2,n}). \tag{4.1}$$

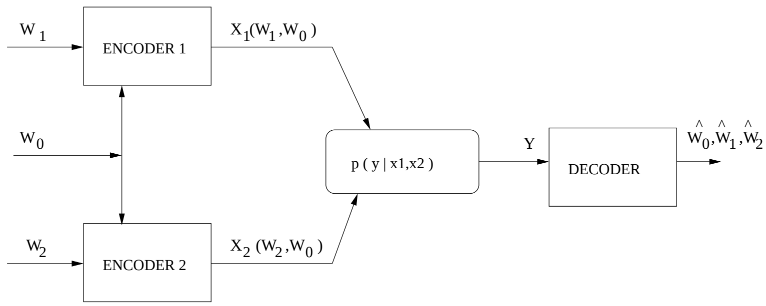

Figure 4.1: MAC with common information.

where $\mathbf{x}_t^n = \begin{bmatrix} x_{t,1}, & \ldots, & x_{t,n} \end{bmatrix}$ and where $p_{Y|X_1X_2}(\cdot)$ is the channel probability distribution. For the remainder of the chapter, we follow the convention of dropping subscripts of probability distributions if the arguments of the distributions are lower case versions of the corresponding random variables. To simplify notation, we also drop the superscript when $n = N$.

Indexes W_0, W_1 and W_2 are independently generated at the beginning of each block of N channel uses. An encoder $t, t = 1, 2$ maps the common message W_0 and the private message W_t into a codeword $\mathbf{x}_t$

$$\mathbf{x}_1 = f_1(W_0, W_1) \tag{4.2}$$

$$\mathbf{x}_2 = f_2(W_0, W_2). \tag{4.3}$$

The decoder estimates the common message W_0 and the private messages W_t based on the received N-sequence $\mathbf{Y}$ as

$$(\hat{W}_0, \hat{W}_1, \hat{W}_2) = g(\mathbf{Y}). \tag{4.4}$$

An (M_0, M_1, M_2, N, P_e) code for the channel consists of two encoding functions f_1, f_2, decoding function g, and an error probability

$$P_e = \sum_{(w_0,w_1,w_2)} \frac{1}{M_0 M_1 M_2} P\left[g(\mathbf{Y}) \neq (w_0, w_1, w_2) | (w_0, w_1, w_2) \text{ sent}\right]. \tag{4.5}$$

A rate triple (R_0, R_1, R_2) is achievable if, for any $\epsilon > 0$, there is an (M_0, M_1, M_2, N, P_e) code such that

$$P_e \leq \epsilon \text{ and } M_i \geq 2^{NR_i} \quad i = 0, 1, 2.$$

The capacity region of the *MAC with common information* is the closure of the set of all achievable rate triplets (R_0, R_1, R_2). The capacity region of this channel, $\mathcal{C}_{\mathrm{MAC}}$, was shown in [3, 15] to be

$$\begin{aligned}\mathcal{C}_{MAC} = \bigcup\{(R_0, R_1, R_2) :\\ &0 \leq R_1 \leq I(X_1; Y|X_2, U)\\ &0 \leq R_2 \leq I(X_2; Y|X_1, U)\\ &R_1 + R_2 \leq I(X_1, X_2; Y|U)\\ &0 \leq R_0 + R_1 + R_2 \leq I(X_1, X_2; Y)\}\end{aligned} \tag{4.6}$$

where the union is over all $p(u, x_1, x_2, y)$ that factor as $p(u)p(x_1|u)p(x_2|u)p(y|x_1, x_2)$. (In [15] the convex hull operation used in [3] was shown to be unnecessary).

4.3 The Discrete Memoryless Compound Multiaccess Channel with Common Information

If in the MAC with common information, instead of one there are *two* receivers that wish to decode both private messages and a common message, the resulting channel becomes a *compound MAC with common information*, denoted $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_1, y_2|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$. A code for the channel is given by encoding functions (4.2)-(4.3) and two decoding functions

$$(\hat{W}_0, \hat{W}_1, \hat{W}_2) = g_t(\mathbf{Y}_t), \qquad t = 1, 2. \tag{4.7}$$

The error probability of the code is

$$\begin{aligned}P_e = \sum_{(w_0, w_1, w_2)} \frac{1}{M_0 M_1 M_2} P\left[\{g_1(\mathbf{Y}_1) \neq (w_0, w_1, w_2)\}\right.\\ \left.\cup \{g_2(\mathbf{Y}_2) \neq (w_0, w_1, w_2)\} \,|(w_0, w_1, w_2) \text{ sent}\right].\end{aligned} \tag{4.8}$$

As in the case of the MAC with common information, a rate triple (R_0, R_1, R_2) is achievable if, for any $\epsilon > 0$, there is an (M_0, M_1, M_2, N, P_e) code such that

$$P_e \leq \epsilon \text{ and } M_i \geq 2^{NR_i} \quad i = 0, 1, 2.$$

The capacity region of the *Compound MAC with common information* is the closure of the set of all achievable rate triplets (R_0, R_1, R_2).

This channel defines two multiple-access channels with common information, MAC_t $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_t|x_1, x_2), \mathcal{Y}_t)$, one for each receiver t where

$$p(y_1|x_1, x_2) = \sum_{y_2 \in \mathcal{Y}_2} p(y_1, y_2|x_1, x_2) \tag{4.9}$$

$$p(y_2|x_1, x_2) = \sum_{y_1 \in \mathcal{Y}_1} p(y_1, y_2|x_1, x_2). \tag{4.10}$$

An $(M_0, M_1, M_2, N, P_{et})$ code for MAC_t, $t = 1, 2$ is then given by the encoding functions (4.2)-(4.3), a decoding function $g_t(\cdot)$, (4.7), and the error probability as in (4.5). Specifically,

$$P_{et} = \sum_{(w_0,w_1,w_2)} \frac{1}{M_0 M_1 M_2} P\left[g_t(\mathbf{Y}_t) \neq (w_0, w_1, w_2)|(w_0, w_1, w_2) \text{ sent}\right], \quad t = 1, 2 \tag{4.11}$$

where $(w_0, w_1, w_2) \in \mathcal{W}_0 \times \mathcal{W}_1 \times \mathcal{W}_2$.

The encoding and decoding strategy proposed by Willems in [15] can be adapted for the compound MAC with common information to guarantee the achievability of rates

$$\mathcal{C}_{\text{CMAC}} = \bigcup \{\mathcal{R}_{\text{MAC1}} \cap \mathcal{R}_{\text{MAC2}}\} \tag{4.12}$$

where $\mathcal{R}_{\text{MACt}}$, $t = 1, 2$ satisfies the bounds (4.6) with Y replaced by Y_t, and the union is over all $p(u)p(x_1|u)p(x_2|u)p(y_1, y_2|x_1, x_2)$. More specifically, the following theorem gives the capacity region of the compound MAC with common information.

Theorem 11 *The capacity region of the compound multiple access channel* $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_1, y_2|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$, $\mathcal{C}_{CMAC}$, *is given by*

$$\begin{aligned}
\mathcal{C}_{CMAC} = \bigcup \{ & (R_0, R_1, R_2) : \\
& 0 \leq R_1 \leq \min\{I(X_1; Y_1|X_2, U), I(X_1; Y_2|X_2, U)\} \\
& 0 \leq R_2 \leq \min\{I(X_2; Y_1|X_1, U), I(X_2; Y_2|X_1, U)\} \\
& R_1 + R_2 \leq \min\{I(X_1, X_2; Y_1|U), I(X_1, X_2; Y_2|U)\} \\
& 0 \leq R_0 + R_1 + R_2 \leq \min\{I(X_1, X_2; Y_1), I(X_1, X_2; Y_2)\}
\end{aligned} \tag{4.13}$$

where the union is over all joint distributions that factor as $p(u)p(x_1|u)p(x_2|u)p(y_1, y_2|x_1, x_2)$.

4.3.1 Converse

Consider an (M_0, M_1, M_2, N, P_e) code for the $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_1, y_2|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$ compound MAC with common information. By comparing (4.11) with (4.8), we conclude that

$$\max\{P_{e1}, P_{e2}\} \leq P_e. \tag{4.14}$$

From (4.14), the necessary condition for $P_e \to 0$ is that $P_{e1} \to 0$ and $P_{e2} \to 0$. Since each $P_{et}, t = 1, 2$ is the error probability in the MAC with a single receiver, it follows from (4.6) that to guarantee that $P_{et} \to 0$ for $t = 1, 2$, the rates (R_0, R_1, R_2) have to belong to the $\mathcal{C}_{\text{MAC}}$ region (4.6) for each receiver t. That is, these rates have to satisfy

$$\begin{aligned} R_1 &\leq I(X_1; Y_t|X_2, U) \\ R_2 &\leq I(X_2; Y_t|X_1, U) \\ R_1 + R_2 &\leq I(X_1, X_2; Y_t|U) \\ R_0 + R_1 + R_2 &\leq I(X_1, X_2; Y_t) \end{aligned} \tag{4.15}$$

for a joint distribution $p(u)p(x_1|u)p(x_2|u)p(y_t|x_1, x_2)$. $\square$

A more detailed proof can be obtained starting from Fano's inequality and following the converse proof in [15]. We will present this approach when proving the converse for the Compound MAC with conferencing in Section 4.4.1.

We can show that the region of Theorem 11 is convex. In particular, we define a region

$$\begin{aligned} \mathcal{R}^8_{\text{CMAC}} \triangleq \{(R_1, \ldots, R_8), t = 1, 2 :& \\ R_{t+2k} \leq I(X_1; Y_t|X_2, U) & \quad k = 0 \\ R_{t+2k} \leq I(X_2; Y_t|X_1, U) & \quad k = 1 \\ R_{t+2k} \leq I(X_1, X_2; Y_t|U) & \quad k = 2 \\ R_{t+2k} \leq I(X_1, X_2; Y_t) & \quad k = 3\} \end{aligned} \tag{4.16}$$

for $p(u)p(x_1|u)p(x_2|u)p(y_1, y_2|x_1, x_2)$.

Applying the approach in [15, Appendix A] we can show that the region $\mathcal{R}^8_{\text{CMAC}}$ is convex. For completeness, we next present the proof.

Theorem 12 *Region $\mathcal{R}^8_{CMAC}$ is convex.*

Proof:

Assume that there are two points that belong to the region $\mathcal{R}^8_{\text{CMAC}}$. That is, $(R^1_1, \ldots, R^1_8) \in \mathcal{R}^8_{\text{CMAC}}$ and $(R^2_1, \ldots, R^2_8) \in \mathcal{R}^8_{\text{CMAC}}$. (4.16) implies that there exist distributions $p(u^1, x^1_1, x^1_2, y^1_1, y^1_2) = p(u^1)p(x^1_1|u^1)p(x^1_2|u^1)p(y^1_1, y^1_2|x^1_1, x^1_2)$ and $p(u^2, x^2_1, x^2_2, y^2_1, y^2_2) = p(u^2)p(x^2_1|u^2)p(x^2_2|u^2)p(y^2_1, y^2_2|x^2_1, x^2_2)$ such that

$$
\begin{aligned}
R^1_{t+2k} &\le I(X^1_1; Y^1_t|X^1_2, U^1) & R^2_{t+2k} &\le I(X^2_1; Y^2_t|X^2_2, U^2) & k &= 0\\
R^1_{t+2k} &\le I(X^1_2; Y^1_t|X^1_1, U^1) & R^2_{t+2k} &\le I(X^1_2; Y^2_t|X^2_1, U^2) & k &= 1\\
R^1_{t+2k} &\le I(X^1_1, X^1_2; Y^1_t|U^1) & R^2_{t+2k} &\le I(X^2_1, X^2_2; Y^2_t|U^2) & k &= 2\\
R^1_{t+2k} &\le I(X^1_1, X^1_2; Y^1_t) & R^2_{t+2k} &\le I(X^2_1, X^2_2; Y^2_t) & k &= 3
\end{aligned}
\tag{4.17}
$$

Let a random variable I have a distribution $P[I = 1] = \alpha$ and $P[I = 2] = 1 - \alpha$. We define new random variables $X_t \triangleq X^I_t$, $Y_t \triangleq Y^I_t$ and $U \triangleq (U^I, I)$ and observe that

$$
\begin{aligned}
p(u,&x_1, x_2, y_1, y_2)\\
&= P(U^i = u', X^i_1 = x_1, X^i_2 = x_2, Y^i_1 = y_1, Y^i_2 = y_2|I = i)P(I = i)\\
&= P(U^i = u'|I = i)P(X^i_1 = x_1, X^i_2 = x_2|U^i = u', I = i)\\
&\qquad P(Y^i_1 = y_1, Y^i_2 = y_2|X^i_1 = x_1, X^i_2 = x_2, U^i = u', I = i)P(I = i)\\
&= P(U = u)P(X^I_1 = x_1|U = u)P(X^I_2 = x_2, |U = u)\\
&\qquad P(Y^I_1 = y_1, Y^I_2 = y_2|X^I_1 = x_1, X^I_2 = x_2)\\
&= p(u)p(x_1|u)p(x_2|u)p(y_1, y_2|x_1, x_2)
\end{aligned}
\tag{4.18}
$$

where $u = (u', i)$. It is now easy to show that

$$
\begin{aligned}
\alpha R^1_{t+2k} + (1-\alpha)R^2_{t+2k} &\le I(X_1; Y_t|X_2, U) & k &= 0\\
\alpha R^1_{t+2k} + (1-\alpha)R^2_{t+2k} &\le I(X_2; Y_t|X_1, U) & k &= 1\\
\alpha R^1_{t+2k} + (1-\alpha)R^2_{t+2k} &\le I(X_1, X_2; Y_t|U) & k &= 2\\
\alpha R^1_{t+2k} + (1-\alpha)R^2_{t+2k} &\le I(X_1, X_2; Y_t) & k &= 3
\end{aligned}
\tag{4.19}
$$

and therefore $\big(\alpha R^1_1 + (1-\alpha)R^2_1, \ldots, \alpha R^1_8 + (1-\alpha)R^2_8\big) \in \mathcal{R}^8_{\text{CMAC}}$. $\square$

4.3.2 Achievability

We adapt the encoding and decoding strategy proposed by Willems in [15] to achieve the rates (4.13). Specifically, we use the codebook in [15, Section 3] constructed as follows:

1. Fix the distribution $p(u, x_1, x_2) = p(u)p(x_1|u)p(x_2|u)$.
2. Generate M_0 sequences $\mathbf{u}$ each with probability $p(\mathbf{u}) = \prod_{n=1}^{N} p(u_n)$. Label them $\mathbf{u}(w_0), w_0 \in \{1, \ldots, M_0\}$.
3. For each $\mathbf{u}(w_0)$, generate M_t sequences $\mathbf{x}_t$ with probability $P(\mathbf{x}_t|\mathbf{u}) = \prod_{n=1}^{N} p(x_{tn}|u_n)$ where $t = 1, 2$. Label them $\mathbf{x}_t(w_0, w_t), w_t \in \{1, \ldots, M_t\}$.

Encoding: To send a common message w_0 and a private message w_t encoder t sends the codeword $\mathbf{x}_t(w_0, w_t)$.

Decoding: At each decoder, we use the decoding scheme of [15]: After receiving $\mathbf{y}_t$, decoder t determines unique $(\hat{w}_0, \hat{w}_1, \hat{w}_2)$ such that

$$(\mathbf{u}(\hat{w}_0), \mathbf{x}_1(\hat{w}_0, \hat{w}_1), \mathbf{x}_2(\hat{w}_0, \hat{w}_2), \mathbf{Y}_t) \in A_\epsilon(\mathbf{U}, \mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_t)$$

where $A_\epsilon(\mathbf{U}, \mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_t)$ is the set of ϵ-typical N-sequences $(\mathbf{u}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{y}_t)$ as defined in [81, Section 14.2].

The probability of error: We apply the union bound to (4.8) to obtain

$$P_e \leq P_{e1} + P_{e2} \tag{4.20}$$

where P_{e1} and P_{e2} are given by (4.11). It was shown in [15] that P_{e1} and P_{e2} can made arbitrarily small when the rates satisfy (4.13). From (4.20) it then follows that the probability of error P_e can be made arbitrarily small. □

4.4 The Discrete Memoryless Compound Multiaccess Channel with Conferencing Encoders

In this section, we now assume that in the two-source, two-destination network under consideration, there exist two communication links with known capacities between the

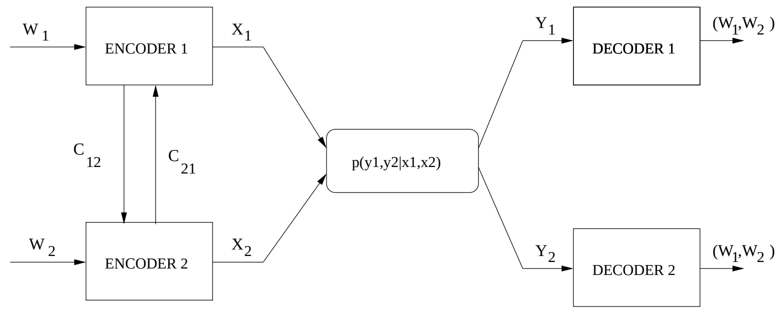

Figure 4.2: Compound MAC with conferencing encoders.

two encoders, allowing them to partially cooperate to send their intended messages. This model was introduced by Willems for multi-access channels [65]. The amount of information exchanged between the transmitters is bounded by the capacities C_{12} and C_{21} of the communication links. The proposed channel model enables investigation of the gains obtained using transmitter cooperation. The communications system is shown in Figure 4.2. As in the previous section, it is assumed that both messages are required to be decoded at both receivers. We refer to this channel as a *compound multiple access channel with conferencing encoders* and denote it $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_1, y_2|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$. We next determine the capacity region of the compound MAC with conferencing encoders. We show that for an input distribution with a specific Markov property, the rate region is an intersection of two rate regions of the MAC with partially cooperating encoders [65]. The capacity region is the union of all such rate regions. For $C_{12} = C_{21} = 0$, it reduces to the capacity region of the two-sender, two-receiver channel with non-cooperating encoders [57].

The channel consists of finite sets $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2$ and a conditional probability distribution $p(y_1, y_2|x_1, x_2)$. Symbols $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$ are channel inputs and $(y_1, y_2) \in \mathcal{Y}_1 \times \mathcal{Y}_2$ are corresponding channel outputs. Each encoder t, $t = 1, 2$, wishes to send a message $W_t \in \{1, \ldots, M_t\}$ to both receivers in N channel uses. The channel is memoryless and time-invariant in the sense that

$$p(y_{1,n}, y_{2,n}|\mathbf{x}_1^n, \mathbf{x}_2^n, \mathbf{y}_1^{n-1}, \mathbf{y}_2^{n-1}) = p(y_{1,n}, y_{2,n}|x_{1,n}, x_{2,n}). \tag{4.21}$$

Encoders use the communication links in the form of a *conference* [65]. A conference is given by two sets of K communicating functions $\{h_{t,1}, \quad \ldots, \quad h_{t,K}\}$, $t = 1, 2$. Each function $h_{t,k}$ maps the message W_t and the sequence of previously received communications from the other encoder into the kth communication $V_{t,k}$, where $V_{t,k}$ ranges over a finite alphabet $\mathcal{V}_{t,k}$, for $k = 1, \ldots, K$,

$$h_{1,k} : \mathcal{W}_1 \times \mathcal{V}_2^{k-1} \to \mathcal{V}_{1,k}, \quad v_{1,k} = h_{1,k}(W_1, V_2^{k-1}) \tag{4.22}$$

$$h_{2,k} : \mathcal{W}_2 \times \mathcal{V}_1^{k-1} \to \mathcal{V}_{2,k}, \quad v_{2,k} = h_{2,k}(W_2, V_1^{k-1}). \tag{4.23}$$

The amount of information that can be exchanged during the conference is bounded by the capacities C_{12} and C_{21}. A conference is (C_{12}, C_{21})-permissible if

$$\sum_{k=1}^{K} \log(|\mathcal{V}_{1,k}|) \leq NC_{12} \tag{4.24}$$

$$\sum_{k=1}^{K} \log(|\mathcal{V}_{2,k}|) \leq NC_{21}. \tag{4.25}$$

An encoding function f_t maps the message W_t and what was learned from the conference into a codeword $\mathbf{x}_t$. An (M_1, M_2, N, K, P_e) code for the channel consists of two sets of K communicating functions (4.22)-(4.23), two encoding functions

$$f_1 : \mathcal{W}_1 \times \mathcal{V}_2^K \to \mathcal{X}_1^N \tag{4.26}$$

$$f_2 : \mathcal{W}_2 \times \mathcal{V}_1^K \to \mathcal{X}_2^N \tag{4.27}$$

generating codewords

$$\mathbf{x}_1 = f_1(W_1, V_2^K) \tag{4.28}$$

$$\mathbf{x}_2 = f_2(W_2, V_1^K) \tag{4.29}$$

and two decoding functions

$$(\hat{W}_1, \hat{W}_2) = g_t(\mathbf{Y}_t) \quad t = 1, 2 \tag{4.30}$$

such that the average probability of error of the code is

$$\begin{aligned} P_e = \sum_{(w_1, w_2) \in \mathcal{W}_1 \times \mathcal{W}_2} & \frac{1}{M_1 M_2} P\left[\left\{g_1(Y_1^N) \neq (w_1, w_2)\right\}\right. \\ & \left. \cup \left\{g_2(Y_2^N) \neq (w_1, w_2)\right\} |(w_1, w_2) \text{ sent}\right]. \end{aligned} \tag{4.31}$$

A rate pair (R_1, R_2) is achievable if, for any $\epsilon > 0$, there exists an (M_1, M_2, N, K, P_e) code such that

$$M_t \geq 2^{NR_t} \quad t = 1, 2, \text{ and } P_e \leq \epsilon. \tag{4.32}$$

The capacity region of the *compound multiple access channel with conferencing encoders* is the closure of the set of all achievable rate pairs (R_1, R_2).

The next theorem presents the capacity region of the compound multiple access channel with conferencing encoders. For an input distribution with a specific Markov property, the rate region is an intersection of two rate regions of the MAC with partially cooperating encoders [65]. The capacity region is the union of all such rate regions.

Theorem 13 *For the compound multiple access channel $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_1, y_2|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$ with communication links with capacities C_{12} and C_{21} the capacity region $\mathcal{R}(C_{12}, C_{21})$ is given by*

$$\begin{aligned}
\mathcal{R}(C_{12}, C_{21}) = \bigcup \{ & (R_1, R_2) : \\
& R_1 \leq \min\{I(X_1; Y_1|X_2, U), I(X_1; Y_2|X_2, U)\} + C_{12} \\
& R_2 \leq \min\{I(X_2; Y_1|X_1, U), I(X_2; Y_2|X_1, U)\} + C_{21} \\
& R_1 + R_2 \leq \min\{I(X_1, X_2; Y_1|U), I(X_1, X_2; Y_2|U)\} + C_{12} + C_{21} \\
& R_1 + R_2 \leq \min\{I(X_1, X_2; Y_1), I(X_1, X_2; Y_2)\}
\end{aligned} \tag{4.33}$$

where the union is over all joint distributions that factor as

$$p(u)p(x_1|u)p(x_2|u)p(y_1, y_2|x_1, x_2).$$

4.4.1 Converse

The exact same reasoning as in the converse proof of Theorem 11 applies in this case.

Alternatively, we can obtain the same result by applying Fano's inequality to the message estimate $(\hat{W}_1, \hat{W}_2)$ at each receiver. We then use the approach of Willems [65, Section III] to obtain

$$R_1 \leq \frac{1}{N} \sum_{n=1,N} I(X_{1n}; Y_{tn}|X_{2n}, U_n) + C_{12} + \epsilon_{tN}$$
$$R_2 \leq \frac{1}{N} \sum_{n=1,N} I(X_{2n}; Y_{tn}|X_{1n}, U_n) + C_{21} + \epsilon_{tN}$$
$$R_1 + R_2 \leq \frac{1}{N} \sum_{n=1,N} I(X_{1n}, X_{2n}; Y_{tn}|U_n) + C_{12} + C_{21} + \epsilon_{tN}$$
$$R_1 + R_2 \leq \frac{1}{N} \sum_{n=1,N} I(X_{1n}, X_{2n}; Y_{tn}) + \epsilon_{tN}. \tag{4.34}$$

We proceed as in [65] and show that the region of Theorem 13 is convex. We define a region

$$\begin{aligned} \mathcal{R}^8 \triangleq \{(R_1, \ldots, R_8), t = 1, 2 : \\ R_{t+2k} \leq I(X_1; Y_t|X_2, U) + C_{12} \qquad & k = 0 \\ R_{t+2k} \leq I(X_2; Y_t|X_1, U) + C_{21} \qquad & k = 1 \\ R_{t+2k} \leq I(X_1, X_2; Y_t|U) + C_{12} + C_{21} \quad & k = 2 \\ R_{t+2k} \leq I(X_1, X_2; Y_t) \qquad & k = 3\} \end{aligned} \tag{4.35}$$

for $p(u)p(x_1|u)p(x_2|u)p(y_1, y_2|x_1, x_2)$. Applying the approach in [15, Appendix A] we can show that the region $\mathcal{R}^8$ is convex, and thus rates (4.34) belong to the region $\mathcal{R}^8$. From the definition of $\mathcal{R}^8$ and $\mathcal{R}$ in (4.35) and (4.33), it then follows that rates (R_1, R_2) belong to $\mathcal{R}(C_{12}, C_{21})$. □

4.4.2 Achievability

In a conference, transmitters cooperate over the communication links with capacities C_{12} and C_{21} using the strategy proposed by Willems [65, Sec. IV]. Specifically, the set $\{1, \ldots, M_1\}$ is partitioned into $2^{NR_{12}}$ cells, labeled $s_1 \in \{1, \ldots, 2^{NR_{12}}\}$, each with $2^{N(R_1 - R_{12})}$ elements labeled $t_1 \in \{1, \ldots, 2^{N(R_1 - R_{12})}\}$. When w_1 belongs to cell s_1, we let $c_1(w_1) = s_1$. The same type of partitioning is done for messages W_2. Rates R_{12}, R_{21} are chosen such that $R_{12} = \min\{R_1, C_{12}\} \leq C_{12}$, $R_{21} = \min\{R_2, C_{21}\} \leq C_{21}$ ensuring that during the communication, the partial information

$$w_0' = (w_{01}', w_{02}') = (c_1(w_1), c_2(w_2))$$

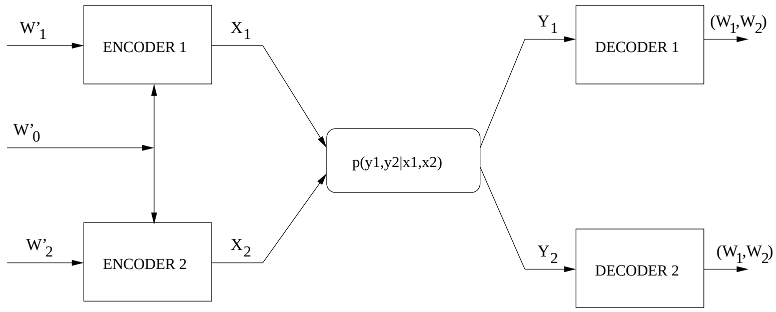

Figure 4.3: After the conference, the compound MAC with conferencing encoders becomes the compound MAC with common information.

$(c_1(w_1), c_2(w_2)) \in \{1, \ldots, 2^{NR_{12}}\} \times \{1, \ldots, 2^{NR_{21}}\}$ can be exchanged between the encoders. We refer to W_0' as a *common* message.

The part of the original message unknown to the other encoder is given by

$$w_1' = t_1(w_1) \in \{1, \ldots, 2^{N(R_1 - R_{12})}\}$$

$$w_2' = t_2(w_2) \in \{1, \ldots, 2^{N(R_2 - R_{21})}\}.$$

The obtained system is shown in Figure 4.3. Thus, after the conference, coding has to be done for a common message with alphabet $\mathcal{W}_0'$

$$w_0' \in \{1, \ldots, M_0'\} \tag{4.36}$$

and private messages

$$w_1' \in \{1, \ldots, M_1'\} \tag{4.37}$$

$$w_2' \in \{1, \ldots, M_2'\} \tag{4.38}$$

with corresponding alphabets $\mathcal{W}_1'$ and $\mathcal{W}_2'$. We use the notation $M_0' = 2^{N(R_{12}+R_{21})}$, $M_1' = 2^{N(R_1 - R_{12})}$ and $M_2' = 2^{N(R_2 - R_{21})}$.

In the case of a single receiver, the MAC channel after the conference reduces to a MAC with common and private messages at the encoders, [65]. The capacity region of this channel (4.6) then guarantees that the rates for the MAC with conferencing in [65, Sec. II.] are achievable. Similarly, the compound MAC with conferencing after the conference is identical to the compound MAC with common information. The common message W_0' and private messages W_1' and W_2' are shown in Figure 4.3. From

Equation (4.36), we observe that the common rate is $R_{12}+R_{21}$. From Equations (4.37)-(4.38), the private rates are $R_1 - R_{12}$ and $R_2 - R_{21}$. It follows from Theorem 11 that these rates $(R_{12} + R_{21}, R_1 - R_{12}, R_2 - R_{21})$ are achievable as long they belong to the region given by (4.13). This guarantees that rates (4.33) in Theorem 13 are achievable. $\square$

4.4.3 Implications

For $C_{12} = C_{21} = 0$, the capacity region (4.33) of the compound MAC with conferencing encoders becomes the capacity region of the two-sender, two-receiver channel established by Ahlswede [57]. Rates (4.33) qualify the improvement due to transmitter cooperation over the dedicated communication links with capacities C_{12} and C_{21}.

Furthermore, the rates (4.33) give inner bounds on the rates achievable in an interference channel in which users partially cooperate and each decoder decodes a message sent from a single encoder. It would be interesting to characterize the class of interference channels for which these rates in fact give the capacity region.

Finally, we apply (4.33) to a Gaussian network with channel outputs

$$y_{1i} = x_{1i} + h_{21}x_{2i} + z_{1i} \tag{4.39}$$

$$y_{2i} = h_{12}x_{1i} + x_{2i} + z_{2i} \tag{4.40}$$

where Z_t is zero-mean, unit-variance noise. The code definition is the same as that given in Section 4.4 with the addition of the power constraints

$$\frac{1}{N}\sum_{i=1}^{N} E[X_{ti}^2] \leq P_t, \qquad t = 1, 2. \tag{4.41}$$

The power expended for the conference is thus not considered. We have the following.

Corollary 1 *The capacity region of the Gaussian compound MAC with conferencing*

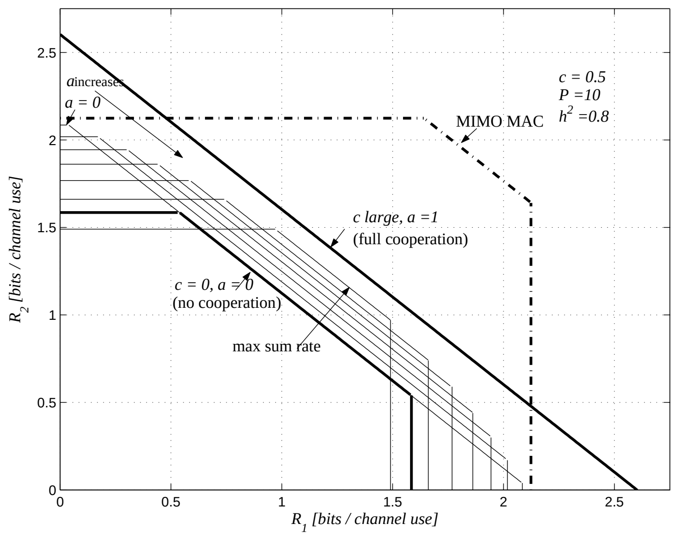

Figure 4.4: The Gaussian Compound MAC with conferencing encoders capacity region.

encoders is given by

$$\mathcal{R}(C_{12}, C_{21}) = \bigcup \{(R_1, R_2) :$$

$$0 \leq R_1 \leq \min \left\{ C(\bar{a}P_1), C(h_{12}^2 \bar{a}P_1) \right\} + C_{12} \tag{4.42}$$

$$0 \leq R_2 \leq \min \left\{ C(\bar{b}P_2), C(h_{21}^2 \bar{b}P_2) \right\} + C_{21} \tag{4.43}$$

$$R_1 + R_2 \leq \min \left\{ C(\bar{a}P_1 + h_{21}^2 \bar{b}P_2), C(h_{12}^2 \bar{a}P_1 + \bar{b}P_2) \right\} + C_{12} + C_{21} \tag{4.44}$$

$$R_1 + R_2 \leq \min \Big\{ C(P_1 + h_{21}^2 P_2 + 2h_{21}\sqrt{aP_1 bP_2}), \\ C(h_{12}^2 P_1 + P_2 + 2h_{12}\sqrt{aP_1 bP_2}) \Big\} \} \tag{4.45}$$

where the union is over all a, b, *for* $0 \leq a \leq 1, 0 \leq b \leq 1$ *and* $\bar{a} = 1 - a$, $\bar{b} = 1 - b$.

Figure 4.4 shows the capacity region for a symmetric case where $C_{12} = C_{21} = c = 0.5$, $P_1 = P_2 = P = 10$, $N_1 = N_2 = 1$, $h_{12} = h_{21} = h = \sqrt{0.8}$. Due to the symmetry, we choose $a = b$. To illustrate the cooperation benefit, also shown are the rates achievable when there is no cooperation ($c = 0$). Furthermore, for $c = 0$ we also consider perfect *receiver* cooperation in which case the considered channel becomes the MIMO MAC with two receive antennas and a single antenna at each of the two transmitters. We plot the capacity region of the MIMO MAC. We observe much higher receiver cooperation gains as compared to the gains from transmitter cooperation.

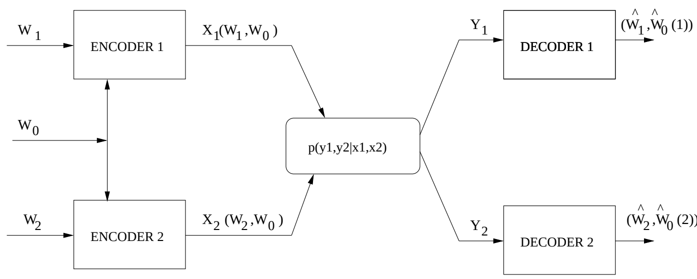

Figure 4.5: Interference channel with common information.

The bounds (4.42), (4.43) and (4.44) are maximized for $a = 0$. As a increases, these bounds decrease, but the bound on the sum rate (4.45) increases. The sum rate is maximized when a is chosen such that (4.44) and (4.45) are the same. The capacity region is the union of all the pentagons obtained for different values of a.

4.5 The Capacity Region of the Strong Interference Channel with Common Information

Consider again the communication situation in which two encoders each have a private message and a common message they wish to send. In this section, we relax the constraints of Section 4.3 that both receivers have to decode both private messages. Instead, we assume that each decoder t is interested in only one private message sent at the corresponding encoder t. Both decoders wish to decode the common message. We refer to this channel as *an interference channel with common information*, denoted $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_1, y_2|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$. The communication system is shown in Figure 4.5.

We determine the capacity region of interference channels with a common message if

$$I(X_1; Y_1|X_2, U) \leq I(X_1; Y_2|X_2, U) \tag{4.46}$$

$$I(X_2; Y_2|X_1, U) \leq I(X_2; Y_1|X_1, U) \tag{4.47}$$

for all joint distributions $p(u, x_1, x_2, y_1, y_2)$ that factor as $p(u)p(x_1|u)p(x_2|u)p(y_1, y_2|x_1, x_2)$. We further show that this class of interference channels is the same as those determined by (1.1) and (1.2) with independent X_1 and X_2.

As in Section 4.3, the channel consists of finite sets $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2$ and a conditional probability distribution $p(y_1, y_2|x_1, x_2)$. Each encoder t, $t = 1, 2$, wishes to send a private message $W_t \in \{1, \ldots, M_t\}$ to receiver t in N channel uses. In addition, a common message $W_0 \in \{1, \ldots, M_0\}$ needs to be communicated from the encoders to both decoders, as shown in Figure 4.5. The channel is memoryless and time-invariant as given by Equation (4.21).

Indexes W_0, W_1 and W_2 are independently generated at the beginning of each block of N channel uses. An encoder $t, t = 1, 2$ maps the common message W_0 and the private message W_t into a codeword $\mathbf{x}_t$

$$\mathbf{x}_1 = f_1(W_0, W_1) \tag{4.48}$$

$$\mathbf{x}_2 = f_2(W_0, W_2). \tag{4.49}$$

Each decoder t estimates the common message W_0 and the private message W_t based on the received N-sequence $\mathbf{Y}_t$ as

$$(\hat{W}_0, \hat{W}_1) = g_1(\mathbf{Y}_1) \tag{4.50}$$

$$(\hat{W}_0, \hat{W}_2) = g_2(\mathbf{Y}_2). \tag{4.51}$$

An (M_0, M_1, M_2, N, P_e) code for the channel consists of two encoding functions f_1, f_2, two decoding functions g_1, g_2 and a maximum error probability

$$P_e \triangleq \max\{P_{e,1}, P_{e,2}\} \tag{4.52}$$

where

$$P_{e,t} = \sum_{(w_0,w_1,w_2)} \frac{1}{M_0 M_1 M_2} P\left[g_t(\mathbf{Y}_t) \neq (w_0, w_t) | (w_0, w_1, w_2) \text{ sent}\right], \quad t = 1, 2. \tag{4.53}$$

A rate triple (R_0, R_1, R_2) is achievable if, for any $\epsilon > 0$, there is an (M_0, M_1, M_2, N, P_e) code such that

$$P_e \leq \epsilon \text{ and } M_i \geq 2^{NR_i} \quad i = 0, 1, 2.$$

The capacity region of the *interference channel with common information* is the closure of the set of all achievable rate triplets (R_0, R_1, R_2).

The next theorem is the main result of this section. It gives conditions under which the capacity region coincides with the capacity region of the channel in which both private messages are required at both receivers.

Theorem 14 *For an interference channel* $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_1, y_2|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$ *with common information satisfying the strong interference conditions [56]*

$$I(X_1; Y_1|X_2) \leq I(X_1; Y_2|X_2) \tag{4.54}$$

$$I(X_2; Y_2|X_1) \leq I(X_2; Y_1|X_1) \tag{4.55}$$

for all joint distributions $p(x_1, x_2, y_1, y_2)$ *that factor as* $p(x_1)p(x_2)p(y_1, y_2|x_1, x_2)$ *the capacity region* $\mathcal{C}$ *is given by*

$$\mathcal{C} = \bigcup \{(R_0, R_1, R_2) :$$

$$0 \leq R_1 \leq I(X_1; Y_1|X_2, U) \tag{4.56}$$

$$0 \leq R_2 \leq I(X_2; Y_2|X_1, U) \tag{4.57}$$

$$R_1 + R_2 \leq \min\{I(X_1, X_2; Y_1|U), I(X_1, X_2; Y_2|U)\} \tag{4.58}$$

$$0 \leq R_0 + R_1 + R_2 \leq \min\{I(X_1, X_2; Y_1), I(X_1, X_2; Y_2)\} \tag{4.59}$$

where the union is over all joint distributions that factor as

$$p(u)p(x_1|u)p(x_2|u)p(y_1, y_2|x_1, x_2). \tag{4.60}$$

When the constraints (4.54) and (4.55) are satisfied, we refer to the considered channel as a *strong interference channel with common information.* We next prove the converse.

4.5.1 Converse: Strong Interference Conditions

Consider a code (M_0, M_1, M_2, N, P_e) for the interference channel with common information. Applying Fano's inequality results in

$$H(W_0, W_1|\mathbf{Y}_1) \leq P_{e1} \log(M_0 M_1 - 1) + h(P_{e1}) \triangleq N\delta_{1,N} \tag{4.61}$$

$$H(W_0, W_2|\mathbf{Y}_2) \leq P_{e2} \log(M_0 M_2 - 1) + h(P_{e2}) \triangleq N\delta_{2,N}. \tag{4.62}$$

where $\delta_{t,N} \to 0$ as $P_{et} \to 0$ (or as $P_e \to 0$). It follows that

$$H(W_0, W_1|\mathbf{Y}_1) = H(W_0|\mathbf{Y}_1) + H(W_1|\mathbf{Y}_1, W_0) \leq N\delta_{1,N} \tag{4.63}$$

$$H(W_0, W_2|\mathbf{Y}_2) = H(W_0|\mathbf{Y}_2) + H(W_2|\mathbf{Y}_2, W_0) \leq N\delta_{2,N}. \tag{4.64}$$

Since conditioning cannot increase entropy, from (4.64) it follows that

$$H(W_2|\mathbf{Y}_2, W_0, W_1) \leq H(W_2|\mathbf{Y}_2, W_0) \leq N\delta_{2,N}. \tag{4.65}$$

To prove the converse, we will use the data processing inequality for the following Markov chains:

Lemma 3 *The following form Markov chains for the interference channel with a common message:*

$$W_1 \to (\mathbf{X}_1, W_0, W_2) \to \mathbf{Y}_1 \tag{4.66}$$

$$W_2 \to (\mathbf{X}_2, W_0, W_1) \to \mathbf{Y}_2 \tag{4.67}$$

$$(W_0, W_t) \to (\mathbf{X}_t, W_0) \to \mathbf{Y}_t \tag{4.68}$$

for $t = 1, 2$.

Proof:

The result follows easily by the problem definition. □

We will need the data processing inequality in the following form:

Lemma 4 *For a Markov chain* $W \to (U, X) \to Y$

$$I(W; Y|U) \leq I(X; Y|U). \tag{4.69}$$

Proof:

We have

$$\begin{aligned} H(Y|U, X) &=^{(a)} H(Y|W, U, X) \\ &\leq^{(b)} H(Y|W, U) \end{aligned} \tag{4.70}$$

where (a) holds because of the Markov property and (b) since conditioning cannot increase entropy. Subtracting both sides from $H(Y|U)$ gives the desired result. □

Applying Lemma 4 to the Markov Chains (4.67)- (4.68) and using (4.68) yields,

$$\begin{aligned} I(W_2;\mathbf{Y}_2|W_0,W_1) &\leq I(\mathbf{X}_2;\mathbf{Y}_2|W_0,W_1) & (4.71)\\ I(W_t;\mathbf{Y}_t|W_0) &\leq I(\mathbf{X}_t;\mathbf{Y}_t|W_0) & (4.72)\\ I(W_0,W_1;\mathbf{Y}_1) &\leq I(W_0,\mathbf{X}_1;\mathbf{Y}_1). & (4.73)\end{aligned}$$

We first consider the bound (4.59) at the decoder 1. We have

$$\begin{aligned} N(R_0{+}R_1+R_2)&\\ &= H(W_0)+H(W_1)+H(W_2)\\ &=^{(a)} H(W_0,W_1)+H(W_2|W_0,W_1)\\ &= I(W_0,W_1;\mathbf{Y}_1)+I(W_2;\mathbf{Y}_2|W_0,W_1)+H(W_0,W_1|\mathbf{Y}_1)+H(W_2|\mathbf{Y}_2,W_0,W_1)\\ &\leq^{(b)} I(W_0,\mathbf{X}_1;\mathbf{Y}_1)+I(\mathbf{X}_2;\mathbf{Y}_2|W_0,W_1)+H(W_0,W_1|\mathbf{Y}_1)+H(W_2|\mathbf{Y}_2,W_0,W_1)\\ &\leq^{(c)} I(W_0,\mathbf{X}_1;\mathbf{Y}_1)+I(\mathbf{X}_2;\mathbf{Y}_2|W_0,W_1)+N\delta_{1,N}+N\delta_{2,N}\\ &=^{(d)} I(W_0,\mathbf{X}_1;\mathbf{Y}_1)+I(\mathbf{X}_2;\mathbf{Y}_2|W_0,W_1,\mathbf{X}_1(W_0,W_1))+N\delta_{1,N}+N\delta_{2,N}\\ &=^{(e)} I(W_0,\mathbf{X}_1;\mathbf{Y}_1)+I(\mathbf{X}_2;\mathbf{Y}_2|W_0,\mathbf{X}_1)+N\delta_{1,N}+N\delta_{2,N} \end{aligned} \tag{4.74}$$

where (a) follows from the independence of W_0, W_1, W_2; (b) from (4.71) and (4.73); (c) from (4.63) and (4.65); (d) and (e) from (4.48). If

$$I(\mathbf{X}_2;\mathbf{Y}_2|W_0,\mathbf{X}_1) \leq I(\mathbf{X}_2;\mathbf{Y}_1|W_0,\mathbf{X}_1) \tag{4.75}$$

then it follows from (4.74) that

$$\begin{aligned} N(R_0+R_1+R_2) &\leq I(W_0,\mathbf{X}_1;\mathbf{Y}_1)+I(\mathbf{X}_2;\mathbf{Y}_1|\mathbf{X}_1,W_0)+N\delta_{1,N}+N\delta_{2,N}\\ &= I(W_0,\mathbf{X}_1,\mathbf{X}_2;\mathbf{Y}_1)+N\delta_{1,N}+N\delta_{2,N}\\ &= I(\mathbf{X}_1,\mathbf{X}_2;\mathbf{Y}_1)+N\delta_{1,N}+N\delta_{2,N}\\ &\leq \sum_{n=1}^{N} I(X_{1n},X_{2n};Y_{1n})+N\delta_{1,N}+N\delta_{2,N}. \end{aligned} \tag{4.76}$$

Applying Willems' notation [15, Section 3]

$$U_n = W_0 \qquad n=1,\ldots,N \tag{4.77}$$

to condition (4.75) yields

$$I(\mathbf{X}_2;\mathbf{Y}_2|\mathbf{X}_1,\mathbf{U}) \leq I(\mathbf{X}_2;\mathbf{Y}_1|\mathbf{X}_1,\mathbf{U}) \tag{4.78}$$

Furthermore, the condition

$$I(X_2;Y_2|X_1,U) \leq I(X_2;Y_1|X_1,U) \tag{4.79}$$

implies (4.75) and (4.78) due to the following theorem.

Theorem 15 *If*

$$I(X_1;Y_1|X_2,U) \leq I(X_1;Y_2|X_2,U) \tag{4.80}$$

for all probability distributions on $\mathcal{U}\times\mathcal{X}_1\times\mathcal{X}_2$ *such that* $p(u,x_1,x_2) = p(u)p(x_1|u)p(x_2|u)$, *then for the strong interference channel with a common message we have*

$$I(\mathbf{X}_1;\mathbf{Y}_1|\mathbf{X}_2,\mathbf{U}) \leq I(\mathbf{X}_1;\mathbf{Y}_2|\mathbf{X}_2,\mathbf{U}). \tag{4.81}$$

Equivalently, $I(X_2;Y_2|X_1,U) \leq I(X_2;Y_1|X_1,U)$ implies $I(\mathbf{X}_2;\mathbf{Y}_2|\mathbf{X}_1,\mathbf{U}) \leq I(\mathbf{X}_2;\mathbf{Y}_1|\mathbf{X}_1,\mathbf{U})$. The proof of Theorem 15 relies on the result in [90, Proposition 1] and follows the same approach as Lemma in [56].

To prove that the bound (4.58) at the decoder 1 is valid, we consider

$$\begin{aligned}
N(R_1&+R_2)\\
&= H(W_1) + H(W_2)\\
&=^{(a)} H(W_1|W_0) + H(W_2|W_0,W_1)\\
&= I(W_1;\mathbf{Y}_1|W_0) + I(W_2;\mathbf{Y}_2|W_0,W_1) + H(W_1|\mathbf{Y}_1,W_0) + H(W_2|\mathbf{Y}_2,W_0,W_1)\\
&\leq^{(b)} I(\mathbf{X}_1;\mathbf{Y}_1|W_0) + I(\mathbf{X}_2;\mathbf{Y}_2|W_0,W_1) + H(W_1|\mathbf{Y}_1,W_0) + H(W_2|\mathbf{Y}_2,W_0,W_1)\\
&\leq^{(c)} I(\mathbf{X}_1;\mathbf{Y}_1|W_0) + I(\mathbf{X}_2;\mathbf{Y}_2|W_0,W_1) + N\delta_{1,N} + N\delta_{2,N}\\
&=^{(d)} I(\mathbf{X}_1;\mathbf{Y}_1|W_0) + I(\mathbf{X}_2;\mathbf{Y}_2|W_0,W_1,\mathbf{X}_1(W_0,W_1)) + N\delta_{1,N} + N\delta_{2,N}\\
&=^{(e)} I(\mathbf{X}_1;\mathbf{Y}_1|\mathbf{U}) + I(\mathbf{X}_2;\mathbf{Y}_2|\mathbf{X}_1,\mathbf{U}) + N\delta_{1,N} + N\delta_{2,N}.
\end{aligned} \tag{4.82}$$

where again (a) follows from the independence of W_0, W_1, W_2; (b) from (4.71) and (4.72); (c) from (4.63) and (4.65); (d) from (4.48); (e) from (4.48) and (4.77). Again, if

(4.75) holds, then (4.82) becomes

$$\begin{aligned} R_1 + R_2 &\le I(\mathbf{X}_1;\mathbf{Y}_1|\mathbf{U}) + I(\mathbf{X}_2;\mathbf{Y}_1|\mathbf{X}_1,\mathbf{U}) + N\delta_{1,N} + N\delta_{2,N} \\ &= I(\mathbf{X}_1,\mathbf{X}_2;\mathbf{Y}_1|\mathbf{U}) + N\delta_{1,N} + N\delta_{2,N} \\ &\le \sum_{n=1}^{N} I(X_{1n},X_{2n};Y_{1n}|U_n) + N\delta_{1,N} + N\delta_{2,N}. \end{aligned} \tag{4.83}$$

The same approach can be used to show that the bounds (4.58) and (4.59) are satisfied at decoder 2 under a condition equivalent to (4.78)

$$I(\mathbf{X}_1;\mathbf{Y}_1|\mathbf{X}_2,\mathbf{U}) \le I(\mathbf{X}_1;\mathbf{Y}_2|\mathbf{X}_2,\mathbf{U}) \tag{4.84}$$

which, due to Theorem 15, reduces to

$$I(X_1;Y_1|X_2,U) \le I(X_1;Y_2|X_2,U). \tag{4.85}$$

Finally, the bounds (4.56) and (4.57) are the single user upper bounds and hence have to be satisfied.

To conclude the proof we need the following lemma that shows that the obtained conditions (4.79) and (4.85) are identical to the strong interference conditions (4.54) and (4.55).

Lemma 5 *The conditions*

$$I(X_1;Y_1|X_2,U) \le I(X_1;Y_2|X_2,U) \tag{4.86}$$

$$I(X_2;Y_2|X_1,U) \le I(X_2;Y_1|X_1,U) \tag{4.87}$$

that hold for all joint $p(u,x_1,x_2,y_1,y_2)$ *that factor as in (4.60), and the strong interference conditions (4.54)-(4.55)*

$$I(X_1;Y_1|X_2) \le I(X_1;Y_2|X_2) \tag{4.88}$$

$$I(X_2;Y_2|X_1) \le I(X_2;Y_1|X_1) \tag{4.89}$$

that hold for all product distributions on the inputs X_1 *and* X_2*, are equivalent, that is, they are satisfied by the same class of interference channels.*

Proof.

We use the following result from [56, Lemma]: If $I(X_1;Y_1|X_2) \leq I(X_1;Y_2|X_2)$ for all product probability distributions on $\mathcal{X}_1 \times \mathcal{X}_2$, then $I(X_1;Y_1|X_2,U) \leq I(X_1;Y_2|X_2,U)$ where $U \to (X_1,X_2) \to (Y_1,Y_2)$ and $X_1 \to U \to X_2$. From this result it follows directly that the strong interference conditions (4.54) and (4.55) imply the conditions in the interference channel with common information (4.86) and (4.87).

To prove that the other direction is also true, we observe that since the conditions (4.86) and (4.87) are satisfied for all input distributions of the form (4.60), the conditions (4.86) and (4.87) must hold also when U is independent of inputs X_1, X_2. Therefore, the strong interference conditions must hold, because for such input distribution $p(u,x_1,x_2)$, conditions (4.86) and (4.87) reduce to the strong interference conditions (4.54)-(4.55). $\square$

From (4.76) and (4.83), it follows that $(R_0,R_1,R_2) \in \mathrm{co}(\mathcal{C})$. Since the region $\mathcal{C}$ is already convex, then $(R_0,R_1,R_2) \in \mathcal{C}$. $\square$

4.5.2 Achievability

The achievability of the rates of Theorem 14 follows from the capacity region of the compound MAC with common information (4.13): from Theorem 11, rates (4.13) are achievable in the case in which both private messages are decoded at the receivers; this further guarantees that these rates are also achieved when a weaker constraint of decoding of a single private message is imposed at the receivers, and are hence achievable in the interference channel with common information. Furthermore, the strong intereference conditions (4.54)-(4.55) imply conditions (4.86)-(4.87) by Lemma 5 and therefore the region (4.13) reduces to (4.56)-(4.59). Hence the proof of achievability in Theorem 14 is immediate. $\square$

In the weak interference case in which the conditions (4.54)-(4.55) do not hold, the capacity region will contain the rates (4.13). The capacity region in the general case is however still an open problem.

4.5.3 Gaussian Channel

We next consider the Gaussian interference channel in the standard form (4.39)-(4.40). The code definition is the same as that given in Section 4.5 with the addition of the power constraints (4.41). From the maximum-entropy theorem [81, Thm. 9.6.5] it follows that Gaussian inputs are optimal. We have the following result.

Corollary 2 *When the strong interference conditions $h_{12}^2 \geq 1$, $h_{21}^2 \geq 1$ are satisfied, the capacity region of the Gaussian strong interference channel with common information is given by*

$$
\begin{aligned}
\mathcal{R} = \bigcup \{ & (R_1, R_2) : \\
& 0 \leq R_1 \leq C\left(\bar{a}P_1\right) && (4.90)\\
& 0 \leq R_2 \leq C\left(\bar{b}P_2\right) && (4.91)\\
& R_1 + R_2 \leq \min\left\{C(\bar{a}P_1 + h_{21}^2\bar{b}P_2), C(h_{12}^2\bar{a}P_1 + \bar{b}P_2)\right\} && (4.92)\\
& 0 \leq R_0 + R_1 + R_2 \leq \min\Big\{C(P_1 + h_{21}^2P_2 + 2h_{21}\sqrt{aP_1bP_2}) \\
& \qquad\qquad C(h_{12}^2P_1 + P_2 + 2h_{12}\sqrt{aP_1bP_2})\Big\}\} && (4.93)
\end{aligned}
$$

where the union is over all a, b, for $0 \leq a \leq 1, 0 \leq b \leq 1$ and $\bar{a} = 1 - a$, $\bar{b} = 1 - b$.

4.5.4 Discussion

Communication systems with encoders that have to send both private and common information naturally arise in the case when encoders can partially cooperate as in [65,66]. After such cooperation, the common information consists of two indexes, each partially describing one of the two original messages. The assumption of our model that the entire common message is decoded simplifies the problem. However, a receiver interested in a message from only one encoder, as is the case in the interference channel, will be interested in only a *part* of the common message. Understanding such communication problems appears to be much more challenging and is the subject of our future work.

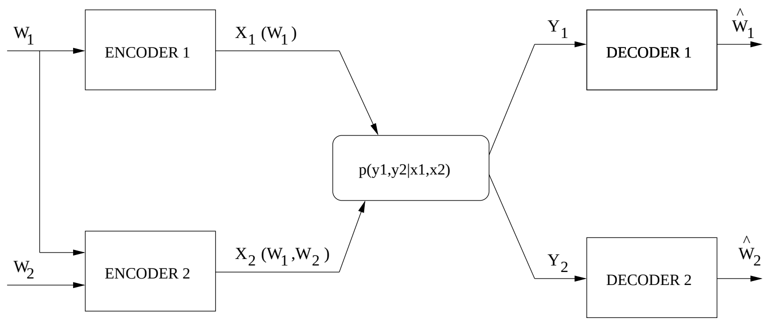

Figure 4.6: Interference channel with unidirectional cooperation.

4.6 The Strong Interference Channel With Unidirectional Cooperation

In this section, we consider the interference channel in which full information about messages sent at one encoder is available to the other encoder, but not vice versa. Such a channel model allows the encoder that knows both messages to exploit that information to improve the achievable rates. The achievable rates for this channel model have been presented in [91]. Furthermore, for the case of weak interference, i.e. $h_{21} < 1$, the capacity region was determined in [92]. The communication system is shown in Figure 4.6.

We derive conditions equivalent to (1.1)-(1.2) under which there is no penalty in decoding both messages at both decoders in the interference channel with unidirectional cooperation. We compare the obtained conditions to the strong interference conditions (1.1)-(1.2).

We consider a memoryless interference channel that consists of finite sets $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2$ and a conditional probability distribution $p(y_1, y_2|x_1, x_2)$. Each encoder t, $t = 1, 2$, wishes to send an independent message $W_t \in \{1, \ldots, M_t\}$ to receiver t in N channel uses. It is assumed that message W_1 is also known at encoder 2, thus allowing for unidirectional cooperation. The channel is memoryless and time-invariant as given by Equation (4.21).

An (M_1, M_2, N, P_e) code for the channel consists of two encoding functions generating codewords

$$\mathbf{x}_1 = f_1(W_1) \tag{4.94}$$

$$\mathbf{x}_2 = f_2(W_1, W_2) \tag{4.95}$$

two decoding functions

$$\hat{W}_t = g_t(\mathbf{Y}_t) \quad t = 1, 2 \tag{4.96}$$

and a maximum error probability

$$P_e = \max\{P_{e,1}, P_{e,2}\} \tag{4.97}$$

where, for $t = 1, 2$

$$P_{e,t} = \sum_{(w_1, w_2)} \frac{1}{M_1 M_2} P\left[g_t(\mathbf{Y}_t) \neq (w_t) | (w_1, w_2) \text{ sent}\right]. \tag{4.98}$$

A rate pair (R_1, R_2) is achievable if, for any $\epsilon > 0$, there is an (M_1, M_2, N, P_e) code such that

$$P_e \leq \epsilon \text{ and } M_i \geq 2^{NR_i} \quad i = 1, 2.$$

The capacity region of the *interference channel with unidirectional cooperation* is the closure of the set of all achievable rate pairs (R_1, R_2).

The next theorem gives conditions under which the capacity region coincides with the capacity region of the channel in which both messages are required at both receivers.

Theorem 16 *For an interference channel* $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_1, y_2|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$ *with unidirectional cooperation satisfying*

$$I(X_2; Y_2|X_1) \leq I(X_2; Y_1|X_1) \tag{4.99}$$

$$I(X_1, X_2; Y_1) \leq I(X_1, X_2; Y_2) \tag{4.100}$$

for all joint input distributions on X_1 *and* X_2*, the capacity region* $\mathcal{C}$ *is given by*

$$\mathcal{C} = \bigcup \Big\{ (R_1, R_2) :$$

$$R_2 \leq I(X_2; Y_2|X_1) \tag{4.101}$$

$$R_1 + R_2 \leq I(X_1, X_2; Y_1) \Big\} \tag{4.102}$$

where the union is over joint distributions $p(x_1, x_2, y_1, y_2)$*.*

In the following section, we prove the converse and determine the strong interference conditions. To complete the proof of Theorem 16, in Section 4.6.2 we discuss how the achievability of the $\mathcal{C}$ rate region follows from the compound MAC with common information.

4.6.1 Converse: Strong Interference Conditions

Consider a code (M_1, M_2, N, P_e) for the interference channel with unidirectional cooperation. Applying Fano's inequality results in

$$H(W_1|\mathbf{Y}_1) \leq P_{e1}\log(M_1-1)+h(P_{e1}) \triangleq N\delta_{1,N}, \tag{4.103}$$

$$H(W_2|\mathbf{Y}_2) \leq P_{e2}\log(M_2-1)+h(P_{e2}) \triangleq N\delta_{2,N}. \tag{4.104}$$

where $\delta_{t,N} \to 0$ as $P_{et} \to 0$ (or as $P_e \to 0$). For notational convenience, we define $\delta_N \triangleq \delta_{1,N}+\delta_{2,N}$ and $R_s = R_1 + R_2$. We now derive the R_s bound (4.102) for receiver $t = 1$.

From independence of W_1 and W_2 and Fano's inequalities (4.103) and (4.104), we have

$$\begin{aligned} NR_s &= H(W_1) + H(W_2|W_1) && (4.105)\\ &= I(W_1;\mathbf{Y}_1) + I(W_2;\mathbf{Y}_2|W_1) \\ &\qquad + H(W_1|\mathbf{Y}_1) + H(W_2|\mathbf{Y}_2, W_1) && (4.106)\\ &\leq I(W_1;\mathbf{Y}_1) + I(W_2;\mathbf{Y}_2|W_1) + N\delta_N. && (4.107)\end{aligned}$$

With the assumption that (4.94) defines $\mathbf{X}_1$ as a deterministic one-to-one function of W_1, it follows that

$$\begin{aligned} NR_s &\leq I(W_1,\mathbf{X}_1;\mathbf{Y}_1) + I(\mathbf{X}_2;\mathbf{Y}_2|W_1) + N\delta_N && (4.108)\\ &\leq I(W_1,\mathbf{X}_1;\mathbf{Y}_1) + I(\mathbf{X}_2;\mathbf{Y}_2|\mathbf{X}_1, W_1) + N\delta_N. && (4.109)\end{aligned}$$

Therefore, if the condition

$$I(\mathbf{X}_2;\mathbf{Y}_2|\mathbf{X}_1, W_1) \leq I(\mathbf{X}_2;\mathbf{Y}_1|\mathbf{X}_1, W_1) \tag{4.110}$$

holds, then it follows from (4.109) that

$$NR_s \leq I(W_1, \mathbf{X}_1; \mathbf{Y}_1) + I(\mathbf{X}_2; \mathbf{Y}_1|\mathbf{X}_1, W_1) + N\delta_N \tag{4.111}$$

$$= I(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_1) + N\delta_N \tag{4.112}$$

$$\leq \sum_{n=1}^{N} I(X_{1n}, X_{2n}; Y_{1n}) + N\delta_N. \tag{4.113}$$

And therefore, we obtain the sum-rate outer bound (4.102). Per-letter conditions follow from the next lemma.

Lemma 6 *If per-letter conditions*

$$I(X_2; Y_2|X_1) \leq I(X_2; Y_1|X_1) \tag{4.114}$$

are satisfied for joint distributions $p(x_1, x_2)$*, then*

$$I(\mathbf{X}_2; \mathbf{Y}_2|\mathbf{X}_1, W_1) \leq I(\mathbf{X}_2; \mathbf{Y}_1|\mathbf{X}_1, W_1). \tag{4.115}$$

Lemma 6 is similar to the lemma by Costa and El Gamal [56]. One difference is that the condition (4.114) in our case has to be satisfied for all the input distributions not only for product ones, as cooperation introduces dependence between inputs.

The bound (4.101) is an immediate single-link outer bound and hence the converse follows. □

In the case when (4.100) does not hold, showing that (4.116) is indeed the capacity region would require proving an outer bound of the form $R_S \leq I(X_1, X_2; Y_2)$. Due to the asymmetry of the problem, the approach (4.105)-(4.113) does not apply.

4.6.2 Achievability

We apply the same reasoning as in Section 4.5.2. The achievability of the rates (4.13) of Theorem 11 in the compound MAC with common information guarantees that these rates are also achieved when a weaker constraint of decoding of a single message is imposed at the receivers. Since in the interference channel with unidirectional cooperation encoder 2 knows the entire message W_1, we can view R_1 as the common rate. For the

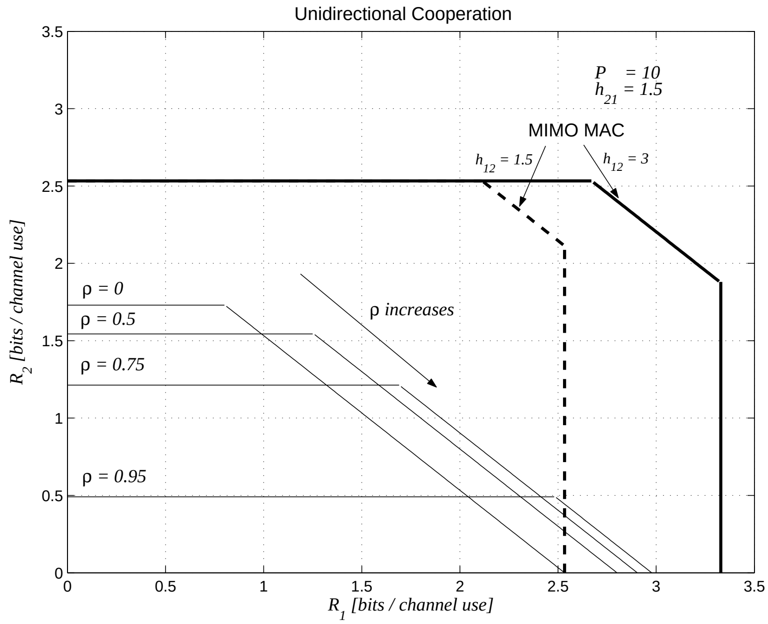

Figure 4.7: Gaussian interference channel with unidirectional cooperation capacity region.

same reason, user 1 in the corresponding compound MAC has zero rate for the private message. We can choose $U = X_1$ and the region (4.13) becomes

$$\mathcal{C}_{CMAC} = \bigcup \Big\{ (R_1, R_2) : \quad R_2 \leq \min_t I(X_2; Y_t | X_1) \quad R_1 + R_2 \leq \min_t I(X_1, X_2; Y_t) \Big\} \tag{4.116}$$

where the union is over all $p(x_1, x_2, y_1, y_2)$. When conditions (4.99)-(4.100) are satisfied, region (4.116) reduces to region (4.101)-(4.102) in Theorem 16. □

4.6.3 Gaussian Channel

We next consider the Gaussian interference channel in the standard form (4.39)-(4.40). The code definition is the same as given in Section 4.6 with the addition of the power constraints (4.41). From the maximum-entropy theorem [81, Thm. 9.6.5] it follows that Gaussian inputs are optimal.

Corollary 3 *When the strong interference conditions*

$$h_{21} \geq 1 \tag{4.117}$$

$$h_{12} \geq \frac{1}{\alpha}\left(\sqrt{\alpha^2 + h_{21}^2 - 1 + 2\rho\alpha h_{21} + \rho^2} - \rho\right) \tag{4.118}$$

hold, where ρ is a correlation coefficient for X_1, X_2 and $\alpha = \sqrt{P_1/P_2}$, the capacity region of the Gaussian interference channel with unidirectional cooperation is given by

$$\begin{aligned} \mathcal{C} = \bigcup_{\rho} \{ & (R_1, R_2) : \\ & R_2 \leq C\left((1-\rho^2)P_2\right) && (4.119) \\ & R_1 + R_2 \leq C\left(P_1 + h_{21}^2 P_2 + 2\rho\sqrt{h_{21}^2 P_1 P_2}\right)\}. && (4.120) \end{aligned}$$

Figure 4.7 shows the capacity region for $P_1 = P_2 = P = 10$ and $h_{21} = 1.5$.

One can show that satisfying conditions (4.117)-(4.118) for all possible values of ρ is more demanding than the conditions (1.1)-(1.2) which in Gaussian case reduce to $h_{21} \geq 1, h_{12} \geq 1$. Still, we observe there will always be some set of values h_{21}, h_{12} that satisfy conditions (4.117)-(4.118), except when $P_1 = 0$. For $P_1 = 0$, the channel reduces to the broadcast channel from encoder 2. As the channel is degraded, there can be no strong interference conditions.

4.6.4 Conclusion

In this chapter, we presented three channel models that incorporate partial transmitter cooperation. For the interference channels presented in Sections 4.5 and 4.6, we determined the capacity region under the strong interference conditions in which decoders can decode all transmitted messages with no rate penalty. For the interference channel with unidirectional cooperation it is possible that weaker conditions exist. Determining the strong interference conditions for more general channel models such as the interference channel with correlated sources is still an open problem.

Chapter 5

Summary and Future Directions

This thesis was motivated by an interest in understanding the mechanisms and gains of relaying and node cooperation in wireless networks. While the crucial cooperative strategies were proposed back in late seventies [10], recent interest in sensor networks has rekindled interest in multi-terminal systems with many nodes. Motivated by sensor applications, we began this work by considering cooperation in large, energy-constrained networks. For such networks, we proposed a cooperative strategy for the multicast and broadcast traffic model. We used insights offered by network information theory on the importance of exploiting the overheard received signals, to propose an accumulative broadcast strategy that increases the network energy-efficiency. We analyzed two problems concerned with energy-efficient data broadcast. First, we formulated the minimum-energy accumulative broadcast problem. We showed that the problem is NP-complete and proposed an energy-efficient heuristic algorithm. We then addressed the maximum lifetime multicast problem and presented the Maximum Lifetime Accumulative Broadcast (MLAB) Algorithm that finds the optimum solution. The power levels found by the algorithm ensure that the lifetimes of the active relays are the same, causing them to fail simultaneously. Several questions regarding accumulative broadcast remain at this time.

- The proposed accumulative broadcast scheme employs a decode-and-forward relay strategy. The potential benefit of unreliable forwarding in the accumulative broadcast problem still needs to be investigated.

- In this work, accumulative broadcast was proposed for the AWGN channel with constant link gains. However, cooperation among the nodes in the fading channel offers additional benefit as a form of diversity [18,78,93]. It would be interesting to consider the implications of time varying channels on the accumulative broadcast problems.

- The maximum lifetime solution found by MLAB is *static* since it stays constant throughout the multicast session. Similar to conventional broadcast, a dynamic strategy with time varying powers can extend the network lifetime [42]. Dynamic switching between schedules, corresponding to routing packets along multiple routes, may yield a network lifetime larger than that of the ASAP distribution, the optimal static policy. A general solution for optimal dynamic cooperative multicast remains an open problem.

We then considered cooperation in a large network with a different traffic model that assumes a single source-destination link. We considered a simple two-hop protocol that precludes communication among relays. Rather, relays process information received from the source and forward it to the destination. For the two-hop AF power/bandwidth problem (3.12) presented in Chapter 3, we characterized the optimum bandwidth by showing that it allows the network to operate in the linear regime. We then determined the optimum relay power allocation for two cases - when relays are signaling in shared bandwidth and for orthogonal signaling. While we considered a single source-destination pair, our results have implications to networks with multiple source-destination pairs. Our view is that, for each such pair, the relay network in between is a resource that we aim to use efficiently. Such a view motivates a total power contraint as the network budget. The optimum power allocation then allows determination of the best subset of relay nodes for each source-destination pair.

In a wireless network, messages are typically expected to travel further than just two hops and the two-hop protocol approach should not be viewed as an obstacle to multihoping protocols. In fact, it is expected that a routing protocol will still exist on the network layer. The cooperative relay strategies will be run on the lower medium-access control layer, allowing for faster network adaptation to changes due to fading or high mobility. In that sense, routing and relaying will work together to increase network performance.

Three immediate open problem become apparent from our analysis.

- Proof that the optimum AF bandwidth problem (3.30) has a unique solution. Given the relay powers (3.26) or (3.33), the AF problem (3.12) reduces to an optimization problem in two-variables: bandwidth and the power fraction allocated to the source. Even when the source power is known, the closed-form solution for the optimum bandwidth does not appear to exist. Proving the uniqueness of the optimum solution reduces to proving the unimodality of the rate function (3.30). The difficulty in the proof seems to come from the optimality of the infinitely large bandwidth for certain values of channel gains. It occurs in the special case when direct transmission at the source is optimal. While the numerical results indicate the uniqueness of the solution, the explicit proof remains elusive.

- Practical design of networks that allow for the form of cooperation described above. These protocols in general require a cross-layer design in which physical layer cooperation is intergrated with MAC and network layer design for impoved performance. Several aspects of practical design have been addressed in [94].

- In low-SNR scenarios, the relay power solutions (3.26) or (3.33) exhibit a clustering behavior by favorizing the relays that are in the vicinity of the source and/or the destination, as shown in Figures 3.1, 3.2 and 3.4. This further suggests solutions that, instead of two-hop, employ a *three-hop* network architecture. In the first hop, the source transmits to the nodes in its vicinity. After decoding the message, these nodes, together with the source, can jointly encode and transmit, forming a transmit cluster. On the receiving side, a receiving cluster performs cooperative detection. Therefore, data is transmitted through two clusters that act as multiple transmit and receive antennas. Preliminary results in that direction have been presented in [95]. The problem involves specifying the exact cooperation strategy among the cluster nodes and investigating the potential gains from such MIMO architectures.

In Chapter 4, we considered the interference channel with limited cooperation, capturing a communication situation in which two sources wish to send messages to two corresponding receivers, while being able to partially cooperate with each other. The capacity region of the basic interference channel has been a long-standing open problem and thus one has little hope in solving the interference channel with cooperation problem, which adds yet another dimension to the problem. In Chapter 4, we determined the capacity region of such channels under very special conditions in which there is no penalty in decoding both messages at both receivers. For transmitter cooperation, we initially assumed that encoders cooperate over links with finite capacities in the form of a conference. In Section 4.4, we presented the capacity region of the compound MAC with conferencing in which both receivers wish to decode both messages. The obtained capacity region is an inner bound on the rates achievable in a channel in which each decoder decodes a message sent from a single encoder. We then determined the strong interference conditions and the capacity region for the strong interference channel with common information in Section 4.5 and for the strong interference channel with unidirectional cooperation in Section 4.6. It would be interesting to investigate whether there exists a weaker set of strong interference conditions for the interference channel with unidirectional cooperation than those determined in [69].

For the conferencing model, after cooperation over the links with finite capacities, encoders will have partial information about each other's messages. The exact amount of such common information is determined by the link capacities and we expect the solution to be parametrized by these capacities. Therefore, channel models with specific assumptions about the knowledge of each other's messages at the encoders, as in the interference channel with unidirectional cooperation, are captured by the channel model with conferencing. We list three immediate interesting problems still ahead.

- The strong interference conditions for more general models such as the interference channel with conferencing or the interference channel with correlated sources.
- Even more challenging, the capacity region of the interference chanel with conferencing for a wider set of channel conditions.
- Even more general, cooperation gains for more general discrete channel models with cooperation.

Appendix A

Additional Proofs

A.1 Proof: Theorem 1

An upper bound to the achievable rate between the source and the destination is the maximum conditional mutual information across a minimum cut [81]. Consider the multiaccess cut in the given network that separates the destination node from the rest of the network. Let X_j denote a symbol transmitted at node j and Y denote the received signal at the destination. The maximum mutual information across this cut is given by

$$C_{\text{MAC}} = I(X_1, \dots X_{m-1}; Y). \tag{A.1}$$

In this network, each orthogonal channel is assigned bandwidth W and hence the mutual information above is given by the sum of rates achieved in each of the channels. For Gaussian channels,

$$C_{\text{MAC}} = W \sum_{k=1}^{m-1} \log_2 \left(1 + \frac{h_{mk} p_k}{N_0 W}\right). \tag{A.2}$$

In the wideband regime, Equation (A.2) becomes

$$C_{\text{MAC}} = \lim_{W \to \infty} W \sum_{k=1}^{m-1} \log_2 \left(1 + \frac{h_{mk} p_k}{N_0 W}\right) \tag{A.3}$$

$$= \frac{1}{N_0 \log 2} \sum_{k=1}^{m-1} h_{mk} p_k, \tag{A.4}$$

which is precisely the rate given by (2.10) achieved using the repetition strategy. Since this rate is achievable, this cut is the minimum cut. No better rate can be achieved since it would violate the condition for the upper bound. □

A.2 Proof: Theorem 2

For the purpose of this proof we represent a solution to the accumulative broadcast problem by a vector with each entry i containing the ith transmitting node n_i and the ith transmitted power level P_i. A solution $\mathbf{S}$ is represented as

$$\mathbf{S} = \left[\begin{array}{cccc} (n_1, P_1) & (n_2, P_2), & \ldots & (n_M, P_M) \end{array} \right]^T \tag{A.5}$$

for some $M \geq N$. We write $(n_i, P_i) = (0, 0)$ if no node transmits at step i.

Assume that $\mathbf{S}$ schedules the same node for a transmission more than once. It is sufficient to show that there is a feasible schedule $\hat{\mathbf{S}}$ that uses the same total transmit power as $\mathbf{S}$, in which that node transmits once. Let l denote the smallest integer such that there exists an integer $m > l$ with $n_m = n_l$. Consider the policy $\hat{\mathbf{S}}$, a vector of length $M - 1$ with elements $(\hat{n}_i, \hat{P}_i)$ such that

$$(\hat{n}_i, \hat{P}_i) = \begin{cases} (n_l, P_l + P_m) & \text{if } i = l, \\ (0, 0) & \text{if } i = m, \\ (n_i, P_i) & \text{if } i \geq m. \end{cases} \tag{A.6}$$

The solution $\hat{\mathbf{S}}$ combines transmissions at steps l and m into a single transmission with power $P_l + P_m$ at step l. The rest of the nodes are scheduled as in $\mathbf{S}$.

For any node j, the energy accumulated by step k in new schedule is $\sum_{i=1}^{k-1} h_{j\hat{n}_i} \hat{P}_i \geq \sum_{i=1}^{k-1} h_{jn_i} P_i$. Therefore, $\hat{\mathbf{S}}$ is a feasible schedule since any node j made reliable by step k in schedule $\mathbf{S}$ is also reliable at step k in the new schedule. $\square$

A.3 Proof: Theorem 4

The proof is by induction on k, where k is the index to a sequence of stages during the ASAP($\tilde{p}$) distribution. We prove by induction that at the start of stage k, nodes $\{x_1, \ldots, x_k\} \subset S_k(\tilde{p})$. In case that the number of stages is $\tilde{\tau} = \tau(\tilde{p}) < M$, we define $S_k(\tilde{p}) = S_{\tilde{\tau}}(\tilde{p})$ for all $\tilde{\tau} < k \leq M$. The idea is that ASAP($\tilde{p}$) makes nodes reliable at least as soon as the schedule $\tilde{\mathbf{x}}$.

Case $k = 1$ is obvious since $S_1(\tilde{p}) = \{1\}$ for any $\tilde{p}$. Next we assume that $\{x_1, \ldots, x_k\} \subset$

$S_k(\tilde{p})$. This implies

$$\tilde{p} \sum_{x_j \in S_k(\tilde{p})} g_{k+1,j} \geq \tilde{p} \sum_{x_j \in \{x_1,\ldots,x_k\}} g_{k+1,j} \geq^{(a)} \overline{P} \tag{A.7}$$

where (a) follows from the feasibility of power $\tilde{p}$ for schedule $\tilde{\mathbf{x}}$, because under schedule $\tilde{\mathbf{x}}$, node x_{k+1} is made reliable by transmissions of $\{x_1, \ldots, x_k\}$. We conclude that $x_{k+1} \in S_{k+1}(\tilde{p})$ and since $\{x_1, \ldots, x_k\} \subset S_k(\tilde{p}) \subset S_{k+1}(\tilde{p})$, it follows that $\{x_1, \ldots, x_{k+1}\} \subset S_{k+1}(\tilde{p})$. Thus, $\{x_1, \ldots, x_M\} \subset S_M(\tilde{p})$, implying the ASAP($\tilde{p}$) distribution makes all the nodes in a schedule $\tilde{\mathbf{x}}$, and thus all destination nodes, reliable. □

A.4 Proof: Theorem 5

Suppose the last restart of the MLAB algorithm occurs when the power is p_0 and the ASAP(p_0) distribution stalls at stage $\tau_0 = \tau(p_0)$. This implies

$$p_0 \sum_{k \in S_{\tau_0}(p_0)} h_{jk} < \overline{P}, \qquad j \in U_{\tau_0}(p_0). \tag{A.8}$$

In this case, we restart MLAB with broadcast power $p_0 + \delta_0$ where $\delta_0 = \min_{j \in U_{\tau_0}(p_0)} \delta_j$ and δ_j satisfies

$$(p_0 + \delta_j) \sum_{k \in S_{\tau_0}(p_0)} h_{jk} = \overline{P}. \tag{A.9}$$

This implies

$$(p_0 + \delta_0) \sum_{k \in S_{\tau_0}(p_0)} h_{jk} \leq \overline{P}, \qquad j \in U_{\tau_0}(p_0). \tag{A.10}$$

Since this is the last restart of MLAB, the ASAP($p_0 + \delta_0$) distribution is a feasible multicast. It follows that $p^* \leq p_0 + \delta_0$ since p^* is the optimal broadcast power. To show that $p^* = p_0 + \delta_0$ requires the following lemma.

Lemma 1 *For any power $p' < p_0 + \delta_0$, the ASAP(p') distribution stalls at stage $\tau' = \tau(p')$ with $S_{\tau'}(p') \subset S_{\tau_0}(p_0)$.*

Lemma 1 implies that if $p^* < p_0 + \delta_0$, then the ASAP(p^*) distribution will stall, which is a contradiction of Theorem 4. Thus, at the final restart of the MLAB algorithm, the power is $p_0 + \delta_0 = p^*$.

Proof: Lemma 1

Let $\mathcal{F} = S_{\tau'}(p') \setminus S_{\tau_0}(p_0)$. First, we show by contradiction that $\mathcal{F}$ is an empty set. Suppose $\mathcal{F}$ is nonempty. Let $\tau_{\mathcal{F}}$ denote the first stage in which a node $j' \in \mathcal{F}$ was made reliable by the ASAP(p') distribution. Thus,

$$\overline{P} \leq p' \sum_{k \in S_{\tau_{\mathcal{F}}}(p')} h_{j'k}. \tag{A.11}$$

Moreover, $S_{\tau_{\mathcal{F}}}(p') \subset S_{\tau_0}(p_0)$ since up to stage $\tau_{\mathcal{F}}$, all nodes that were made reliable by ASAP(p') belong to $S_{\tau_0}(p_0)$. Hence,

$$\overline{P} \leq p' \sum_{k \in S_{\tau_0}(p_0)} h_{j'k} \tag{A.12}$$

$$<^{(a)} (p_0 + \delta_0) \sum_{k \in S_{\tau_0}(p_0)} h_{j'k} \tag{A.13}$$

$$\leq^{(b)} \overline{P} \tag{A.14}$$

since (a) follows from $p' < p_0 + \delta_0$ and (b) follows from Equation (A.10). Thus we have the contradiction $\overline{P} < \overline{P}$ and we conclude that $\mathcal{F}$ is empty, $S_{\tau'}(p') \subset S_{\tau_0}(p_0)$, and $U_{\tau_0}(p_0) \subset U_{\tau'}(p')$. Second, we observe that ASAP(p') stalls at stage τ' since for all $j \in U_{\tau'}(p')$,

$$p' \sum_{k \in S_{\tau'}(p')} h_{jk} \leq p' \sum_{k \in S_{\tau_0}(p_0)} h_{jk} \tag{A.15}$$

$$< (p_0 + \delta_0) \sum_{k \in S_{\tau_0}(p_0)} h_{jk} \leq \overline{P}. \tag{A.16}$$

□

A.5 Proof: Theorem 6

In a network that is connected under power Δ, there is a path from the source node to every other node in the network. Consider a path from node 1 to some node K. We relabel the nodes such that the path is given by $[1, 2, \ldots K]$. For any reliable node $1 \leq k \leq K - 1$ such that $k + 1$ is unreliable, it holds that $k + 1 \in U_k$. By distributed MLAB, node k will increase its transmit power whenever it overhears no transmissions for T_o, until $k+1$ is reliable. Thus, eventually all the nodes on the path will be reliable. This holds for any path for any node K.

We next find an upper bound on $\mathcal{T}_i$, time it takes for a node i to make all of its neighbors reliable. An upper bound on the number of transmissions needed at a node i to make node $j \in N_i(\Delta)$ reliable, neglecting the energy node j may have collected from transmission from other nodes, is $\lceil \overline{P}/h_{ji}\Delta \rceil$. In the worst case, node i will wait for T_o between any two consecutive broadcasts. Thus,

$$\mathcal{T}_i \leq \max_{j \in N_i(\Delta)} \left\{ T_o \left\lceil \frac{\overline{P}}{\Delta h_{ji}} \right\rceil \right\} \tag{A.17}$$

$$= T_o \left\lceil \frac{\overline{P}}{\Delta \min_{j \in N_i(\Delta)}\{h_{ji}\}} \right\rceil. \tag{A.18}$$

Since $j \in N_i(\Delta)$, it follows that $h_{ji} \neq 0$ and therefore $\mathcal{T}_i$ is finite for every node i. Since there is only a finite number of nodes, all nodes will be made reliable in finite time. □

A.6 Proof: Theorem 7

To prove Theorem 7, we next upper bound the time $\mathcal{T}(k)$ it takes for maximum transmit power $\overline{q} = k\Delta$, $k > 0$ to propagate through the network.

Lemma 2 *Let T be the duration of a single transmission and let $\tau(k) = N(k+1)T$. Then, $\mathcal{T}(k) < \tau(k)$.*

Proof: Lemma 2

The time it takes for one node to transmit with $\overline{q}$ is upper bounded by kT, the case when the node never previously transmitted. Since the node may have to wait for NACKs for additional time T, the total time at a node is upper bounded by $(k+1)T$, Since the propagation cannot take more than N hops, the total time is upper bounded by $N(k+1)T$. □

To prove Theorem 7, we first observe that power q is lower bounded by p^*: before the power p^* is reached, there are always nodes that are unreliable and the distributed MLAB does not stop at the reliable nodes. The power p^* is reached for $\overline{q} = \lceil p^*/\Delta \rceil \Delta <$

$p^* + \Delta$ and no further increase in power is neccessary. By Lemma 2, $\overline{q}$ will propagate in less than $\tau(\Delta) = N(\lceil p^*/\Delta \rceil + 1)T$ time. If $T_o \geq \tau(\Delta)$, no node will increase $\overline{q}$ before all reliable nodes transmitted with $\overline{q}$. However, at that point all network nodes will be reliable and distributed MLAB will stop at all nodes with $q = \overline{q}$. $\square$

References

[1] R. Ahlswede, "Multi-way communication channels," in *IEEE Int. Symp. Inf. Th.*, 1971, pp. 23–52.

[2] H. Liao, "Multiple access channels," *Ph.D. dissertation, University of Hawaii*, 1972.

[3] D. Slepian and J. K. Wolf, "A coding theorem for multiple access channels with correlated sources," *Bell Syst. Tech J.*, vol. 52, pp. 1037–1076, 1973.

[4] G. Caire and S. Shamai(Shitz), "On the achievable throughput of the multi-antenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, July 2003.

[5] S. Vishwanath, N. Jindal, and A. J. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[6] H. Weingarten, Y. Steinberg, and S. Shamai(Shitz), "The capacity region of the Gaussian MIMO broadcast channel," in *IEEE Int. Symp. Inf. Theory*, June 2004, p. 174.

[7] ——, "The capacity region of the Gaussian MIMO broadcast channel," *IEEE Trans. Inf. Theory, submitted*, July 2004.

[8] E. C. V. D. Meulen, "Transmission of information in a T-terminal discrete memoryless multiple access channel," *Ph.D. dissertation, University of California, Berkley, CA*, June 1968.

[9] ——, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, pp. 120–154, June 1971.

[10] T. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sept. 1979.

[11] P. Gupta and P. R. Kumar, "Towards an information theory of large networks: An achievable rate region," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1877–1894, Aug. 2003.

[12] M. Gastpar, G. Kramer, and P. Gupta, "The multiple-relay channel: Coding and antenna-clustering capacity," in *Proc. of Int. Symp. Inf. Theory*, June 2002.

[13] G. Kramer, M. Gastpar, and P. Gupta, "Capacity theorems for wireless relay channels," in *Proc. of the Allerton Conference on Communications, Control and Computing, Monticello, IL*, Oct. 2003.

[14] ——, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sept. 2005.

[15] F. M. J. Willems, "Informationtheoretical results for the discrete memoryless multiple access channel," *Ph.D. dissertation, Katholieke Universiteit Leuven, Belgium*, Oct. 1982.

[16] L.-L. Xie and P. R. Kumar, "A network information theory for wireless communication: Scaling laws and optimal operation," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 748–767, May 2004.

[17] ——, "An achievable rate for the multiple level relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1348–1358, Apr. 2005.

[18] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 3062–3080, Dec. 2004.

[19] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: The relay case," in *Proc. of INFOCOM 2002*, June 2002.

[20] ——, "On the asymptotic capacity of Gaussian relay networks," in *Proc. of Int. Symp. Inf. Theory (ISIT'02)*, June 2002, p. 195.

[21] A. Dana and B. Hassibi, "On the power efficiency of sensory and ad-hoc wireless networks," *IEEE Trans. Inf. Theory, accepted.*

[22] ——, "On the power efficiency of sensory and ad-hoc wireless networks," in *Proc. of IEEE Int. Symp. Inf. Theory*, June 2003.

[23] M. Katz and S. Shamai(Shitz), "Relaying protocols for two col-located users," in *IEEE Int. Symp. Inf. Theory*, Sept. 2005, pp. 936–940.

[24] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[25] J. Wieselthier, G. Nguyen, and A. Ephremides, "Algorithms for energy-efficient multicasting in static ad-hoc wireless network," *Mobile Networks and Applications*, pp. 251–263, 2001.

[26] ——, "On the construction of energy-efficient broadcast and multicast trees in wireless networks," in *Proc. of INFOCOM 2000*, Mar. 2000.

[27] C. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity.* Prentice Hall, Englewood Cliffs, NJ, 1982.

[28] F. Li and I. Nikolaidis, "On minimum-energy broadcasting in all-wireless networks," in *Proc. of Local Computer Networks (LCN 2001)*, Nov. 2001.

[29] A. Ahluwalia, E. Modiano, and L. Shu, "On the complexity and distributed construction of energy-efficient broadcast trees in static ad hoc wireless networks," in *Proc. of Conf. on Information Science and Systems*, Mar. 2002.

[30] M. Cagalj, J. Hubaux, and C. Enz, "Energy-efficient broadcast in all-wireless networks," *Wireless Networks, Springer Science*, vol. 11, no. 1, pp. 177–188, 2005.

[31] W. Liang, "Constructing minimum-energy broadcast trees in wireless ad hoc networks," in *Proc. of Int. Symp. on Mobile Ad Hoc Networking and Computing (MobiHoc)*, June 2002.

[32] D. Bertsekas and R.G.Gallager, *Data Networks.* Prentice Hall, Englewood Cliffs, NJ, 1992.

[33] P.-J. Wan, G. Calinescu, X.-Y. Li, and O. Frieder, "Minimum-energy broadcasting in static ad hoc wireless networks," *Wireless Networks*, vol. 8, pp. 607–617, 2002.

[34] N. Li and J. Hou, "BLMST: A scalable, power-efficient broadcast algorithm for wireless sensor networks," in *First Int. Conf. on Quality of Service in Heterogenous Wired/Wireless Networks (QSHINE)*, Oct. 2004, pp. 44–51.

[35] N. Li and C. Hou, "BLMST: a scalable, power efficient broadcast algorithm for wireless sensor networks," in *IEEE INFOCOM 2004.*

[36] J. H. Chang and L. Tassiulas, "Routing for maximum system lifetime in wireless ad-hoc networks," in *Proc. of 37-th Annual Allerton Conference on Communication, Control and Computing*, Sept. 1999.

[37] I. Kang and R. Poovendran, "Maximizing static network lifetime of wireless broadcast adhoc networks," in *Proc. of ICC'03*, May 2003.

[38] ——, "Maximizing network lifetime of broadcasting over wireless stationary adhoc networks," *Mobile Networks and Applications*, vol. 10, pp. 879–896, 2005.

[39] R. J. M. II, A. K. Das, and M. El-Sharkawi, "Maximizing lifetime in an energy constrained wireless sensor array using team optimization of cooperating systems," in *Proc. of the Int. Joint Conf. on Neural Networks, IEEE World Congress on Computational Intelligence*, May 2002.

[40] P. Floreen, P. Kaski, J. Kohonen, and P. Orponen, "Lifetime maximization for multicasting in energy-constrained networks," *IEEE JSAC Special Issue on Wireless Ad Hoc Networks*, vol. 23, no. 1, Jan. 2005.

[41] ——, "Multicast time maximization in energy constrained wireless networks," in *Proc. of ACM/IEEE Mobicom 2003 Workshop on Foundations of Mobile Computing*, Sept. 2003.

[42] I. Maric and R. D. Yates, "Efficient multihop broadcast for wireless networks," *IEEE JSAC Special Issue on Fundamental Performance Limits of Wireless Sensor Networks*, vol. 22, no. 6, pp. 1080–1088, Aug. 2004.

[43] ——, "Efficient multihop broadcast for wideband systems," in *Multiantenna Channels: Capacity, Coding and Signal Processing, DIMACS Workshop on Signal Processing for Wireless Transmission, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, G.J. Foschini and S. Verdu, eds.*, vol. 62, Oct. 2002, pp. 285–299.

[44] ——, "Efficient multihop broadcast for wideband systems," in *Proc.of the Allerton Conference on Communications, Control and Computing, Monticello, IL*, Oct. 2002.

[45] ——, "Cooperative broadcast for maximum network lifetime," in *Conference on Information Sciences and Systems, Invited paper*, Mar. 2004.

[46] ——, "Cooperative multicast for maximum network lifetime," *IEEE JSAC Special Issue on Wireless Ad Hoc Networks*, vol. 23, no. 1, Jan. 2005.

[47] ——, "Forwarding strategies for Gaussian parallel-relay networks," in *Proc. of IEEE Int. Symp. Inf. Theory*, June 2004.

[48] ——, "Bandwidth and power allocation for cooperative strategies in Gaussian relay networks," in *Proc. of the Asilomar Conference On Signals, Systems and Computers*, Nov. 2004.

[49] S. Servetto, "On the feasibility of large scale wireless sensor networks," in *Allerton Conference on Communication, Control and Computing*, 2002.

[50] A. B. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. 24, no. 1, p. 60, Jan. 1978.

[51] H. Sato, "Two user communication channels," *IEEE Trans. Inf. Theory*, vol. 23, no. 3, p. 295, May 1977.

[52] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.

[53] H.-F. Chong, M. Motani, and H. K. Garg, "A comparison of two achievable rate regions for the interference channel," in *Information Theory and Applications (ITA), Inaugural Workshop*, Feb. 2006.

[54] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 53, pp. 581–586, Mar. 2004.

[55] A. B. Carleial, "Outer bounds on the capacity of interference channels," *IEEE Trans. Inf. Theory*, vol. 29, no. 1, pp. 602–606, July 1983.

[56] M. H. M. Costa and A. A. E. Gamal, "The capacity region of the discrete memoryless interference channel with strong interference," *IEEE Trans. Inf. Theory*, vol. 33, no. 5, pp. 710–711, Sept. 1987.

[57] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *Annals of Probability*, vol. 2, no. 5, pp. 805–814, 1974.

[58] E. Telatar, "Capacity of multi-antenna Gaussian channels," in *Europ. Trans. Telecommunications*, vol. 10, Nov. 1999, pp. 585–595.

[59] S. V. G. Kramer, S. Shamai(Shitz), S. Jafar, and A. Goldsmith, "Capacity bounds for Gaussian vector broadcast channels," *Multiantenna Channels: Capacity, Coding and Signal Processing, DIMACS Workshop on Signal Processing for Wireless Transmission, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, G.J. Foschini and S. Verdu, eds.*, vol. 62, pp. 107–122, Oct. 2002.

[60] N. Jindal, U. Mitra, and A. Goldsmith, "Capacity of ad-hoc networks with node cooperation," in *IEEE Int. Symp. Inf. Theory*, 2004, p. 271.

[61] C. Ng and A. Goldsmith, "Transmitter cooperation in ad-hoc wireless networks: Does dirty-payer coding beat relaying?" in *IEEE Information Theory Workshop*, Oct. 2004.

[62] A. Høst-Madsen, "On the capacity of cooperative diversity," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1522–1544, apr 2006.

[63] ——, "A new achievable rate for cooperative diversity based on generalized writing on dirty paper," in *IEEE Int. Symp. Inf. Theory*, June 2003, p. 317.

[64] ——, "On the achievable rate for receiver cooperation in ad-hoc networks," in *IEEE Int. Symp. Inf. Theory*, June 2004, p. 272.

[65] F. M. J. Willems, "The discrete memoryless multiple channel with partially cooperating encoders," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 441–445, May 1983.

[66] I. Maric, R. D. Yates, and G. Kramer, "The discrete memoryless compound multiple access channel with conferencing encoders," in *IEEE Int. Symp. Inf. Theory*, Sept. 2005.

[67] ——, "The strong interference channel with common information," in *Allerton Conference on Communications, Control and Computing*, Sept. 2005.

[68] ——, "The capacity region of the strong interference channel with common information," in *Asilomar Conference On Signals, Systems and Computers*, Nov. 2005.

[69] ——, "The strong interference channel with unidirectional cooperation," in *Information Theory and Applications (ITA), Inaugural Workshop*, Feb. 2006.

[70] ——, "The strong interference channel with unidirectional cooperation," in *MSRI Workshop: Mathematics of Relaying and Cooperation in Communication Networks, poster session*, Apr. 2006.

[71] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, June 2002.

[72] K. Sohrabi, J. Gao, V. Ailawadhi, and G. Pottie, "Protocols for self-organization of a wireless sensor network," *IEEE Personal Communications*, pp. 16–27, Oct. 2000.

[73] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 1019–1030, Sept. 1990.

[74] G. Caire and D. Tuninetti, "The throughput of hybrid-ARQ protocols for the Gaussian collision channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1971–1988, July 2001.

[75] I. Maric and R. D. Yates, "Performance of repetition codes and punctured codes for accumulative broadcast," in *Proc. of the Modeling and Optimization in Mobile, Ad Hoc and Wireless networks Workshop (WiOpt'03), INRIA Sophia Antipolis, France*, Mar. 2003.

[76] A. Goldsmith and P. Varaiya, "Capacity of fading channels with channel state information," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.

[77] L. H. Ozarow, S. S. (Shitz), and D. Wyner, "Information theoretic consideration for celluar mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 359–378, May 1984.

[78] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity - part I: System description," *IEEE Trans. on Communications*, vol. 51, no. 11, pp. 1927–1938, Nov. 2004.

[79] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1073–1096, May 2003.

[80] K. Azarian, H. E. Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4152–4172, Dec. 2005.

[81] T. Cover and J. Thomas, *Elements of Information Theory.* John Wiley Sons, Inc., 1991.

[82] G. Caire, D. Tuninetti, and S. Verdú, "Suboptimality of TDMA in the low-power regime," *IEEE Trans. Inf. Theory*, vol. 50, no. 4, pp. 608–620, Apr. 2004.

[83] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity in wireless networks," *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming*, 1998.

[84] S. Narayanaswamy, V. K. R. S. Sreenivas, and P. R. Kumar, "Power control in ad-hoc networks: Theory, architecture, algorithm and implementation of the COMPOW protocol," in *Proc. of European Wireless 2002*, Feb. 2002, pp. 179–186.

[85] O. Oyman and A. J. Paulraj, "Spectral efficency of relay networks in the power-limited regime," in *Proc. of Allerton Conference on Communication, Control and Computing*, Sept. 2004.

[86] B. Schein, "Distributed coordination in network information theory," in *Ph.D dissertation, Massachusetts Institute of Technology*, Sept. 2001.

[87] G. Strang, *Linear Algebra and its applications.* Harcourt Brace Jovanovich College Publishers, 1988.

[88] J. A. Thomas, "Feedback can at most double Gaussian multiple access channel capacity," *IEEE Trans. Inf. Theory*, vol. 33, pp. 711–716, Sept. 1987.

[89] S. Avestimehr and D. Tse, "Outage optimal relaying in the low SNR regime," in *Proc. of IEEE Int. Symp. Inf. Theory*, Sept. 2005.

[90] J. Körner and K. Marton, *Comparison of two noisy channels*, topics in information theory ed., I. Csiszár and P. Elias, Eds. Colloquia Mathematica Societatis Janos Bolyai, Amsterdam, The Netherlands: North Holland, 1977.

[91] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. on Inf. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.

[92] W. Wu, S. Vishwanath, and A. Arapostathis, "On the capacity of Gaussian weak interference channels with degraded message sets," *Conf. on Inf. Sciences and Systems (CISS)*, Mar. 2006.

[93] A. Catovic, S. Tekinay, and T. Otsu, "Reducing transmit power and extending network lifetime via user cooperation in the next generation wireless multihop networks," *Journal on Communications and Networks*, vol. 4, no. 4, pp. 351–362, Dec. 2002.

[94] B. Zhao and M. Valenti, "Practical relay networks: A generalization of hybrid-ARQ," *IEEE J. Selec. Area in Comm.*, vol. 23, no. 1, pp. 1089–98, Jan. 2005.

[95] S. Cui, A. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques for sensor networks," *IEEE JSAC Special Issue on Fundamental Performance Limits of Wireless Sensor Networks*, vol. 22, no. 6, pp. 1089–98, Aug. 2004.