SIGNAL SUPERPOSITION STRATEGIES IN WIRELESS NETWORKS: THEORY AND EXPERIMENTS

A Dissertation

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

by

Sundaram Vanka

Martin Haenggi, Director

Graduate Program in Electrical Engineering
Notre Dame, Indiana
May 2012

SIGNAL SUPERPOSITION STRATEGIES IN WIRELESS NETWORKS: THEORY AND EXPERIMENTS

Abstract

by

Sundaram Vanka

Physical layer models for node cooperation study small groups of nodes whose operation is analyzed and optimized in isolation from the rest of the wireless network. Current analytical tools provide limited insight into how such techniques impact large wireless networks, where interference from other nodes in the network is a significant concern. Consequently, very little is known about how these techniques impact higher layers, which, in practice, manage this interference.

A recurring theme in many such techniques is signal superposition, where different transmissions on a common communication medium mutually interfere at one or more receivers. Each receiver recovers its message(s) of interest by optimally exploiting its knowledge of the codebooks of different interferers. Carefully designed signal superposition techniques are in fact optimal for certain types of one-to-many ("broadcast"), many-to-one ("multiple-access") and certain cases of relay-aided communication. These techniques stand in contrast to more traditional orthogonal schemes that are designed specifically to avoid such interference.

We examine signal superposition strategies in two canonical cases: broadcast and multiple-access. Our investigations involve a combination of theoretical analysis and experimental prototyping. In our theoretical study we employ tools from

stochastic geometry to analyze the medium access problem in networks composed of node clusters with local broadcast or multiple-access influencing each another through interference. We show that in networks composed of many randomly-placed clusters, each with local broadcast or multiple-access, orthogonal schemes offer useful properties such as their flexibility in adapting their spatial re-use to each receiver (broadcast) or a smaller spatial contention (multiple-access).

We show how a single broadcast cluster can be realized by designing the first known prototype of a superposition-coded wireless system using off-the-shelf channel codes and experimentally demonstrate the spectral efficiency gains over time division multiplexing for a fixed error rate. Furthermore, we use this prototype to show that the coding gain (rather than the spectral efficiency gain) from superposition codes significantly improves link reliability without the need to increase transmit power or bandwidth, opening up the possibility of novel medium access protocols that can leverage superposition codes.

To my parents,

the late Rajeswara Rao Tatagaru,

and my brother Rajeshwar

CONTENTS

FIGURES

TABLES

ACKNOWLEDGMENTS

First, I express my sincere gratitude to Professor Martin Haenggi for his guidance and encouragement. Starting from his class on communication networks, the past five years of our interactions have been one of the most intellectually rewarding periods of my life. They have increased my appreciation of clear thinking, attention to detail, and mathematical rigor. His immense faith in my abilities is reflected in the diverse research topics I was able to work on. And yet, on each of these topics, he has always found the time to patiently listen to my ideas and provide invaluable inputs and advice. Along with his friendly disposition, this has greatly enriched my experience as a graduate student. I couldn't have asked for a better advisor.

I thank Professors Thomas Fuja, Nicholas Laneman and Paolo Minero for serving on my doctoral committee and for their feedback on this thesis. I also thank Professor Fuja for his help and advice during my move to ND and in my first year. I am also grateful to Professor Laneman for his excellent class on basic information theory, and to Professor Minero for letting me attend his insightful lectures on multiuser information theory. I also thank Professor Vijay Gupta for his useful inputs during our collaboration on the consensus problem.

Looking back, I will always cherish the discussions with my good friends and labmates Sunil Srinivasa, Radha Krishna Ganti, Zhen Tong, and Kostas Stamatiou. Our discussion topics included, but were certainly not restricted to, wireless

networks :) I also thank the entire software-radio team for their efforts. In particular, I thank Sunil and Radha for helping me with the Matlab model and acknowledge Sunil's invaluable help in conducting experiments. Thanks to Peter Vizi and Zhenhua Gong for helping us implement and debug the prototype on GNURadio/USRP—Peter has been a great source of inspiration for rookie programmers like me! I also thank David Tisza, Chia-han Lee, Daniele Puccinelli, and Xinchen Zhang for their insights from time to time. I have many happy memories from the times I spent with my (past and current) roommates and friends—Srinath, Vel, Suu, Rahul, Quyet, Yuan, Krishnan, Utsaw, Mallu, Danny, Tiwari. I will always remember my many discussions with Shivaprasad Kotagiri during my first year.

Last, but not the least, I am forever grateful to my parents Professors Sastry and Satyavathi Vanka. Needless to say, what I am today is the result of their efforts to nurture my interest in science and mathematics from an early age. Nannagaru, I will always remember our discussions about operational amplifiers, linear systems and Hilbert spaces. I thank my grandfather the late Professor Rajeswara Rao Nishtala for his encouragement; his work ethic was truly inspiring. Along with my brother (soon-to-be-Dr.) Rajeshwar Vanka, my family has been the source of great love, support, and encouragement throughout my life.

CHAPTER 1

INTRODUCTION

1.1 The Cross-Layer Implications of Cooperation

The design of wireless networks with ever-growing size and complexity requires a certain amount of collaboration among nodes that share a common communication to achieve the design objective. Such colloboration strategies are typically tailored to improve specific layers in the protocol stack. Consequently, these protocols are studied based on models and assumptions that are deemed valid for that layer. For instance, cooperative diversity [1–3] or cooperative beamforming [4] are often studied from a physical layer perspective assuming small groups of fixed backlogged nodes with fixed destinations. This view implicitly assumes that the desired topology has been established and maintained for several rounds of physical layer communication in a larger network. It is not immediately clear how non-trivial higher layer concerns, such as, for example, medium access delays or the route formation overhead can be incorporated into this framework. Conversely, very little is known about how specific tasks at higher layers could leverage cooperative techniques at the physical layer.

On a different level of abstraction, cooperative algorithms running at the application layer—such as distributed algorithms for computation (e.g., [5], detection, estimation and control (e.g., [6, 7])—often use simplified models to characterize

the behavior of the lower layers in the protocol stack. For instance, [5] studies in-network function computation in sensor networks with one or more sink nodes. Using a simple model for unrelibale inter-node communication, the authors derive scaling laws for the computational throughput (i.e., the rate at which a function can be "refreshed" using new informaton). The relevance of this scaling result to actual wireless neworks depends crucially on whether the communication model provides the correct level of abstraction to model delays from coding, routing, medium access or ARQ. A similar problem arises in sensor-actuator systems where delay-sensitive information is carried over wireless networks, in applications ranging from real-time environmental monitoring to the control of vehicle formations or industrial plants.

As the above examples illustrate, there is a need to move beyond, or at least refine, the "separation-of-concerns" approach, if only to analyze the impact of choosing a cooperative scheme at one layer on other layers. A more holistic picture also opens up the possibility of exploiting cooperation at one layer in the design of other layers. Our work has been motivated by the need to study these problems in large wireless networks. In view of the impracticality of conveying the information required for collaboration, such as channel state information, at negligible delay and high reliability over large distances, we will assume that cooperation is restricted to small clusters of nearby nodes. Hence, for the rest of the network, the signals emitted by a *cooperative cluster* have to be treated as interference. Modeling this interference in turn requires modeling network geometry, as we discuss next.

1.2 Physical Layer Cooperation Schemes based on Signal Superposition

In this thesis we focus on physical layer cooperation schemes that require signal superposition, where different transmissions on a common communication medium mutually interfere at one or more receivers. Each receiver recovers its message(s) of interest by optimally exploiting its knowledge of the codebooks of different interferers. Carefully designed signal superposition techniques are in fact optimal for certain types of one-to-many ("broadcast"), many-to-one ("multiple-access") and certain cases of relay-aided communication. These techniques stand in contrast to more traditional orthogonal schemes that are designed specifically to avoid such interference.

We examine signal superposition strategies in two canonical cases: broadcast and multiple-access. Our investigations involve a combination of theoretical analysis and experimental prototyping. In our theoretical study we employ tools from stochastic geometry to analyze the medium access problem in large networks composed of node clusters with local broadcast or multiple-access influencing each another through interference. Confronted with the difficulties of analyzing large wireless networks such as cellular, mesh, ad hoc, cognitive, and sensor networks, there has been significant interest in developing random spatial models to model the locations of interferers. This is analogous to how fading models capture the randomness in the locations of different reflectors. As a consequence *stochastic geometry* [8] has emerged as an essential tool in the analysis and design of large wireless systems. Instead of making statements about networks with a specific geometry, which limits their expressivity, stochastic geometric models assume the node placement to be governed by a stochastic point process and allow the designer to make predictions about the performance averaged over many network

realizations.

This approach is often more tractable if the network has more than a few nodes. It has been used to gain insights into some fundamental limits as well as guidelines for protocol design, see [9–12] and references therein. However these analyses have been restricted to networks with point-to-point communication (where all simultaneous transmissions result in mutual interference)—stochastic geometry has not been applied to analyze cooperative networks. Our main contribution is the use of stochastic geometric tools to analyze the cooperative techniques in large networks. In the course of our work, we have examined the following three important special cases of cooperation within each cluster:

1. Broadcast Clusters [13, 14] (Fig. 1.1). Each cluster has a single transmitter (Tx) that transmits different information to a set of receivers (Rxs) simultaneously. The cluster uses coding schemes for the information-theoretic broadcast channel.

2. Multiple-Access Clusters [13, 15]. Each cluster consists of a Rx that receives different information from a set Txs. The cluster uses coding schemes for the information-theoretic multiple-access channel.

3. Information Exchange Clusters [16–20] (Fig. 1.2). Each node exchanges its information with all its neighbors as specified by the network topology. This combines broadcast and multiple-access.

Our work also has an important experimental component that concerns the design and performance characterization [21–24] of a single cluster prototype for (1) on a GNURadio/USRP platform [25].

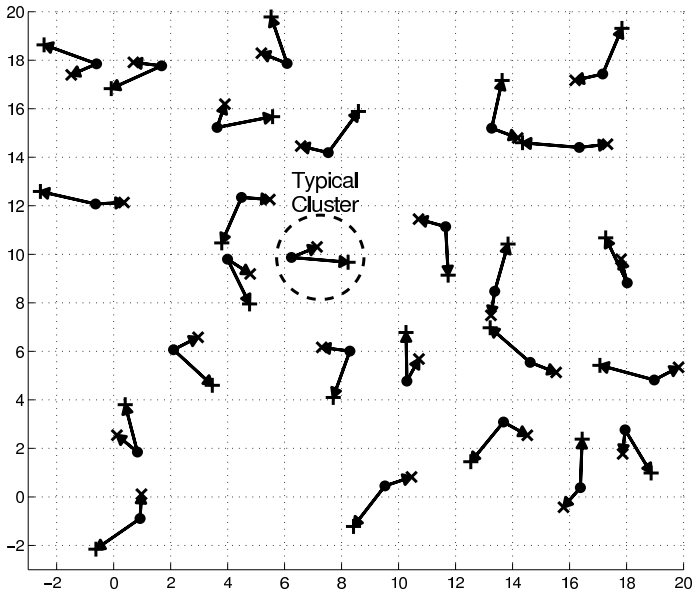

Figure 1.1. A network of two-Rx broadcast clusters. Each Tx (black circle) communicates with two Rxs (+ and ×), and the arrows denote the direction of communication. In a network of multiple-access clusters the positions of the Txs and Rxs are reversed. The cluster enclosed in the dashed circle denotes the typical cluster.

1.3 Organization of the Thesis

The detailed organization of this thesis is as follows.

In Chapter 2 we present the design and implementation of a software-radio system for two-Rx Superposition Coding (SC), the capacity-achieving coding scheme for the Gaussian broadcast channel [26], using a library of off-the-shelf point-to-point channel codes. We experimentally determine the set of rate-pairs achieved by this transmission scheme under a packet-error constraint. Our results suggest that SC can provide substantial gains in spectral efficiencies over those achieved by orthogonal schemes such as Time Division (TD). Our findings also question the practical utility of the Gaussian approximation for the inter-user interference

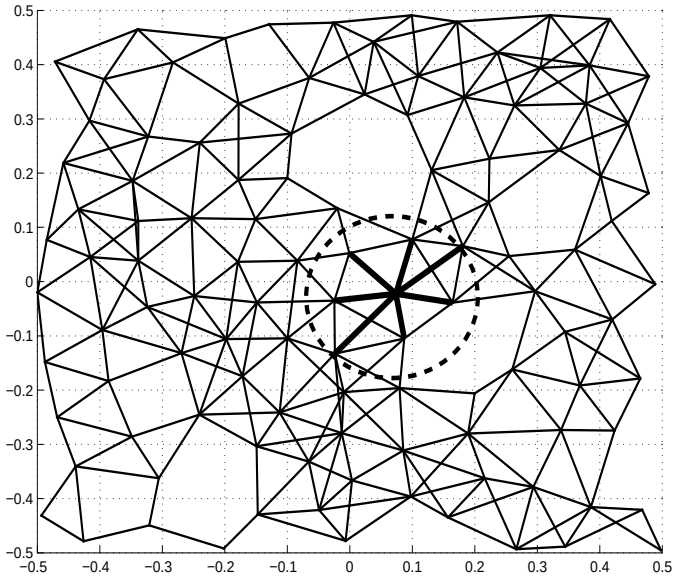

Figure 1.2. A network with information-exchange clusters. All edges denote bidirectional links. The bold lines denote the edges in the typical cluster (enclosed in a dashed circle). Clearly, this communication requires broadcast *and* multiple-access in the presence of interference.

in superposition-coded systems. This forms [23].

In contrast to Chapter 2, Chapter 3 (forming [22]) compares SC and TD as two channel coding schemes that encode at the same transmission rates. In particular, using the platform in Chapter 2 we show how SC can improve the coverage of far-away Rxs in an experimental scenario that approximates a Gaussian broadcast channel. Although this "coding-theoretic" view is equivalent to the more traditional view of using SC to increase transmission rates, it nonetheless provides a powerful yet practical approach to use channel codes based on signal superposition in networks with interference. This is explored in detail in [13].

In Chapter 4 we study the problem of medium access in a network of many mutually intefering broadcast clusters. In such a network, each link that a trans-

mitter communicates on has a different level of noise and interference depending on the link distance and the presence of nearby interferers. A medium access protocol must thus manage the interference seen on each link. When a transmitter orthogonalizes its signals to its receivers, the medium access problem decouples into multiple point-to-point medium access problems, thereby permitting the protocol to adjust the spatial re-use of each link independently of others. This is no longer true if transmitters use a non-orthogonal scheme such as SC. We model this effect by introducing a network utility network function to measure the effective rates of information transfer in space offered by the two schemes in the presence of interference. Due to the decoupling achieved by orthogonal schemes, maximizing this utility function reduces to individually maximizing each sub-function. With SC, however, this maximum is decided by the relative contribution of each link to the utility function. Since optimal spatial reuse is a function of the network geometry, the relative benefit of SC depend on the receiver placement.

In Chapter 5 we study the medium access problem for a network of mutually-interfering multiple-access clusters. In addition to the differing levels of noise on each link as in the broadcast case, allowing signal superposition in many multiple access clusters also increases network interference. In this chapter, we model this latter effect using a signal superposition strategy based on the capacity-achieving scheme for the Gaussian multiple access channel. We compare it with a more traditional TD scheme where the transmitters in a cluster take turns to communicate with their common receiver. Due to the increased interference from signal superposition, this form of superposition has little or no benefit when compared to TD. We also find that the same problem reduces the efficacy of successive decoding.

CHAPTER 2

PRACTICAL SUPERPOSITION CODING STRATEGIES

2.1 Introduction

2.1.1 Motivation and Prior Work

The problem of communicating with many receivers arises in many "downlink" scenarios such as communication from an access point to stations in WiFi or from a base station to mobile users in cellular systems. The conventional approach is to set up orthogonal channels to each user by time/frequency/code-division multiplexing. Although this approach eliminates interference between transmissions, it does not in general achieve the highest possible transmission rates for a given packet error rate (or reliability) [27]. In fact, Superposition Coding (SC) [26] is a well-known non-orthogonal scheme that achieves the capacity on a scalar Gaussian broadcast channel.

We motivate the use of SC for the two-receiver case. Consider a cellular downlink with several active users. Given the user density in typical networks, it is always possible to pick two users N (the "near" user) and F (the "far" user), as shown in Fig. 1. The key observation here is that N being geographically closer to the base station (BS) has a "stronger" (less noisy) link to the BS than F; thus any packet that can be decoded at F can most probably be decoded at N as well (but not vice versa). The idea behind SC is to optimally exploit this channel ordering.

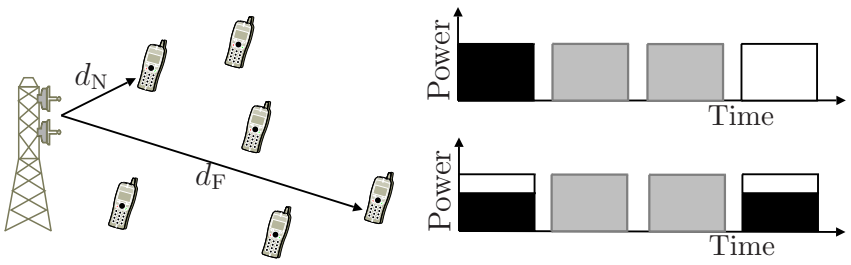

Figure 2.1. Illustration of two-user SC. (Left) The users N and F picked are at distances d_{N} and d_{F} respectively with $d_{\mathrm{N}} < d_{\mathrm{F}}$. (Right) Typical transmission timelines with and without SC. The gray slots represent transmissions to other active users which can remain unchanged. With Time-Division (TD, top), N and F are served in different slots (black and white). With SC (bottom), the BS transmits a linear combination of individually-coded user waveforms.

A BS that uses two-receiver SC can transmit *superimposed* F and N packets (or more precisely, the far and near user codewords) in *both* F's and N's time slots (see Fig. 1). Thus both links enjoy the *combined* degrees of freedom available to N and F, while sharing the transmit power. For large blocklengths, it can be shown that it is possible to encode F's packets such that they can be decoded in the presence of interference from N's packets. Since N has a stronger link to the BS, N can replicate this step to regenerate and thereby cancel F's signal from its received signal. It can then decode its own packet. This is the well-known *Successive Decoding* (SD) or *Successive Interference Cancellation* (SIC) procedure [27].

We can extend this two-user scheme to any number of users. In fact, SC (combined with SD) achieves the capacity on a scalar Gaussian broadcast channel. This implies that *any* TD-achievable rate-pair (i.e., the pair of *spectral efficiencies* on a Gaussian channel) can also be achieved using SC, with the rate gain over TD increasing with the disparity in the user link qualities.

While information theory sufficiently motivates the use of SC, it is largely

silent on practical issues such as finite block length codes, finite encoding and decoding complexity, hardware non-idealities (e.g., carrier frequency offset, phase noise) that one would encounter while designing such a system. This motivates the experimental study of SC.

For rapid prototyping and streamlining the design effort, we adopt a software-defined radio (SDR) [28] paradigm using the well-known open-source GNU Radio platform in conjunction with the Universal Software Radio Peripheral (USRP) hardware board that serves as an analog and RF front-end [25]. A well-known prototyping system [29, 30], it has been recently used in testbed design, including UT Austin's Hydra [31] and by Bell Labs and Microsoft Research [32]. In [32], the authors analyze how the benefits of superposition coding and successive cancellation could be harnessed at higher layers in the network stack. This difference in perspective has the unfortunate consequence that key physical layer details are neither sufficiently precise nor comprehensive enough to reproduce the results being reported.

2.1.2 Main Contributions

The author has led a team in building the first known prototype of a superposition-coded wireless system. In doing so, the author has:

1. Proposed practical superposition coding strategies using point-to-point channel codes.

2. Designed a complete system architecture that implements these strategies for BICM codes on the GNURadio/USRP platform.

3. Developed Matlab models to realize this architecture and provided extensive

assistance to the GNURadio programming team in code integration, testing and debugging.

4. Developed efficient experimental procedures that measure the achieved set of spectral efficiencies under a packet-error constraint.

5. Proposed experimental procedures that quantify the effect of non-Gaussian interference from signal superposition.

6. Proposed practical approaches and their corresponding experimental procedures to leverage the coding gain (rather than the spectral efficiency gain) from superposition codes.

2.1.3 Chapter Overview

The remainder of the chapter is organized as follows. In Section 2.2 we briefly summarize how SC achieves capacity and discuss some implications of restricting the library of codes to a finite set of finite-blocklength codes. In Section 2.3, while retaining the rate-centric approach, we propose a design technique for SC with such a finite code library and specialize this technique to a library comprised of a well-known family of codes designed using the Bit Interleaved Coded Modulation (BICM) technique [33], and predict the theoretically achievable rate region. In Section 2.4, we describe the system architecture that uses these BICM codes to implement SC. In Section 2.5, we present an experimental setup that emulates a Gaussian BC and use it to experimentally determine the achievable spectral efficiency pairs for a two-receiver BC under a packet-error constraint. The resulting rate region is the finite-library analog of the information-theoretic rate region. We also discuss some practical issues that arise in the design of superposition-

coded systems, including the validity of treating inter-user interference as Gaussian noise. In Section 2.6 we conclude the chapter and suggest possible avenues for future work.

2.2 SC: From Theory to Practice

We will briefly summarize relevant results from [26, 27] on achieving the capacity of the (scalar) Gaussian broadcast channel (BC) using SC with SD. In addition to making this chapter self-contained, this discussion identifies the key architectural building blocks of a superposition-coded system. A closer examination of the blocks allows us to identify some key practical issues in implementing this ideal scheme.

We use calligraphic fonts (e.g., $\mathcal{C}$) to represent sets and sans-serif fonts (e.g., $\mathsf{f}(\cdot)$) to denote the encoding/decoding maps. Also, we use $[M]$ to represent $\{1, \cdots, M\}$ for $M \in \mathbb{Z}^+$, and occasionally use the short-hand Tx for a transmitter and Rx for a receiver.

2.2.1 Achieving the Capacity on the Gaussian BC

Consider a BS that wants to communicate with two receivers N and F. The broadcast nature of the wireless medium is captured by the broadcast channel model $X \rightarrow (Y_{\mathrm{N}}, Y_{\mathrm{F}})$ where X denotes the channel input and Y_{N} and Y_{F} are the channel outputs at N and F[1]. Let $(X(n))$ be a sequence of channel inputs indexed by the channel use $n \in [L]$. Clearly $(X(n))$ must encode information relevant to each user. The *capacity region* of this channel is the closure of the set

[1]In practical terms, $X(n)$ can be understood as a (coded) symbol stream from the BS, and the Y's as the corresponding noisy and/or distorted observations of this symbol stream at N and F.

of all possible pairs of transmission rates at which the BS can reliably send two independent information streams, one each to N and F (allowing $L \to \infty$).

For a Gaussian BC, we have

$$Y_{\mathrm{N}}(n) = h_{\mathrm{N}} X(n) + Z_{\mathrm{N}}(n); \quad Y_{\mathrm{F}}(n) = h_{\mathrm{F}} X(n) + Z_{\mathrm{F}}(n) \tag{2.1}$$

where N (resp. F) has a complex channel gain h_{N} (resp. h_{F}) and $Z_{\mathrm{u}}, \mathrm{u} \in \{\mathrm{N}, \mathrm{F}\}$ denote the WGN processes. We assume the BS operates with an average power constraint P [W] and a (baseband) bandwidth W [Hz], and denote the noise power spectral density by N_0 [W/Hz]. From the above, the power constraint per channel use is P/W and $\mathbb{E}[|Z(n)|^2] = N_0 W$. From the definition of N and F, $|h_{\mathrm{N}}|^2 > |h_{\mathrm{F}}|^2$.

One way for $(X(n))$ to encode information is to communicate with each user in turns by partitioning the total number of channel uses into time slots (as in TD). For a given n, $X(n)$ contains information pertaining to just one user. This is the well-studied point-to-point communication problem, for which good practical encoding and decoding schemes exist. However, for a BC it is known that TDM is suboptimal in general; the root cause lies in its inability to fully exploit the fact that $|h_{\mathrm{N}}| > |h_{\mathrm{F}}|$: N has a "stronger" channel to BS, and hence can always decode information that can be decoded at F. This makes the scenario ideal for the SC scheme which achieves every pair of transmission rates in the capacity region.

The key architectural elements of an SC system are:

1. A *superposition encoder* f that consists of

   (a) Two point-to-point encoders, $\mathsf{f}_{\mathrm{N}} \colon \{0,1\}^{\lfloor 2^{LR_{\mathrm{N}}} \rfloor} \to \mathbb{C}^L$ (which we call the *near-encoder*) and $\mathsf{f}_{\mathrm{F}} \colon \{0,1\}^{\lfloor 2^{LR_{\mathrm{F}}} \rfloor} \to \mathbb{C}^L$ (which we call the *far-encoder*), that map their respective inputs (the *near-* and *far*-messages)

to complex-valued sequences $(X_{\mathrm{N}}(n))$ and $(X_{\mathrm{F}}(n))$, each of *block length* L. Here R_{N} and R_{F} denote the bandwidth-normalized transmission rates (or spectral efficiencies) of N and F (the *near-* and *far-rates* for short).

(b) A summation device that outputs a sequence

$$X(n) = \sqrt{1-\alpha}\, X_{\mathrm{F}}(n) + \sqrt{\alpha}\, X_{\mathrm{N}}(n), \tag{2.2}$$

where a fraction $\alpha \in [0,1]$ of the power is assigned to N (the *near-fraction* for short).

2. A *single-user* decoder $\mathsf{g}_{\mathrm{F}} : \mathbb{C}^L \to \{0,1\}^{2^{LR_{\mathrm{F}}}}$ that estimates the far packet from the observations $(Y_{\mathrm{F}}(n))$ by treating $(X_{\mathrm{N}}(n))$ as Gaussian noise.

3. A *successive cancellation* decoder $\mathsf{g}_{\mathrm{F,N}} : \mathbb{C}^L \to \{0,1\}^{2^{LR_{\mathrm{N}}}}$ that is used to recover N's packet in the following steps:

   (a) Decode F's packet using the single-user decoder g_{F}.

   (b) Cancel $\sqrt{1-\alpha}h_{\mathrm{N}}X_{\mathrm{F}}(n)$ from $Y_{\mathrm{N}}(n)$ by regenerating $X_{\mathrm{F}}(n)$ using the far-encoder f_{F} and the knowledge of h_{N} and α:

$$Y'_{\mathrm{N}}(n) = Y_{\mathrm{N}}(n) - h_{\mathrm{N}}\sqrt{1-\alpha}\, X_{\mathrm{F}}(n) = h_{\mathrm{N}}\sqrt{\alpha}\, X_{\mathrm{N}}(n) + Z_{\mathrm{N}}(n). \tag{2.3}$$

   (c) Decode N's packet using the single-user decoder $\mathsf{g}_{\mathrm{N}} : \mathbb{C}^L \to \{0,1\}^{2^{LR_{\mathrm{N}}}}$.

It is well known that as $L \to \infty$, for all α there exist $\mathsf{f}_{\mathrm{N}}, \mathsf{f}_{\mathrm{F}}, \mathsf{g}_{\mathrm{F}}, \mathsf{g}_{\mathrm{N}}$ such that communication can occur arbitrarily reliably for all pairs of transmission rates

satisfying

$$R_{\mathrm{N}} < W \log_2 \left(1 + \alpha\gamma_{\mathrm{N}}\right); \quad R_{\mathrm{F}} < W \log_2 \left(1 + \frac{(1-\alpha)\gamma_{\mathrm{F}}}{\alpha\gamma_{\mathrm{F}} + 1}\right), \qquad \alpha \in [0,1], \tag{2.4}$$

where

$$\gamma_{\mathrm{N}} \triangleq \frac{P|h_{\mathrm{N}}|^2}{N_0 W} \quad \text{and} \quad \gamma_{\mathrm{F}} \triangleq \frac{P|h_{\mathrm{F}}|^2}{N_0 W}$$

represent the *near-* and *far-SNRs* respectively. We are interested in the spectral efficiencies

$$r_{\mathrm{u}} = R_{\mathrm{u}}/W, \qquad \mathrm{u} \in \{\mathrm{N}, \mathrm{F}\} \tag{2.5}$$

which we will simply call *rates.* Clearly, making α discrete would also make the rate region boundary discrete. The following subsections elaborate on this issue.

2.2.2 Practical Design Issues

As noted above, a discrete α results in a discrete set of rate-pairs defining the corner points of the achievable rate region[2]. The lowest (resp. highest) value of α restricts the minimum (resp. maximum) power that can be assigned to a user for superposed transmission. For finite block lengths there is also a non-zero probability of decoding error.

With these practical constraints factored in, the rate benefits from SC over TD will depend on α, the chosen discrete set of codes and the system implementation. Using a combination of theory, simulations and experiments, we show that it is indeed possible to build efficient superposition-coded systems using off-the-shelf

[2]Time-sharing can be used to convexify this boundary.

single-user coding and decoding techniques.

2.3 Designing a Superposition-Coded System

Based on our observations in the previous section, we first describe our design approach in its full generality in Section 2.3.1. In Section 2.3.2 we illustrate this approach for a well-known single-user coded modulation technique known as Bit-Interleaved Coded Modulation (BICM). Indeed, good practical codes designed using BICM techniques form the basis for error correction in many real-world wireless networks (see, e.g., [33] and the references therein).

2.3.1 Practical SC using a Finite Code Library

We assume all codes have a *block length* $L < \infty$ [channel uses] and fix a *target Packet Error Rate (PER)* $\epsilon \ll 1$. The latter is the probability that a user cannot decode its packet[3]. Define a *code library* as a collection $\mathcal{C}$ of $M > 1$ *single-user* encoder-decoder function pairs[4] $(\mathsf{f}(\cdot), \mathsf{g}(\cdot))$ (a "code" for short). These codes are ordered by their rates $r_1 < r_2 < \cdots < r_M$ [bps/Hz]. We label each code by its *rate index* $i \in [M]$. We say a *rate index* i is ($\epsilon-$)*feasible* on a link if the receiver can decode a packet encoded at rate r_i with a PER no larger than ϵ. In the following, denote the user rate, the corresponding rate index and the PER by r_{u}, i_{u} and $\mathrm{PER}_{\mathrm{u}}$ for $\mathrm{u} \in \{\mathrm{N}, \mathrm{F}\}$.

Consider a BS communicating with N alone. When assigned the full BS power,

[3]We assume one packet is encoded as one codeword. Hence packet error is equivalent to codeword error.

[4]In practical terms, an encoding function is a mapping from the packet bits to the signal waveforms induced by the encoder and the modulator. A decoding function is the detection rule to perform the inverse operation, and is induced by the demodulator and decoder.

we assume this link has an SNR γ_{N} that allows the largest rate in the code is feasible[5]. Since lower rates require a smaller link SNR to remain feasible, the BS can backoff from this full power to support these rates. We define $\beta(l) \leq 1$ to be the *minimum* fraction of the full power for a rate index l to remain feasible at N, $l \in [M]$. To communicate with N at rates $r_l < r_M$, its link requires only a fraction $\beta(l)$ of the full BS power. Therefore, without violating the power constraint the BS can superpose a signal with power $1-\beta(l)$ to communicate with another user, assuming that N can perfectly cancel this interfering signal before decoding its own message[6].

We use this superposed signal to encode F's packets. The link to F has an SNR $\gamma_{\mathrm{F}} < \gamma_{\mathrm{N}}$ when assigned the full BS power. Suppose this allows rates no larger than r_K, $K < M$, to be feasible with full BS power. Clearly γ_{F} must be chosen with some care. If the far link is too noisy (γ_{F} is too small, perhaps because F is too far away from the BS), even a small amount of interference from N's signal (from superposition) can render even r_1 infeasible. If $\gamma_{\mathrm{F}} \lessapprox \gamma_{\mathrm{N}}$, there is not enough disparity between the links to N and F to take full advantage of signal superposition[7]. For comparison when BS serves N and F using TD, it uses rate index M to serve N and rate index K to serve F. In this case, the rate-pair is determined by the fraction of time slots assigned to each link.

Given $\mathcal{C}$ we would like to find the set of largest simultaneously achievable rates at N and F. Since $\mathcal{C}$ has only M possible rate indices, all of which are feasible at N, we only need to choose (at most) M values of power assignments to N. For each

[5]This models a proximate user in an urban WiFi or cellular network.

[6]In Section 2.5.3.2, we find that that this cancellation, although not perfect, does not require increasing N's signal power significantly beyond $\beta(l)$ for rate-indices $l \in [M]$.

[7]We study the implications of a poor choice of F in Section 2.5.3.1.

near-user rate r_l, $l \in [M]$, we maximize the far-user rate r_{F} such that the PER constraints at N and F are simultaneously satisfied. Given the near-user rate r_l, this can be stated as the optimization problem

$$\max_{\{r_1,\cdots,r_M\}} r_{\mathrm{F}} \quad \text{s.t.} \quad r_{\mathrm{N}} = r_l, \mathrm{PER}_{\mathrm{N}}(l) \leq \epsilon, \mathrm{PER}_{\mathrm{F}}(l) \leq \epsilon. \tag{2.6}$$

Clearly, the feasibility of any far-user rate depends on the distribution of power to N and F. Now, in theory, if interference from F's signal could be cancelled perfectly at N, the power assigned to N could be chosen just large enough to sustain r_l, i.e., no more than $\beta(l)$. However, in practice, due to imperfect cancellation, this power should be slightly larger. Hence, for each l, we could start by assigning $\beta(l)$ to N (and $1-\beta(l)$ to F) and increase the power assigned to N step-by-step until the PER constraint at N is satisfied (this procedure can be made more efficient, as explained in Sec. V-B). By this procedure, we determine the smallest near-user power $\alpha(l)$ that ensures that a rate r_l is feasible at N when F is assigned a power $1-\alpha(l)$. For this power assignment, we now increase the far-user rate until the PER constraint at F is satisfied. We denote by $i^*_{\mathrm{F}}(l)$ the rate index of the solution to (2.6). Assuming small levels of residual interference after cancelling F, $\alpha(l) \gtrapprox \beta(l), l \in [M]$[8].

The solution set of (2.6) is the set of optimal rate pairs $\{(r_l, r_{i^*_{\mathrm{F}}(l)}) : l \in [M]\}$. Combining the achievable endpoints $(r_M, 0)$ and $(0, r_K)$, we obtain the *rate region* of this code library $\mathcal{R}^*(\mathcal{C}; \epsilon, L, \gamma_{\mathrm{N}}, \gamma_{\mathrm{F}}) \triangleq \mathcal{R}^*$ as

$$\mathcal{R}^* \triangleq \mathrm{Conv}(\{(0,0), (r_M, 0), \{(r_l, r_{i^*_{\mathrm{F}}(l)}) : l \in [M]\}, (0, r_K)\}). \tag{2.7}$$

[8]We find that this is indeed the case in Sec. V-C2.

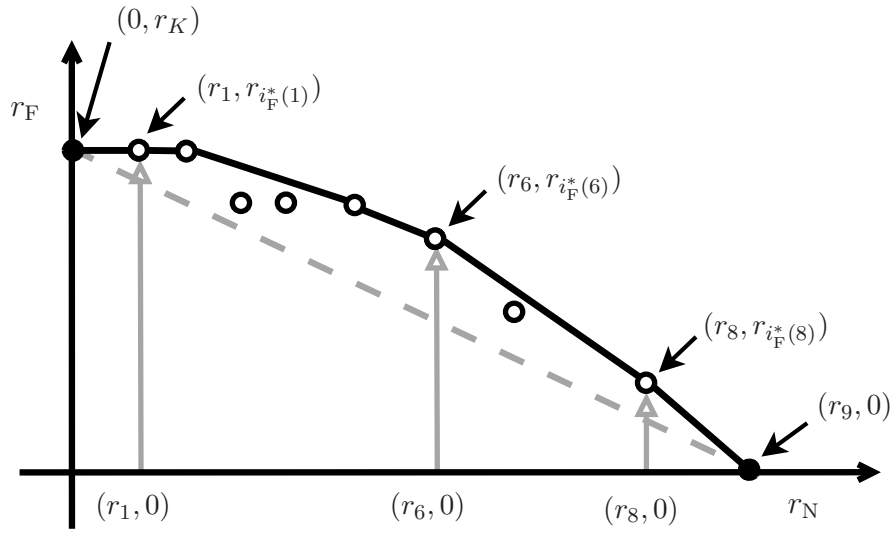

Figure 2.2. A graphical illustration of the rate region obtained by the design process in Section 2.3.1 for a hypothetical code library with $M = 9$ codes. The single-user points (filled dark circles) are $(r_9, 0)$ and $(0, r_K)$, for some $K \leq 9$. The gray dashed line joining these two points represents the rate region achievable by TD. The open circles represent the optimal rate-pairs obtained by solving (2.6) as N's rate is varied from r_1 through r_8. To obtain $\mathcal{R}$ using (2.7) we find the convex hull of this optimal set (along with the single-user points and the origin) to obtain a convex polygon with vertices $\{(r_l, r_{i^*_{\mathrm{F}}(l)}) : k \in \{2, 5, 6, 8\}\}$ inside the positive quadrant. The solid black lines represent the rate region boundary in this quadrant.

where $\mathrm{Conv}(\cdot)$ denotes the convex hull operator. Fig. 2.2 graphically summarizes the design procedure. In the following subsection we pose (2.6) for a code library consisting of point-to-point codes based on BICM.

2.3.2 SC using a BICM Code Library

Based on the framework developed so far, we revisit the key elements discussed in Section 2.2.1 to map them to specific subsystems in a Superposition-Coded-BICM (SC-BICM) system; we also introduce some terminology specific to BICM.

Subsequently we solve (2.6) numerically for BICM codes in the high-reliability regime ($\epsilon \to 0$). For reliabilities of practical interest (e.g., $\epsilon \lessapprox 0.1$), we solve this problem via Monte Carlo simulations for a specific class of BICM codes we have implemented on our testbed.

2.3.2.1 The SC-BICM System

The channel coding in a canonical BICM system [34] is specified by:

- c: Convolutional code with code-rate ρ_{c}, free distance d_{c} and number of free-distance error events w_{c}.
- x: A gray-coded constellation mapper that maps b_{x} interleaved coded bits mapped to each symbol. The symbols have unit average energy and a minimum intersymbol Euclidean distance of d_{x}[9]. We will refer to the range $\mathcal{S}_{\mathsf{x}}$ of x as its *constellation*. For example, the constellation of a BPSK mapper is $\{-1, 1\}$.
- Π: An interleaver matrix that specifies the order in which code bits are read by the constellation mapper. The corresponding de-interleaver is denoted by Π^{-1}.

Combining a convolutional code c with a modulator x results in a spectral efficiency of $r = \rho_{\mathsf{c}} b_{\mathsf{x}}/W$ [bps/Hz]. Thus the near-encoder f_{N} is the composition $\mathsf{x}_{\mathrm{N}} \circ \Pi_{\mathrm{N}} \circ \mathsf{c}_{\mathrm{N}}$. A block length L can encode $B = Lr$ information bits per codeword or packet. With a slight abuse of notation, we denote the Viterbi decoder and the constellation demapper by c^{-1} and x^{-1} respectively.

[9] The gain of the subsequent Tx stages controls the overall transmit power.

We can now describe the key functional units in an SC-BICM system: the superposition encoder f finds the weighted sum of the outputs of two separate BICM encoders f_{N} and f_{F} to construct X which we will now call the *composite* symbol stream. The composite symbol stream is mapped to a waveform and transmitted over the wireless medium. At F, Y_{F} in (2.1) represents the noisy observations of X seen by the constellation demapper (that is the first stage in g_{F}). From Y_{F} the demapper estimates the reliabilities of each encoded bit of the far-packet, treating the symbol stream $(X_{\mathrm{N}}(n))$ as interference. The remaining steps are the same as in standard BICM decoding.

Since g_{F} is part of the successive cancellation decoder $\mathsf{g}_{\mathrm{F,N}}$, these steps are reproduced at N; thereafter the far-encoder f_{F} (at N) reconstructs X_{F} from this decoded packet which is then subtracted from Y_{N} to yield Y'_{N} as in (2.3). In the final step g_{N} estimates the near-packet from Y'_{N} using standard BICM decoding.

2.3.2.2 A Theoretical Estimate of the Rate Region of an SC-BICM System in the High-Reliability Regime

Given $\mathcal{C}$, (2.6) can be solved by checking the feasibility[10] of every candidate rate pair via (time-consuming) Monte Carlo simulation. One could reduce this computational overhead by reducing the search space of possible far-rates for a given near-rate. In fact, a formula to compute the PER for a given rate pair and near-fraction would obviate the need for simulations.

Unfortunately, accurate formulas for the PER are difficult to obtain even for the point-to-point case, although there exist well-known upper bounds that are asymptotically tight in the high-reliability regime (PER $\to 0$) [35]. In the high-

[10]To any desired confidence interval.

reliability regime, these upper bounds can be suitably modified, as we will show.

The key difference between SC-BICM and point-to-point BICM lies in the far-demodulator $\mathsf{x}_{\mathrm{F}}^{-1}$ that estimates the reliabilities of the far-code bits from observations of the form

$$Y_{\mathrm{F}}(n) = \underbrace{h_{\mathrm{F}}\sqrt{1-\alpha}X_{\mathrm{F}}(n)}_{\text{Signal}} + \underbrace{h_{\mathrm{F}}\sqrt{\alpha}X_{\mathrm{N}}(n) + Z_{\mathrm{F}}(n)}_{\text{Perturbation}}, \quad n \in [L]. \tag{2.8}$$

Unlike in point-to-point BICM, the perturbation term is not Gaussian; its statistics depend on N's constellation, which we assume is known to the demodulator[11]. Whether or not x_{N} is known, the finiteness of the interference constellation raises an interesting question about the validity of the Interference-As-Gaussian-Noise (IAGN) model that assumes such Gaussianity. We investigate this question in greater detail in Section 2.5.3.3.

When N's constellation is known, it is useful to treat each composite symbol $X(n) = \sqrt{1-\alpha}X_{\mathrm{F}}(n) + \sqrt{\alpha}X_{\mathrm{N}}(n)$ as a member of a superconstellation with $2^{b_{\mathsf{x}_{\mathrm{N}}}+b_{\mathsf{x}_{\mathrm{F}}}}$ points (see also Fig. 2.7). Viewed from the demodulator, interference perturbs each original far-symbol (the *parent* point) to a randomly chosen *daughter* point. For each parent point, define the set of all possible daughter points to be its *potential daughter cluster* ("cluster" for short). The size and shape of this cluster depends on the interferer's constellation. Thus a maximum-likelihood demodulator interested only in the far-packet infers the most probable parent point of the observed (noisy) daughter point by identifying the most probable cluster to which an observation belongs. Identifying successively less probable clusters helps the demodulator refine its reliability estimate of each detected code

[11] In practice sending this information entails a small overhead, which we neglect in this chapter.

bit of F. Analogous to the single-user case, the reliability of the k^{th} bit in the n^{th} symbol can be approximated using the max-log-MAP approximation:

$$L_n^{(k)} \approx N_0 \min_{s \in \sqrt{1-\alpha}\mathcal{S}_{\mathsf{x}_\mathrm{F}}^{(k-)} \times \sqrt{\alpha}\mathcal{S}_{\mathsf{x}_\mathrm{N}}} |Y_\mathrm{F}(n) - h_\mathrm{F} s|^2 - \min_{s \in \sqrt{1-\alpha}\mathcal{S}_{\mathsf{x}_\mathrm{F}}^{(k+)} \times \sqrt{\alpha}\mathcal{S}_{\mathsf{x}_\mathrm{F}}} |Y_\mathrm{F}(n) - h_\mathrm{F} s|^2 \quad (2.9)$$

where $\mathcal{S}_{\mathsf{x}_\mathrm{F}}^{(k-)} \subset \mathcal{S}_{\mathsf{x}_\mathrm{F}}$ and $\mathcal{S}_{\mathsf{x}_\mathrm{F}}^{(k+)} \subset \mathcal{S}_{\mathsf{x}_\mathrm{F}}$ comprise symbols whose k^{th} bits are 0 and 1, respectively.

In the high-reliability regime, the dominant error events in such a demodulator are events where a daughter point is incorrectly identified with a neighboring cluster. Analogously to the point-to-point case (when these "clusters" are just points), the probability of these error events is controlled by the effective cluster separation

$$d_{\text{eff}}^{(\mathrm{F,N})} \equiv d_{\text{eff}}(\alpha, \mathsf{x}_\mathrm{F}, \mathsf{x}_\mathrm{N}) = \min_{\substack{p_1, p_2 \in \sqrt{1-\alpha}\mathsf{x}_\mathrm{F} \\ p_1 \neq p_2}} \min_{d_1, d_2 \in \sqrt{\alpha}\mathsf{x}_\mathrm{N}} |p_1 + d_1 - p_2 - d_2|. \quad (2.10)$$

Here the superscript (F,N) on the left hand side emphasizes that the parent points are drawn from x_F and the interferer is drawn from x_N. Using arguments similar to those in [34, 35], PER_F can be approximated as

$$\mathrm{PER}_\mathrm{F} \approx B_\mathrm{F} w_{\mathsf{c}_\mathrm{F}} Q\left(d_{\text{eff}}^{(\mathrm{F,N})} \sqrt{\frac{d_{\mathsf{c}_\mathrm{F}} \gamma_\mathrm{F}}{2}} \right) \quad (2.11)$$

in the high-reliability regime, where $Q(x)$, $x \geq 0$ is the Q-function. Assuming perfect cancellation at N,

$$\mathrm{PER}_\mathrm{N} \approx B_\mathrm{N} w_{\mathsf{c}_\mathrm{N}} Q\left(d_{\mathsf{x}_\mathrm{N}} \sqrt{\frac{d_{\mathsf{c}_\mathrm{N}} \gamma_\mathrm{N}}{2}} \right), \quad (2.12)$$

which can be plugged into (2.12) to approximate $\mathrm{PER}_{\mathrm{N}}$.

However, for PERs of practical interest (say PER $\lessapprox 0.1$) these Q-function bounds are too loose to predict the correct solutions to (2.6). In this regime we use these bounds as estimates that reduce the search space of the achievable far-rates, and refine them further via simulation. In the following subsection we illustrate these ideas with a design example. The code library chosen in the example is the same as the one used in our system design.

2.3.3 A Design Example

We use a BICM code library with a decoder structure explained in Section 2.3.2.2. These BICM codes were implemented in our testbed in a point-to-point setting.

The library consists of all possible pairings of 4 convolutional codes $\{\mathsf{c}^{(1)}, \mathsf{c}^{(2)}, \mathsf{c}^{(3)}, \mathsf{c}^{(4)}\}$ with three constellation mappers. There is no interleaving[12]. Table 2.1 summarizes the details of these convolutional codes. Note that $\mathsf{c}^{(2)}, \mathsf{c}^{(3)}$ and $\mathsf{c}^{(4)}$ are obtained by appropriately puncturing $c^{(1)}$, which is the standard rate-1/2 constraint length 7 convolutional code with the generator matrix $[133, 171]$. The constellations are $\mathsf{x} \in \{\mathsf{BPSK}, \mathsf{QPSK}, \mathsf{16QAM}\}$. These are all QAM constellation with even b_{x}, so $d_{\mathsf{x}} = \sqrt{6/(2^{b_{\mathsf{x}}} - 1)}$ [36]. With this code library the available set of spectral efficiencies is the sequence $(r_1, r_2, \ldots, r_{12})$ obtained by ordering the elements of the set $\{1, 2, 4\} \times \{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{5}{6}\} = \{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{5}{6}, 1, \frac{4}{3}, \frac{3}{2}, \frac{5}{3}, 2, \frac{8}{3}, 3, \frac{10}{3}\}$ in ascending order.

We now solve (2.6) via Monte Carlo simulation (using the simplified proce-

[12]Code performance can be optimized by suitably tailoring the interleaver Π. However its absence does not change the main message of this example (or indeed, that of the chapter), which is to show substantial gains from SC even for finite blocklengths and constellations.

TABLE 2.1

KEY PARAMETERS OF THE CONVOLUTIONAL CODES IN THE CODE LIBRARY.

Code	Rate ρ_{c}	Free Distance d_{c}	#Free Dist. Error Events w_{c}
$\mathsf{c}^{(1)}$	1/2	10	11
$\mathsf{c}^{(2)}$	2/3	6	1
$\mathsf{c}^{(3)}$	3/4	5	8
$\mathsf{c}^{(4)}$	5/6	4	14

dure outlined in Section 2.5.2) for $\epsilon = 0.1$ and $L = 1536$ with perfect receiver CSI and perfect interference cancellation. For our implementation this choice of L strikes a balance between code performance and implementation constraints. Besides illustrating the design procedure for a concrete example, the far-rates from this simulation provide upper bounds for a practical system (where neither of these conditions holds). The simulation procedure closely follows the experimental procedure described in Section 2.5.2. The first step involves obtaining the single-user PER curves similar to those in Fig. 2.5 by simulating a link operating over a point-to-point Gaussian channel. As with the experiment, these results yield $\{\beta(l)\}$, $l \in [M]$ that are given to another Matlab program that simulates a Gaussian BC and numerically solves (2.6) using the procedure described in Section 2.5.2. Fig. 2.3 shows the simulated rate region for $\gamma_{\mathrm{N}} = 15$ dB and $\gamma_{\mathrm{F}} = 8$ dB. For these parameters, it takes about 4 hours to complete this procedure on Matlab version 7.13 running on a dual-core Linux workstation running at 2.4GHz and with a 2 GB RAM.

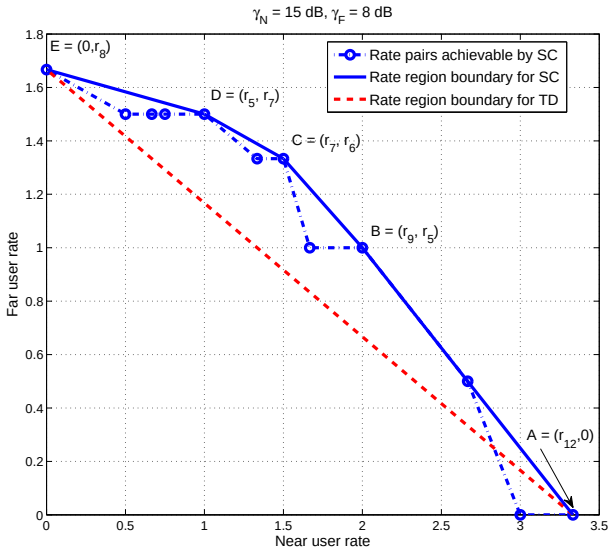

Figure 2.3. Optimal rate pairs (solution to (2.6)) for $\gamma_{\mathrm{N}} = 15$ dB and $\gamma_{\mathrm{F}} = 8$ dB for the library of $M = 12$ codes described in Section 2.3.3. The TD rates are obtained by time-sharing between the single-user operating points. The values of α at corner points A-E are $1, 0.2138, 0.1259, 0.0631, 0$ respectively.

2.4 Implementing a Superposition-Coded System

2.4.1 The Platform

We implemented all physical layer processing steps at BS, N and F by suitably modifying an existing point-to-point wireless testbed. The testbed uses Orthogonal Frequency Division Multiplexing (OFDM), and its design parameters are similar to those in the IEEE 802.11a standard. The testbed runs on GNU Radio (revision 10923) on a Linux PC. GNURadio provides driver functions that interface the PC with the USRP board that functions as the analog frontend and the RF.

2.4.2 Packet Structure

Transmissions occur in *frames*. As in WiFi, each frame consists of a preamble followed by a header and a payload. The preamble assists the receiver in frame acquisition and channel estimation. The header encodes the BICM code type and α. The preamble and the header designs are left unchanged from the single-user case[13]. The only "multiuser" section in the frame is the payload. For lack of space, we will not provide more details here; they can be found in our technical report [38].

The block of $L = 1536$ coded symbols is transmitted over 192 OFDM symbols with 8 subcarriers for the payload in each OFDM symbol. We retain the 4 pilot subcarriers used by the testbed for frequency and phase tracking and 4 null subcarriers for spectral shaping. This brings the total number of subcarriers to $8+4+4 = 16$. The message bandwidth of 2 MHz is limited by the USRP, and the cyclic prefix is made commensurate with the (relatively) flat frequency response of the channel for this bandwidth. We summarize the system parameters in Table 2.2 (see p. 28).

2.4.3 Single-User Characterization

Creating experimental conditions that ensure a time-invariant wireless propagation loss can be complicated: although the indoor radio channel and the USRP boards have a reasonably flat response over a 2 MHz bandwidth, the propagation loss is quite sensitive to changes in the environment (e.g., those caused by motion). One approach would be to compensate for such changes by appropriate

[13]Consequently, known techniques (e.g., the use of windowed correlators as in [37]) and low-rate header encoding (as in IEEE 802.11a) can be used.

TABLE 2.2

SYSTEM PARAMETERS.

Center Frequency	903 MHz
Message Bandwidth	2 MHz
Modulation	16-tone OFDM (8 data, 4 pilot, 4 null)
CP Length	1μs

power control at the BS. Doing so would require SNR feedback from the users on a control channel; designing such feedback links would be worth the effort only if SC could provide substantial rate gains in a *perfectly power-controlled* environment. We focus on the latter question in the chapter.

While manual power adjustments work at smaller time scales lasting a few minutes [22], they are quite cumbersome for longer experiments (e.g., measuring the PER for all the rates in the library, or solving (2.6), see Section 2.5.1).

To circumvent this problem, we emulate perfect power control by connecting the Tx and Rx with a coaxial cable (see Fig. 2.4), resulting in controllable experimental conditions and reproducible outcomes[14]. With this coaxial setup, we measure the point-to-point PER as a function of the transmit power P (which implicitly determines the SNR γ) for each rate index in the code library using parameters from Table 2.2.

Starting from a value of P chosen such that the received power (at the antenna port, as measured by a spectrum analyzer) is at a fixed level above the theoretically predicted thermal noise floor, we change the transmit power in 1 dB steps and

[14]Indeed, the same practical considerations motivate the use of channel emulators.

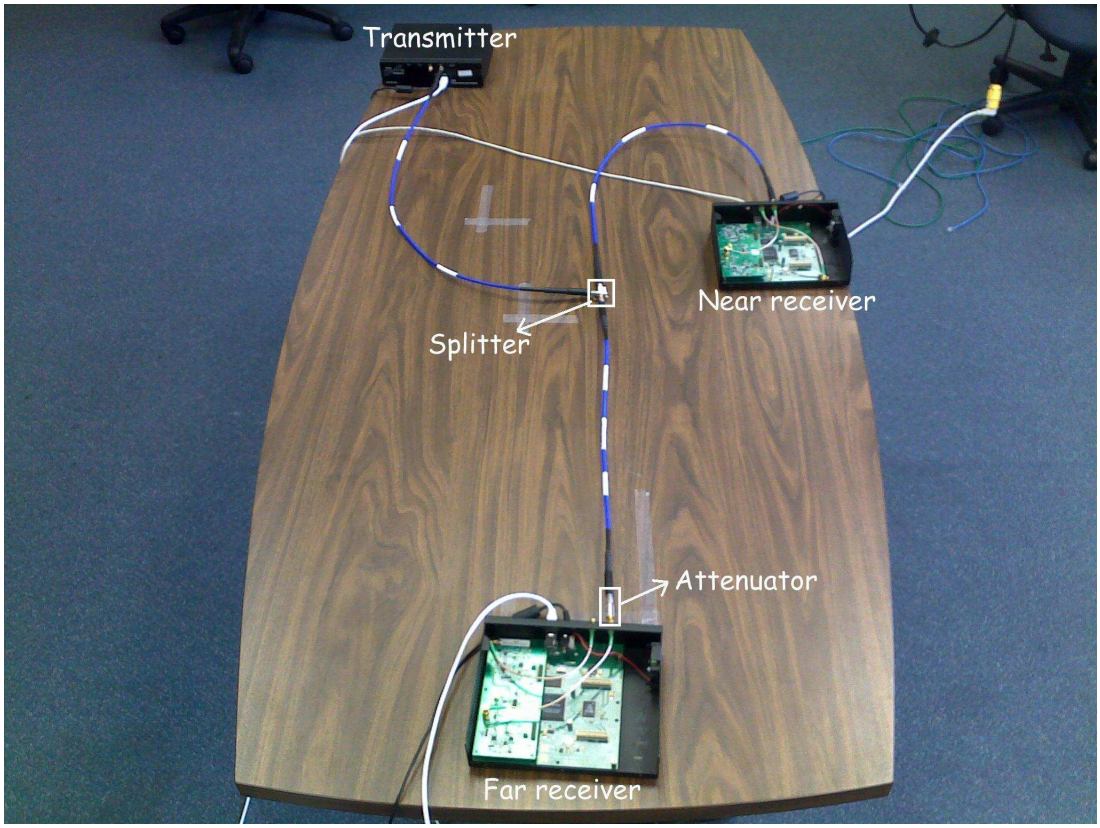

Figure 2.4. The setup used to approximate a Gaussian BC. The USRP boards are connected via cables. A splitter is used to split the transmitted signal to the two receivers, while an attenuator is used to (virtually) create the presence of a far user.

measure the PER for each rate index[15]. For comparison we overlay the PER plots for a simulated Gaussian system with the same theoretical noise floor but without additional sources of noise and distortion in Fig. 2.5. For a rate index l, let $P_{\min}(l;\epsilon) \equiv P_{\min}(l)$ be the smallest experimentally obtained power level that makes a link $\epsilon-$feasible for rate index l. Denote the corresponding power level by $\tilde{P}_{\min}(l)$ on the simulated Gaussian link.

Observe that the slopes of the ideal and experimental waterfall curves are similar up to PER $\gtrapprox$ 1%. At PER $\lessapprox$ 10%, the combined non-ideality of the hardware and implementaion result in a maximum power loss of $10\log\frac{P_{\min}(l)}{\tilde{P}_{\min}(l)} \leq 3.5\,\text{dB}, i \in [12]\ (= 3{\times}4)$ from the ideal results. Thus we take this "coaxial channel" to be a reasonable approximation of a Gaussian channel in our experiments.

Using these results we obtain the smallest feasible near-fractions $\beta(l)$ for N as

$$\beta(l) = \frac{P_{\min}(l)}{P_{\min}(12)}, \qquad l \in [12]. \tag{2.13}$$

We use $(\beta(l))$ as starting points to find $\alpha(l)$ in the rate region experiment in Section 2.5.2.

2.4.4 Modifications to Implement an SC-BICM System

The SC encoder f is implemented as two instances of a point-to-point BICM (sub-)encoder and a combiner. The encoded data is then transmitted via a standard OFDM modulator. The receiver is implemented as a single GNU Radio

[15]This mimics a system with a 1 dB granularity in power control. Also note that the effective signal distortion seen at the receiver is usually higher than that due to thermal noise alone, due to additional sources of noise and distortion such as imperfect receive implementation and hardware imperfections. In our experiments an initial power level of 20 dB above the theoretical value was found to be adequate.

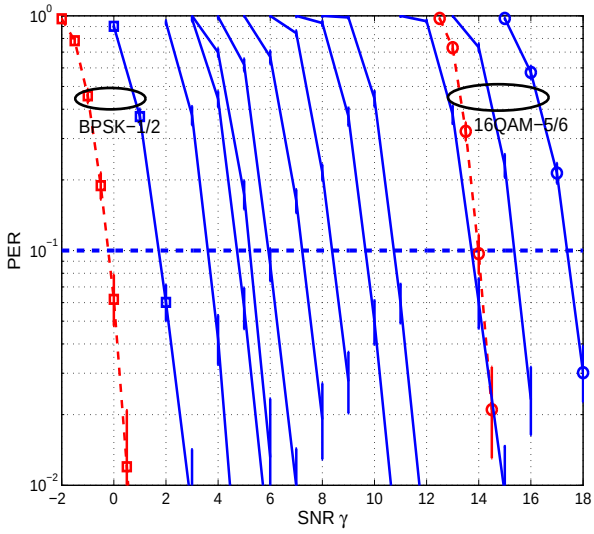

Figure 2.5. Single-user PER versus SNR, along with the 95% confidence intervals for all the 12 rates in the library. The solid lines depict the experimental results; the ideal curves for BPSK-1/2 and 16QAM-5/6 are also shown (dashed lines). Observe that at 10% PER the implementation loss is in the range 2dB − 3.5dB.

signal processing block and contains a successive decoding block to decode the near user packet. More details on the transceiver operation including their block diagrams may be found in our technical report [38].

2.4.5 The Runtime Requirement

The GNU Radio runs on a general purpose computer. In our experiment, we use the desktop with two quad-core Intel Xeon CPU E5520@2.27 GHz with 8 GB RAM. The transmitter code takes approximately 2.1 ms for the transmitter to generate a single packet. For the far receiver, the decoding procedure takes approximately 8.7 ms. For the near receiver, it needs approximately 19.7 ms to decode its packet. The near user needs to do the decoding twice and encoding once so that the complexity is more than doubled compared to the far user. The time is measured by recording the timestamps before and after encoding(decoding) a packet. The near user needs to do the decoding twice and encoding once so that the complexity is more than doubled compared to the far user. In the experiment, the transmitter has an idle gap of 100 ms between two packets. Hence the receiver has enough time to decode. More detailed discussion about implementation complexity is available in [38].

2.5 Experimental Results

2.5.1 Emulating a Gaussian BC

The time variation of the propagation loss is a bigger problem for the BC because of the presence of *two* Tx-Rx paths in the BC as opposed to just one in the point-to-point setting of Section 2.4.3. Moreover, checking the feasibility of candidate rate-pairs while solving (2.6) requires repetitive PER measurements.

Thus we emulate a Gaussian BC using a combination of coaxial cables, a splitter, and an attenuator bank, as shown in Fig. 2.4. The USRPs shown communicate with three Linux PCs via USB 2.0. The PCs are configured to run the appropriate Tx or Rx code in GNURadio.

The second step is to fix the single-user SNRs $\gamma_{\mathrm{N}}, \gamma_{\mathrm{F}}$. In our experiment we fix them implicitly by choosing a transmit power P_{N} and an attenuator setting a_{F}[16]. In the single-user mode, P_{N} is chosen to be $P_{\min}(12)$ (plus some additional loss due to the splitter) using the single user results from Section 2.4.3. The measured γ_{N} was found to be ≈ 18 dB (see [38] for details on SNR measurement). With this value of P_{N}, we set F in the single-user mode to operate at a desired rate (e.g., QPSK, rate-5/6), and increase the attenuation in steps of size $\eta = 1$ dB until the far link violates the PER constraint. γ_{F} is the receive SNR for the largest attenuation setting that supports a 10% (or lower) PER at F (for QPSK rate-5/6, this was found to be $\approx 10\,\mathrm{dB}$). With this approximation to the Gaussian BC, we are now ready to obtain the rate region for this system by solving (2.6) experimentally.

2.5.2 The Rate Region Experiment

Given $P_{\mathrm{N}}, a_{\mathrm{F}}$ and a code library $\mathcal{C}$ with rate-indices $l \in [M]$, we use the above experimental setup to solve (2.6) to find $(l, i_{\mathrm{F}}^*(l))$ using the following procedure:

1. Initialize:

   (a) $\alpha(l) = \beta(l)$ where $\beta(l)$ is calculated from (2.13).

   (b) $i_{\mathrm{F}}(l) = \arg\max_l \{P_{\min}(l) < (1 - \beta(l)) a_{\mathrm{F}} P_{\mathrm{N}}\}$.

[16] It is possible to obtain an estimate of γ_{N} by measuring the signal and noise powers in the digital domain (see [38]).

2. Calculate the PER for $(l, i_{\mathrm{F}}(l))$ for stream weights $\sqrt{\alpha(l)}$, $\sqrt{1-\alpha(l)}$.

3. If

   (a) (feasible at N) AND (feasible at F): set $i^*_{\mathrm{F}}(l) = i_{\mathrm{F}}(l)$, $k \mapsto l+1$, go to Step 1.

   (b) (feasible at N) AND (infeasible at F): $i_{\mathrm{F}}(l) \mapsto i_{\mathrm{F}}(l) - 1$, go to Step 2.

   (c) (infeasible at N) AND (feasible at F): $\alpha(l) \mapsto \alpha(l) \times 10^{\frac{\eta}{10}}$, go to Step 2.

   (d) (infeasible at N) AND (infeasible at F): $\alpha(l) \mapsto \alpha(l) \times 10^{\frac{\eta}{10}}$, $i_{\mathrm{F}}(l) \mapsto i_{\mathrm{F}}(l) - 1$, go to Step 2.

Here η is the step size for the α parameter (in this chapter, we set $\eta = 1$ dB).

2.5.3 Results

We use the procedure in Section 2.5.2 to study three interesting problems: (a) How does the measured rate region change with a_{F} (i.e., the far-link is made stronger or weaker)? (b) how much does imperfect interference cancellation at N affect the rate region?, and (c) what are good models to account for N's interference at F, and, in particular, how useful is the popular Interference-As-Gaussian-Noise (IAGN) model in predicting its impact? We discuss these problems in the following.

2.5.3.1 Changing the Strength of the Far Link

To study this problem, we find the rate region for two possible far-link SNRs: $\gamma_{\mathrm{F}} = 5$ dB and $\gamma_{\mathrm{F}} = 10$ dB, which correspond to single user rate indices $K = 3$ and $K = 8$ respectively. The near-user SNR for both cases is kept at $\gamma_{\mathrm{N}} = 18$

dB. These scenarios are emulated by using suitable attenuator values $a_{\mathrm{F}} = 9, 4$ respectively[17]. The results are shown in Fig. 2.6. Here, we used a transmit power $P_{\mathrm{N}} = -43$ dBm and step size $\eta = 1$ dB. We clearly see dependence on the choice of γ_{F}. With $\gamma_{\mathrm{F}} = 5$ dB, there is not enough disparity between the near- and far-links to fully benefit from superposition (F is "too close" to BS). This is in fact predicted by theory [27]. On the other hand, we see the effect of a finite code library when $K = 3$ (F is "too far" from BS): since its single-user rate is too small to begin with, interference from N's symbols rapidly degrades its link quality so as to make any far-rate infeasible.

Therefore, the far-user modulation and rate pair may be appropriately chosen based on the rate that the near user's traffic demands. For instance, when the near user's spectral efficiency is 1 (QPSK-1/2), choosing BPSK-3/4 for the far user provides a rate gain of about 28% over TD (as compared to a gain of about 21% over TD for QPSK-5/6), while when the near user's spectral efficiency is 3, it is preferable to choose QPSK-5/6 for the far user (over BPSK-3/4).

2.5.3.2 Impact of Imperfect Interference Cancellation at N

The deviation of $\alpha(l)$ from its ideal value $\beta(l)$ is a measure of the residual far-user interference seen at N due to the imperfect cancellation of F's symbols (even when the far-packet is decoded correctly). The β's are calculated from the single-user results using (2.13). The α's are determined from the experiment. For $K = 8$, we find that when $l = 6$, $\alpha(l) = 0.13 \geq \beta(l) = 0.1$. At small near-rates the desired symbol stream (of near-symbols) has a much lower power than the interference symbol stream (of far-symbols). In this regime, even small regeneration errors

[17]Note that a_{F} is not simply equal to $\gamma_{\mathrm{N}} - \gamma_{\mathrm{F}}$ owing to the two different boards and cables used, which had to be calibrated separately.

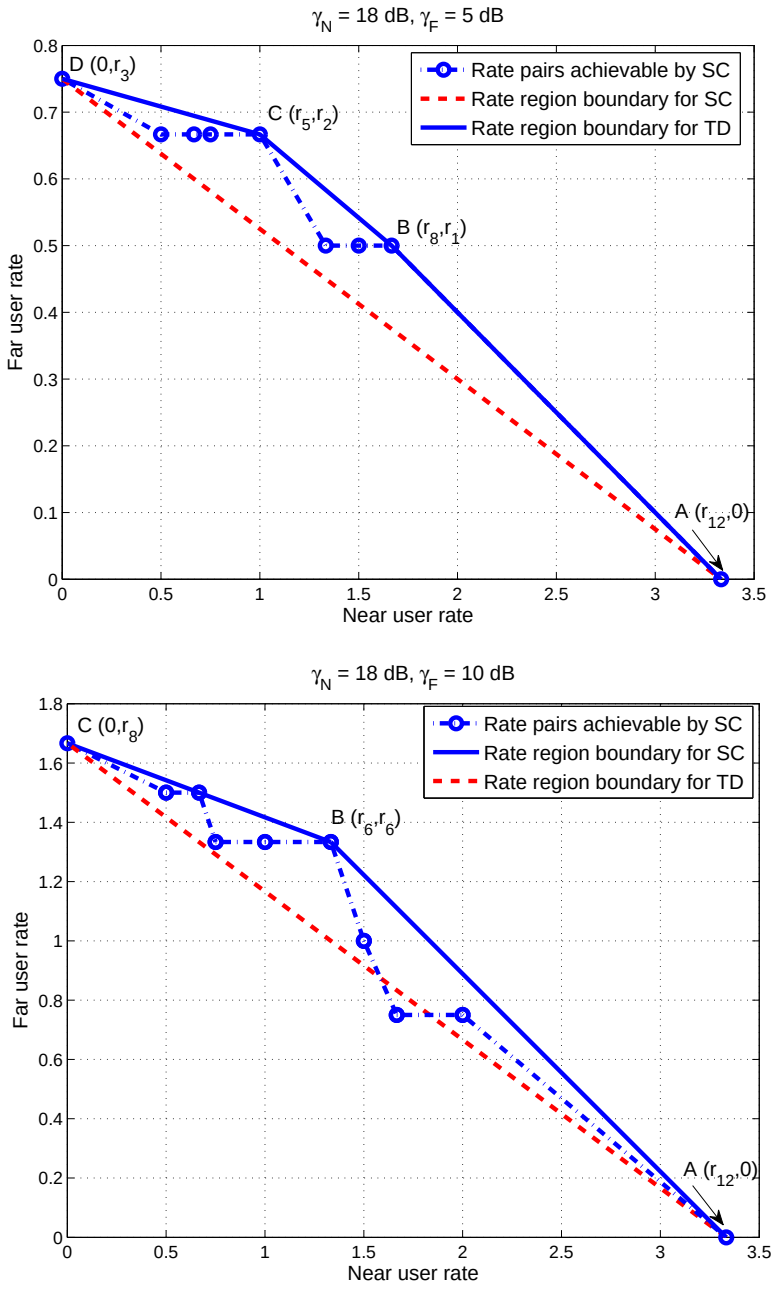

Figure 2.6: Experimentally obtained rate region for the library of $M = 12$ codes using the setup shown in Fig. 2.4 for two different choices of F. (Top) The values of α at corner points A-D are respectively $1, 0.126, 0.079, 0$. (Bottom) The values of α at corner points A-C are respectively $1, 0.126, 0$.

manifest themselves as large residual interference, necessitating an increase of $\alpha(l)$ beyond $\beta(l)$. For example, $1-\beta(1) \approx 90\%$ of the transmit power is assigned to F. Even if only 10% of this power remains after cancellation, it is still about the same as the signal power. As $\beta(l)$ increases, so does the near-rate, thereby making the near-link susceptible to even smaller levels of residual interference. as well.

The root cause of this problem lies in the small (but non-zero) estimation error in the channel frequency response and a small error in compensating the carrier frequency and phase offsets. Although this level of inaccuracy may result in relatively small losses in a single-user system (as shown in Fig. 2.5), a multiuser system is much less tolerant to these errors as our results show. Despite this inaccuracy, we find that SC can still provide rate gains using a reasonably well-designed single-user building blocks.

2.5.3.3 Modeling the Near-User Interference at the Far User

We study the performance of F's (maximum-likelihood) demodulator for the three interfering signal constellations (BPSK,QPSK,16QAM) and and for two different interferer strengths $\alpha = 0.2$ and $\alpha = 0.8$. F's rate is maintained at BPSK-1/2. Indeed, as explained in Section 2.3.2.2, the performance of the demodulator is dictated by the superconstellation with $2^{b_{\mathrm{x_N}}+b_{\mathrm{x_F}}}$ points; its error probability critically depends on the effective minimum distance (2.10). Fig. 2.8 depicts the far user PERs versus the SINR at F which we define as

$$\mathsf{SINR} \triangleq \frac{(1-\alpha)\gamma_{\mathrm{F}}}{1+\alpha\gamma_{\mathrm{F}}}. \tag{2.14}$$

We observe the following:

- In the weak-interference regime (for $\alpha = 0.2$, i.e., SIR ≈ 6 dB), it is seen that BPSK is the *worst interferer*. For small α, the error probability arises primarily from the separation between the clusters (see Fig. 2.7 (left)).

- In the strong interference regime (e.g., $\alpha = 0.8$, i.e., SIR ≈ -6 dB), it is seen that 16QAM is the worst interferer. When α is large, the clusters overlap and the error probability may be attributed to the high density of the clusters (see Fig. 2.7 (right)).

Validity of the IAGN model: The rate constraint on the far-user in (2.4) is of the form $R = W \log(1+\mathsf{SINR})$. However, *conditioned on N's codebook*, this interference is clearly not Gaussian. When the channel codes and the detection process at F are allowed to be arbitrarily complex (e.g., allowing Gaussian signaling, ML decoding), the *combined* effect of interference and noise can indeed be shown to be Gaussian for $L \to \infty$ when N's codewords cannot be jointly decoded with those of F [39]. It is not clear if this remains valid when the choices of channel codes are restricted (e.g., to those with finite L using finite constellations) and/or the receiver architecture (e.g., to demodulate-and-decode). Under such constraints, it is of fundamental interest to check the validity of the Gaussian approximation. We clarify that the interference considered here is from N's symbols only, and not from other transmitters in the network.

Our results show that for SC-BICM systems with a demodulator structure described in Section 2.3.2.2, treating the interference-plus-noise term as Gaussian perturbation can be quite inaccurate. For a small near-user power (i.e., small α), the cluster centers are well separated and each daughter point is likely to be close to its parent (see Fig. 2.7 (left)). In this case, a BPSK interference distribution can place *every* daughter point at the farthest possible distance ($\sqrt{\alpha}$) from its parent.

For Gaussian interference with the same power, 68% of all daughter points will lie *within* a ball of radius $\sqrt{\alpha}$ centered at the parent point[18]. Thus for a given noise level, Gaussian interference would result in fewer demodulation errors compared to BPSK interference: a Gaussian interference-plus-noise term *underestimates* the number of decoding errors.

For large near-user power, the cluster centers become closer and the clusters begin to overlap (see Fig. 2.7 (Right)), making the precise form of the interference distribution important. For a given interferer power, higher-order interferer constellations (such as 16QAM) result in more densely packed clusters, which upon overlapping result in a smaller effective minimum distance $d_{\text{eff}}^{(\text{F,N})}$ from (2.10). Viewed from this perspective, it is clear that a Gaussian interferer would be the worst than either BPSK, QPSK, or 16QAM in that it results in infinitely dense clusters. Thus, in the low SIR regime, a Gaussian interference-plus-noise term *overestimates* the number of decoding errors. These trends are apparent from Fig. 2.8.

2.6 Concluding Remarks

2.6.1 Summary

We have presented a software-radio implementation of Superposition Coding using off-the-shelf single-user coding and decoding blocks. We experimentally determine the set of achievable rate-pairs for this system under a packet error constraint. Our results suggest that SC can provide substantial gains in spectral efficiencies over those achieved by orthogonal schemes such as Time Division

[18] About 68% of values drawn from a normal distribution are within one standard deviation of the mean.

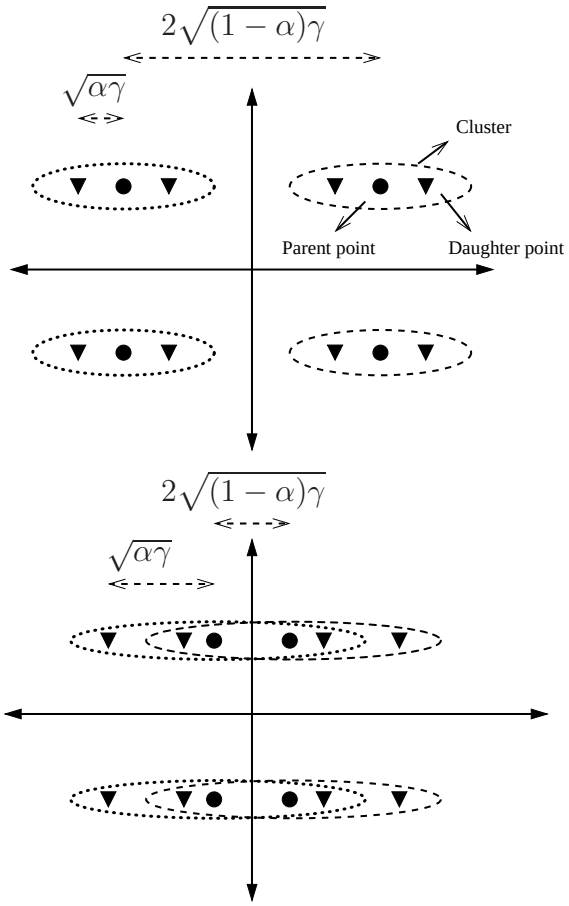

Figure 2.7. Depiction of the superconstellation points in the weak (left) and strong (right) interference cases. Evidently, when α is small, the PER depends on the cluster separation, while when α is large, the PER is determined by the cluster density.

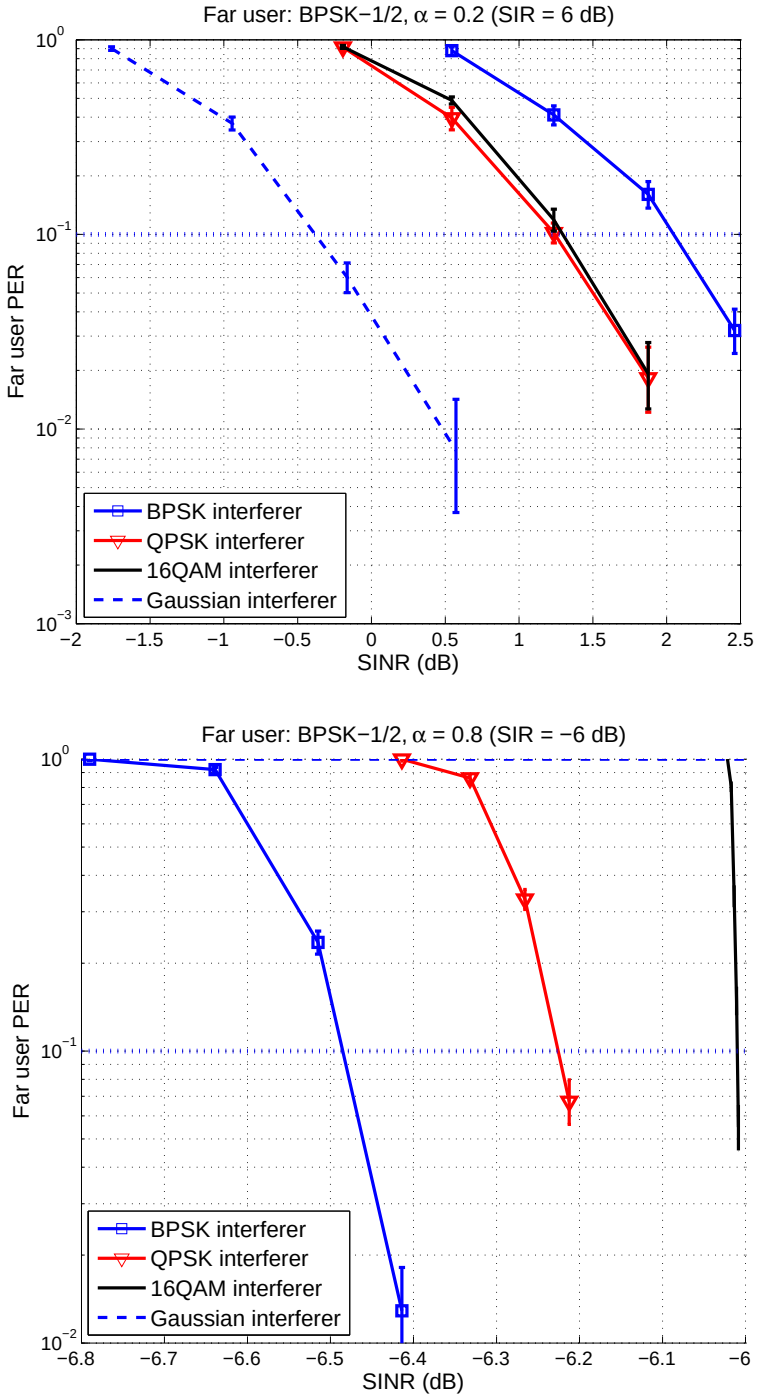

Figure 2.8: Far user PER versus SINR for different near-user constellations.

Multiplexing. Our findings also question the validity of treating inter-user interference as Gaussian noise to measure system performance in practical systems, and thereby the validity of the IAGN model for these systems. To the best of our knowledge, this is the first such attempt.

2.6.2 Discussion

In our investigation, we made some assumptions to simplify the design process and performance characterization. We now discuss how our framework can be modified to investigate the value of SC even when these assumptions are relaxed.

1. The channel ordering exploited for two-user SC in this chapter can be extended to three or more users, although finding the rate-region experimentally can be more cumbersome due to the presence of two power variables α_1 and α_2.

2. The code library that was held fixed in our design can in general be a design parameter by itself. This becomes relevant when an SC system is to be built from scratch rather than on top of an existing single-user system as in this chapter. Nonetheless, the current framework can be adopted to evaluate the performance of a given library.

3. The BS determines α from its knowledge of the individual link SNRs. In practice, these can be obtained via a feedback link. For every such SNR pair, the BS can choose an SC scheme by adopting the procedure described in this chapter.

2.6.3 Future Work

By experimentally demonstrating the benefits of SC using off-the-shelf single-user techniques, we have shown that SC is a potentially valuable transmission scheme. Our work can be extended in several directions. For example, the code library selection can leverage both advanced coding techniques such as turbo or LDPC codes and advanced receiver architectures based on iterative interference cancellation (see, e.g., [40, 41] and the references therein). Also, the principle of superposition can applied to other cases where signal superposition is known to be theoretically optimal, e.g., in multiple-access channels and for certain classes of relay channels. In these cases, the superposition process occurs at the *receiver* rather than at the transmitter in the BC; thus signals that interfere have different propagation paths. This difference opens up a new set of problems in system design, such as node synchronization, channel code selection and receiver design. Allowing multiple antennas at the TX and/or RX adds another dimension to the design space. Another possible line of investigation is to analyze the implications of SC for higher layers in the network stack. For example, the problem of scheduling multiple users with an SC-enabled physical layer involves many interesting tradeoffs [24].

The "rate-centric" view in this chapter focuses on the spectral efficiency gains from SC (for a given reliability). Viewing SC as a multiuser coding scheme, it is also possible to adopt a "reliability-centric" view wherein one measures the reliability gains (due to SC's coding gains) for a given pair of spectral efficiencies. Although both viewpoints are equivalent, they differ greatly in their experimental complexity. In the next chapter we use the platform developed here to demonstrate the benefits of SC in an actual wireless environment.

CHAPTER 3

SUPERPOSITION FOR MULTIUSER CHANNEL CODING

In this chapter, we propose and experimentally demonstrate a novel approach to improve the packet delivery efficiency on a vulnerable downlink (e.g., from a transmitter to a far-away receiver) using superposition coding, a multiuser transmission scheme that forgoes orthogonal transmission and deliberately introduces interference among signals at the transmitter. On a software-radio platform that uses off-the-shelf point-to-point channel codes, we show that a transmitter serving multiple links can use simple two-user superposition codes to dramatically improve (compared to time division multiplexing) the packet delivery efficiency on its most vulnerable links. Interestingly, our results suggest that superposing signals of far-away users onto those of high-traffic users yields the maximum benefits—implying that the degrees-of-freedom gain in doing so can more than compensate for the increased interference from signal superposition.

3.1 Introduction

As we saw in Chapter 2 it is possible to design superposition-coded systems using practical point-to-point codes. The main benefit offered by superposition codes is increased spectral efficiency. For a fixed spectral efficiency, this suggests

that there exist superposition codes that provide coding gain over orthogonal strategies. However, practical approaches that leverage this coding gain remain largely unexplored. In this chapter, we take the first step towards filling this gap by developing simple experimental procedures that build the previous chapter to show that two-user superposition codes for the downlink can achieve dramatic improvements in packet delivery efficiency in static wireless environments. This is quite unlike other experimental work on multiuser systems, e.g., [42, 43], that focus entirely on the gains in throughputs or transmission rates. Moreover, this approach opens up the possibility of designing novel protocols that leverage signal superposition in medium acess and power control (both higher layer functionalities), which we explore further in [13].

3.2 Problem Setup

3.2.1 A Case for Two-User Superposition Codes

As in Chapter 2 consider a BS serving several active users[1]. Given the user density in typical urban cellular networks, it is always possible to pick two active users N (the "near" user) and F (the "far" user), as shown in Fig. 3.1. The use of good superposition codes can exploit this channel ordering.

With two-user SC the BS transmits a weighted sum of the waveforms resulting from individually-coded user packets (or more precisely, codewords) (see Fig. 3.1). Thus both links enjoy the *combined* degrees of freedom available to N and F, while sharing the transmit power. The idea is to encode F's data such it can be decoded in the presence of interference from N's signal. N can decode its message

[1]The set of active users and each user's share of the BS's resources is determined by the users' traffic profiles and the BS's scheduling policy.

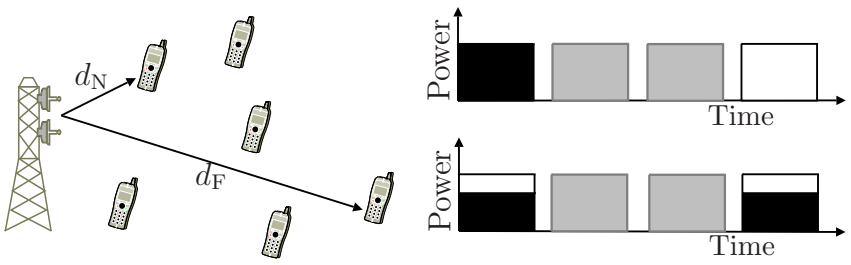

Figure 3.1: Illustration of two-user SC.(Left) The users N and F picked are at distances d_N and d_F respectively with $d_N < d_F$. (Right) Typical transmission timelines with and without SC. The gray slots represent transmissions to other active users which can remain unchanged. With Time-Division (TD, top), N and F are served in different slots (black and white). With SC (bottom), the BS transmits a linear combination of individually-coded user waveforms.

via successive decoding (SD): since it has a stronger link to the BS, N can first decode and subsequently regenerate and cancel F's contribution to its received signal. N can then decode its own packet. As noted in the previous chapter, simple superposition codes built from off-the-shelf point-to-point channel codes can substantially increase the spectral efficiency on each link (for a fixed PER) compared to orthogonal scheduling.

For a fixed spectral efficiency, SC can thus provide *coding gain* over orthogonal multiplexing schemes such as time division (TD)[2]. This coding gain can be leveraged either as a lower link SNR to realize a given link reliability[3], or as increased reliability for a fixed link SNR. In particular, this suggests that SC makes it possible to improve the reliability of far link (relative to TD) without changing the transmission rate on either link or degrading the reliability of the near link. In the following we examine this idea in greater detail.

[2]Hereafter we take TD to be the reference orthogonal scheme.

[3]The link reliability is measured as the probability of successfully decoding a codeword.

3.2.2 Improving Link Reliability using SC

As in Chapter 2, the BS has a code library $\mathcal{C}$ of $M > 1$ point-to-point channel codes, and we will (also) refer to a code's spectral efficiency as its *rate*[4]. Irrespective of the code, each *packet* supplied by the link layer is encoded as one *codeword* of *blocklength* $L < \infty$ [channel uses]. Define the *Packet Error Rate (PER)* $\epsilon < 1$ as the probability that the intended receiver cannot decode its packet. The BS transmits with an average power P [W] over a bandwidth W [Hz] (normalized to 1 without loss of generality (w.l.o.g)). When all of this power is assigned to one link (as in TD), the user enjoys a *single-user SNR* γ. In the following, we denote the single-user SNR and PER of a user u by γ_{u} and $\mathrm{PER}_{\mathrm{u}}$ respectively, for $\mathrm{u} \in \{\mathrm{N}, \mathrm{F}\}$. We assume $\gamma_{\mathrm{N}} > \gamma_{\mathrm{F}}$ w.l.o.g.

To communicate at a rate r_{N} with N and a rate r_{F} with F, the BS has at least two choices. With TD it can assign a fraction $u \in [0, 1]$ of the total slots to N and the remaining to F. In each slot, the entire power is assigned to the packet being transmitted. To sustain the desired rates, the BS must encode N's packets using a code with rates r_{N}/u and those of F at a rate $r_{\mathrm{F}}/\bar{u}$ where $\bar{u} \triangleq 1 - u$. Thus TD eliminates interference between N and F at the cost of increasing the encoding rates of individual packets.

On the other hand, with SC, BS individually codes N's and F's packets *exactly* at r_{N} and r_{F} respectively, and transmits a superposition of these waveforms in every slot. To meet the power constraint, a fraction α of the available power is assigned to N's waveform (and a fraction $\bar{\alpha} \triangleq 1 - \alpha$ to F). Thus SC allows the encoding of individual packets at exactly the desired rate, albeit at the cost of interference between the waveforms.

[4]Not to be confused with the code-rate of a binary code.

In particular, for a given γ_{N} and γ_{F}, the PER at F with SC depends on the interference from N's signal, as measured by the signal-to-interference ratio

$$\mathrm{SIR} \triangleq (1-\alpha)/\alpha. \tag{3.1}$$

We now explain how SC can improve F's PER with the help of Fig. 3.2, that idealizes the dependence of $\mathrm{PER}_{\mathrm{F}}$ on γ_{F} as a straight-line waterfall curve (for each SIR). Consider two codes with rates $r_{\mathrm{F}}/\bar{u}$ and $r_{\mathrm{F}},\bar{u} \in [0,1]$ from a hypothetical code library. With TD, only the solid PER curves ($\mathrm{SIR} = \infty$) are accessible. Thus for any $\bar{u}$, the PERs with TD are controlled by the waterfall curve of the code with rate $r_{\mathrm{F}}/\bar{u}$. For each finite SIR, each solid curve gives rise to a dashed curve (or a dash-dotted curve, depending on the SIR, which is in turn determined by the choice of N and α). Clearly, only the latter PER curves are accessible to SC.

For any given $r_{\mathrm{N}}, r_{\mathrm{F}}, u$, provided N can cancel most of F's signal upon decoding it correctly[5], α (and thereby the SIR) can be kept close to the minimum that would be necessary to maintain N's rate and reliability in TD. In this case (dashed curve), good channel codes can help SC leverage the increase in the time slots available to F (more generally, the increased degrees of freedom) to more than compensate for its reduced share of transmit power *and* the interference from N's signal, thereby outperforming TD[6]. There are two ways to measure this performance gain. For the same rate and PER, moving from TD to SC reduces the required γ_{F} by an amount determined by the coding gain (shown as CG). For a fixed operating SNR point (shown as γ_F^o) this transition would provide a PER gain (shown as PG),

[5]We did not observe imperfect decoding in our experiments.

[6]Indeed, SC systems built with simple off-the-shelf codes can exhibit this property, see Section 3.4.2.

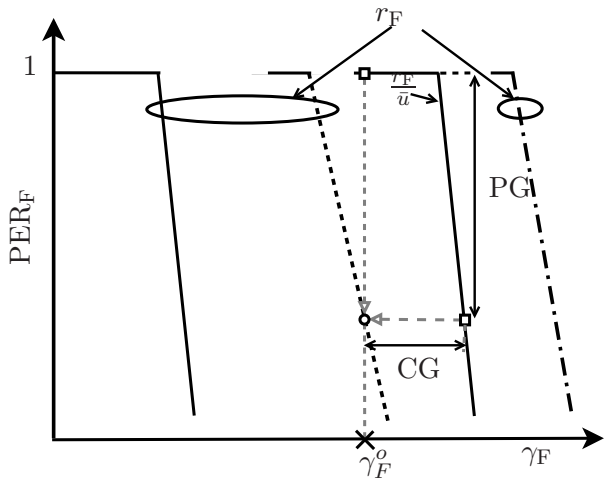

Figure 3.2. Improving the PER at F using SC, measured at the operating points for TD (squares) and SC (circle). The coding gain (CG) is the reduction in the single-user SNR at F resulting from the coding gain of SC over TD. This gain can be also measured as a PER drop (PG) for a fixed SNR γ_F^o.

that translates to a *reliability gain* (RG). Indeed, for a PER of ϵ, a typical packet will require an average of $1/(1-\epsilon)$ transmissions before it is received correctly at F. Thus, given a large number of packets to be sent to F, a BS using TD would require

$$\mathrm{RG} = \frac{1-\epsilon_{\mathrm{SC}}}{1-\epsilon_{\mathrm{TD}}} \tag{3.2}$$

more transmissions than a BS using SC. If ARQ is used, RG can also be interpreted as a link-level throughput gain. On the other hand, when channel estimation errors prevent N from regenerating F's signal accurately enough despite having decoded it correctly, the residual interference may be large enough to require an increased α to maintain the same rate and reliability as in TD. Apart from increasing N's interference at F, this also reduces F's share $1-\alpha$. In such cases switching to SC may not result in performance gains (dash-dotted curves).

The above framework is quite general and can be used to compare any superposition scheme with an orthogonal scheme. To derive quantitative results, in the following subsection we will focus on the specific code library and a receiver architecture used in our testbed.

3.3 Effect of Near-User Interference on the Far-User PER

We first describe the code library and the transmitter and receiver operation. Our description will focus on the details that are most relevant to the problem at hand. A more detailed discussion of the testbed's architecture and operation as well as some relevant details about our calibration and measurement procedures can be found in our technical report [38].

For channel coding, the testbed uses off-the-shelf point-to-point channel codes designed using the well-known Bit Interleaved Coded Modulation (BICM) technique [33]. It maps these symbols on to waveforms using Orthogonal Frequency Division Multiplexing (OFDM). The encoder is realized as a binary convolutional encoder followed by an bit-interleaver and a QAM modulator. The decoder is implemented as a maximum-likelihood demodulator, a bit de-interleaver and a Viterbi decoder. The convolutional codes are constructed from a mother code with a generator polynomial $[133, 171]$ using four puncturing patterns to yield code-rates $\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{5}{6}\}$. Each of the four binary codes can be paired with one of three possible interleaver-QAM constellation pairs, one each for gray-coded BPSK, QPSK, or 16QAM. This process results in $4 \times 3 = 12$ BICM codes with rates $\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{5}{6}, 1, \frac{4}{3}, \frac{3}{2}, \frac{5}{3}, 2, \frac{8}{3}, 3, \frac{10}{3}\}$.

In SC mode, the BS first encodes each near- and far-packet using its separate point-to-point BICM encoder into streams of $L = 1536$ symbols each. An adder

then sums up N's stream weighted by $\sqrt{\alpha}$ and F's stream by $\sqrt{1-\alpha}$ to produce a single composite symbol stream, which is fed to the OFDM modulator. At the receivers, after standard OFDM pre-FFT processing, both N and F try to recover F's symbol stream first using a maximum-likelihood (ML) demodulator. The demodulator computes the reliability of each of F's code bits in the presence of interference from N's symbols. The de-interleaver and Viterbi decoder process this reliability information in turn to estimate F's data bits. N regenerates F's symbol stream by re-encoding these data bits with a standard point-to-point BICM encoder. After appropriately weighting these symbols by $\sqrt{1-\alpha}$ and the channel coefficient estimates, N cancels F's interference from its received signal and then proceeds to decode its own message using point-to-point BICM decoding.

We now examine the demodulation process at F. Suppose the channel to F has gain h_{F}[7]. Then the demodulator differs from a point-to-point demodulator in that it observes a noisy symbol stream

$$Y_{\mathrm{F}}(n) = \underbrace{h_{\mathrm{F}}\sqrt{\bar{\alpha}P}X_{\mathrm{F}}(n)}_{\text{Signal}} + \underbrace{h_{\mathrm{F}}\sqrt{\alpha P}X_{\mathrm{N}}(n) + Z_{\mathrm{F}}(n)}_{\text{Interference+Noise}}, \tag{3.3}$$

where $n \in \{1, \ldots L\}$ denotes the symbol index and $Y_{\mathrm{F}}(\cdot), X_{\mathrm{F}}(\cdot), X_{\mathrm{N}}(\cdot), Z_{\mathrm{F}}(\cdot)$ denote the observations, the far- and near-symbol streams, and the white noise sequence respectively. The detection rule clearly depends on the distribution of the perturbation term, that in turn depends on N's constellation. We assume this is known to the demodulator[8]. Now each interfering symbol from N perturbs the original far-symbol (the *parent* point) to a randomly chosen *daughter* point. For

[7]The analysis that follows can be generalized to each subcarrier in a frequency selective channel due to the use of OFDM.

[8]In practice sending this information entails a small overhead, which we neglect in this paper.

each parent point, define the set of all possible daughter points to be its *potential daughter cluster* ("cluster" for short). The shape of this cluster depends on N's constellation, and its spread increases with a decrease in SIR. The demodulator at F infers the most probable parent point of the observed (noisy) daughter point by identifying the most probable cluster to which an observation belongs. Identifying successively less probable clusters helps refine its reliability estimate of each detected code-bit. Analogous to the single-user case, the reliability of the k^{th} bit in the n^{th} symbol is approximated using the max-log-MAP approximation (see [23, Sec. III-B2] for details). The probability of the dominant error events is controlled by the inter-cluster separation

$$d_{\text{eff}} \triangleq \sqrt{\bar{\alpha}P}|h_{\text{F}}| \min_{\substack{p_1,p_2\in\mathcal{X}_{\text{F}},p_1\neq p_2 \\ d_1,d_2\in\mathcal{X}_{\text{N}}}} \left| p_1 - p_2 + \frac{d_1 - d_2}{\sqrt{\text{SIR}}} \right|. \tag{3.4}$$

Here $\mathcal{X}_{\text{u}}$ denotes the constellation points of user $\text{u} \in \{\text{N}, \text{F}\}$. For the point-to-point case (when these "clusters" are just points) the constellation minimum distance can be recovered by allowing $\text{SIR} \to \infty$ in (3.4). It is evident from (3.4) that the choice of $\mathcal{X}_{\text{N}}$ determines the effect of interference at F, suggesting the possibility of modifying its geometry to mitigate this interference. For example, when both N and F are BPSK-modulated, a *rotated*-BPSK constellation for N is preferable to a standard BPSK constellation.

From (3.4) it is clear that $\mathcal{X}_{\text{N}}$ affects PER_{F}. To gain more insight into this problem by constraining both F and N to be BPSK-modulated[9], which we denote

[9]Although not particularly relevant for certain values of SIR (e.g., at low SIR, $\mathcal{X}_{\text{N}}$ is more likely to be a 16QAM constellation, see Section IV-B), this case does help understand some key geometric aspects of d_{eff} while permitting simple closed-form results.

as BPSK/BPSK (Signal/Interference). Now d_{eff} has the closed-form

$$d_{\text{eff}}^{\text{BPSK/BPSK}} = 2P|h_{\text{F}}| \min\left(\sqrt{1-\alpha}, \left|\sqrt{1-\alpha} - \sqrt{\alpha}\right|\right), \tag{3.5}$$

which is plotted in Fig. 3.3 for $\alpha \in [0, 1]$. For a given P and h_{F}, its dependence on α can be understood geometrically. Each cluster consists of two daughter points, one each to the left and to the right of the parent point. For $\alpha < \frac{1}{2}$, the cluster interiors[10] do not overlap; hence the nearest points from neighboring clusters lie on the *opposite* sides of their parent points. Increasing α brings these points closer to one another, making them overlap with the origin for $\alpha = \frac{1}{2}$. As α is increased beyond $\frac{1}{2}$, these overlapped points separate, increasing their mutual distance. Here, the value of α that maximizes d_{eff} in (3.5) can be obtained as the solution to the saddle point equation $\sqrt{1-\alpha} = \sqrt{\alpha} - \sqrt{1-\alpha}$ in $\alpha \in (\frac{1}{2}, 1)$, which is $\alpha = \frac{4}{5}$. For $\alpha \geq \frac{4}{5}$, the nearest points lie on the *same* side of their parent points; their mutual distance is therefore the same as that of their parent points, that in turn vanishes as $\alpha \to 1$. Based on this behavior we will refer to $\alpha = \frac{1}{2}, \frac{4}{5}$ as the *inflection points* for BPSK/BPSK. Arguing similarly, it is easy to see that BPSK/QPSK will have two inflection points (evaluated to be $\alpha = \frac{2}{3}, \frac{8}{9}$) and that BPSK/16QAM will have six inflection points ($\alpha = \frac{40}{76}, \frac{40}{65}, \frac{40}{56}, \frac{40}{49}, \frac{40}{44}, \frac{40}{41}$). A similar analysis can be done for higher order constellations for F (by directly evaluating (3.4) numerically as necessary).

The inflection points capture the non-monotonic dependence of d_{eff} (and therefore that of PER_{F}) on α. Given that (3.4) also depends on $P|h_{\text{F}}|$, this behavior may not be apparent at small values of P (when the noise is then large enough to

[10] As defined by a cluster's convex hull.

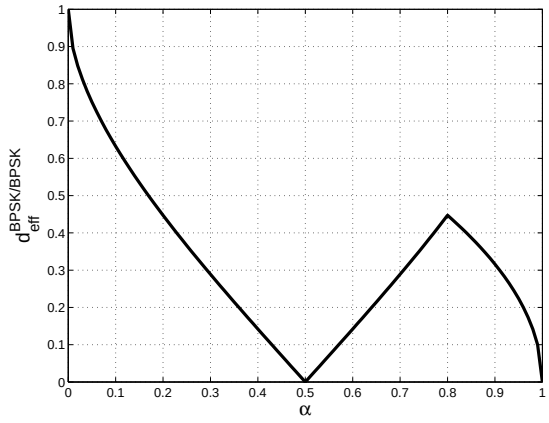

Figure 3.3. The distance in (3.5) as a function of α for $2P|h_{\text{F}}| = 1$.

cause decoding failures by itself) or when the number of trials limits the statistical reliability of the estimate (which is the case, for example, when the number of observed error events is quite small at larger values of P). In the next section, we experimentally validate the insights obtained so far.

3.4 On-Air Experimental Procedure and Results

We perform our experiments on a testbed developed in-house [38]. The testbed software is built on the well-known open-source GNU Radio platform [25] and interfaces with a Universal Software Radio Peripheral (USRP) hardware board that serves as an analog and RF front-end. Some key parameters of the experiment are listed in Table 3.1.

The experimental setup consists of three USRPs as shown in Fig. 3.4, where $d_{\text{N}} = 0.6$ m, and $d_{\text{F}} = 1.2$ m. Note that the goal of the experiment is to fix γ_{N} and γ_{F}—the apparent relationship between d_{N} and d_{F} is due to the table geometry.

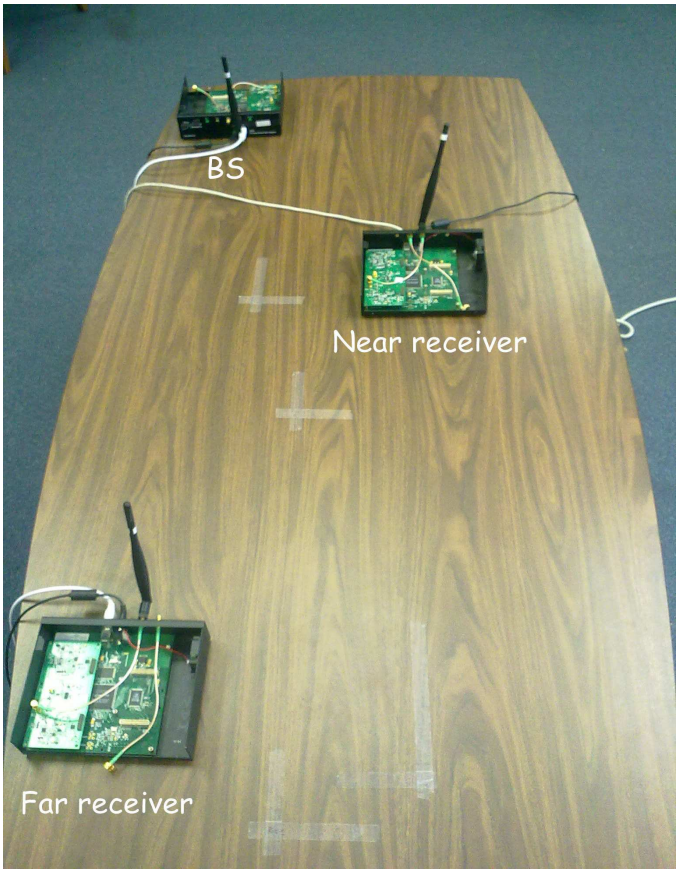

Figure 3.4. The on-air setup with the three USRPs used to study the efficacy of SC over TD. The BS transmits the private packet to N and the broadcast packet to F. Throughout our experiments, the near and far user distances are fixed at $d_{\mathrm{N}} = 0.6$ m and $d_{\mathrm{F}} = 1.2$ m.

TABLE 3.1

PARAMETERS USED IN THE EXPERIMENT.

Center Frequency	903 MHz
Message Bandwidth	2 MHz
Modulation	16-tone OFDM (8 data, 4 pilot, 4 null)
CP Length	1μs
Average Tx power	-31 dBm

Similar results can be obtained as long as one link has a much lower SNR than the other. Due to indoor scattering, the disparity between γ_{N} and γ_{F} can be quite different from what large-scale path-loss models would predict (see Table 3.2 on 60) for measured SNR values).

3.4.1 Impact of N's constellation on the PER at F

In light of the discussion in Section 3.3, we first study how $\mathcal{X}_{\mathrm{N}}$ and α affect $\mathrm{PER}_{\mathrm{F}}$. Keeping in mind our experiment in Section 3.4.2, we focus only on BPSK/$\cdot$ and adopt the following procedure: first, we adjust the total transmit power P to obtain a given single-user SNR γ_{F} at F. Second, we transmit $K = 1000$ packets and estimate the PER. We found this value of K to be sufficient for the PER range of interest. The testbed sends a packet every ~ 0.05s, so the channel was kept coherent during a single PER measurement (~ 50 s) . The results are summarized in Fig. 3.5 (on p. 63). We observe the following:

- When γ_{F} is small[11] (noise-limited regime), the noise is large enough to cause

[11]Evidently, the adjectives small, moderate and large are constellation-dependent. For in-

decoding errors by itself, thus F's PER is small only at small α.

- For moderate γ_{F}, the PER generally follows the trend in d_{eff} (for instance compare Fig. 3.3 with the curve for BPSK/BPSK (Fig. 3.5 (left)). In some cases (see Fig. 3.5 (right)), many errors can be corrected despite $d_{\mathrm{eff}} = 0$.

- When γ_{F} is large (interference-limited regime), the PER is high only in the vicinity of those inflection points for which $d_{\mathrm{eff}} = 0$; at other values of α, it is very small. Indeed, as long as the signal clusters do not overlap, perfect decoding is possible[12].

Note that F's PER curve is not monotonically increasing with α as one might expect; in fact, it is highly non-monotonic in most cases. Also, depending on $\mathcal{X}_{\mathrm{N}}$ and α, the reliability seen at F can be drastically different.

3.4.2 Measuring the Reliability Gain of SC over TD

We now design novel experimental procedures to measure the reliability gain of SC. The basic aim of the experiment is to show that for the same single-user SNR (fixed via the transmit power in a static environment), SC improves $\mathrm{PER}_{\mathrm{F}}$ over TD without significantly degrading the $\mathrm{PER}_{\mathrm{N}}$.

There are five important parameters that determine the performance gain of SC (see Section 3.2.2): (a) F's encoding rate r with TD (b) The share u of F's slots in the slots pooled from N and F, (c) $\mathcal{X}_{\mathrm{N}}$ (as seen from 3.4.1), (d) F's single-user SNR γ_F, and (e) The near power allocation fraction α (that determines

stance, $\gamma_{\mathrm{F}} = 13.73$ dB qualifies as a large value of γ when N's constellation BPSK, but is only a moderate value when N's signal constellation is QPSK (see Fig. 3.5).

[12]This is only an experimental artifact.

the SIR at F). We now relate these parameters by shrinking this (rather large) parameter space when the BS chooses a nearby N and a distant F, perhaps the most interesting case in practice.

When N is close to the BS, its packets are most likely encoded using spectrally efficient codes (e.g., the 16QAM-5/6 code in our library) in the TD mode. As a result, when the BS tries to maintain the same rate as in TD to a nearby N that contributes most of the combined slots (i.e., small u), it requires a large α (i.e., low SIR) *and* a spectrally efficient ("dense") $\mathcal{X}_{\mathrm{N}}$. The opposite is true for u close to 1. Therefore, choosing a nearby N implies a progressively denser $\mathcal{X}_{\mathrm{N}}$ with increasing α (or equivalently, with decreasing u)[13]. Along similar lines, a distant F means a (relatively) noisy link that operates at a small rate $r' = ur$ to achieve an acceptable PER.

Putting it all together, the experiment would involve measuring $\mathrm{PER_F}$ for the two approaches that both achieve a rate r': TD, which encodes F's packets at rate r but assigns them only a fraction u of the combined slots, and SC, which can encode at a rate r' by assigning all the slots to F but subjects F's signal to interference from a constellation $\mathcal{X}_{\mathrm{N}}$ that becomes denser as u decreases. We can fix the relatively small rate r' (e.g., BPSK-1/2, the smallest code rate in the library), so that r and $\mathcal{X}_{\mathrm{N}}$ are now controlled by a single parameter u. Each experiment thus involves picking a value of u and then choosing a pair of operating single-user SNRs $\gamma_{\mathrm{F}}^o, \gamma_{\mathrm{N}}^o$ (implicitly done by choosing the total transmit power P^o and α) that achieve a reliability gain at F. Instead of showing how changing from TD to SC provides reliability gain (as explained in Section 3.2.2), we find it easier

[13]Also note from Fig. 3.5 on p. 63 that starting from moderate α a dense $\mathcal{X}_{\mathrm{N}}$ results in a worse $\mathrm{PER_F}$ than a "sparse" (e.g., BPSK) $\mathcal{X}_{\mathrm{N}}$ for a given γ_{F} (ignoring the behavior at the inflection points). Using these points to improve the performance further is a subject of future work.

to show that changing from SC to TD implies a reliability *loss*.

To this end, we set up the USRPs as shown in Fig. 3.4. The near-receiver shown therein depicts the N that is paired with F. It is always served using 16QAM-5/6 in the TD mode. Recall that F is always served using $r' = 1/2$ (BPSK-1/2 from the library). In SC mode, for each value of u BS selects N's channel code using its knowledge of N's rate $4 \times 5/6 \times (1-u)$. This fixes $\mathcal{X}_{\mathrm{N}}$, and the BS must now set α and P^o such that both links operate at acceptable PER $\leq 10\%$ (w.l.o.g.).

To achieve this objective, we first define

$$(P_{\mathrm{N}}, P_{\mathrm{F}}) = (\alpha P^o, (1-\alpha)P^o), \tag{3.6}$$

where P_{N} and P_{F} denote N's and F's signal power respectively. Now, starting from $P_{\mathrm{N}} = P_{\mathrm{F}} = 0$, we adopt the following four-step procedure to find (α, P^o):

1. Keeping $P_{\mathrm{N}} = 0$ and increase P_{F} until the condition $\mathrm{PER}_{\mathrm{F}} \leq 0.1$ is met.

2. With the value of P_{F} from 1), increment P_{N} until[14] $\mathrm{PER}_{\mathrm{N}} < 0.1$.

3. Keeping the ratio $P_{\mathrm{N}}/P_{\mathrm{F}}$ the same, increase P_{N} and P_{F} until $\mathrm{PER}_{\mathrm{F}} < 0.1$. Note that N can still decode its packets at least as reliably as in 2).

4. Use (3.6) to find (α, P^o).

With this set up we consider the total BS power to be $P^o = P^o_{\mathrm{N}} + P^o_{\mathrm{F}}$ and estimate the single-user SNRs γ^o_{N} and γ^o_{F} at N and F respectively using standard windowed correlator methods (details in [38]).

[14]Due to the impact of imperfect cancellation of F's symbols at N (discussed in detail in [23, Sec. V C. 2)]), P_{N} is, in general, higher than the power that is required to meet that constraint for N alone (when $P_{\mathrm{F}} = 0$).

TABLE 3.2

ON-AIR PER MEASUREMENTS WITH SC.

Superposition Coding						
u	N's rate	F's rate r'	N's PER	F's PER	α	γ_F^0 (dB)
0.1	16QAM-3/4	BPSK-1/2	2.6%	6.8%	0.44	8.8
0.2	16QAM-2/3	”	5.3%	6.4%	0.39	7.4
0.4	16QAM-1/2	”	3.7%	2.6%	0.24	5.5
0.5	QPSK-5/6	”	3.8%	6.6%	0.29	4.5
0.55	QPSK-3/4	”	1.6%	5.1%	0.24	4.3
0.8	BPSK-2/3	”	3.5%	5.6%	0.2	2.7
0.85	BPSK-1/2	”	1.1%	5.1%	0.14	2.6

For TD, we set the transmit power to $P^o = P_{\rm N} + P_{\rm F}$. Now N's packets are encoded using 16QAM-5/6, while F's packets are encoded at $r = (1/2)/u$, and the PER[15] at F is measured. We thus make an *apples-to-apples* comparison: the total transmit power at the BS, the bandwidth and the spectral efficiencies of the two users all remain the same as in the case when SC is employed.

Tables 3.2 and 3.3 summarize our results. For each experiment $\gamma_N^o = \gamma_F^o + 12.8$ dB, and the standard deviation of the SNR estimation method is $\approx 0.3\,$dB. The values of γ_F^0 and α at the operating points are also listed. The last column in 3.3 depicts F's RG (computed using (3.2)) that SC offers for F's packets. Note that at every value of u, we obtain reliability gains for F using SC. Interestingly, higher

[15]If r does not exist in the code library, the PER at F may be evaluated using the time-sharing principle. Accordingly, if $r_i < r < r_j$ wherein $r_i, r_j \in \mathcal{C}$, and $\gamma r_i + (1-\gamma) r_j = r$, $0 < \gamma < 1$, we have that the PER at rate $r = \gamma$(PER at rate r_i) $+ (1-\gamma)$(PER at rate r_j) .

TABLE 3.3

ON-AIR PER MEASUREMENTS WITH TD.

Time Division Multiplexing					
γ_{F}^{0} (dB)	N's peak rate	F's peak rate	N's PER (%)	F's PER (%)	**RG**
8.8	16−QAM-5/6	5	0.1	Infeasible	N/A
7.4	”	2.5	0.1	100	∞
5.5	”	1.25	0.2	74.6	3.83
4.5	”	0.5	14.7	36.7	1.47
4.3	”	0.45	18.9	38.1	1.53
2.7	”	0.63	77.5	36.6	1.49
2.6	”	0.59	80.4	29.2	1.34

gains are realized at smaller values of u—implying that the *degrees-of-freedom* gain derived in doing so can more than compensate for the increased inter-user interference. Of course, this would also require a larger γ_{F}^{o}, as shown. Moreover, in this regime, achieving the equivalent TD rates for F may also be infeasible with our code library (e.g., when $u = 0.1$). Table 3.3 also indicates that a huge benefit in N's reliability is seen at large values of u. Reversing the roles of F and N, our results also suggest that a moderate-to-high rate F needs to be paired-up with a moderate-to-low-rate N.

3.5 Concluding Remarks

We have experimentally validated the efficacy of superposition coding on a software radio platform. Our contributions are two-fold: First, we clearly de-

scribe the dependence of N on F's packet error performance, thus motivating the need for joint code optimization. Second, we show how transmitting F's message at a reduced rate in the presence of deliberately-introduced interference can dramatically improve F's PER (and also its link-layer throughput, if ARQ is used).

In the next chapter, we discuss the problem of medium access when many broadcast clusters interfere with each other. In such a network, each noisy link that a transmitter communicates on has a different level of interference depending on the locations of nearby interferers. A medium access protocol must thus manage the interference seen on each link. When each transmitter orthogonalizes its signals to its receivers, the medium access problem decouples into multiple point-to-point problems, thereby permitting the protocol to adjust the spatial reuse of each link independently of others. This is no longer true if transmitters use a non-orthogonal scheme such as SC. We model this effect by introducing a network utility network function to measure the effective rates of information transfer in space offered by the two schemes in the presence of interference. Due to the decoupling in orthogonal schemes, maximizing this utility function reduces to individually maximizing each sub-function. With SC, however, this maximum is decided by the relative contribution of each link to the utility function. Since optimal spatial reuse is a function of network geometry, the relative benefit of SC depends on the geometry of node placement.

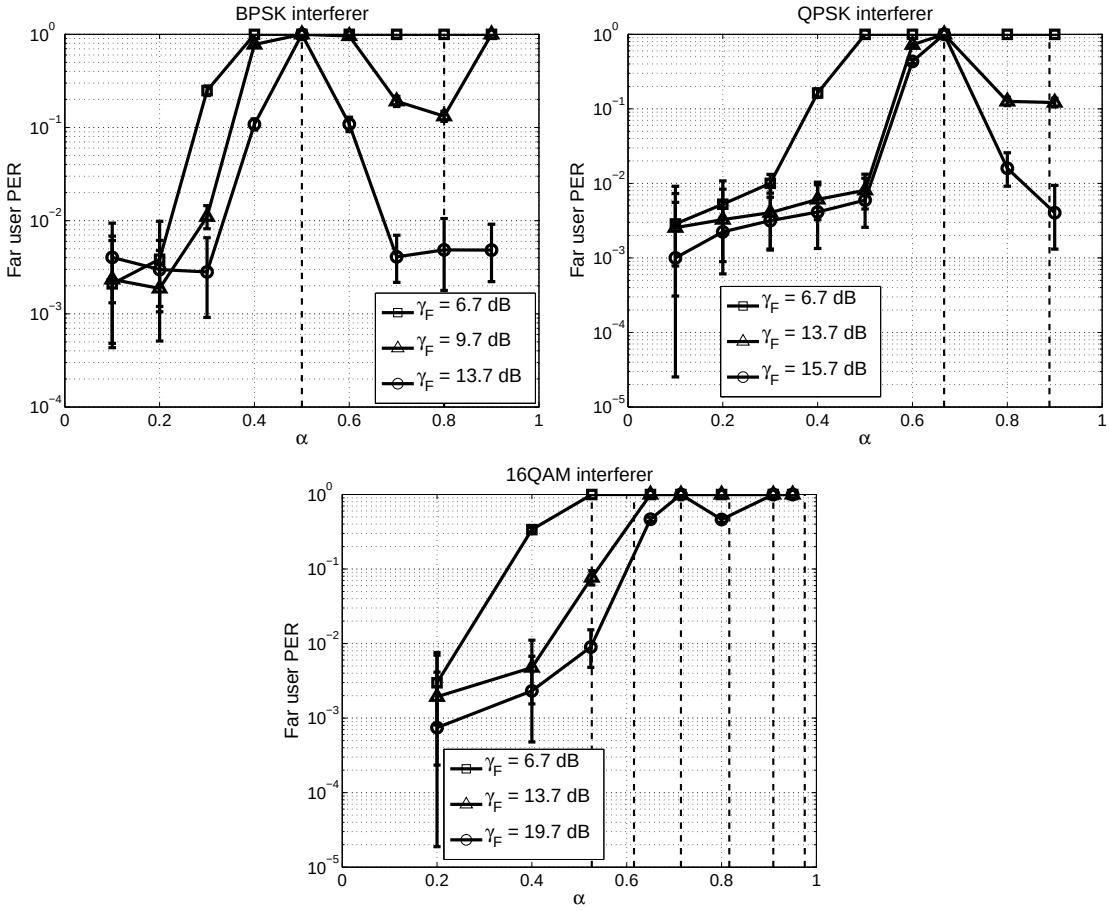

Figure 3.5. Far user PER (and their 95% confidence intervals) versus the power allocation parameter α at different values of γ for BPSK/BPSK, BPSK/QPSK and BPSK/16QAM respectively. The inflection points are also marked (by dashed lines) in the figure.

CHAPTER 4

SUPERPOSITION CODING IN RANDOM NETWORKS

In this chapter, we study networks of mutually interfering broadcast cluster using tools from point process theory. In particular, we compare the benefits of superposition coding (SC) and traditional time division (TD) with this interference via a utility function that measures the rate of information transfer per unit area. For a fixed subslot allocation across all transmitters, TD allows the spatial reuse to be independently optimized in each subslot. On the other hand, with SC for a fixed power allocation, the optimal spatial reuse depends on the relative contribution of each link to the utility function. Since optimal spatial reuse is a function of the network geometry, the gains provided by SC depend on the geometry of the receiver placement. Ours is the first study to analyze the effect of superposition codes (a physical layer scheme) on interference management (a link layer concern).

4.1 Introduction

Conventional link-layer abstractions assume a collision model for packet reception, i.e., a receiver is capable of decoding only one packet at any given time. As a result, scheduling at the transmitter requires orthogonality among transmissions to different users at the physical layer. However, when receivers are capable of sophisticated physical-layer processing such as successive decoding or receive

beamforming, such an assumption may overly restrict the design space of the scheduler, whether in a base station or in a relay node serving multiple routes.

As mentioned in the earlier chapters, superposition coding (SC) deliberately introduces interference among user signals that originate at each transmitter. Although information-theoretic models such as degradedness may not be applicable in networks with interference, little is known about how such local signal superposition impacts higher layers in the network stack. In this chapter, we take the first steps towards such an investigation by analyzing the impact of SC on medium access. Medium access protocols manage the interference among different broadcast clusters. In networks composed of multiple mutually interfering broadcast clusters, medium access protocols depend on the local cooperation scheme. With orthogonal transmission schemes such as TD the one-to-many medium access problem decomposes into separate point-to-point problems, at the cost of assigning a smaller bandwidth to each user. This decoupling is no longer valid in SC, since assigning power to one link changes the interference seen on others.

In this chapter, we present some results of our investigation of the implications of using SC as a multipacket transmission scheme in a network. To the best of our knowledge, this is the first such study. We investigate this problem using a stochastic geometry framework. In particular, we study a network consisting of many randomly placed three- (in general multi-) node *broadcast clusters*, as shown in Fig. 4.1. Our approach allows us to compare throughputs obtained with SC and TD, averaged over four important sources of uncertainty: node placement (modeled as a homogeneous Poisson point process (PPP)), channel access, link distances and fading. We quantify the benefit offered by each scheme in terms of a utility function that accounts for both the local throughput and the distance of

successful packet transmissions, averaged over all these sources of uncertainty.

We present results for a two-user broadcast; extending these results to a greater number of users is straightforward. Our results show that the benefits from SC depend on the receiver geometry. This is a direct consequence of the medium access mechanism. Unlike SC, TD allows independent spatial reuse among the non-interfering sub-networks, as noted in [44]. Medium access protocols for small link distances permit greater spatial reuse; the opposite is true for large link distances. In SC the near-far disparity in channel quality is optimally exploited to maximize the rate of communication when there is just one broadcast cluster.

4.2 System Model

4.2.1 Network

The transmitters using multipacket transmission are points drawn from a unit intensity homogeneous Poisson point process (PPP) $\Phi \triangleq \{x_i\} \in \mathbb{R}^2$, which we call the *ground process*. Using the ground process we define a marked homogenous PPP $\Phi \triangleq \{(x, t_{x,1}, t_{x,2}) : x \in \Phi\}$ where $t_{x,k} \in \{0, 1\}$ denotes the transmit decision to user k (defined below) of the transmitter $x \in \Phi$. Signal propagation is subject to power-law path loss with exponent $\beta > 2$ and (frequency-) flat fading. All point-to-point channels are iid Rayleigh block fading over unit time slots. Each broadcast cluster consists of a transmitter $x \in \Phi$ and its two intended users or receivers: a "near" user at a distance $r_{x,1}$ and a "far" user at a distance $r_{x,2}$. The sequence $\{(r_{x,1}, r_{x,2}) : x \in \Phi\}$ is iid drawn from a distribution F_{r_1,r_2}. This distribution is known at the transmitters. Denote the typical transmitter by T centered at the origin communicating with its k^{th} typical user at a distance r_k at

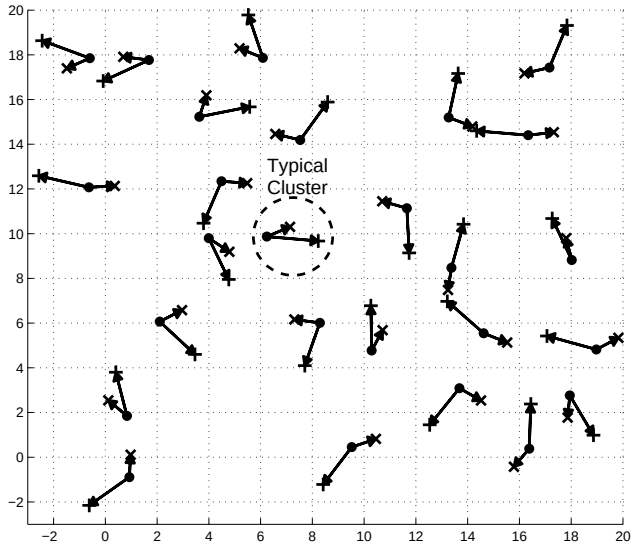

Figure 4.1. A network consisting of many three-node clusters. Transmitters (filled circles) are a realization of a homogeneous PPP. Each transmitter communicates with two intended receivers, as shown by the arrows. The receivers closer to (resp. farther from) to their transmitter (resp. the near and far receivers) are marked as $\times$-marks (resp. $+$-marks). The typical cluster in this network is highlighted.

a transmission rate R_k.

4.2.2 Physical Layer

4.2.2.1 Transmission

The physical layer uses Gaussian signalling over long blocklengths over a unit bandwidth channel and at unit power. Suppose single-user communication from a transmitter T to its k^{th} receiver is established using a single-user code of rate $C(\theta_k)$, where $C(x) \equiv \log(1+x)$, $k = 1, 2$. Let $\theta_k = 1/N'_k$, where N'_k can be viewed as the *presumed* noise-variance at receiver k. We assume $\theta_2 < \theta_1$, i.e., the transmitter presumes that the channel to user 2 is noisier than that to user 1. We

will use the superscripts sc and td to identify quantities pertaining to SC and TD.

SC is implemented by assigning the near (resp. far) receiver's packets a power $0 \leq \alpha \leq 1$ (resp. $\bar{\alpha} = 1-\alpha$) of the transmit power and simultaneously transmitting both the encoded messages during the same slot. R1 is assumed to implement SD: R2's message is decoded first, its contribution to the received signal subtracted, and its own message is then decoded. Thus we have a transmission rate $C(\alpha\theta_1)$ to R1. On the other hand, SD is presumed to be not possible at R2 - which means a fraction α of the received power causes self-interference. Therefore T assumes a received signal-to-interference-plus-noise-ratio (SINR) of $\frac{\bar{\alpha}}{\alpha+N_2'} = \frac{\bar{\alpha}\theta_2}{\alpha\theta_2+1}$ at R2, and transmits at a rate $C(\frac{\bar{\alpha}\theta_2}{\alpha\theta_2+1})$. TD is implemented by assigning a fraction u_k to user k, with $\sum_k u_k = 1$. Let $u_1 \equiv u$, $u_2 \equiv 1-u$. We define:

$$R_1^{\text{sc}} \triangleq C(\alpha\theta_1) \tag{4.1}$$

$$R_2^{\text{sc}} \triangleq C\left(\frac{\bar{\alpha}\theta_2}{\alpha\theta_2+1}\right) \tag{4.2}$$

$$R_k^{\text{td}} \triangleq u_k C(\theta_k). \tag{4.3}$$

4.2.2.2 Reception

Receivers have CSI of their intended transmitter and decode the signal from their intended transmitter while treating all signals from outside the cluster as noise. Such a strategy is optimal in the weak-interference regime [45]. The *actual* noise variance at all the near (resp. far) users is N_1 (resp. N_2). A receiver decodes packets from its intended transmitter on a per-slot basis, and the decoding process is approximated by the well-known SINR model: decoding is successful iff the SINR exceeds the SINR threshold of the message transmission rate.

4.2.3 Link Layer

We assume the packet queues at all transmitters are backlogged to ensure their participation in medium access. We focus on ALOHA-type random access where each transmitter decides independently on the links it activates in each slot. Different from the point-to-point case, a transmitter can activate more than one link in each slot.

Suppose all transmitters in the network either follow TD or SC. Due to the orthogonality in TD, a multipolar network can be broken down into multiple mutually decoupled bipolar subnetworks, and traditional (i.e., point-to-point) ALOHA can be separately applied to each subnetwork. Let p_k be the attempt probability for the k^{th} such subnetwork. This independence no longer holds with SC, since activating one link can causes interference to other links originating from the same transmitter. Therefore, for K links, a random access protocol is defined by a $K-$dimensional joint distribution that specifies the probability of each of the 2^K possibilities for link activation, which we term as *Super-ALOHA*, in that it relies on signal superposition. As a result the correlations between the activation of different links become important. In particular, for a transmitter $x \in \mathbb{R}^2$ we can write

$$\mathbb{P}(t_{x,1}, t_{x,2}) = \mathbb{P}(t_{x,2} \mid t_{x,1})\mathbb{P}(t_{x,1}) \tag{4.4}$$

from Bayes theorem. The conditional probability term models the correlation. In this chapter we study the special case where

$$\mathbb{P}(t_{x,2} = b_2 | t_{x,1} = b_1) = 1_{b_1 = b_2}, \tag{4.5}$$

so the links are either simultaneously active or inactive with a (common) prob-

ability p. The general case is a topic of current investigation [13]. Denote the success probability at the k^{th} typical user by $p_{s,k}$. The local throughput T on the k^{th} typical link is defined as

$$T_k = p_k p_{s,k} R_k. \tag{4.6}$$

4.3 Success Probabilities

We find the expected local throughput seen at the typical transmitter T when it communicates with its near and far receivers R1 and R2 respectively. The throughput is derived for the special case of fixed link distances using an extension of the bipolar model in [46].

Proposition 1. *If each transmitter uses SC, the success probability on the k^{th} typical link for a fixed r_k is*

$$p_{s,k}^{\text{sc}} = \exp(-(p\gamma_k r_k^2 + N_k \theta_k r_k^{\beta})) \tag{4.7}$$

where $\gamma_k = \pi\Gamma(1+\delta)\Gamma(1-\delta)\theta_k^{\delta}$ for $k = 1, 2$ and $\delta \triangleq 2/\beta$.

Proof. We derive the throughput to R2 first. If g_2 denotes the channel gain from T to R2, I_2 the interference power, and N_2 the noise power, the SINR at R2 is given by

$$\mathsf{SINR}_2 \;=\; \frac{\bar{\alpha} g_2 r_2^{-\beta}}{\alpha g_2 r_2^{-\beta} + I_2 + N_2}$$

Since fading states are assumed to be spatially iid, from standard arguments (e.g., [47, Lemma 3.1]) we get (4.7) when $k = 2$.

Denote the interference power at R1 by I_1. Using the SD condition

$$\begin{aligned} p_{s,1}^{\text{sc}} &= \mathbb{P}\left(\frac{\alpha g_1 r_1^{-\beta}}{I_1+N_1} \geq \alpha\theta_1, \frac{\bar{\alpha} g_1 r_1^{-\beta}}{\alpha g_1 r_1^{-\beta} + I_1 + N_1} \geq \frac{\bar{\alpha}\theta_2}{\alpha\theta_2 + 1}\right) \\ &= \mathbb{P}\left(\frac{g_1 r_1^{-\beta}}{I_1+N_1} \geq \theta_1, \frac{g_1 r_1^{-\beta}}{I_1+N_1} \geq \theta_2\right) \\ &= \mathbb{P}\left(\frac{g_1 r_1^{-\beta}}{I_1+N_1} \geq \theta_1\right), \end{aligned}$$

since $\theta_1 > \theta_2$. Again using standard results, we can show that (4.7) holds for $k = 1$. □

For TD, the results are just the single-user success probabilities specialized to each band. Therefore, for TD we have for $k = 1, 2$:

$$p_{s,1}^{\text{td}} = \exp(-p_k \gamma_k r_k^2 - N_1 \theta_k r_k^{\beta}). \tag{4.8}$$

Using (4.6) the local throughput on the k^{th} typical link for SC and TD is respectively:

$$T_k^{\text{sc}} = p p_{s,k}^{\text{sc}} R_k^{\text{sc}}. \tag{4.9}$$

$$T_k^{\text{td}} = u_k p_k p_{s,k}^{\text{td}} R_k^{\text{td}}. \tag{4.10}$$

In the next section, we propose a utility function to compare SC and TD.

4.4 Transport Density

We would like the utility function of each broadcast cluster to account for both the rate of successful packet transmissions and the (possibly random) distance

over which these packets are transmitted. One such metric at each cluster is the expected product of the link distances and the number of packets that can be successfully transmitted per time slot. This expectation is computed over all possible spatial interferer configurations, fading channel states and link distances. A natural extension of this idea is to define the network utility function as the average of individual cluster utilities.

Define the transport density U as

$$U = \frac{\mathbb{E}\left[\sum_{x\in\Phi\cap B(0,a)} \sum_k R_k r_{x,k} t_{x,k} 1_{\text{Rx } k \text{ decodes its packet}}\right]}{\pi a^2} \tag{4.11}$$

where the summation is over transmitters $x \in \Phi$ in a ball $B(o, a)$ of radius $a > 0$ centered at the origin o, R_k is the spectral efficiency on the k^{th} link, and the indicator counts only those receivers that decode their packet. Note that U is independent of $a > 0$. From the assumptions in Section 4.2, all clusters have the same utility; hence the transport density may also be viewed as the utility of the typical cluster of the network. As a result,

$$U = U_1(p_1; \Lambda) + U_2(p_2; \Lambda),$$

where

$$U_k(p_k; \Lambda) = \mathbb{E}[r_k T_k] \tag{4.12}$$

is the k^{th} *utility sub-function* for link $k = 1, 2$. As noted earlier in Section 4.2.3, for SC $p_k = p$. The parameter vector Λ includes the transmission rates T_1 and T_2, path loss exponent β, and all the parameters related to the distributions of link distances r_1 and r_2. Let $\Lambda^{\text{sc}} \equiv (\Lambda^c, \alpha)$ for SC and $\Lambda^{\text{td}} \equiv (\Lambda^c, u)$ for TD, where the

common parameter vector Λ^c contains the parameters common to both. We state and prove some properties of the utility function defined in (4.12).

4.4.1 Some Properties of the Utility Function

The first result is an observation concerning SC and TD when both schemes have the same spatial reuse that is constant over the entire bandwidth.

Proposition 2. *Suppose Λ^{sc} is given. Then for a fixed attempt rate p across all transmitters and across the entire bandwidth, for every fraction $0 \leq u \leq 1$ of the bandwidth assigned to near receivers, there exists a fraction α of the transmit power that can be assigned to the near receivers such that $U^{\mathrm{sc}} \geq U^{\mathrm{td}}$.*

Proof. Ignoring the effect of interference and fading, the result follows from the optimality of SC over a Gaussian BC. With interference and fading, from (4.9), (4.10) we find the respective pre-factor terms that multiply the transmission rates to be the same for both SC and TD, thus preserving this inequality. □

Therefore for a fixed spatial density of interfering transmitters, SC provides greater average throughput than an orthogonal scheme such as TD. Before proving the second property, we give the following auxiliary result that follows from Proposition 1:

Corollary 3. *The success probabilities for SC in (4.7) (resp. TD in (4.8)) are log-concave in p, r_k (resp. p_k and r_k) on $[0,1] \times \mathbb{R}^+ \cup \{0\}$.*

Proof. Taking logarithms on both sides in (4.7) for $k = 1$ yields

$$\ln p_{s,1}^{\mathrm{sc}} \;=\; -p\gamma_1 r_1^2 - N_1\theta_1 r_1^{\beta},$$

which is clearly concave in p and r_1. A similar result holds $p^{\text{sc}}_{s,2}$. The TD case is similar. □

The second result in this section is a property of the utility sub-functions U_1 and U_2:

Proposition 4. *The utility sub-function U_k in (4.12) is log-concave in p_k if the marginal density of r_k is also log-concave.*

Proof. For SC, if r_k has a log-concave marginal density $f_k(\cdot)$ over a (convex) support $\mathcal{S}_k \subseteq \mathbb{R}^+ \cup \{0\}$ we use (4.7) and (4.9) in (4.12) to get

$$U^{\text{sc}}_k = pR^{\text{sc}}_k \int_{\mathcal{S}_k} r_k p^{\text{sc}}_{s,k}(p; r_k) f_k(r_k) dr_k.$$

Since the integrand is log-concave in both r_k and p over $[0,1] \times \mathcal{S}_k$ we apply the general result [48, p. 105] and infer the log-concavity of U_k in p over $[0,1]$. The proof for TD is identical. □

The log-concavity condition is satisfied by a large family of densities encountered in practice: exponential, uniform, gamma distribution to name a few. We will hereafter assume this condition is satisfied.

4.4.2 Optimizing Spatial Reuse

Denote the *unconstrained* maximizer of U_k as

$$\pi_k \triangleq \arg\max_{p_k} U_k,$$

From the log-concavity of these functions, we have at $p_1 = \pi_i$

$$\frac{\partial U_i}{\partial p_i} = 0,$$

for $i = 1, 2$. Let $\bar{\pi}_i \triangleq \min(1, \pi_i)$ be the corresponding constrained maximizers of the utility sub-functions U_i. The optimization problem is discussed for TD and SC separately.

4.4.2.1 TD

Orthogonalization decouples the network into two non-interfering sub-networks. Hence for each link its attempt rate p_k can be chosen independently. Thus we have

$$\begin{aligned} \max_{(p_1,p_2)\in[0,1]^2} U^{\mathrm{td}} &= \max_{p_1\in[0,1]} U_1^{\mathrm{td}} + \max_{p_2\in[0,1]} U_2^{\mathrm{td}} \\ &= U_1^{\mathrm{td}}(\bar{\pi}_1; \Lambda^{\mathrm{td}}) + U_2^{\mathrm{td}}(\bar{\pi}_2; \Lambda^{\mathrm{td}}). \end{aligned}$$

4.4.2.2 SC

In general, the maximizer $\bar{\pi}$ of U does not necessarily maximize U_1 or U_2. However the log-concavity of U_1 and U_2 implies $\bar{\pi}$ lies in $[\bar{\pi}_1, \bar{\pi}_2]$, as shown in the following result:

Proposition 5. *If π_1 and π_2 are the unconstrained maximizers of U_1^{sc} and U_2^{sc}, then if $\bar{\pi}$ is a constrained maximizer of U, there exists $t_\alpha \in [0, 1]$ such that $\bar{\pi} = t_\alpha \bar{\pi}_1 + \bar{t}_\alpha \bar{\pi}_2$, where $\bar{t}_\alpha = 1 - t_\alpha$.*

Proof. Recall that for SC, $p_1 = p_2 = p$. Without loss of generality, assume $\pi_1 < \pi_2$. Then

$$\frac{\partial U^{\mathrm{sc}}}{\partial p} = \frac{\partial U_1^{\mathrm{sc}}}{\partial p} + \frac{\partial U_2^{\mathrm{sc}}}{\partial p}.$$

Since $\ln U_1^{\text{sc}}$ is a differentiable concave function of p,

$$\frac{\partial \ln U_1^{\text{sc}}}{\partial p} = \frac{1}{U_1}\frac{\partial U_1^{\text{sc}}}{\partial p} > 0,$$

for $p < \pi_1$. Similarly one can argue that $\frac{\partial U_1^{\text{sc}}}{\partial p} < 0$ for $p > \pi_1$. Thus $\frac{\partial U^{\text{sc}}}{\partial p} > 0$ for $p < \pi_1$ and $\frac{\partial U^{\text{sc}}}{\partial p} < 0$ for $p > \pi_2$. We have the following possibilities:

1. $\pi_1 > 1$. This implies $\pi_2 > 1$. Therefore $\frac{\partial U^{\text{sc}}}{\partial p} > 0$ for $0 \le p \le 1$, i.e., $p = 1$ is a feasible maximizer of U^{sc}. In this case increasing the intensity of the PPP also increases U until $\pi_1 = 1$.

2. $\pi_1 < 1, \pi_2 > 1$. Then a feasible maximizer of U^{sc} should lie in $[\pi_1, 1]$, since $\frac{\partial U^{\text{sc}}}{\partial p} > 0$ for $0 \le p < \pi_1$.

3. $\pi_1 < 1, \pi_2 < 1$. Then a feasible maximizer of U^{sc} should lie in $[\pi_1, \pi_2]$, since $\frac{\partial U^{\text{sc}}}{\partial p} > 0$ for $p \in [0, \pi_1)$ and $\frac{\partial U^{\text{sc}}}{\partial p} < 0$ for $p \in (\pi_2, 1]$.

In all these cases a feasible maximizer can be written as $\bar{\pi} = t_\alpha \min(1, \pi_1) + \bar{t}_\alpha \min(1, \pi_2)$ for some $t_\alpha \in [0, 1]$ since any point in an interval can be written as a convex combination of its end points. □

Corollary 6. *For any fixed $0 \le \alpha \le 1$ the utility function $U^{\text{sc}}(p; \Lambda)$ can be maximized by the following ALOHA protocol: In each time slot, each node independently tosses a coin of bias t_α obtained from the optimization in Proposition 5. If the outcome is heads, it transmits with probability $\bar{\pi}_1$. Else it transmits with probability $\bar{\pi}_2$.*

4.5 Numerical Results

We present some numerical studies to gain more insight into our results. We compare the transport densities offered by both TD and SC. The network is assumed to be interference-limited, i.e., $N_1 = N_2 = 0$. The single-user SINR thresholds (or SIR, in this case), are chosen as $\theta_1 = 10$ dB, $\theta_2 = 0$ dB. The path-loss exponent $\beta = 3$.

For reference, in Fig. 4.2, we show the transport density for the near and far receivers with an attempt rate $p = 1$, for $r_1 = 0.1$ and $r_2 = 0.6$ (which are scaled by $\lambda^{-1/2}$ for a PPP with intensity λ). As Proposition 2 predicts, SC offers a greater overall utility but in terms of individual utility sub-functions, we find a Pareto improvement by switching from TD to SC by choosing an appropriate power α to the near user and using the entire bandwidth for communication.

We now discuss the implications of optimizing spatial re-use for SC and TD for both fixed and randomized link distances.

4.5.1 Fixed Link Distances

Each transmitter has a pair of designated receivers at fixed distances r_1 and $r_2 > r_1$. For the simulation, $r_1 = 0.1$ and $r_2 = 0.3, 0.6$. Fig. 4.3 shows the optimized transport densities for SC and TD for each far receiver distance. SC always has greater transport density compared to TD. Interestingly, this is *not always* the result of improved transport densities to both receivers, as we find from the utility sub-function plot in Fig. 4.4, which can be interpreted as the throughput-distance product seen at the typical transmitter in the network.

While the transport densities for *both* links are improved for $r_2 = 0.3$, when $r_2 = 0.6$ this gain comes from improving the throughput-distance product to the

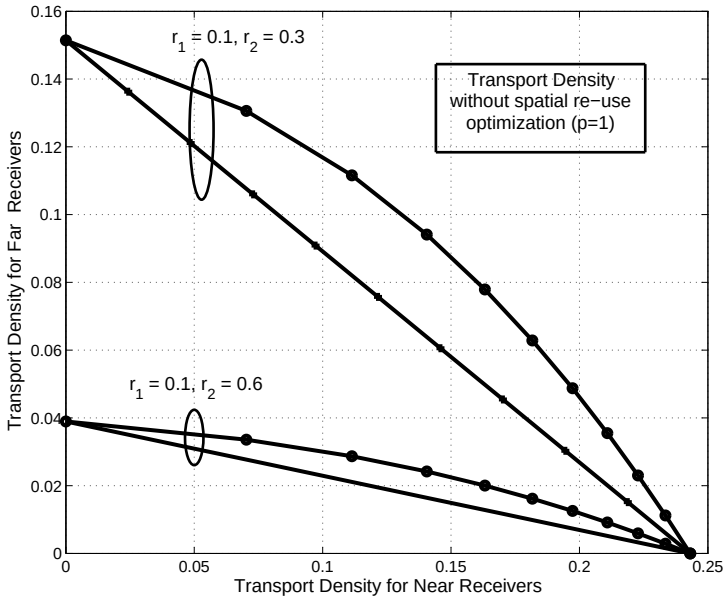

Figure 4.2. Individual utility sub-functions that constitute the utility function for fixed attempt rate $p = 1$ for near receiver distance $r_1 = 0.1$ and far receiver distance $r_2 = 0.6$. These functions are compared for TD (solid black lines), SC (lines with circular markers).

near receivers at the expense of the far receivers. This is a result of increased disparity in optimal spatial reuse among the near and far receivers for 0.6, as discussed in the following.

For fixed r_1 and SINR thresholds θ_1 and θ_2 , the optimal attempt rate $\bar{\pi}_1$ remains unchanged. In our case this is $\bar{\pi}_1 = 1$. For $r_2 = 0.3$, this is also the optimal attempt rate for the far receiver, i.e., $\bar{\pi}_2 = 1$. Thus an attempt rate of 1 simultaneously maximizes both U_1 and U_2, resulting in a large gain from SC over TD. On the other hand, when $r_2 = 0.6$, the optimal attempt rate falls to about 0.37, much below 1. Here SC chooses a rate between 0.37 and 1, depending on the power allocation parameter α, as suggested by Proposition 5, but TD uses the optimal attempt rate for each subslot.

4.5.2 Random Link Distances

Here we provide results for a specific model of randomness that is a natural extension to the case with fixed link distances discussed above. For some $0 < a < b$, assume that the near receiver distance $r_1 \sim \text{Unif}(0, a)$ and the far user distance $r_2 \sim \text{Unif}(a, b)$, independent of each other. Clearly, these are log-concave probability density functions; thus Proposition 5 applies. For the plots we set $a = 0.2$ and $b = 1$, so that $\mathbb{E}[r_1] = 0.1$ and $\mathbb{E}[r_2] = 0.6$. The results are shown in Fig. 4.5.

For this model, it can be shown that the maximizer for U_1 is $\bar{\pi}_1 = 1$, while that for U_2 is $\bar{\pi}_2 = \min(1, \frac{\ln(b^2/a^2)}{\gamma_2(b^2-a^2)}) \neq \bar{\pi}_1$ in general. Thus for fixed transmission rates, the utility gain from SC is actually a function of the parameters a and b that determine the receiver placement.

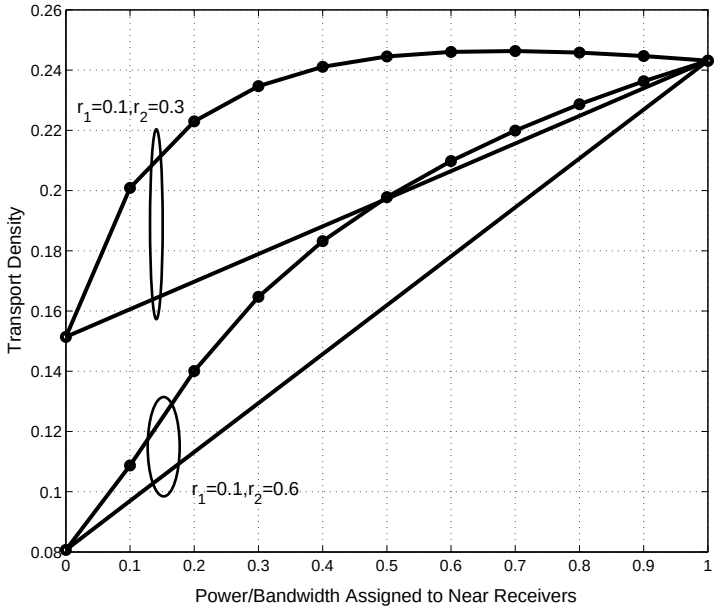

Figure 4.3. Optimized utility function for fixed near receiver distance $r_1 = 0.1$ for two far receiver distances $r_2 = 0.3$ and $r_2 = 0.6$. For each case, these functions are compared for TD (solid black lines), SC with perfect SD (dashed lines) and SC without perfect SD (lines with circular markers).

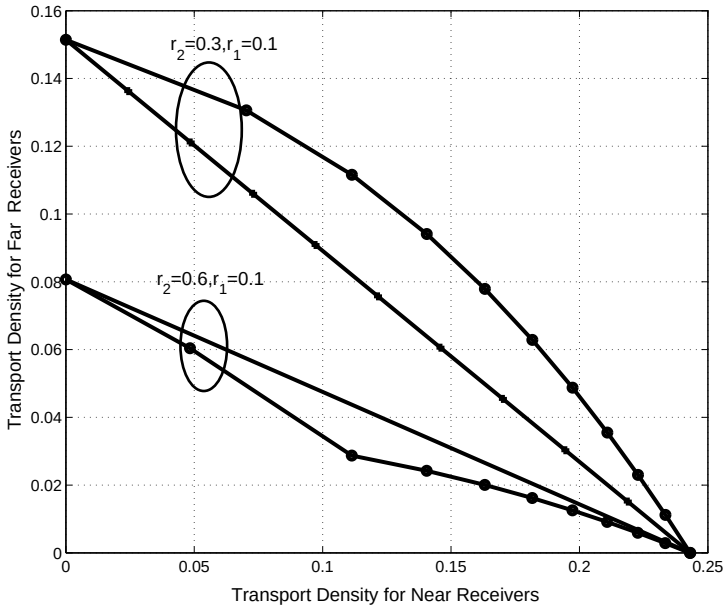

Figure 4.4. Individual utility sub-functions that constitute the optimized utility function in Fig. 4.3, with the same legend.

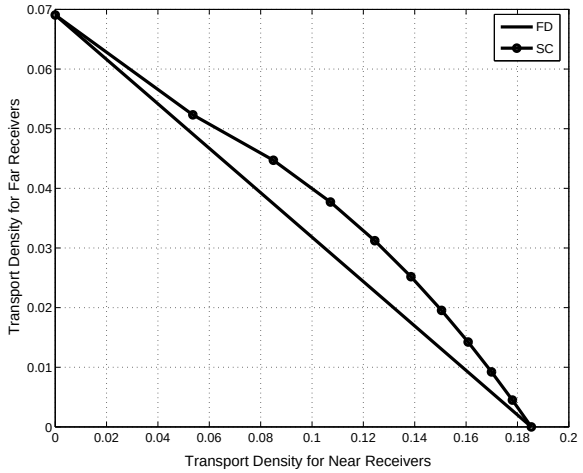

Figure 4.5. Individual utility sub-functions that constitute the optimized utility function for near receiver distance $r_1 \sim \text{Unif}(0, a)$ and $r_2 \sim \text{Unif}(a, b)$ for $a = 0.2$ and $r_2 = 1$.

4.6 Conclusions

We have analyzed SC—an information-theory inspired multipacket transmission scheme—with conventional TD in a stochastic geometric setting. We compared the these schemes by introducing a utility function that measures the effective rate of information transfer in space. While TD can adapt its spatial reuse independently for each link, the utility-maximizing spatial reuse for SC is always a compromise between maximizing the utility sub-function to each receiver separately. Since optimal spatial reuse is a function of the network geometry, the utility seen at the typical receivers from SC depends on the geometry of the receiver node placement and the chosen transmission rates. To obtain benefits from SC, for a given a set of transmission rates and a fixed near receiver distance, the far receivers must be placed at a distance far enough from their intended transmitters to provide long-range connectivity but close enough to ensure that the optimal spatial reuse to serve them is not very different from that of the near receivers.

CHAPTER 5

COORDINATED TRANSMISSION IN RANDOM NETWORKS

5.1 Introduction

Traditional scheduling algorithms assume a simple collision model to decide on the set of active links. In particular, the use of this model will not permit more than one node to transmit to a given receiver in any given time slot or time subslot. However, this restriction can be relaxed for those receivers capable of multi-packet decoding (MPD) of transmissions from a cluster of transmitters. Such nodes can be built, for example, by MIMO processing at the receiver [49], or by successive/joint decoding that are reminiscent of capacity-achieving schemes for the information-theoretic multiple access channel [26].

When many such MPD-capable nodes exist in a network, the problem of scheduling becomes important. In [50] the authors propose a random scheduling algorithm with MPD-capable nodes. However, their packet reception model does not model interference from transmitters communicating to other MPD receivers.

A more realistic model for an ad hoc setting needs to incorporate such inter-cluster interference from other clusters (possibly due to uncoordinated transmissions) as well as the network geometry. When these inter-cluster interactions are factored in, the multiple-access scheme that each transmitter cluster adopts locally can have a network-wide impact in the form of interference. If each cluster

adopts an orthogonal scheme such as Time Division Multiple Access (TDMA), we locally break down the multiple access problem into the formation of several non-interfering links, at the cost of poor bandwidth efficiency. However, more sophisticated approaches (such as the capacity-achieving scheme for the Gaussian MAC) that rely on coordinated transmissions and joint/successive decoding at the receiver make more degrees of freedom available per transmitter at the cost of increasing the overall density of interferers. If the objective is to increase the throughputs seen in a typical cluster in the network, it is not clear which of these considerations - increasing bandwidth efficiency or managing interferer density - should take precedence in protocol design.

We study this trade-off by studying a network consisting of many randomly placed *multiple-access* clusters. Each cluster consists of a number of transmitters that wish to communicate with a single receiver equidistant from each of its intended transmitters. We take the first step towards understanding the trade-off described above and analyze two well-known schemes for a symmetric multiple-access problem: TDMA and coordinated transmission inspired by the capacity-achieving scheme for a symmetric Gaussian MAC (GMAC). Using a combination of analytical and numerical approaches, we find that the performance of these schemes depends on the link distance and the transmission rate. Specifically, for a given set of transmission rates, the coordinated scheme is found to be useful only at small link distances. When nodes following TDMA are allowed to concentrate their power, i.e., increase their power in their subslots while satisfying the average power constraint, TDMA can in fact offers a competitive design choice. The increased interference from coordination also degrades the performance of the low-complexity successive decoding strategy.

5.2 System Model

5.2.1 Network Geometry

The set of receivers is a unit intensity homogeneous Poisson Point Process (PPP) $\Phi = \{x_i\} \in \mathbb{R}^2$, which we refer to as the ground process as in Chapter 4. For each receiver $x \in \Phi$, we place a cluster of K transmitters marked $1, 2, \ldots, K$ respectively, at $x + r_{x,k}$, $k = 1, 2, \ldots K$, where $r_{x,k}$ are iid random variables (in both x and k) drawn from a distribution F_r. The transmitter marked k in a cluster is called the k^{th} transmitter or user in the cluster. Denote the transmit decision of the k^{th} node attached to receiver node i by a binary variable t_{ik}. Thus the set of transmitters Φ_t is a Poisson cluster process [11] formed by the union of K unit-intensity, marked homogeneous PPPs $\Phi_t^{(k)} = \{(x + r_{x,k}, t_{x,k}) : x \in \Phi\}$, $k = 1, 2, \ldots, K$. It is worth noting Φ_t is the union of K *correlated* point processes $\Phi_t^{(k)}$ sharing a common randomness Φ. In this chapter, we assume $|r_{x,k}| = r$ is fixed and is known. We label the nodes in the *typical cluster* by $\{\mathrm{D}, \mathrm{S}_1, \mathrm{S}_2, \ldots, \mathrm{S}_K\}$, where D is the receiver node located at the origin and S_k is the k^{th} *typical* transmitter or user in the cluster located at r_{0k}. For ease of exposition we derive results for $K = 2$.

5.2.2 Communication Model

5.2.2.1 Medium Access

As in the broadcast case, we assume packet queues at all transmitters are backlogged to ensure their participation in medium access. As before we assume super-ALOHA. The marks for each transmit cluster are drawn from a common $K-$dimensional joint distribution, independently from other clusters. The mark of the k^{th} transmitter in each cluster has a marginal distribution which is Bernoulli

with parameter p_k.

In this chapter, we study the special case is when all links in a cluster are scheduled simultaneously, i.e., $t_{ik} \equiv t_i$ with some probability p. For orthogonal multiple-access, each link uses an independent ALOHA protocol.

5.2.2.2 Packet Transmission

Transmitters have a unit average power constraint per degree of freedom and use Gaussian signaling. The noise power spectral density at each receiver is N_0 (in W/Hz). The path-loss model follows a power law with exponent $\beta > 2$. The fading between any two nodes is iid block Rayleigh fading. Each receiver has full CSI only from all its intended transmitters. We further assume that transmitters have no CSI and do not use power control. All clusters use a common transmission scheme, the parameters of which are fixed during design time.

Packet transmissions are slotted and encoding and decoding are done on a per-slot basis, and immediate error-free ACK/NACK is available (i.e., we adopt a per-slot outage-based model). The number of channel uses during each time slot is large enough to permit the use of information-theoretic results. Each receiver treats inter-cluster interference as noise, which is optimum in the weak-interference regime [45].

5.3 Multiple Access Strategies

When user k is assigned the entire bandwidth (or all the time slots), it communicates using a capacity-achieving single-user AWGN channel code with an SNR threshold θ, which we call the *single-user threshold*.

5.3.1 Orthogonal Multiple Access

Users transmit in non-overlapping time slots (TDMA) or frequency bands (FDMA). This partition is common throughout the network. Without loss of generality, we assume TDMA-type multiple access, where a single time slot of unit duration is divided into two subslots, of durations $u_1 = u$ and $u_2 = 1 - u$. If transmitters marked k use ALOHA with transmit probability p_k and encode their packets using a channel code with SNR threshold $\tilde{\theta}_k$, the transmission rate R_k, packet success probability $p_{s,k}$ and the local throughput T_k at the typical cluster are, respectively, defined as

$$R_k \triangleq C(\tilde{\theta}_k) \tag{5.1}$$

$$p_{s,k} \triangleq \mathbb{P}(\mathsf{SINR}_{\mathrm{S}_k \to \mathrm{D}} \geq \tilde{\theta}_k) \tag{5.2}$$

$$T_k \triangleq p_k p_{s,k} R_k \tag{5.3}$$

where $C(x) \equiv \log(1 + x)$ for $x \geq 0$. Note that in general $\tilde{\theta}_k$ is a function of user k's bandwidth u_k. We study two approaches:

1. Naive TDMA, where transmitters marked k use their allotted subslot with unit power spectral density and continue to use codes with the single-user threshold θ.

2. TDMA with power concentration (PC-TDMA), where transmitters marked k boost their psd in their allotted subslot to $1/u_k$ and use a Gaussian channel code with SNR threshold θ/u_k.

We use subscripts n and pc respectively for naive TDMA and PC-TDMA for the parameters defined in (5.1)-(5.3).

5.3.2 Coordinated Multiple Access

5.3.2.1 Coordinated Multiple Access in a Single Cluster Network

We use a scheme inspired by the capacity-achieving scheme for a two-user symmetric GMAC with single-user threshold θ [26]. The scheme has two modes:

1. Single-User Mode: Only one of the two users transmits at a rate $C(\theta)$.

2. Coordinated Mode: The transmitters communicate using the entire bandwidth, using the rate pairs

$$M_1' = (C(\theta), C(\theta/(1+\theta))) \tag{5.4}$$

$$M_2' = (C(\theta/(1+\theta)), C(\theta)). \tag{5.5}$$

   We will call the user transmitting at the single-user rate by the *full-rate user* and the other low-rate user as the *overlaid user*.

Any other operating point can be obtained by time-sharing between these points. A procedure of practical interest is successive decoding (SD), that achieves capacity for GMAC [26]. The receiver decodes the message encoded at the lowest rate first. If unsuccessful, an error is declared. Else the decoded bits are re-encoded, and their contribution to the receiving signal is removed. The message with the next lowest rate is decoded next, until messages from all users are decoded. It is also known that for a $K-$user GMAC, the capacity region is defined by the convex hull of $K!$ vertices and the origin, where each of the $K!$ vertices corresponding to a different order of successive decoding.

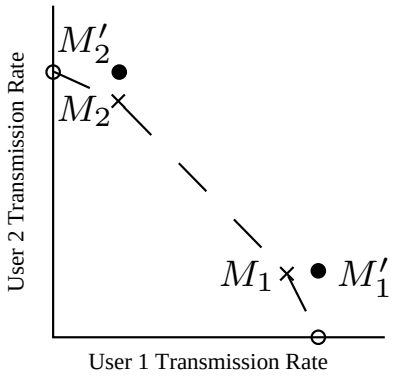

Figure 5.1. Transmission rates chosen for coordinated multiple access. The hollow circles represent the single-user mode.For the coordinated mode, we show the transmission rate-pairs chosen for a network with just one cluster (black circles), and with many clusters ($\times-$marks). The dashed line represents the set of effective transmission rates achievable by time-sharing among adjacent points.

5.3.2.2 Coordinated Medium Access in a Network with Many Clusters

We will capture the essence of the above scheme - that of overlaid transmission and successive decoding - to devise a scheme in a network with many clusters. As before, it has two modes:

1. Single-User Mode: Only one user per cluster transmits using a code with SNR threshold θ. The single-user mode for the k^{th} user corresponds to TDMA with $u_k = 1$.

2. Coordinated Mode:

   (a) Corner Point 1: User 2 is the overlaid user. User 1 is called the *high-rate* user. $M_1 = (C(\xi_1), C(\xi_2))$.

   (b) Corner Point 2: User 1 is the overlaid user. User 2 is the high-rate user. $M_2 = (C(\xi_2), C(\xi_1))$.

As before, any other operating point can be obtained by time-sharing between these points. When every cluster operates in the coordinated mode, there will be a greater spatial density of interferers resulting in a higher level of interference. Unlike in the single-cluster case, single-user and coordinated modes operate at different levels of interference. This difference in the chosen transmission rates shown in Fig. 5.1.

We thus pose the question: Given a channel access mechanism across clusters, what is the throughput on each typical link $S_k \to D$, for $k = 1, 2$ in the coordinated mode? Without loss of generality, we analyze the first corner point M_1 where $R_1^{\mathrm{c}} \triangleq C(\xi_1)$, $R_2^{\mathrm{c}} \triangleq C(\xi_2)$.

If the SD procedure at this operating point, user 2 (the overlaid user) is decoded first before decoding user 1. Thus at the typical receiver D, the packet success probability from S_2 is

$$p_{s,2}^{\mathrm{c}} \triangleq \mathbb{P}(\mathsf{SINR}_{\mathrm{S}_2 \to \mathrm{D}} \geq \xi_2). \tag{5.6}$$

If decoded correctly, the packets from S_1 are decoded. Therefore

$$p_{s,1}^{\mathrm{c}} \triangleq q_{12}^{\mathrm{c}} p_{s,2}^{\mathrm{c}}, \tag{5.7}$$

where q_{12}^{c} is the conditional success probability for decoding high-rate user's packet given that overlaid user's packet have been decoded correctly.

5.4 Average Throughput

5.4.1 Orthogonal Multiple Access

With orthogonal multiple access, from the Poisson property of Φ and the system model in Section 5.2.1, the interference I_k at the typical receiver D when decoding its k^{th} user S_k is

$$I_k = \sum_{(x+r_{x,k},t_{x,k})\in\Phi_t^{(k)}\setminus\{r_{0,k},t_{0,k}\}} t_{ik} g_{ik}(x_i)^{-\beta},$$

which is determined by only *one* constituent point process of $\Phi_t = \cup_{k=1}^{K}\Phi_t^{(k)}$. Since $\Phi_t^{(k)}$ is Poisson for each k, so is its reduced Palm distribution (by Slivnyak's theorem, see [8]). We can thus apply known results [47] to derive the success probabilities.

Proposition 7. *(Success Probabilities with naive TDMA, PC-TDMA). For a transmit probability p_k, the success probabilities $p_{s,k}^{\mathrm{n}}$,$p_{s,k}^{\mathrm{pc}}$ and for naive TDMA, PC-TDMA and are respectively*

$$p_{s,k}^{\mathrm{n}} = \exp(-p_k\gamma r^2 - \theta r^{\beta} N_0) \tag{5.8}$$

$$p_{s,k}^{\mathrm{pc}} = \exp(-p_k u_k^{-\delta}\gamma r^2 - \theta r^{\beta} N_0) \tag{5.9}$$

for $k = 1, 2$, $\delta \triangleq 2/\beta$ and

$$\gamma \triangleq \pi\theta^{\delta}\Gamma(1+\delta)\Gamma(1-\delta).$$

Proof. Readily obtained by specializing (5.2) to a homogeneous PPP (see e.g., [47], [12]). □

Comparing PC-TDMA and naive TDMA, we find that interference limits the benefits of power concentration. In fact for a homogeneous Poisson-distributed transmitter nodes with uncoordinated transmissions, naive TDMA can outperform PC-TDMA in average throughput at small bandwidth allocations, as shown in Corollary 8 below.

Corollary 8. *For any transmit probabilities p_k^{n} and p_k^{pc} chosen for naive TDMA and PC-TDMA respectively, there exists a $u_k^* > 0$ such that $T_k^{\mathrm{n}} > T_k^{\mathrm{pc}}$ for $u_k < u_k^*$.*

Proof. Using the expressions for success probabilities from Proposition 7 in the throughput expression (5.3) we can write for all $u_k > 0$

$$\frac{T_k^{\mathrm{n}}}{T_k^{\mathrm{pc}}} \propto \frac{\exp(\gamma r^2 (p_k^{\mathrm{pc}} u^{-\delta}))}{C(\theta_k/u_k)}.$$

Since $\lim_{u_k \to 0} T_k^{\mathrm{n}}/T_k^{\mathrm{pc}} = \infty$, $\exists u_k^* > 0$ such that $T_k^{\mathrm{n}}/T_k^{\mathrm{pc}} > 1 \ \forall u_k < u_k^*$. □

Corollary 8 also holds for the respective throughput-maximizing transmit probabilities $\bar{p}_k^{\mathrm{n}}$ and $\bar{p}_k^{\mathrm{pc}}$. As a result, for fixed link distances and link design SNRs, there exists $u^* = \min_k u_k^*$ for all classes of transmitters, where a Pareto improvement is possible if transmitters marked k switch to naive TDMA from PC-TDMA. Intuitively, this happens because at small u_k, PC-TDMA concentrates power in a very small subslot and allocates a correspondingly large transmission rate (SINR threshold) for this subslot. When the thresholds become too large, outage events become frequent enough to negate the benefit of using a higher spectral efficiency. The average throughputs can now be evaluated from the definition (5.3).

5.4.2 Coordinated Multiple Access

5.4.2.1 Co-location Approximation

The interference I at the typical receiver due to transmitters that do not belong to the typical cluster is

$$I = \sum_{k} \sum_{x \in \Phi_t^{(k)} \setminus \{r_{0,k}, t_{0,k}\}} t_{x,k} g_{x,k} (x + r_{x,k})^{-\beta}. \tag{5.10}$$

Thus the interferers form a cluster process $\Phi_t = \bigcup_k \Phi_t^{(k)}$. To retain the analytical simplicity of our treatment and yet gain insight into the effect of increased interference, we restrict our discussion to a regime where the intra-cluster transmitter node separation is small compared to the average distance between receiver nodes of the network (which is $1/2\sqrt{\lambda}$ for a homogeneous PPP of intensity λ). Here each transmitter cluster can be approximated by a single multi-antenna virtual transmitter node located at an arbitrarily chosen transmitter (say $x + r_{x,1}$) in the cluster. The antenna separation at this virtual node is assumed to be sufficient to create independent fading paths. The resulting transmitter point process is thus a homogeneous PPP with unit intensity, resulting in the approximation

$$I \approx \sum_{x \in \Phi_t^{(1)} \setminus \{r_{0,1}, t_{0,1}\}} t_i \left(\sum_{k} g_{x,k} \right) (x + r_{x,1})^{-\beta}, \tag{5.11}$$

assuming super-ALOHA. Although co-location of transmitters captures the increase in interference from concurrent transmissions, it does not precisely capture its effect in the vicinity of each interferer cluster where the geometry of the interferers also becomes important. This limits the utility of the co-location approximation in a more general case. In the next subsection we use this approximation

to derive packet success probabilities for coordinated transmission with super-ALOHA. A numerical validation of this approximation is presented in Section 5.5.1 (see Fig. 5.2)

5.4.2.2 Success Probabilities using the Co-location Approximation

Proposition 9. *(Success Probability with Coordinated Transmissions and super-ALOHA). If every cluster operates at the first corner point M_1 for super-ALOHA with transmit probability p, the success probabilities (5.6) and (5.7) at the typical receiver are, respectively,*

$$p_{s,2}^{\mathrm{c}} = \frac{\exp(-p\gamma_2 r^2 - \xi_2 r^\beta N_0)}{1+\xi_2} \tag{5.12}$$

$$p_{s,1}^{\mathrm{c}} = \frac{\exp(-p\gamma_1 r^2 - (\xi_2 + \xi_1 + \xi_1\xi_2) r^\beta N_0)}{1+\xi_2} \tag{5.13}$$

where $\gamma_1 \triangleq |b(0,1)|\Gamma(2+\delta)\Gamma(1-\delta)(\xi_1+\xi_2+\xi_1\xi_2)^\delta$, $\gamma_2 \triangleq \pi\Gamma(2+\delta)\Gamma(1-\delta)\xi_2^\delta$.

Proof. Suppose g_k $(k = 1, 2)$ denote the fading gains from each of the typical transmitters. Recall from (5.7) that

$$\begin{aligned} p_{s,2}^{\mathrm{c}} &= \mathbb{P}(\mathsf{SINR}_{\mathrm{S}_2\to\mathrm{D}} \geq \xi_2) \\ &= \mathbb{P}\left(\frac{g_2 r^{-\beta}}{g_1 r^{-\beta} + I_{\Phi_t\setminus\{S_1,S_2\}} + N_0} \geq \xi_2\right). \end{aligned}$$

Since $g_2 \sim \exp(1)$, using standard arguments (see e.g., [47] for single-user decoding) we can show that $p_{s,2}^{\mathrm{c}}$ can be written as the Laplace transform evaluated at $\xi_2 r^\beta$ of the sum distribution of the three denominator terms. Given that these random variables are mutually independent, the Laplace transform of their sum distribution is the product of the Laplace transforms of the marginal distributions.

The latter are known to be respectively:

$$\begin{aligned}\mathcal{L}_1(s) &= 1/(1+sr^{-\beta}) \\ \mathcal{L}_2(s) &= \exp(-p|b(0,1)|\mathbb{E}[h_2^{\delta}]\Gamma(1-\delta)s^{\delta}) \\ \mathcal{L}_3(s) &= \exp(-sN_0),\end{aligned}$$

where h_2 is fading variable representing Nakagami-2 fading. Using the properties of gamma functions it is easy to show that $\mathbb{E}[h_2^{\delta}] = \Gamma(2+\delta)$[1]. Setting $s = \xi_2 r^{\beta}$ we get (5.12). From (5.6) we know that $p_{s,1}^{\mathrm{c}} = q_{12}^{\mathrm{c}} p_{s,2}^{\mathrm{c}}$. Writing $\tilde{I} = I_{\Phi_t \setminus \{S_1,S_2\}} + N_0$, we expand this using Bayes' rule as the joint probability

$$p_{s,1}^{\mathrm{c}} = \mathbb{P}\left(\frac{g_1 r^{-\beta}}{\tilde{I}} \geq \xi_1, \frac{g_2 r^{-\beta}}{g_1 r^{-\beta} + \tilde{I}} \geq \xi_2\right).$$

Utilizing the mutual independence of g_1, g_2 and $\tilde{I}$, the right hand side can be expressed as

$$\int_0^{\infty} \underbrace{\mathbb{P}(g_2 r^{-\beta} \geq \xi_2(g_1 r^{\beta} + x), g_1 r^{-\beta} \geq \xi_1 x)}_{\text{Term1}} d\mathbb{P}(\tilde{I} \leq x). \tag{5.14}$$

Term1 can be expressed as

$$\text{Term1} = \int_{\xi_1 x}^{\infty} \mathbb{P}(g_2 r^{-\beta} \geq \xi_2(y+x)) \exp(-y r^{\beta}) r^{\beta} dy,$$

where we have used the fact that $g_1 r^{-\beta}$ has a probability density. Since $g_2 \sim \exp(1)$, the integrand reduces to $\exp(-\xi_2 r^{\beta}(y+x))$. Combining the two exponen-

[1]For coordinated transmission with K users, this can be generalized to $\mathbb{E}[h_K^{\delta}] = \Gamma(K+\delta)$ (Nakagami$-K$ fading).

tials in y we obtain

$$\begin{aligned} \text{Term1} &= \exp(-\xi_2 r^\beta x) \int_{\theta x}^{\infty} \exp(-(1+\xi_2) r^\beta y) r^\beta dy \\ &= \frac{\exp(-(\xi_2 + \xi_1 + \xi_1\xi_2) r^\beta x)}{1+\xi_2}. \end{aligned}$$

Plugging this result into the first step (5.14) yields

$$p_{s,1}^{\mathrm{c}} = \frac{\mathcal{L}_{\tilde{I}}((\xi_1 + \xi_2 + \xi_1\xi_2) r^\beta)}{1+\xi_2}.$$

Since Φ_t is well approximated by a homogeneous PPP with intensity p, we get (5.13). □

5.5 Numerical Results

5.5.1 Validating the Co-location Approximation

Relative to their intended receiver, suppose the two transmitters of each cluster have uniformly random and independent orientations but are located at a fixed distance r. Conditioned on the location of one transmitter, the other transmitter is located uniformly randomly inside a ball of radius $2r$ centered at the known transmitter. In general if an angular spread of $\omega \leq \pi$ is permitted between the transmitter orientations within a cluster (i.e., the orientations are no longer iid within a cluster), the radius of this ball is $2r \sin(\omega/2)$. The co-location approximation assumes that the distance between transmitters in a cluster is small.

We validate this approximation as follows. We create realizations of the point process $\Phi_t^{(1)}$, with unit intensity without loss of generality. We fix a small link distance $r \ll 1$ and an ω. Centered at each point in this process, we place the

point marked 2 uniformly randomly inside a ball of a radius $2r\sin(\omega/2)$. These latter points correspond to the second transmitter point process $\Phi_t^{(2)}$.

For each realization, we measure the interference at the origin using the exact locations from (5.10) and from the approximation (5.11), and compare the empirical complementary (cumulative) distribution functions (CCDFs) of the interference for both these cases. Some results are shown in Fig. 5.2 for $r = 0.1, \omega = \pi$ (independent orientations). We find that the approximation is a good fit as long as r remains much smaller than the characteristic distance $1/2\sqrt{\lambda}$ of the network.

5.5.2 Comparing Orthogonal and Coordinated Transmission

We present numerical results to gain insight into the results presented in Section 5.4. The effect of inter-cluster interference is the most apparent in the interference-limited regime ($N_0 \to 0$). We use results derived in Section 5.4 to analyze a system of two-user symmetric multiple-access clusters with link distance $r = 0.05, 0.1$ ($\ll 1$) for two values of a single-user threshold $\theta = 0\,\mathrm{dB}$. The path-loss exponent $\beta = 3$. For the coordination scheme we let $\xi_1 = \theta$ and $\xi_2 = \theta/(1+\theta)$.

We use the throughput-maximizing transmit probability for both users for TDMA. By plugging in the result from Proposition 7 into the throughput definition (5.3), the optimum transmit probability is $\min(1, a^{-1})$, where $a > 0$ is a function of link distance, path-loss exponent and the SNR threshold. For the chosen set of parameters above, $a^{-1} < 1$; hence the optimal transmit probability is 1.

When the users adopt the coordinated scheme described in Section 5.3.2.2, the optimum transmit probability depends on whether full-rate users' or the overlaid

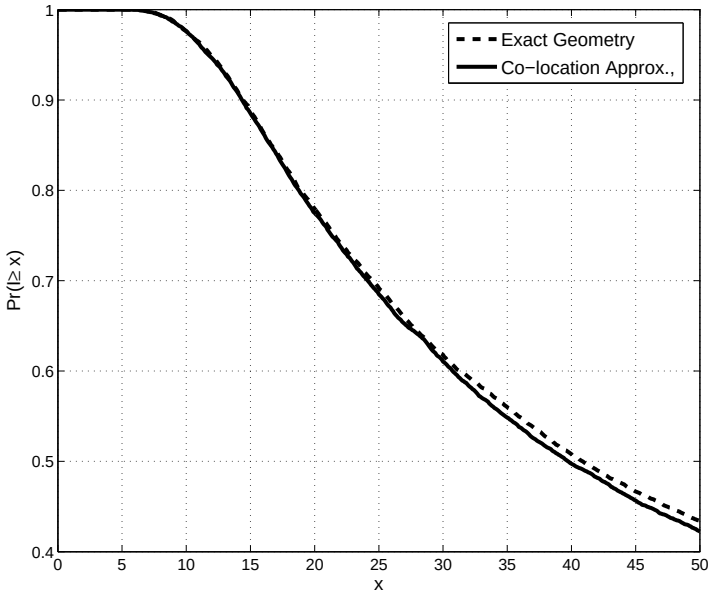

Figure 5.2. CCDF of I from (5.10) and its co-location approximation (5.11) for $r = 0.1$, $\omega = \pi$, $\lambda = 1$.

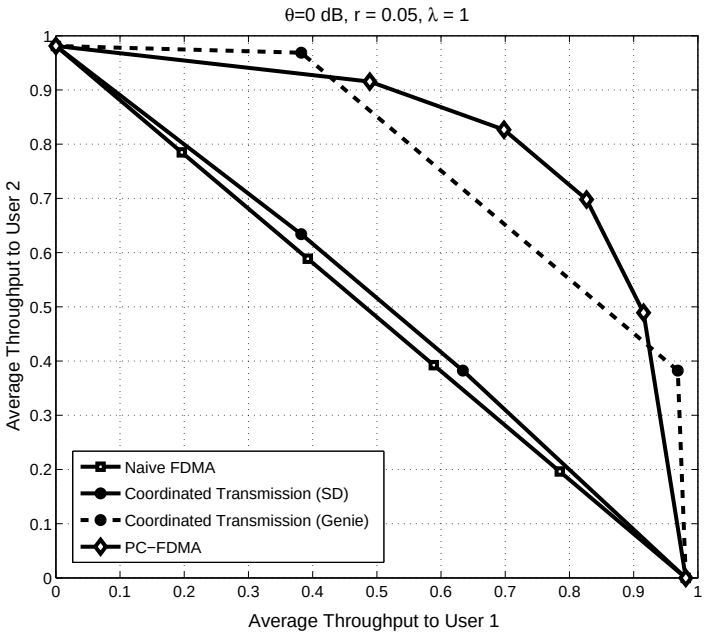

Figure 5.3. Average Link Throughputs for $\theta = 0\,\mathrm{dB}$, $r = 0.05$.

users' throughput is to be maximized. Since these users transmit at different rates, these probabilities are in general different. However, when both the link distances and the transmission rates are small (as in the present parameter set), both these probabilities will be equal to 1.

We compare the average throughputs per link for naive TDMA, PC-TDMA and coordinated transmission. For coordinated transmission we plot the throughputs obtained with SD and with genie-aided cancellation of the overlaid user. These results are shown for in Fig. 5.3 ($r = 0.05$) and Fig. 5.4 ($r = 0.1$). For small link distances (high SINR regime), a moderate increase in the transmission rate increases the throughput without an appreciable loss in reliability. This explains the PC-TDMA throughput gain over naive TDMA at reasonable bandwidth allocations. Error propagation from successive decoding restricts the gains

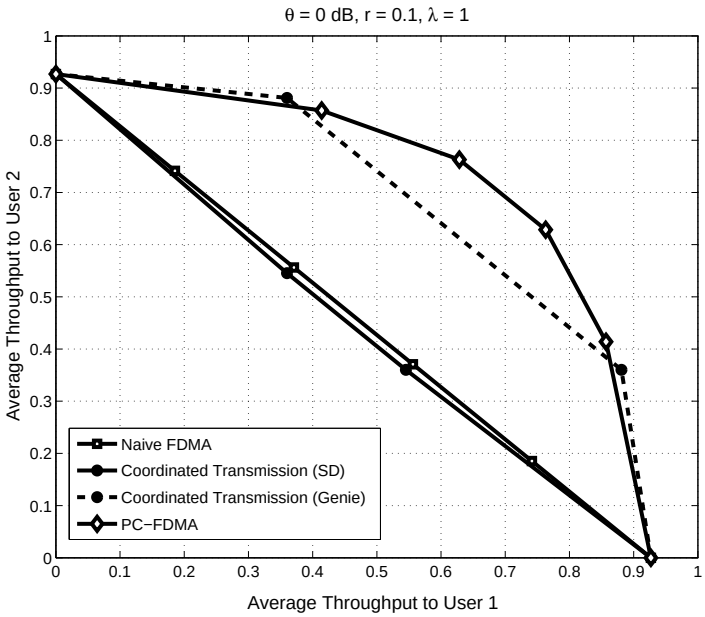

Figure 5.4. Average Link Throughputs for $\theta = 0\,\mathrm{dB}$, $r = 0.1$.

from coordinated transmission - evident only over naive TDMA - to small link distances (and for a given link distance, for small SNR thresholds). Increasing the link distance reduces the received signal power relative to the interference power (which remains fixed), worsening the error propagation problem. We find this in Fig. 5.3 and 5.4.

Even with perfect SD, for a wide range of throughputs there is a Pareto improvement by switching to PC-TDMA, i.e., trading bandwidth efficiency for lower interferer density is beneficial.

5.6 Conclusions

We have investigated the average link throughput in a network consisting of randomly placed multiple-access clusters where channel access within each cluster can be made to be coordinated but is uncoordinated across clusters. We compared two distributed schemes, namely orthogonal multiple access and a coordinated multiple access scheme inspired by the capacity-optimal scheme for the two-user Gaussian symmetric multiple access channel. We found that for a given transmission rate, the average throughput of each link with this coordinated scheme depends on the link distance. Specifically, this scheme is useful only at small link distances and even then only if peak transmit power constraints make power concentration infeasible. This is fundamentally due to an increase in interference power that restricts the throughput achievable even when there is no error propagation. The same problem also reduces the efficacy of the relatively low-complexity successive decoding strategy. Note that our comparison does not even account for the overhead necessary to establish coordinated transmission compared to orthogonal access (especially in an ad hoc setting). Our results thus show that in terms of average link throughput, orthogonal schemes are a competitive design option.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

In this chapter, we summarize the work presented in the previous chapters and place them in the broader context of this thesis, namely, the implications of using signal superposition strategies in wireless networks. In doing so, we also discuss possible avenues for future work, both theoretical and experimental.

6.1 Theoretical Results

Perhaps the main contributions of the theoretical work presented in the thesis are to identify the impact of superposition schemes on medium access and to propose a framework to quantify these effects. In both Chapters 4 and 5, we find that although superposition can encode individual packets at a higher spectral efficiency, the lack of orthogonality among signals within each cluster implies that activating one link impacts others: this is not so in orthogonal schemes. This rises a fairness question in how to choose the active links. While shorter links can tolerate more interference, activating them simultenaously with longer links raises the important fairness question of how to choose the set of active transmitters.

A smaller number of active transmitters lowers the outage on longer links at the expense of letting fewer short links be active. In Chapter 4, we illustrate this issue by comparing the disparity in the optimum access probabilties and show

by comparing the utility subfunctions of superposition and orthogonal schemes that the gains from superposition depend on the receiver geometry. An additional problem in multiple-access is the increased interference when all the links within each cluster are simultaneously activated. We study this in Chapter 5 where we show that orthogonal schemes can in fact offer a between local throughput than superposition schemes, purely because they generate less interference.

It is worth noting that in both these analyses, the medium access protocols treat the links in each cluster within each cluster as a single *hyperlink*. In this model, either the hyperlink is active or inactive: thus individual link states within each cluster are completely correlated. A natural generalization would be to relax this restriction and allow each hyperlink to be only *partially* active. This raises interesting medium access possibilities that we explore further in [13]. Another interesting possibility is superposition-enabled routing multihop networks, where common relays for two or more routes provides natural opportunities for both superposition-based broadcast *and* multiple-access schemes, giving rise to even richer problems.

6.2 Experimental Results

Although the theoretical models presented assume the use of long block length point-to-point Gaussian codes to gain insights into some fundamental design issues, such schemes can only be approximated in certain conditions (such as a relaxed delay constraint). However, as we show in Chapters 2 and 3, it is possible to exploit the benefits of signal superposition even at small to moderate payload sizes using simple off-the-shelf point-to-point channel codes. For example, the smallest payload size in Chapter 2 is 93 bytes (for BPSK, rate-1/2) while the

largest is 640 bytes (for 16-QAM, rate-5/6), suggesting that signal superposition can be promising even in practice.

This work can be extended in several directions. One natural extension would be to leverage both advanced coding techniques such as turbo or LDPC codes and advanced receiver architectures based on iterative interference cancellation (see, e.g., [40, 41] and the references therein). A different direction would be to apply build superposition-coded systems for scenarios where such superposition is known to be theoretically optimal, e.g., in multiple-access channels (models for cellular uplinks). An important difference in these cases is that the superposition process occurs at the *receiver* rather than at the transmitter in the BC; thus signals that are superposed have different propagation paths. This difference opens up a new set of problems in system design, such as node synchronization, channel code selection and receiver design. Allowing multiple antennas at the TX and/or RX adds another dimension to the design space.

Another promising line of investigation would be to build upon the existing physical layer to design medium access protocols that draw on the insights developed from Section 6.1. Our preliminary study involving just one transmitter has shown there exist many interesting tradeoffs [24]. Experiments would permit a more direct study of larger networks of mutually interfering broadcast clusters.

BIBLIOGRAPHY

1. J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Info. Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

2. A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. part i. system description," *IEEE Trans. Communications*, vol. 51(11), pp. 1927–1938, 2003.

3. ——, "User cooperation diversity. part ii. implementation aspects and performance analysis," *IEEE Trans. on Comm.*, vol. 51(11), pp. 1939–1948, 2003.

4. R. Mudumbai, D. Brown, U. Madhow, and H. Poor, "Distributed transmit beamforming: Challenges and recent progress," *IEEE Comm. Magazine*, vol. 47(2), pp. 102–110, 2009.

5. A. Giridhar and P. Kumar, "Toward a theory of in-network computation in wireless sensor networks," *IEEE Communications Magazine*, vol. 44(4), pp. 98–107, 2006.

6. Y. Hatano and M. Mesbahi, "Agreement over random networks," *IEEE Trans. Automatic Control*, vol. 50, No. 11, pp. 1867–72, Nov. 2005.

7. M. Porfiri and D. J. Stilwell, "Consensus seeking over random weighted directed graphs," *IEEE Trans. Automatic Control*, vol. 52, No. 9, pp. 1767–73, Sept. 2007.

8. D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and its Applications*. John Wiley & Sons, 2/e, 1995.

9. F. Baccelli and B. Blaszczyszyn, *Stochastic Geometry and Wireless Networks*, Foundations and T. in Networking, Eds. NOW, 2009.

10. R. K. Ganti and M. Haenggi, *Interference In Large Wireless Networks*. NOW Publishers, 2009.

11. M. Haenggi, *Stochastic Geometry for Wireless Networks*. Cambridge, 2012.

12. M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic Geometry and Random Graphs for the Analysis and Design of Wireless Networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1029–1046, Sep. 2009.

13. S. Vanka and M. Haenggi, "Superposition-coded medium access," *IEEE Trans. Info. Theory*, 2012 (In preparation).

14. ——, "Analysis of the benefits of superposition coding in random wireless networks," in *IEEE International Symposium on Information Theory (ISIT '10), (Austin, TX)*, June 2010.

15. ——, "Coordinated packet transmission in random wireless networks," in *IEEE Global Communications Conference (GLOBECOM'10), (Miami, FL), Dec. 2010*, 2010.

16. S. Vanka, V. Gupta, and M. Haenggi, "On consensus over stochastically switching directed topologies," in *2009 American Control Conference (ACC'09)*, St. Louis, MO, Jun. 2009.

17. ——, "Effect of Network Geometry and Interference on Consensus in Wireless Networks," in *2009 International Conference on the Dynamics of Information Systems (ICDIS'09)*, Gainesville, FL, Jan. 2009.

18. S. Vanka, M. Haenggi, and V. Gupta, "Distributed averaging in dense wireless networks," in *2009 IEEE Global Communications Conference (GLOBECOM'09)*, Honolulu, HI, Dec. 2009.

19. S. Vanka, V. Gupta, and M. Haenggi, "Power-Delay Analysis of Consensus Algorithms on Wireless Networks with Interference," *International Journal of Systems, Control, and Communications*, vol. 2, no. 1/2/3, pp. 256–274, 2010.

20. S. Vanka, M. Haenggi, and V. Gupta, "Convergence speed of the consensus algorithm with interference and sparse long-range connectivity," *IEEE J. Sel. Top. Sig. Proc., Spl. Issue on Signal Processing in Gossiping Algorithms Design and Applications*, vol. 5, pp. 855–865, Aug. 2011.

21. R. K. Ganti, Z. Gong, M. Haenggi, C. Lee, S. Srinivasa, D. Tisza, S. Vanka, and P. Vizi, "Implementation and experimental results of superposition coding on software radio," in *International Conference on Communications (ICC)*, 2010.

22. S. Vanka, S. Srinivasa, and M. Haenggi, "A practical approach to strengthen vulnerable downlinks using superposition coding," in *IEEE International Communications Conference (ICC'12)*, (Ottawa, Canada), June 2012.

23. S. Vanka, S. Srinivasa, Z. Gong, P. Vizi, K. Stamatiou, and M. Haenggi, "Superposition coding strategies: Design and experimental evaluation," *IEEE Trans. Wireless Comm.*, 2012. Accepted. [Online]. Available: www.nd.edu/~mhaenggi/pubs/twc12.pdf

24. P. Vizi, S. Vanka, S. Srinivasa, M. Haenggi, and Z. Gong, "Scheduling using superposition coding: Design and software radio implementation," in *IEEE Radio Wireless Week (RWW)*, Jan. 2011.

25. GNU Radio, http://gnuradio.org/.

26. T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley Interscience, Inc., 2006.

27. D. Tse and P. Vishwanath, *Fundamentals of Wireless Communications*. Cambridge University Press, 2005.

28. J. Mitola-III, "Software radios: Survey, critical evaluation and future directions," *IEEE Aerospace and Electronic Systems Magazine*, vol. 8(4), pp. 25–36, 1993.

29. E. Bayraktaroglu, C. King, X. Liu, G. Noubir, R. Rajaraman, and B. Thapa, "On the performance of ieee 802.11 under jamming," in *Proceedings of The 27th IEEE Conference on Computer Communications (INFOCOM), p. 1265-1273*, Apr. 2008.

30. X. Li, W. Hu, H. Yousefizadeh, and A. Qureshi, "A case study of a MIMO SDR implementation," in *Proc. IEEE Military Communications Conference (MILCOM)*, Nov. 2008.

31. K. Mandke, S. H. Choi, G. Kim, R. Grant, R. C. Daniels, W. Kim, R. W. Heath, and S. M. Nettles, "Early results on hydra: a flexible MAC/PHY multihop testbed," in *Proceedings of IEEE 65th Vehicular Technology Conference (VTC2007-Spring), pp. 1896-1900*, Apr. 2007.

32. R. Alimi, L. Li, R. Ramjee, H. Viswanathan, and Y. R. Yang, "ipack: in-network packet mixing for high throughput wireless mesh networks," in *Proceedings of The 27th IEEE Conference on Computer Communications (INFOCOM), p. 66-70*, Apr. 2008.

33. A. G. Fàbregas, A. Martinez, and G. Caire, *Bit-Interleaved Coded Modulation*. NOW Publishers, 2008.

34. G. Caire and E. Viterbo, "Upper bound on the packet error probability of terminated trellis codes," *IEEE Communication Letters*, vol. 2, pp. 2–4, Jan. 1998.

35. G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Transactions on Information Theory*, vol. 44, pp. 927–946, May 1998.

36. J. G. Proakis, *Digital Communications*. McGraw-Hill, 2000.

37. T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for ofdm," *IEEE Transactions on Communications*, vol. 45, pp. 1613–1621, 1997.

38. S. Vanka, Z. Gong, S. Srinivasa, P. Vizi, K. Stamatiou, and M. Haenggi, "Superposition coding strategies: Design and experimental evaluation," Emerging Wireless Architectures Laboratory, University of Notre Dame,

Tech. Rep. EWA2011-1, Sept. 2011, available at http://www.nd.edu/~mhaenggi/pubs/techreport_sc11.pdf. [Online]. Available: http://www.nd.edu/~mhaenggi/pubs/techreport_sc11.pdf

39. F. Baccelli, A. E. Gamal, and D. Tse, "Interference networks with point-to-point codes," *IEEE Trans. Info. Theory*, vol. 57(5), pp. 2582–2596, May 2011. [Online]. Available: http://arxiv.org/pdf/1102.2868v1

40. J. G. Andrews, "Interference cancellation for cellular systems: A contemporary overview," *IEEE Wireless Comm.*, vol. 12(2), pp. 19–29, Apr. 2005.

41. R. Zhang and L. Hanzo, "A unified treatment of superposition coding aided communications: Theory and practice," *IEEE Communications Surveys and Tutorials*, vol. 99, pp. 1–18, July 2010.

42. I. Seskar and N. B. Mandayam, "Software defined radio architectures for interference cancellation in ds-cdma systems," *IEEE Personal Communications Magazine*, vol. 6, pp. 26–34, 1999.

43. J. Zhang, J. Jia, Q. Zhang, and E. M. K. Lo, "Implementation and evaluation of cooperative communication schemes in software-defined radio testbed," in *Proc. IEEE INFOCOM*, 2010.

44. N. Jindal, J. G. Andrews, and S. Weber, "Bandwidth partitioning in decentralized wireless networks," *IEEE Trans. on Wireless Comm.*, vol. 7(12), pp. 5408–5419, 2008.

45. X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels," *IEEE Trans. Info. Theory*, vol. 55(2), pp. 689–699, Feb. 2009.

46. F. Baccelli, P. Mühlethaler, and B. Blaszczyszyn, "Stochastic analysis of spatial and opportunistic Aloha," *IEEE J. Sel. Areas in Comm.*, vol. 27(7), pp. 1105–1119, 2009.

47. F. Baccelli, B. Blaszczyszyn, and P. Mühlethaler, "An Aloha protocol for multihop mobile wireless networks," *IEEE Trans. Info. Theory*, vol. 52, pp. 421–436, 2006.

48. S. Boyd and L. Vanderberghe, *Convex Optimization*. Cambridge University Press, 2004.

49. L. Tong, Q. Zhao, and G. Mergen, "Multipacket reception in random access wireless networks: From signal processing to optimal medium access control," *IEEE Comm. Magazine*, vol. 39(11), pp. 108–112, Nov. 2001.

50. G. Mergen and L. Tong, "Random scheduling protocol for medium access in ad hoc networks," in *Proceedings of 2002 MILCOM, Anaheim, CA*, Oct. 2002.

This document was prepared & typeset with LaTeX 2ε*, and formatted with* NDdiss2ε *classfile (v3.0[2005/07/27]) provided by Sameer Vijay.*