Iterative Detection for Overloaded OFDMA with Spatial Diversity

by

Eugene Baik

S.B., Electrical Science and Engineering
Massachusetts Institute of Technology (2005)

Submitted to the Department of Electrical Engineering and Computer
Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2006

© Massachusetts Institute of Technology 2006. All rights reserved.

Author . . . .
Department of Electrical Engineering and Computer Science
August 22nd, 2006

Certified by . . . .
Moe Z. Win
Associate Professor
Thesis Supervisor

Accepted by . . . .
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Iterative Detection for Overloaded OFDMA with Spatial Diversity

by

Eugene Baik

Submitted to the Department of Electrical Engineering and Computer Science
on August 22nd, 2006, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

Orthogonal Frequency Division Multiple Access (OFDMA) systems have built-in mechanisms to mitigate the effects of the wireless multipath channel but are limited in system capacity to available bandwidth. This shortcoming can be worked around through the process of "overloading," where users are additionally multiplexed in the spatial domain to each frequency resource. To efficiently resolve non-orthogonally multiplexed users within the system, sophisticated multiple antenna receivers with multiuser detection methods are necessary.

The focus of this thesis will be the formulation of an iterative multiple antenna receiver framework for overloaded uplink OFDMA systems. Specifically, we formulate optimal MAP and reduced complexity MMSE symbol detection algorithms for the multiuser detection and single user decoding turbo loop. We verify the performance of each algorithm through Monte Carlo simulation with randomly generated multipath MIMO channels. From the results we determine the tradeoffs of algorithm complexity with performance and the effect of channel correlation on the supportable user load. Our MMSE algorithm with soft interference cancellation is observed to closely approach single user performance in low to moderately correlated MIMO channels after turbo loop iteration. Additionally, we observe that increasing the number of antennas relative to the number of overloaded users can mitigate the effects of moderate correlation to provide acceptable error performance.

Thesis Supervisor: Moe Z. Win
Title: Associate Professor

Acknowledgments

This thesis would not have been possible without the help and dedication of numerous people. I would like to start by thanking Professor Moe Z. Win and Dr. Henk Wymeersch for advising me throughout the course of my work. Professor Win has been incredibly supportive and patient of me through the ups and downs of the entire process and I can't imagine working for any other advisor. Dr. Wymeersch has helped me with every aspect of this work and I cannot thank him enough for his technical guidance and moral support. It has been a great privilege to work with Professor Win and Dr. Wymeersch and I enjoyed my time with the research group.

I thank the fellow engineers I worked with during my internships at Qualcomm, Inc. in San Diego, California. I learned a great deal during my time there and the experiences have been an influencing factor in shaping the direction of this thesis work.

Finally, I would like to thank my family for their unwavering love and support. Their words of encouragement have kept me going through the personal odyssey that my past five years at MIT have been.

Contents

List of Figures

Chapter 1

Introduction

Orthogonal Frequency Division Multiple Access (OFDMA) is rapidly being adopted as the system framework of choice for cutting-edge fourth generation (4G) cellular data networks. With the advent of increasingly sophisticated and powerful digital signal processors for communications applications through the years, the theoretical benefits possible with OFDMA are now becoming more realizable in practical systems [1]. OFDMA is a multi-carrier system that transmits its data representations in the frequency domain and uses advanced signal processing techniques to efficiently mitigate the effects of the wireless multipath channel.

A major aspect in the design of multiple access cellular systems is the simultaneous sharing of channel resources to avoid multiple access interference between users. OFDMA systems prevent this type of inter-user interference by allocating channel resources orthogonally in the frequency domain. Mobile users within a cell operate on disjoint sets of these frequency carriers such that the base station receiver of the cell can resolve each of the individual signals without worry of multiple access interference.

However, a downside of OFDMA-based systems is the hard limit on system capacity and supportable user load of orthogonal multiple access. The total number of supportable users is directly proportional to the frequency bandwidth of operation. Frequency bandwidth is tightly regulated by the Federal Communications Commission (FCC) and is hence a limited resource that network carriers expend large amounts

of financial resources to obtain.

Code Division Multiple Access (CDMA) is an alternative, and also very much prevalent, system framework based on a different non-orthogonal scheme for resource sharing. The supportable system capacity for CDMA is interference limited instead of bandwidth limited, and declines gracefully as the system is not hard-limited in the number of users it can serve. To make the overall benefits of OFDMA worthwhile, it must remain competitive with rival systems such as CDMA in the acheivable system capacity and supportable users for comparable allocations of bandwidth.

A possible remedy for this situation is to relax the constraint for inter-user orthogonality in frequency and instead exploit the degrees of freedom of the wireless channel by alternate means to allow for an increase in the number of supportable users through "overloading." One known approach is to increase the number of receiver antennas at the base station to transform the wireless channel into a multiple-input multiple-output (MIMO) channel. Increasing the number of receiver antennas increases the spatial diversity and the degrees of freedom within the wireless channel from the mobile users to the receiving base station [2, 3, 4]. By exploiting principles behind antenna array assisted multiuser detection and Space Division Multiple Access (SDMA), we can overload our system to multiplex (and at the receiver, resolve) multiple users of our system in the "spatial domain" in addition the frequency domain [5].

Unfortunately, this form of spatial multiplexing is susceptible to wide variations in inter-user interference caused by situations when there is a lack of spatial diversity within the channel. Similar to layered MIMO systems [6], analogous optimal antenna combining methods and interference cancellation methods are used in conjunction in these quasi-orthogonal, spatially multiplexed systems to allow for robustness against channels with poor spatial diversity.

The process of resolving the spatially multiplexed and interfering users at the base station receiver can be classified as a multiuser detection and decoding problem [7, 8]. Previous works in technical literature have explored the application of turbo processing loops [9, 10, 11] for iterative detection and decoding algorithms for

9

interference-level limited multiple access systems such as CDMA [12, 13, 14, 15]. Such turbo loops are not of primary interest in standard OFDMA setups because of the inherent orthogonality of the users. However, iterative receiver algorithms provide the possibility of harnessing large iterative processing gains in a multiple antenna equipped base stations for overloaded OFDMA systems [16].

The focus of this thesis will be to develop a framework for iterative receiver implementations in the context of overloaded uplink OFDMA systems with available spatial diversity. We are concerned with understanding the tradeoffs between implementation complexity and error performance of the receivers with respect to standard receivers. Additionally, we are interested in the robustness of receiver designs to the ranges of possible multipath MIMO channels. Finally, we want to generalize the costs in average user performance and complexity associated with overloading standard OFDMA systems.

This thesis document is structured as follows: Chapter 2 will cover the background material regarding cellular communications and explain the general multiple antenna OFDMA system model to be used as a foundation for the remainder of the thesis work. Additionally, the chapter will explain the derivations behind the transmission and reception of OFDM modulated signals.

Chapter 3 will detail the MIMO signal models to be used to represent received signals at the a base station. It will also provide justification for the choice of signal modelling in the context of the iterative algorithms to be developed. Chapter 4 will review the basic problem of detection and decoding in the multiuser scenario of interest. In preparation for the formulation of the iterative algorithms, the chapter will cover existing approaches to implementing detection algorithms in non-iterative receivers. Chapter 5 will contain the main contributions of this thesis work and will derive and detail the iterative detection and decoding algorithms that will make up our iterative receiver framework. We will present the system parameters and implementation details for testing the iterative processing loop in Chapter 6, which describes the framework for numerical simulation. Chapter 7 will present the error performance results of the numerical simulation and analysis of the results. The

remainder of the final chapter will present areas for possible future study and provide a brief summary of the contributions of the thesis work with concluding remarks.

Chapter 2

OFDMA System Model

2.1 Communication System Basics

A cellular system is composed of multiple, geographically separated mobile users communicating with a fixed, centrally located base station. The mobile users do not cooperate with one another and each is transmitting an independent stream of data to the base station. The communication path between the user and the base station is bi-directional. The path from the base station to the user is called the downlink while the path from the user to the base station is referred to as the uplink. There are many necessary differences in the designs of the downlink and the uplink, some of which include channel estimation implementations, time and frequency synchronization requirements, and resource allocation schemes. The uplink must deal with the multiuser detection problem of resolving multiple incoming transmitted signals at the base station while the downlink must broadcast signals containing data for all users. We will be focused on issues related to the uplink transmission and reception of information within the system. A simplified illustration of the scenario can be seen in Figure 2-1.

The system of interest has multiple users actively in transmission to the base station. These users must share the overall channel to the base station in such a way that their transmissions do not interfere with one another after propagation through the channel. Interference seen at the base station negatively affects the base station's
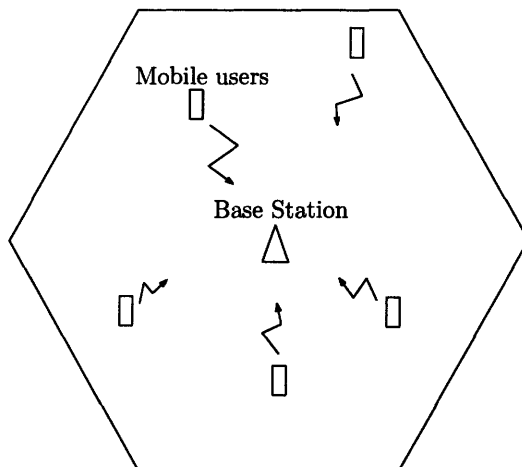
Figure 2-1: Illustration of mobile users transmitting to a fixed base station in the cellular uplink scenario.

ability to resolve the individual signals of the users. For reliable communication to be possible among the mobile users, some coordination between the users and base station is necessary. In other words, a multiple access scheme is needed within the system.

There are many different approaches to multiple access (e.g. Code Division Multiple Access (CDMA), Frequency Division Multiple Access (FDMA), Time Division Multiple Access (TDMA), Space Division Multiple Access (SDMA)) and additional variants to those approaches. Each multiple access scheme has its advantages and its disadvantages, with regards to complexity, performance, supportable system load, etc. The previously mentioned multiple access schemes differ in their approaches to sharing the available degree of freedom, the channel resources, of the underlying physical channel seen at the receiving base station. The multiple access scheme of interest in our system is a variant of frequency division multiple access referred to as OFDMA, or Orthogonal Frequency Division Multiple Access.

The multipath channel between the mobiles and base station has the effect of causing phase and amplitude altered copies and reflections of a signal to arrive at the receiver. These multipath effects can cause destructive interference to the observed signal and result in a creating a frequency selective fading channel for the transmitted

signals. Additionally the multiple path reflections cause the transmitted symbol observations to spillover into neighboring symbol periods at the receiver. This effect is known as intersymbol interference (ISI) and contributes to the frequency dispersiveness of the channel. OFDMA is an attractive approach for cellular communication because of built-in mechanisms to its signaling structure that minimize or eliminate these adverse channel affects [1].

In OFDMA, the frequency bandwidth is partitioned into narrowly spaced, effectively flat-fading subcarriers (i.e. frequency tones) and data is modulated as complex constellation symbols in the frequency domain onto these subcarriers. Each subcarrier is orthogonal to the other subcarriers and in a standard OFDMA system setup, individual users are orthogonalized by being assigned to communicate on disjoint sets of these subcarriers. Since the assignments are non-overlapping, there is no multiple access interference (MAI) between the intra-cellular users. We will be considering the overloaded system scenario in which the requirement of complete inter-user orthogonality is relaxed and where multiple users can be assigned to each frequency subcarrier resource. A given user's channel assignment will, at the discretion of the base station allocating channel assignments, directly overlap with possibly multiple other users' assignments, causing mutual interference between the users on that channel. This introduces the problem of multiuser detection at the receiver, which will now need to process the signals to mitigate the MAI. The remainder of this chapter will focus on the system aspects related to the transmission and reception of data under the OFDMA framework.

2.2 Transmitter Back-end

In this section, we will focus on the transmitter side for one user within the system. Shown in Figure 2-2 is a schematic diagram of a typical transmitter back-end that will be used as reference for this section. Information for transmission for a given user is source coded into binary sequences of 0's and 1's. The information bit stream is segmented into lengths of N_b bits, i.e. $\mathbf{b} = [b[0], b[1], \ldots, b[N_b - 1]]^T$, which is the
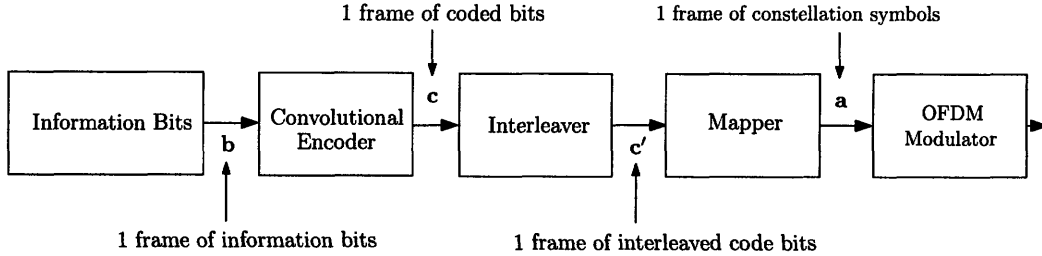
Figure 2-2: Back-end of transmitter chain.

quantity of information in one information word for transmission. We will focus on the synchronous setup where the users in the system send their information on a frame-by-frame basis and where the base station processes the received information in these same frame increments.

The information bits in the frame are protected with an Error-Correcting Code (ECC) before transmission over the wireless channel. We will focus on the use of convolutional codes but the system can be easily adapted for other coding schemes. Convolutional codes can be specified by their constraint lengths and their generator polynomials. We will denote the rate of the convolutional encoder as R, $(R < 1)$ such that the bits $\mathbf{b}$ pass through the encoder block to produce a "codeword" of coded bits $\mathbf{c} = [c[0], c[1], \ldots, c[N_b/R - 1]]^T$. The sequence of code bits produced are then passed to an interleaver block, denoted as $\prod$. Errors induced by the channel are often "bursty" and hence interleaving is performed to shuffle the code bits within the frame to spread out the locations of those bursty errors on the actual, non-interleaved coded sequence. The interleaving pattern is pseudo-random, reversible, and known at both the transmitting and receiving ends.

The sequence of interleaved code bits, denoted $\mathbf{c}'$ is passed to a mapper block to be mapped to modulation symbols for transmission over the channel. In this system, the modulation symbols are chosen from sets of M-QAM constellations, where M is a power of 2 (e.g. 4-QAM, 16-QAM, etc.). Each symbol to be transmitted is determined via some mapping function $\phi(\cdot)$ applied to a $\log_2 M$ bit segment of the incoming code sequence to produce the constellation symbol vector $\mathbf{a} = [a[0], a[1], \ldots, a[N_s - 1]]^T$ where $N_s = (N_b/R)/\log_2 M$. The choice of the constellation set (i.e. value of M)

to use is adaptive (with the aim of maximizing the data rate) and depends on the channel quality information inferred by the transmitter. In practice, this can be accomplished with auxiliary control channels between the receiver and transmitter that monitor channel quality.

The vector of complex constellation symbols **a** from the mapper block are modulated onto individual subcarriers (i.e. tones) in the frequency domain. To understand how this is accomplished, it is helpful to first understand the signal processing aspects behind OFDM (Orthogonal Frequency Division Multiplexing) modulation.

2.3   OFDM Transmission and Reception

In a cellular system, each user transmits a signal occupying a certain band of frequency. Because of FCC regulations, cellular transmissions consist of baseband signals modulated up to occupy some spectral mask at pre-determined carrier frequencies. Assuming proper modulation to the carrier frequency at the transmitter and corresponding demodulation from that carrier frequency at the receiver, the data carrying signals of interest can be viewed and processed in their equivalent baseband representations. In this system, the users' uplink transmissions lie within a baseband frequency range of width W. Thus, it is this frequency bandwidth W that must be shared among the users via a multiple access scheme.

As the name suggests, OFDMA uses Orthogonal Frequency Division Multiplexing (OFDM) methods to effectively partition the transmission bandwidth into many narrow parallel subchannels [5, 8, 17]. If we consider a sampled N-point representation $x[f]$ of the bandwidth W, we know that it can be equivalently represented as an N-point time domain sequence $\bar{x}[n]$ through the application of the N-point IDFT. The square brackets [·] are used to index the discrete sequences $x[f]$ and $\bar{x}[n]$ at frequency sample index f and time sample index n, respectively.

In OFDM, each point of $x[f]$ represents a frequency slice, or subcarrier, within the frequency band of width W/N. The signal of bandwidth W has N assigned degrees of freedom in the frequency domain on which it can modulate data; specifically, the

Multipath Channel

a → Buffer | S/P → IFFT + CP → **x** → $\bar{\mathbf{x}}^{cp}$ → P/S → $\bar{x}(t)$ → $\bar{p}(t)$ → carrier freq. mod. → $\bar{h}(t)$ → $\oplus$ → $\bar{y}(t)$

$n(t)$

$\bar{y}^{cp}$ **y**

$\bar{y}(t)$ → carrier freq. demod. → $\bar{p}(-t)$ → nT_p → S/P → FFT - CP
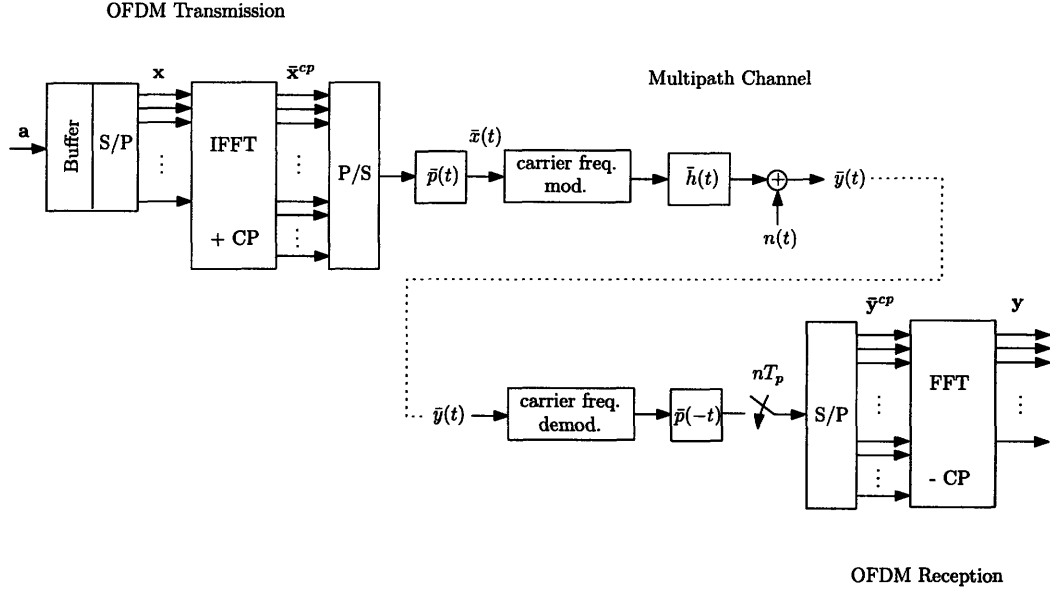
OFDM Reception

Figure 2-3: Schematic diagram of OFDM transmission and reception for one user through a multipath channel.

sample values of $x[f]$. The N-point sequence in frequency can be represented in vector form as $\mathbf{x} = [x[0], \ldots, x[N-1]]^T$ and corresponds to one "OFDM symbol." The equivalent time sequence vector (after the N-point IDFT) is $\bar{\mathbf{x}} = [\bar{x}[0], \ldots, \bar{x}[N-1]]^T$ and thus one OFDM symbol of frequency data takes at least N time samples to transmit. The time duration of OFDM one symbol period is referred to as a "time slot." In a synchronous system, the users's OFDM symbol transmissions arrive at the base station receiver in regular time slot intervals. The details on the synchronization requirements will be explained in Chapter 6.

For our system, we will assume that the N degrees of freedom in frequency, or subcarriers, are the channel resources to be shared among the users in the system. We will first start by looking at transmission from the perspective of a single user k transmitting over the entire available frequency W of the channel. As a reference, the OFDM transmitter and receiver to be described is shown in Figure 2-3.

The OFDM symbol to be transmitted by the user at a given system time slot index m consists of complex constellation symbols from $\mathbf{a}_k$ modulated on the subcarriers. In other words, $x_{k,m}[f] \in \Omega_k$ for $0 \leq f < N$ where Ω_k is the constellation being used

by user k.

An N-point IDFT is taken of the symbol sequence $x_{k,m}[f]$ to get the equivalent time-domain sequence:

$$\bar{x}_{k,m}[n] = \frac{1}{N} \sum_{f=0}^{N-1} x_{k,m}[f] \exp\left(\frac{j2\pi nf}{N}\right), \quad \text{for } 0 \leq n < N. \tag{2.1}$$

In practice, the DFT and IDFT are implemented with computationally efficient FFT and IFFT algorithms and consequently, the values of N are normally constrained to powers of 2. To account for the ISI channel, the signal $\bar{x}_{k,m}[n]$ is augmented with a cyclic prefix of length L, which periodizes the signal, as it is a cyclic extension of the original sequence. The cyclic prefix is the linchpin behind the OFDM modulation concept that facilitates the signal processing techniques used by the receiver to eliminate the ISI and ICI effects caused by the channel. The length L is chosen to be at least as long as the maximum delay spread of the channel and long enough to allow for adequate flexibility in user synchronization [18]. The sufficient value for L is often dependent on environment characteristics of operation, which can give us indications on the range of channel delay spreads during operation. The delay spread is often defined as the time interval between the main multipath component (e.g. the line-of-sight component) and the latest arriving significant multipath component.

The cyclic prefix is, however, a costly system overhead as it contains redundant information, and is therefore limited in length. In an idealized system, the overall data rate of the OFDM link can be represented as

$$\text{data rate} = \log_2 M \cdot \frac{N}{N+L} \text{ bits/time sample duration} \tag{2.2}$$

and it is clear that larger values of L in the denominator will lower the effective data rate.

The resulting (cyclically extended) $N + L$ length time domain sequence for trans-

mission $\bar{x}_{k,m}^{cp}[n]$ is given by

$$\bar{x}_{k,m}^{cp}[n] = \begin{cases} \bar{x}_{k,m}[N-L+n] & \text{for } 0 \leq n < L \\ \bar{x}_{k,m}[n-L] & \text{for } L \leq N < N+L. \end{cases} \tag{2.3}$$

The sequence $\bar{x}_{k,m}^{cp}[n]$ represents the information transmitted for one OFDM symbol. The (baseband) continuous time signal to be transmitted over the channel is constructed by interpolating $\bar{x}_{k,m}^{cp}[n]$ with an appropriately shaped (square root Nyquist) pulse $\bar{p}(t)$ of duration T_p. The time domain signal transmitted, over multiple OFDM symbol periods (e.g. $m = 1,\ldots,\infty$), is

$$\bar{x}_k(t) = \sum_{m=1}^{\infty} \sum_{s=0}^{N+L-1} \bar{x}_{k,m}^{cp}[m]\bar{p}(t - msT_p). \tag{2.4}$$

The signal $x_k(t)$ is up-converted to the carrier frequency and propagates through the ISI channel of impulse response $h_k(t)$. The exact energy profile of $h_k(t)$ will depend again on the environment and reflectors within the channel. The front end of the receiver will down-convert the total received signal back down to baseband and matched filter it with the known shaping pulse $\bar{p}(t)$ at sample intervals T_p to get discrete time observations. The pulse $\bar{p}(t)$ was chosen to be square root Nyquist so that the additive white Gaussian noise process contribution to the signal remains white after the receiver matched filtering and sampling. This simplifies the receiver processing and eliminates the need for additional whitening filters. We will denote the total observed sequence for one OFDM symbol at slot m as $\bar{y}_m^{cp}[n]$ for $0 \leq n < N+L-1$. The relation between the received sequence and the transmitted sequence $\bar{x}_{k,m}^{cp}[n]$ can be represented as the following familiar convolution relation [19]:

$$\bar{y}^{cp}[n] = \sum_{m=1}^{\infty} \sum_{l=0}^{L-1} \bar{h}_{k,m}[l]\,\bar{x}_{k,m}^{cp}[l-n] + \bar{n}_m[n], \tag{2.5}$$

where $\bar{h}_{k,m}[n]$ is the equivalent discrete-time sampled representation of the channel $\bar{h}_k(t)$ and $\bar{n}_m[n]$ is the sampled white Gaussian noise contribution during the slot m.
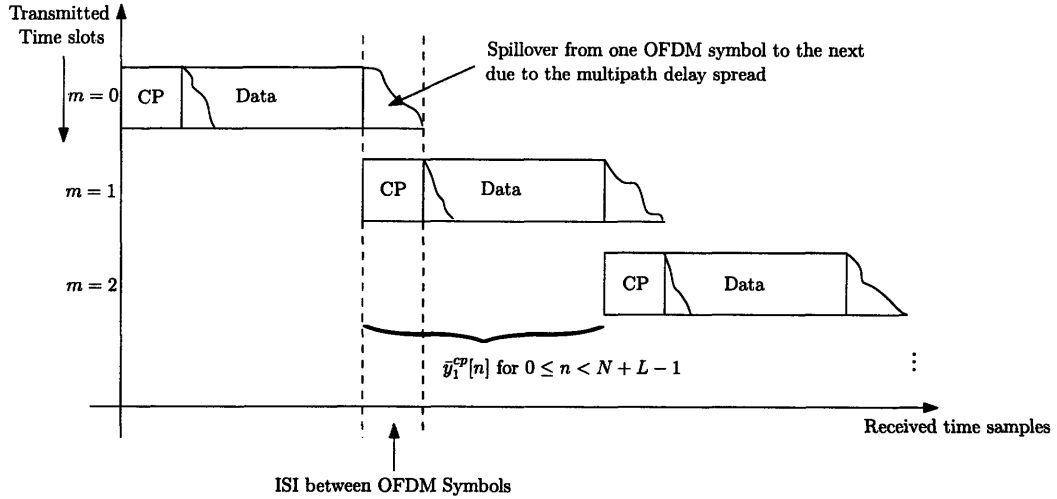
Figure 2-4: Example of the effects of ISI seen at the receiver between consecutive OFDM symbols. ISI contribution from previous symbol at time slot $m = 1$ is contained within the interval of the cyclic prefix of the OFDM symbol at $m = 1$.

The total number of non-negligible taps D (the length of the delay spread) in the sampled channel response must be less than the length of the cyclic prefix L that was determined earlier to avoid ISI problems. The equivalent representation of the channel in vector notation is:

$$\bar{\mathbf{h}}_{k,m} = [\underbrace{\bar{h}_{k,m}[0], \ldots, \bar{h}_{k,m}[D-1]}_{D<L}, \underbrace{0, \ldots, 0}_{N-D}]^{T}. \tag{2.6}$$

An important but reasonable assumption is that the channel impulse responses $\bar{h}_{k,m}[n]$ can be treated as time-invariant for the duration of one OFDM symbol period (or equivalently, for the $N + L$ time sample window). Additionally, we should assume the users are synchronized to the base station and transmitting such that the start of each OFDM symbol of different users arrive within the cyclic prefix duration of the base stations time reference.

An illustration depicting the effects of ISI seen at the receiver between OFDM symbols transmitted in consecutive slots can be seen in Figure 2-4. We will denote the observation vector for an OFDM symbol period (i.e. a single time slot m) as $\bar{\mathbf{y}}_{m}^{cp} = [\bar{y}_{m}^{cp}[0], \ldots, \bar{y}_{m}^{cp}[L-1], \bar{y}_{m}^{cp}[L], \ldots, \bar{y}_{m}^{cp}[L-1+N]]^{T}$. At the receiver, the portion

of the sequence corresponding to the cyclic prefix transmission is removed before further processing. The first L samples of the received signal contain ISI components from the previous OFDM symbol period spilt over into the cyclic prefix of the current OFDM symbol period. The resulting N-point observation is given by

$$\bar{y}_m[n] = \begin{cases} \bar{y}_m^{cp}[n+L] & \text{for } 0 \le n < N \\ 0 & \text{for } n \ge N. \end{cases} \tag{2.7}$$

The observation written in vector notation is $\bar{\mathbf{y}}_m = [\bar{y}_m[0], \dots, \bar{y}_m[N-1]]^T$.

Our intention is to convert our observations to the frequency domain using the DFT operation. Because of how we added and removed the cyclic prefix, the linear ISI channel can now be viewed as a circulant channel, which allows us to take advantage of the known properties between the DFT and circular convolution. The sum of convolutions over the time slots $m = 1, \dots, \infty$ in (2.5) can be separated so that we can consider each convolution term individually without ISI. The linear convolution becomes a circular convolution as a result of the circulant channel [19, 20] and the N-point time domain observation can be written concisely in vector notation as

$$\bar{\mathbf{y}}_m = \bar{\mathbf{h}}_{k,m} \otimes \bar{\mathbf{x}}_{k,m} + \bar{\mathbf{n}}_m, \tag{2.8}$$

where $\otimes$ represents the circular convolution of the two N-element vectors and $\bar{\mathbf{n}}_m$ is the N-element AWGN contribution for the slot.

The final observable for the receiver, $y_m[f]$ is obtained by taking the N-point DFT of the time domain sequence $\bar{y}_m[n]$:

$$y_m[f] = \sum_{n=0}^{N-1} \bar{y}_m[n] \exp\left(\frac{-j2\pi f n}{N}\right). \tag{2.9}$$

Circular convolution in time is equivalent to scalar, element-wise multiplication in frequency after application of the DFT. This gives us the equivalent expression in the

frequency domain between the transmitted and received sequences:

$$y_m[f] = h_{k,m}[f]x_{k,m}[f] + n_m[f] \tag{2.10}$$

where $h_{k,m}[f]$ are elements of $\mathbf{h}_{k,m}$, the N-point DFT of the channel $\bar{\mathbf{h}}_{k,m}$, and $x_{k,m}[f]$ is the original symbol sequence of the user k during time slot m. We see that we can now treat the frequency-selective channel as a set of decoupled narrowband frequency flat-fading channel gains on the frequency subcarriers. The observation on a given subcarrier can be considered free of inter-carrier interference (ICI) from the transmissions on neighboring subcarriers. Additionally, high-complexity time-domain and frequency-domain equalization procedures found in other narrowband time-domain based systems can now be avoided, as the channel can be considered free of ISI and ICI.

We have so far looked at the transmission of an OFDM symbol for a single user k. If there were N_u total users in the system all synchronously transmitting over the entire bandwidth, the base station receiver's observation would be a superposition of the contributions from all the N_u users. Because the DFT and IDFT are linear operations, the total observation at the receiver can be written as:

$$y_m[f] = \sum_{k=1}^{N_u} h_{k,m}[f]x_{k,m}[f] + n_m[f]. \tag{2.11}$$

In a standard OFDMA system, users are assigned disjoint subsets of orthogonal subcarriers in the interests of minimizing multiple access interference. Thus, a user will only modulate constellation symbols on the subcarriers to which it is assigned and assign 0 values on the remaining subcarriers. The observation at the receiver for a given subcarrier f and time slot m will therefore have the signal component of only one user, as the remaining terms in the summation in (2.11) have zero contribution. This however is no longer true in the overloaded OFDMA scenario. More than one non-zero signal contribution on a subcarrier means that there is now MAI within the system. The transmissions of individual users are no longer orthogonal and this

creates the need for multiuser detection methods and processing.

For the purposes of detection and decoding, we will assume that all observations have been converted to their frequency domain form, unless otherwise noted. Because we have transformed our signals to the frequency domain through OFDM, the received data symbols can be considered to be free of ISI (between OFDM symbols) and ICI (between subcarriers within an OFDM symbol).

2.4 OFDMA Resource Allocation

In the OFDMA uplink scenario, each user will receive a channel assignment from the base station that consists of a subset of subcarriers that make up an OFDM symbol in the available frequency band. In our system, for reasons of implementation feasibility and for channel estimation purposes, each user's subset corresponds to a contiguous block of subcarriers in frequency. Additionally, the user will be assigned this block of subcarriers for a designated number of system time slots (i.e. OFDM symbol periods). This *time* × *frequency* block of subcarriers "tiles" is referred to as an data *frame,* and in our system model it will be the basic channel resource that is being shared non-orthogonally among overloaded users in our overloaded system. Figure 2-5 shows an example of a single channel assignment over multiple OFDM symbols. Within this block, the majority of the tiles (subcarriers) are assigned to modulate the aforementioned data-carrying constellation symbols of a and a small subset of the subcarriers are generally assigned to modulate known pilot symbols to aid channel estimation at the receiver.

Each data frame corresponds to a physical channel resource. Standard OFDMA systems partition the entire *time* × *frequency* space into these smaller blocks of the overall grid and orthogonalize users by assigning each user to a unique block. Overloaded OFDMA systems will load each resource block with multiple users. The base station will determine the assignments based on channel conditions and available resources. The overloading factor per block Q can vary from frame assignment period to frame assignment period. Figure 2-6 compares the assignment schemes of the
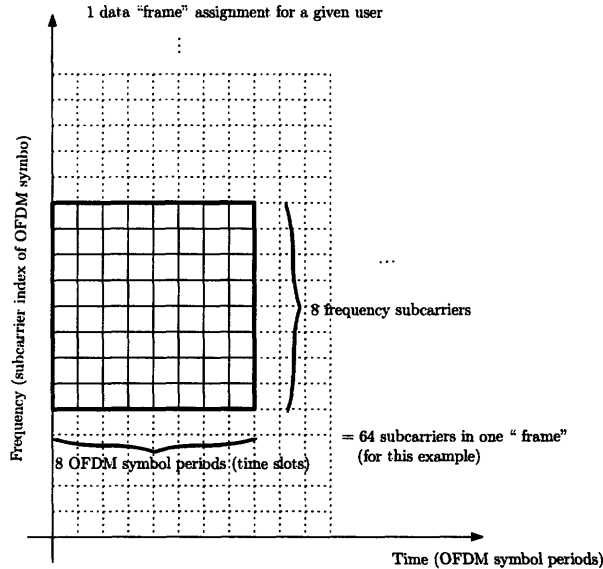
Figure 2-5: An example of a basic channel resource unit, a data frame, within the signal space for one frame assignment period. Each column of the total grid corresponds to an N-point OFDM symbol.

standard and overloaded OFDMA cases. Generally, for the purposes of frequency diversity and interference diversity, channel resource assignments of each user are "hopped" or shuffled independently between frames assignment periods [1]. We are currently focused on the transmission and reception model of a single OFDM frame in the overloaded scenario and thus this hopping procedure does not need to be explicitly considered.

2.5 Multiple Antenna Receiver Front-end

Orthogonal multiple access in standard OFDMA puts a hard limit on overall system capacity, constraining it to the amount of available frequency bandwidth and number of discrete subcarriers. A key element of the overloaded OFDMA system is the multiple receiver antennas available at the base station. It is known that link capacity increases proportionally with the number of antennas at the receiver, as the spatial diversity increases the average received SNR/SINR and additional spatial degrees-of-freedom are created [2, 3]. This gain in capacity is linear for interference

Channel Resource Assignments
for Non-Overloaded OFDMA
(3 users)

Channel Resource Assignments
for Overloaded OFDMA
(6 users)

User 1	User 3
User 2	User 1
User 3	User 2

Frequency → Time

Users 1, 4	Users 3, 5
Users 2, 5	Users 1, 6
Users 3, 6	User 2, 4
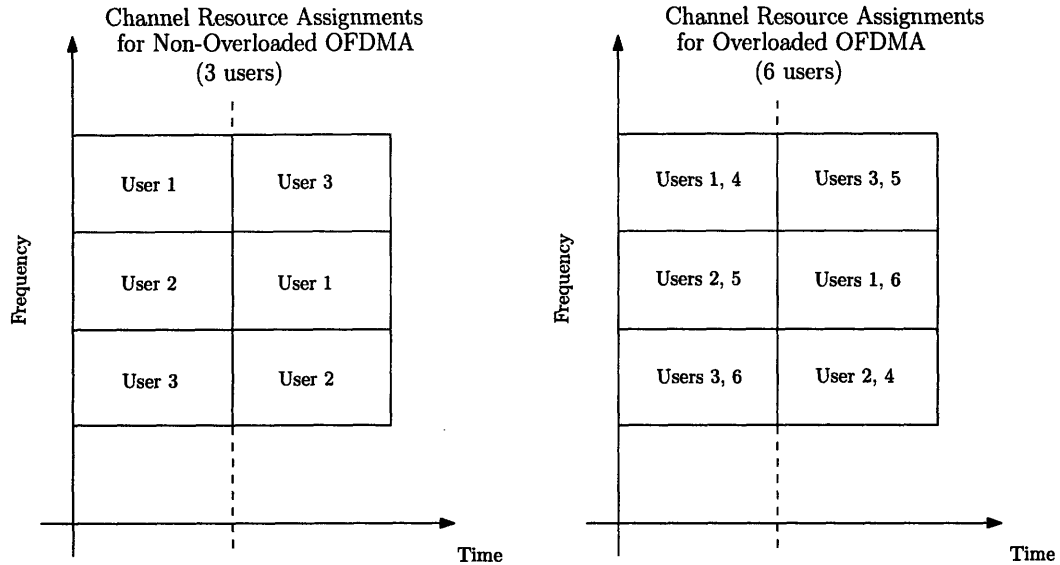
Frequency → Time

Figure 2-6: Contrast of the channel resource assignment schemes in standard and overloaded OFDMA systems. Note that in both cases, channel resource assignments are "hopped" between assignment periods for frequency and interference diversity.

limited systems such as CDMA, which operate in the low SNR/SINR regions. However, systems based on OFDMA normally operate in the higher SNR/SINR regions where the capacity gains are only logarithmic in SNR/SINR. In this region, system capacity is bandwidth (i.e. degree-of-freedom) limited. Hence the immediate gains from multiple antennas are not as significant as the CDMA case. The addition of multiple antennas at the receiver does, however, increase the dimensions of the received signal space, effectively allowing us to multiplex multiple users on these newly induced spatial degrees of freedom. In this case, strictly orthogonal channel assignments (i.e. one user per channel resource) may be an under-utilization of the overall channel and hence sub-optimal [21]. Similar to V-BLAST and other BLAST methods for MIMO channels [6], we can leverage this spatial multiplexing gain at the receiver through overloading the subcarrier tiles. The analogy to the separate transmit layers in V-BLAST are the independently transmitting single antenna users.

The multiple receiver antennas allow us to take advantage of the spatial diversity amongst individual, geographically separated users to aid multiuser detection. In this system, we will relax the constraint of orthogonal resource allocation and instead allow

25

up to Q users to transmit on each OFDM frame (i.e. block assignment of subcarriers), where Q is a subset of the total N_u users in the system. The receiver will use multiple antenna array processing (i.e. receiver beamforming) to separate and resolve the interfering users based on channel knowledge of their estimated spatial signatures [22, 17, 23]. Assuming the correlation between antennas is not excessively high, there are N_r resolvable degrees of freedom in the channel, where N_r is the number of receive antennas [2]. The general rule of thumb is that, with proper multiuser detection techniques, it is possible for Q users to communicate reliably on a given resource when $Q \leq N_r$. In one of the detection approaches we will present in Chapter 5, the receiver will use optimal combining methods based on its $N_r \times Q$ channel estimates to perform multiuser detection. This approach is similar to combined uplink OFDMA/SDMA implementations [5, 16].

Again, the parallels to BLAST-related MIMO transmission schemes are apparent. The constraint for reliable V-BLAST communication is generally that $N_t \leq N_r$ where N_t is the number of effective transmit antennas, which is analogous to the spatial multiplexing and overloading factor constraint in our system design. Successive Interference Cancellation (SIC) algorithms have been studied and developed for V-BLAST systems (a summary of interference cancellation frameworks can be found in [24]) and similarly, we will also be exploring interference cancellation schemes at the receiver to mitigate cochannel interference of the overloaded users during detection and decoding. Our specific focus will be on iterative soft interference cancellation of inter-user interference, which has previously been explored for different contexts in [25, 26, 27].

We will be considering the multiuser detection and decoding case of Q interfering mobile users transmitting synchronously on an OFDM frame to a single receiving base station. We will also assume the following design parameters:

- The Q users will be assigned to the same time-frequency block (i.e. OFDM frame) of subcarriers and hence mutually interfere with each other. As mentioned before, detection and decoding of the users will be done on a frame-by-frame basis.

26

- The Q users will be assumed to be synchronized, to within an allowable fraction of the total cyclic prefix, relative to the time reference of the base station. The transmitted OFDM symbols by each user will arrive within intervals less than the cyclic prefix to the base station's reference time. Channel estimation and the choice of cyclic prefix length allow us some flexibility in this constraint [18].

- Each user will be transmitting with one antenna without any space-time coding. However, each user's information bits will be convolutionally encoded, bit-interleaved, and mapped across the tiles in the frame. Because of the pseudo-random bit-interleaving, the symbols in one data frame can be considered approximately uncorrelated.

Chapter 3

Received Signal Model

3.1 Total Received Signal

In our multiple antenna receiver scenario, a single observation (in the frequency domain) on subcarrier f, receiver antenna n_r and time slot m can be denoted as $y_m^{(n_r)}[f]$. Similarly, the transmitted constellation symbol by a user k on a subcarrier f and time slot m can be denoted as $x_{k,m}[f]$. The *total* received space for the OFDM signal model consists of the N subcarriers (over bandwidth W), for some duration of time slots M (i.e. OFDM symbol periods), over the N_r total receiver antennas. We can stack the observations for the entire space to produce a total observation vector $\mathbf{y}$. We will similarly stack the transmitted symbol vector $\mathbf{x}_k$ for each user k over all frequency and time (we are reminded that unassigned subcarriers are modulated with zero values) such that the following expression for the total observation vector holds:

$$\mathbf{y} = \sum_{k=1}^{N_u} \mathbf{H}_k \mathbf{x}_k + \mathbf{n} \tag{3.1}$$

where $\mathbf{H}_k$ is some block matrix of channel coefficients for user k's channel to the N_r antennas, stacked in appropriate fashion over the N subcarriers and M time slots. The vector $\mathbf{n}$ is the AWGN contributions stacked in the same manner as $\mathbf{y}$.

The observation vector is $N \times M \times N_r$ in size and quite large and we would rather perform detection and decoding over independent, smaller sized observations.

28

However, in order to do so, we must be sure that observations for each subcarrier tile are truly decoupled from other subcarrier tiles within the total received signal space. Because of our OFDM transmission and reception model, our signal space has been partitioned into a grid of discrete frequency subcarriers and time slots. We are hoping to perform detection and decoding for the users transmitting on a given tile by only taking into account the received observations for that tile. Once we can be sure that the subcarrier observations over the received signal space (and the OFDM frame of interest) have been decoupled from one another, the processes of detection and decoding can be done independently over the subcarriers and be parallelized for implementation efficiency.

Multiuser symbol detection is performed based on the a posteriori distribution information of the observations [7]. It is proved in the following section that a posteriori distributions on each subcarrier tile are functions of only the users' transmitted symbols on that tile resource. We show that the observations $y_m[f]$ and $y_{m'}[f']$ for a given $x_{k,m}[f]$ are uncorrelated for $m \neq m', f \neq f', \forall m, m', f, f'$ in the received signal space and conclude that subcarrier tiles within a given data frame can be approximately processed independently.

3.2 Justification for Decoupled Detection

Each element of the stacked vector observation **y** in (3.1) can be written as

$$y_m^{(n_r)}[f] = \sum_{k=1}^{N_u} h_{k,m}^{(n_r)}[f]\, x_{k,m}[f] + n_m^{(n_r)}[f] \tag{3.2}$$

where $h_{k,m}^{(n_r)}[f]$ is the channel from the k-th user to the n_r-th antenna, and $n_m^{(n_r)}[f]$ is the additive Gaussian noise seen at the n_r-th antenna. Because of the no ISI/ICI conditions of OFDM, the transmissions on a tile resource only affect that tile at the receiver after propagation through the channel. Furthermore, because of square root Nyquist pulse shaping at the transmitter and matched filtering (with that pulse at pulse period intervals) at the front end of the receiver, we can assume that there is no

smearing of transmitted signals contributions to other tiles [28]. The autocorrelation of the discrete sampled additive Gaussian noise process can be concluded to be

$$\mathbb{E}\{n_m^{(n_r)}[f]\,(n_{m'}^{(n_r')}[f'])^*\} = \sigma^2 \delta_{m-m',n_r-n_r',f-f'} \quad \forall\, m, n_r, f, m', n_r', f' \tag{3.3}$$

because of the Nyquist pulse filtering. Therefore the AWGN contribution remains white and uncorrelated across tiles after front-end receiver processing.

The OFDM modulation and Nyquist pulse filtering have decoupled the effects of subcarriers on adjacent subcarriers (i.e. no ISI/ICI, with AWGN noise) such that the conditional distribution of the entire signal space observation can be factored as follows:

$$p(\mathbf{y}|\mathbf{x}) = \prod_{\forall f, m} p(\mathbf{y}_m[f] | \mathbf{x}_m[f]) \tag{3.4}$$

where the vector $\mathbf{y}_m[f]$ is the observations stacked over the N_r antennas and $\mathbf{x}_m[f]$ is the transmissions stacked over the Q users, on subcarrier f at time slot m.

We are interested in the a posteriori probabilities for each transmitted symbol $x_{k,m}[f]$ conditioned on the observation $\mathbf{y}$. Using Bayes' rule [29], we can write the a posteriori probability as the following:

$$p(x_{k,m}[f]|\mathbf{y}) \propto p(\mathbf{y}|x_{k,m}[f])\,p(x_{k,m}[f]) \tag{3.5}$$

$$\propto \left(\sum_{\forall \grave{\mathbf{x}}} p(\mathbf{y}, \grave{\mathbf{x}}|x_{k,m}[f])\right) p(x_{k,m}[f]) \tag{3.6}$$

where the expression in (3.6) is the expanded marginalization summation that is equivalent to $p(\mathbf{y}|x_{k,m}[f])$. This summation is over all possible instances of what we define as $\grave{\mathbf{x}}$, the stacked vector of all transmitted symbols over all tiles except $x_{k,m}[f]$ (of which we are calculating the a posteriori probability). We can then rewrite

$p\left(\mathbf{y}, \grave{\mathbf{x}} \mid x_{k,m}[f]\right)$, using the rules of joint distributions as

$$p\left(\mathbf{y}, \grave{\mathbf{x}} \mid x_{k,m}[f]\right) \;=\; p\left(\mathbf{y} \mid \grave{\mathbf{x}}, x_{k,m}[f]\right) p\left(\grave{\mathbf{x}} \mid x_{k,m}[f]\right) \tag{3.7}$$

$$\approx \; p\left(\mathbf{y} \mid \mathbf{x}\right) p\left(\grave{\mathbf{x}}\right) \tag{3.8}$$

$$\approx \; \left(\prod_{\forall f',m'} p\left(\mathbf{y}_{m'}[f'] \mid \mathbf{x}_{m'}[f']\right) \right) p\left(\grave{\mathbf{x}}\right) \tag{3.9}$$

where in (3.8), we stacked $\grave{\mathbf{x}}$ and $x_{k,m}[f]$ to produce $\mathbf{x}$, the vector of *all* symbols. Additionally, we make the simplification that the conditional probability $p\left(\grave{\mathbf{x}} \mid x_{k,m}[f]\right) \approx p\left(\grave{\mathbf{x}}\right)$ in (3.8). This approximation is possible in bit-interleaved coded modulation systems such as the one under consideration. In this coded system, the coded bits out of the channel coder of a given user are temporally correlated because of the internal state of the coder. However, after pseudo-random interleaving and mapping to constellation symbols, the resulting symbol values within a user's frame can be approximately considered as uncorrelated and independent from one another [11], which allows us to make the previous simplification. We replace the conditional distribution $p\left(\mathbf{y} \mid \mathbf{x}\right)$ in (3.8) with the equivalent expression from the identity in (3.4). The indexing variables f' and m' of the product in (3.9) are over all subcarriers and all time slots.

After substituting (3.9) back into (3.6), we get the following (approximate) expression for the a posteriori probability:

$$p\left(x_{k,m}[f] \mid \mathbf{y}\right) \propto \sum_{\forall \grave{\mathbf{x}}} \left(\prod_{\forall f',m'} p\left(\mathbf{y}_{m'}[f'] \mid \mathbf{x}_{m'}[f']\right) p\left(\grave{\mathbf{x}}\right) \right) p\left(x_{k,m}[f]\right). \tag{3.10}$$

The a posteriori probability is a function of $x_{k,m}[f]$ and hence terms in (3.10) independent of $x_{k,m}[f]$ remain constant over the summation and can be factored out to be absorbed into the proportionality relation. Additionally, the terms in summation over elements not related to not related to m and f can be also be factored out. The

a posteriori probability can be simplified to:

$$p\left(x_{k,m}[f] \mid \mathbf{y}\right) \;\propto\; \sum_{\forall \grave{\mathbf{x}}_m[f]} p\left(\mathbf{y}_m[f] \mid \mathbf{x}_m[f]\right) p\left(\grave{\mathbf{x}}_m[f]\right) p\left(x_{k,m}[f]\right) \tag{3.11}$$

$$\propto\; \sum_{\forall \grave{\mathbf{x}}_m[f]} p\left(\mathbf{y}_m[f] \mid \grave{\mathbf{x}}_m[f], x_{k,m}[f]\right) \prod_{\substack{i \in Q, \\ i \neq k}} p\left(x_{i,m}[f]\right) p\left(x_{k,m}[f]\right) \tag{3.12}$$

where the summation is now over all possible instances of the stacked $(Q-1) \times 1$ vector $\grave{\mathbf{x}}_m[f]$ of other users' symbols (excluding user k) at time slot m and frequency subcarrier f. Users' transmitted symbols can be assumed independent of one another and hence $p\left(\grave{\mathbf{x}}_m[f]\right)$ can be expanded as a product over the individual probabilities of the Q overlapping users on subcarrier f and time slot m.

Our final result shows us that, at the receiver after all front-end processing, we have converted the OFDM resource grid of subcarriers and time slots into a set of fully parallel channels that can be treated individually during multiuser detection. The a posteriori probabilities for the Q users' symbols on the tile at subcarrier f and time slot m is only a function of the observation vector $\mathbf{y}_m[f] = [y_m^{(1)}[f], \ldots, y_m^{(N_r)}[f]]^T$ for that tile. When we perform detection on a subcarrier in the grid, all the statistical properties can be determined from the antenna observations on only that subcarrier tile.

3.3 Simplified Received Signal Model

As was demonstrated in the previous section, we can unstack the total vector observation $\mathbf{y}$ and approximately process each subcarrier tile separately. Figure 3-1 shows an example scenario of the processing for one OFDM frame where $N_r = 2$ and $Q = 2$. For ease of notation, and without loss of generality, we will drop the indices m and f when considering detection a single subcarrier tile. Thus, the resulting simplified observation model for a subcarrier at a single time slot is

$$y^{(n_r)} = \sum_{k=1}^{Q} h_k^{(n_r)} x_k + n^{(n_r)} \tag{3.13}$$

32
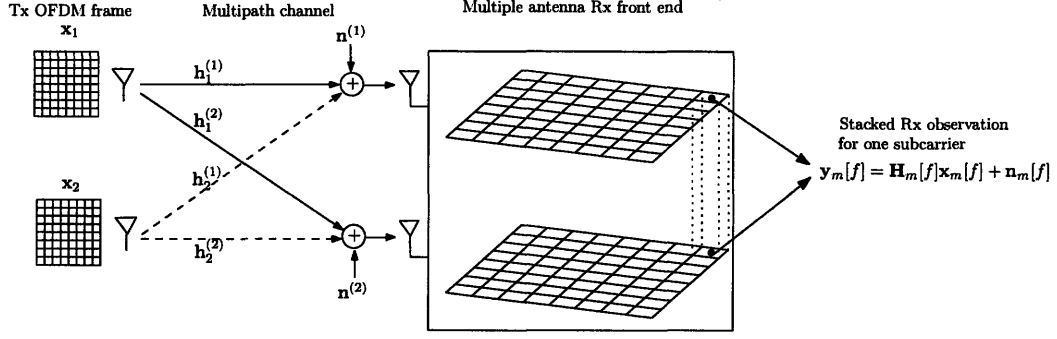
Figure 3-1: Transmission and reception of 1 OFDM frame through multiple antenna channel for $Q = 2$ and $N_r = 2$. Observations have been converted to the frequency domain.

where x_k is the transmitted signal from the k-th user, $h_k^{(n_r)}$ is the channel gain coefficient from the k-th user to the n_r-th antenna, and $n^{(n_r)}$ is the complex additive white Gaussian noise (AWGN) for this observation.

With N_r total receive antennas at the base station, the overall received signal observation can be stacked over the antennas and written in vector notation as $\mathbf{y} = \mathbf{Hx} + \mathbf{n}$ given by:

$$
\begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(N_r)} \end{bmatrix} = \begin{bmatrix} h_1^{(1)} & h_2^{(1)} & \cdots & h_Q^{(1)} \\ h_1^{(2)} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ h_1^{(N_r)} & \cdots & \cdots & h_Q^{(N_r)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_Q \end{bmatrix} + \begin{bmatrix} n^{(1)} \\ n^{(2)} \\ \vdots \\ n^{(N_r)} \end{bmatrix} \tag{3.14}
$$

where the $\mathbf{H}$ is a $N_r \times Q$ matrix of channel gain coefficients. The element in the i-th row and j-th column of $\mathbf{H}$ is $h_j^{(i)}$ and corresponds to the channel gain coefficient from the j-th user to the i-th antenna, where $1 \leq j \leq Q$ and $1 \leq i \leq N_r$. In this setup, $\mathbf{n} \in \mathbb{C}^{N_r}$ and we will assume $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})$ such that $\Re\{\mathbf{n}\} \sim \mathcal{N}(0, \frac{\sigma^2}{2}\mathbf{I})$ and $\Im\{\mathbf{n}\} \sim \mathcal{N}(0, \frac{\sigma^2}{2}\mathbf{I})$.

Chapter 4

Standard Detection and Decoding

The received signal observation passes through four stages at the receiver: multiuser detection, symbol de-mapping, code bit de-interleaving, and codeword decoding. The goal of our multiuser detection block is to determine $\mathbf{x} = [x_1, x_2, \ldots, x_{N_u}]^T$ from the observed $\mathbf{y} = [y^{(1)}, y^{(2)}, \ldots, y^{(N_r)}]^T$, as noted in (3.14). This information gets passed to a de-mapping block that uses the inverse mapping function $\phi^{-1}(\cdot)$ to convert the symbol information to code bit information and a de-interleaving block that restores the original ordering of the code bits. The decoding process will take the output from the de-mapper and determine an estimate $\hat{\mathbf{b}}$ of the transmitted information bits for that user. In the non-iterative receiver, the output of the decoding process is generally the terminal end of the communication link from the mobile user to the base station. It is clear that accurate detection and decoding are essential for a reliable communication link and they therefore account for a significant portion of the receiver back-end processing and complexity.

4.1 Optimum Bit-by-Bit Detection and Decoding

The optimal approach to multiuser detection and decoding, without regard to computational complexity, is to perform them jointly as one process and to maximize the a posteriori probability based on the observation $\mathbf{y}$. This approach satisfies the maximum a posteriori (MAP) criterion in that it minimizes the probability of error

on the information bits for each user [12].

If we wish to do MAP detection for user k's original sequence of information bits $\mathbf{b}_k$ of length N_b, the resulting expression that must be computed is:

$$\hat{b}_{k,i} = \arg \max_{b_{k,i} \in \{0,+1\}} p\left(b_{k,i} | \mathbf{y}\right) \tag{4.1}$$

where $\mathbf{b}_k$ is the transmitted information bit vector. For a given user, this is done for all bit positions i of the corresponding information frame at the time slot of interest. This approach amounts to computing the marginal probability of $b_{k,i}$ by summing the probabilities for each and all possible information codewords (i.e. all possible information sequences of length N_b) that contain $b_{k,i}$ against all possible interfering codewords of the other users. The probabilities are conditioned on the observation $\mathbf{y}$ and take into account the corresponding code constraints of each of the Q users. This is essentially a brute force comparison approach and is very computationally intensive with complexity is on the order $\mathcal{O}(2^{N_b \cdot Q})$.

4.2 Separate MAP Detection and Decoding

Conventional approaches to detection and decoding separate the two operations in order to avoid the overall computational complexity involved with joint processing. The optimal MAP symbol detection scheme is to choose the estimate $\hat{x}$ such that it minimizes the probability of error for that transmitted symbol. This is done by choosing the candidate constellation symbol a to maximize the (a posteriori) probability $p(\cdot)$ of the transmitted symbol x_k for the given observed received signal $\mathbf{y}$[28, 30], which can be expressed as the following:

$$\hat{x}_k = \arg \max_{a_k \in \Omega_k} p\left(x_{k,} = a_k | \mathbf{y}\right) \tag{4.2}$$

where $\hat{x}_k$ is the estimate of the transmitted symbol from user k, at a given subcarrier tile. The symbol that is chosen will be an element of that user's constellation set Ω_k. As we have justified in the previous chapter, we can perform multiuser detection
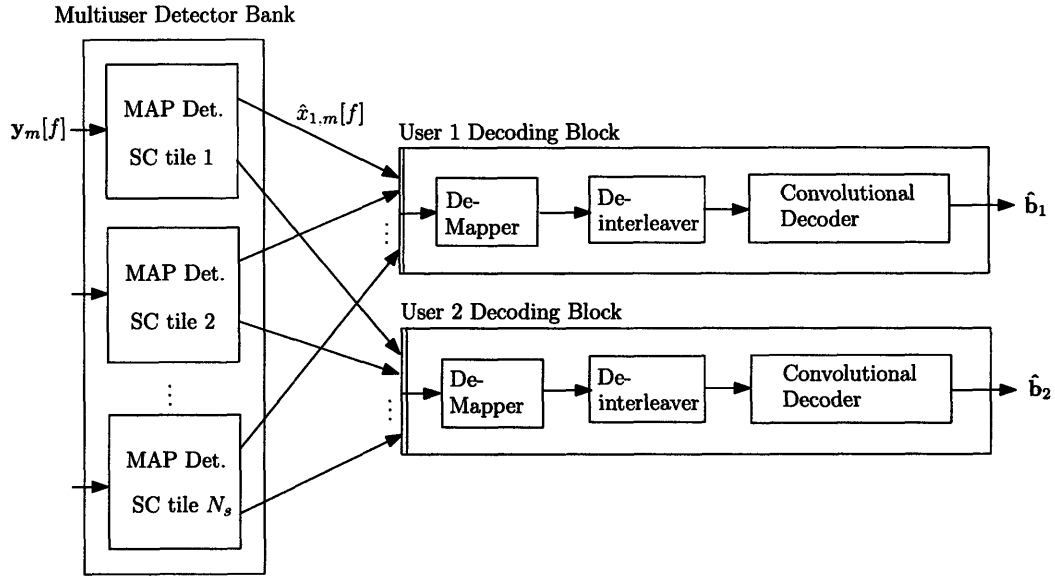
Figure 4-1: Receiver chain implemented at the base station. The order of processing steps in the decoding blocks is the reverse of the corresponding steps in the transmitter chain. This example shows the detection and decoding process for one data frame with N_s tiles and $Q = 2$ users.

on a subcarrier observation without regard to its absolute time or frequency indices. From an intuitive standpoint, this separate "optimal" symbol detection approach is sub-optimal to the joint detection and decoding approach in that detection is done individually across subcarriers without regard to the underlying code structure of the bits.

The MAP estimated symbols for each user on each tile are then de-mapped to form estimates of the transmitted code bits sequences, which are then de-interleaved and sent to the appropriate user's decoder to finally produce estimates on the information bit sequences of each user. Dynamic programming algorithms such as the Viterbi algorithm [31] minimize computational complexity of maximum likelihood sequence estimation and are often used in this non-iterative, separate detection and decoding framework . Figure 4-1 depicts the described receiver chain flow, starting from the signal observations into the parallel bank of multiuser detectors (for each tile) and ending with decoded information bit estimates output from the parallel bank of single user decoders.

The detector just described produces hard decisions on the symbols of the users. Hard decisions amount to thresholding calculated continuous values (such as on the a posteriori probabilities in (4.2)), which results in a loss of information. For better performance, the MAP symbol detector can alternatively be designed to produce soft output values, which in this case would be the a posteriori values for each of the possible constellation symbols for each user (i.e. $p(x_{k,} = a_k | \mathbf{y})$, $\forall a_k \in \Omega_k$ and $\forall k$). We can see that in order to produce a hard decision, these sets of values must be computed regardless. As an example, if user k were transmitting using a 4-QAM constellation where $\Omega_k = \{a_{(0)}, a_{(1)}, a_{(2)}, a_{(3)}\}$, the multiuser detector would calculate and output the set of probabilities $p(x_k = a_{(0)} | \mathbf{y})$, $p(x_k = a_{(1)} | \mathbf{y})$, $p(x_k = a_{(2)} | \mathbf{y})$, $p(x_k = a_{(3)} | \mathbf{y})$ for that user. This set would then get de-mapped and translated to a posteriori probabilities, or equivalently log-likelihood ratios, on each code bit of the code sequence. For this soft information framework, the decoder would be designed to accept soft valued likelihoods for the bits of the code sequence. One common implementation for efficient soft-input decoding of convolutional codes is the BCJR algorithm [32].

4.3   Linear MMSE Detection and Decoding

Oftentimes, the computational complexity of the required processing for hard and soft MAP symbol detection is also too high for practical implementation. The order of complexity in a multiuser optimal MAP symbol detector is typically exponential in the number of users and their constellation sizes. Similar to the optimal joint detector/decoder, the optimal symbol detection approach essentially resorts to comparing the sets of all possible combinations of transmitted symbols across all interfering users. As a result of this high computational complexity, linear (and non-linear) multiuser detection methods have been developed that trade off much of this complexity for acceptable decreases in performance [28, 30]. These linear methods have the attractive property that complexity that scales linearly with the number of interfering users Q.

One such approach to detection is linear Minimum Mean-Squared Estimation

(MMSE). Multiuser detection via MMSE is additionally well-suited for multiple antenna receiver systems with spatial diversity. The MMSE detector in the multiple user and multiple receiver antenna scenario uses optimum combining methods across the antennas based on the estimated spatial signatures of the interfering users. The spatial signatures to each of the antennas provide a basis for the received signal space and the filtering operation with the MMSE filter amounts to projecting the observation in the space to optimally suppress multiple access interference and maximally combine signal components along the given user's spatial signature [2, 3, 22]. Estimates can be produced based on these scalar filtered observations for each user. The base station receiver is essentially taking advantage of the spatial diversity provided by the multiple antenna channel.

To optimally combine the received observations across the multiple antennas, the receiver must adaptively (i.e. for each tile) calculate a filter $\mathbf{w}_k$ for each user k, satisfying the MMSE criterion. The filter is designed to minimize the mean-squared error of the filtered output and the actual transmitted symbol x_k:

$$\mathbf{w}_k = \arg\min_{\mathbf{w}_k \in \mathcal{C}^{N_r}} \mathbb{E}\left\{ |x_k - \underbrace{\mathbf{w}^H \mathbf{y}}_{z_k}|^2 \right\} \tag{4.3}$$

$$= \arg\min_{\mathbf{w}_k \in \mathcal{C}^{N_r}} \mathbb{E}\left\{ (x - \mathbf{w}^H \mathbf{y})^H (x - \mathbf{w}^H \mathbf{y}) \right\} \tag{4.4}$$

$$= \arg\min_{\mathbf{w}_k \in \mathcal{C}^{N_r}} \left[|x_k|^2 + \mathbf{w}^H \mathbb{E}\{\mathbf{y}\mathbf{y}^H\}\mathbf{w} - 2\mathcal{R}\{w^H \mathbb{E}\{x_k \mathbf{y}\}\} \right] \tag{4.5}$$

We define $\mathbf{w}_k$ to be of dimension N_r and z_k to be the complex MMSE filtered output for the kth user. The complex filter $\mathbf{w}_k$ that satisfies the above expression is of the form $\mathbf{w}_k = \mathbb{E}\{\mathbf{y}\mathbf{y}^H\}^{-1}\mathbb{E}\{x_k^* \mathbf{y}\}$. We can use the fact that $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$ to simply as

follows:

$$\mathbb{E}\{\mathbf{y}\mathbf{y}^H\} = \mathbb{E}\left\{(\mathbf{Hx} + \mathbf{n})(\mathbf{x}^H\mathbf{H}^H + \mathbf{n}^H)\right\} \tag{4.6}$$

$$= E_s\mathbf{HH}^H + \sigma^2\mathbf{I} \tag{4.7}$$

$$\mathbb{E}\{x_k^*\mathbf{y}\} = \mathbb{E}\left\{x_k^*(\mathbf{Hx} + \mathbf{n})\right\} \tag{4.8}$$

$$= \mathbf{H}\mathbb{E}\left\{x_k^*\mathbf{x}\right\} + \mathbb{E}\{\mathbf{n}\}\mathbb{E}\{x_k\} \tag{4.9}$$

$$= E_s\mathbf{H}\mathbf{e}_k \tag{4.10}$$

The vector $\mathbf{e}_k$ is $Q \times 1$ with the k-th element set to 1 and the others set to 0. The expression $\mathbb{E}\{x_k^*\mathbf{x}\}$ in (4.9) can be rewritten as $\mathbb{E}\{x_k^2\}\mathbf{e}_k$ if we make the reasonable assumption that different users' symbols are uncorrelated. The expression $\mathbb{E}\{x_k^2\}$ is the variance of the transmitted signal and hence also the energy of the transmitted signal. We will alternately define it as E_s and assume it is be normalized to 1 unless otherwise noted. Using the expressions in (4.7) and (4.10), the form for the output of the MMSE filter, z_k for user k, can be expressed as:

$$z_k = \mathbf{w}_k^H\mathbf{y} \tag{4.11}$$

$$= \mathbf{h}_k^H(\mathbf{HH}^H + \sigma^2\mathbf{I})^{-1}\mathbf{y} \tag{4.12}$$

where $\mathbf{h}_k^H$ is the conjugate transpose of the k-th column of the channel matrix $\mathbf{H}$. In a hard detection scheme, the multiuser detection block consists of a bank of MMSE filters on each of the tiles in the data frame of interest. For each tile, filtered outputs are calculated for the Q interfering users according to (4.11). Each output z_k will be used to produce hard decisions for the estimated symbols $\hat{x}_k$ on that tile. This is done by demodulating the output z_k using the minimum distance criterion to the constellation symbol set Ω_k. The remainder of the receiver chain process is similar to that of the framework initially shown in Figure 4-1.

The MMSE detector can also be implemented to produce soft output values. The output of the filter z_k can be translated to a set of a posteriori probabilities for each possible transmitted symbol $a_k \in \Omega_k$ similar to the soft optimal symbol

detector described in the previous section. The filter output-to-a posteriori conversion leverages the property that the output of the MMSE filter has an approximately Gaussian distribution [30]. This process will be further elaborated in Chapter 5.

Chapter 5

Iterative Detection and Decoding

In the standard, non-iterative receiver, the multiuser detection block and single user decoding blocks work in a serial fashion to produce estimates of the information bits in a one-shot approach. The general goal of the receiver is to minimize the probability of error on the information bits and it is known that the iterative framework originally developed for the "turbo" decoding approach for parallel concatenated codes [33] can be extended to the multiuser detection and decoding problem we are concerned with to greatly improve performance [34]. Frameworks for iterative (turbo) detection and decoding have been previously developed for multiuser CDMA systems, narrowband TDMA systems, and single-user MIMO links [12, 14, 13, 26, 27, 25, 15]. The purpose of this chapter is to describe the details of a iterative multiuser detection and decoding receiver developed for our OFDMA scenario.

5.1 Application of Turbo Principle to Detection, Decoding

The "turbo principle" [9]is a general framework for iterative message-passing algorithms designed for statistical inference problems. A turbo system consists of an iterative processing loop made up of smaller constituent algorithms. The individual constituent algorithm blocks take in observations and produce statistical inferences

on those observations. These blocks will then pass their inferences in the form of messages amongst themselves so that each block can incorporate the new incoming messages with their observation input. Through each iteration of this process, the individual blocks can recalculate their inferred information outputs based on the original observations and the received information messages of the other users. The idea here, and the foundation for turbo loop processing, is that there is a statistical interdependency between the inference calculations of each of the algorithms and that it can be exploited through message passing. The result is that the inferred information produced by each block will become increasingly refined through the successive iterations until the total inferred information is shared between the blocks and the statistical inferences converge.

An important point about well-formed turbo loop systems, which affects convergence behavior, is that the messages passed amongst constituent algorithms should only contain "new" information inferred in the most recent iteration by the block sending the message [34]. Similarly, a constituent block should not receive messages that were derived based on messages originally calculated and output by that constituent block. If this were the case, that block would then be using its previously calculated message to refine its new inference message. These positive feedback paths, which can introduce limit cycles and induce unreliable performance, are undesirable and hence avoided by requiring constituent blocks to exclude input information from its output inference message [11]. From the viewpoint of a constituent algorithm block, the input messages contain the "a priori" information of the inference calculation and the output messages will contain "extrinsic" information of the inferred information.

Before detailing the how this turbo loop is applied to our receiver, we will first review the standard non-iterative receiver framework described by Figure 4-1. In the non-iterative receiver operating with soft information, the multiuser detection blocks produce a posteriori probabilities (APPs) of the constellation symbols of the individual users based on inferences from the received structure of the multiple antenna signal observation. The APPs on the symbols are de-mapped and translated to prob-
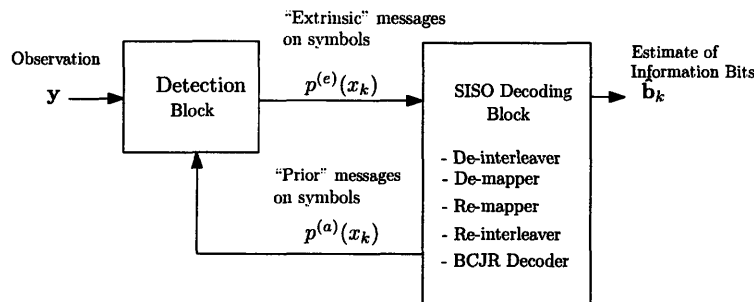
Figure 5-1: Illustration of the basic framework for an iterative detection/decoding turbo loop for a user k. The detection block incorporates all observations pertaining to user k and the decoding block is individualized for user k.

abilities on code bits, based on the constellation in use by the given user. The code bit probabilities are then converted to equivalent likelihood ratios and are de-interleaved to undo the original interleaving at the transmitter end. The single user SISO decoders receive these soft-valued likelihoods on the code bits and, through its own statistical inference and knowledge of the code structure, produce soft-valued APPs on the bits of the original information sequence. Hard decisions can then be made on the information bit LLRs to obtain the estimated information bit sequence. The decoder block is oftentimes implemented with the soft-in, soft-out BCJR algorithm when operating on convolutional codes. This type of setup is commonly known as bit-interleaved coded modulation and can be adapted into an iterative turbo processing loop[12, 14].

The inference messages travel in a one-way fashion, from the detector block to the decoder block in the non-iterative receiver. The detector and de-mapper operate with no knowledge of the coded structure of the transmission. However, there is clearly revealing information within the code that is being used by the decoder. The SISO decoder can be adapted to produce updated likelihoods (or equivalently APPs) on the code bits, in addition to the likelihoods on the information bits. These code bit APP messages contain some amount of "new" information inferred by the decoder block based on the code structure not known at the moment to the decoder block. This "new" information can passed back from the decoder to the detection block in the opposite direction so that it can be used to improve upon the initial calculations

provided by the detection block.

At this point it becomes clear how the turbo principle can be applied to the receiver end of our overloaded OFDMA system. The constituent algorithms of the turbo loop in the context of our joint multiuser detection and decoding problem are the bank of multiuser detector blocks for each of the subcarrier tiles and the set of single user SISO (soft-in soft-out) decoder blocks as seen in Figure 4-1 from the previous chapter. The general message-passing procedure just described implements the turbo loop shown in Figure 5-1. The messages are labeled from the viewpoint of the detection block and follow the principles of a well-formed turbo loop.

The task of the detector block is to produce and pass inference messages of the symbols conditioned on the observation and the "prior" information on the symbols produced by the SISO decoding block. In our system, the decoder for user k should not use likelihood information derived from the prior probability of symbol x_k when that information was calculated by the detector block from that same decoder's previous output messages. Hence the extrinsic message produced by the detector contains "new" information being gleaned from the current iteration that is separate from the "prior" input information.

The depiction in Figure 5-1 combines the specifics of single user decoding into the SISO decoding constituent block. The de-mapping and de-interleaving processes take the incoming extrinsic symbol probability messages and convert them to extrinsic likelihoods on the code bits in the order the decoder expects. The main component of the block is the BCJR decoding algorithm, which produces extrinsic likelihoods on the code bits and the corresponding information bits. For the next stage of the iteration, the extrinsic likelihoods on the code bits are converted to a priori symbol probabilities. These are the output messages of the block that are passed back to the detector block for additional iteration.

This iteration of information passing and re-calculation between the constituent blocks can ultimately improve the joint multiuser detection and decoding performance to approach that of parallel single user performance [12, 23]. We will now present the detailed derivations for the detection and decoding blocks for our OFDMA system in

44

the remainder of the chapter.

5.2 Optimal (Turbo) MAP Detection

This work will mainly focus on the formulation and implementation of the multiuser detectors for the subcarrier tiles within this iterative joint detection and decoding framework. In this section, we will formulate the optimal MAP detector for use within this turbo loop.

Multiuser detection in our OFDMA can be parallelized and done on a subcarrier tile-by-tile basis, as we had demonstrated in a previous chapter. Hence, the task of the multiuser detection process for one subcarrier tile will be to calculate the extrinsic probabilities of each of the symbols $a_k \in \Omega_k$ for each of the k users, $1 \leq k \leq Q$ where Q is the overloading factor for that subcarrier tile. To get the extrinsic probability, it is first useful to derive an expression for the a posteriori probability $p(x_k = a_k | \mathbf{y})$, conditioned on the observation of one subcarrier tile $\mathbf{y}$:

$$p(x_k|\mathbf{y}) \; = \; \frac{\sum_{\forall \mathbf{x}_{\bar{k}}} p(\mathbf{y}, \mathbf{x}_{\bar{k}} | x_k) p(x_k)}{p(\mathbf{y})} \tag{5.1}$$

$$\propto \; \sum_{\forall \mathbf{x}_{\bar{k}}} p(\mathbf{y}|\mathbf{x}_{\bar{k}}, x_k) p(\mathbf{x}_{\bar{k}}) p(x_k) \tag{5.2}$$

$$\propto \; \underbrace{\left(\sum_{\forall \mathbf{x}_{\bar{k}}} p(\mathbf{y}|\mathbf{x}) \prod_{i \neq k} p(x_i) \right)}_{} \cdot \underbrace{p(x_k)}_{} \tag{5.3}$$

where $p(x_k = a_k | \mathbf{y})$ has been written as $p(x_k|\mathbf{y})$, implicitly assuming that the instances of random variable for the transmitted signal x_k are constrained to the candidate values $a_k \in \Omega_k$. For ease of notation, the other probabilities are also denoted in this shortened notation. The expression in (5.1) is obtained by re-writing the left-hand side using Bayes' rule and then expanding the conditional probability $p(\mathbf{y}|x_k)$ to it's marginalization summation. The marginalization is the summation of $p(\mathbf{y}, \mathbf{x}_{\bar{k}}|x_k)$ over all possible vectors $\mathbf{x}_{\bar{k}}$, the vector of the stacked transmitted symbols of all users except user k, i.e. $\mathbf{x}_{\bar{k}} \in \Omega_{\bar{k}}^{(Q-1)}$. An identical formulation of this expression through

the application of the Sum-Product Algorithm for factorgraphs is described in [34].

The resulting expression for the a posteriori probability $p(x_k|\mathbf{y})$ in (5.3) is intentionally written as the two terms separated by the braces to note the extrinsic and a priori contributions. The first term (with a scaling constant) is referred to as the extrinsic probability of the symbol x_k and the second term (also with a scaling constant) is known as the *a priori* probability of symbol x_k. The extrinsic and a priori probabilities will be expressed in notation as $p^{(e)}(x_k)$ and $p^{(a)}(x_k)$, respectively. In an iterative framework, the a priori probabilities on the symbols (obtained from the previous iteration) are inputs to the multiuser detector blocks at the start of the current iteration. As can be seen in (5.3), the extrinsic probability for a symbol for a user k takes into account the priors of all the other (interfering) users, but excludes the prior about that symbol itself. This is to avoid the affect of positive feedback described earlier, as the prior for symbol x_k was derived based on the extrinsic of x_k of the previous iteration.

The multiuser detection block will output a set of extrinsic probabilities (for each of the candidate symbols in Ω_k) of the form:

$$p^{(e)}(x_k) \propto \sum_{\forall \mathbf{x}_{\bar{k}}} p(\mathbf{y}|\mathbf{x}) \prod_{i \neq k} p^{(a)}(x_i), \qquad (5.4)$$

which leaves us with the task of calculating the conditional probability $p(\mathbf{y}|\mathbf{x})$ of each term of the summation. Using the received model $\mathbf{y} = \mathbf{Hx} + \mathbf{n}$, we can rewrite the conditional probability as

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{(\pi)^{N_r}\sigma} \exp\left(-\frac{1}{\sigma^2}(\mathbf{y} - \mathbf{m_y})^H(\mathbf{y} - \mathbf{m_y})\right) \qquad (5.5)$$

$$= \frac{1}{(\pi)^{N_r}\sigma} \exp\left(-\frac{1}{\sigma^2}|\mathbf{y} - \mathbf{Hx}|^2\right) \qquad (5.6)$$

where $\mathbf{y}$ can be considered a Gaussian random vector observation over the N_r antennas distributed as $\mathbf{y} \sim \mathcal{CN}(\mathbf{m_y}, \sigma^2\mathbf{I})$. When conditioned on the transmitted symbol $\mathbf{x}$ (i.e. $\mathbf{x} = \mathbf{a}$) and a given channel gain matrix $\mathbf{H}$, the mean vector is $\mathbf{m_y} = \mathbf{Hx}$, which is substituted into (5.5) to give the more familiar form [27, 28] for the conditional

For each tile (i.e. $\forall f, m$):
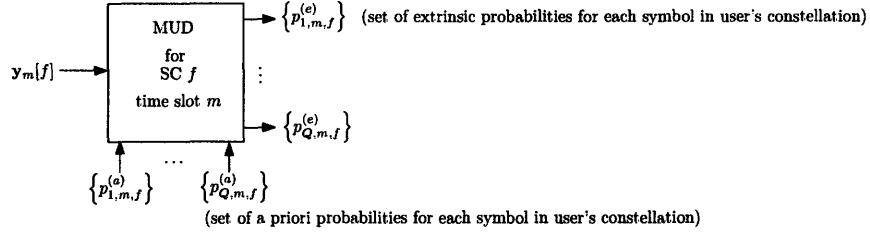
Figure 5-2: Schematic for a single multiuser detector block for a given subcarrier tile. It receives the signal observation vector and priors as inputs and outputs a set of extrinsic probabilities for each of the Q interfering users.

probability seen in (5.6).

We can insert the expression in (5.6) into the extrinsic probability expression in (5.4) to get:

$$p^{(e)}(x_k) = \alpha \sum_{\forall \mathbf{x}_{\bar{k}}} \left\{ \exp\left(-\frac{1}{\sigma^2}|\mathbf{y} - \mathbf{Hx}|^2\right) \prod_{i \neq k} p^{(a)}(x_i) \right\} \tag{5.7}$$

The normalization factors and constant factors have been lumped together and are expressed in the constant α such that the set of $p^{(e)}(x_k)$ for all symbols in Ω_k form a valid probability distribution. In the optimal MAP detector, each multiuser detector block on each tile will receive a set of priors for the symbols of each user and in turn, produce a set of extrinsic probabilities for each of the Q users on that tile. A schematic for this input and output relation of a single multiuser detector block operating on a single tile is shown in Figure 5-2.

5.3 Reduced Complexity MMSE with Soft I.C.

The complexity involved in calculating each extrinsic probability for each possible symbol of the constellation can get to be quite large and unfeasible in actual system implementation [12, 7]. Therefore it is of interest to explore reduced complexity detection, using linear and non-linear methods of calculating $p^{(e)}(x_k)$, to fit in our iterative framework.

We start with the received signal model $\mathbf{y} = \mathbf{Hx} + \mathbf{n}$ where $\mathbf{y}$ is a $N_r \times 1$ vector of the received signal, $\mathbf{x}$ is a $Q \times 1$ vector of the transmitted symbols of the users we are trying to detect and decode, and $\mathbf{H}$ is a $N_r \times Q$ matrix of the channel coefficients estimated by the receiver. Similar to the non-iterative scenario, we are looking to apply MMSE filtering to the observation at each detector block for every iteration. It is not immediately clear however, how the multiuser detector blocks will incorporate the a priori data from the decoders from the previous iterations.

When detecting for a user k, the available a priori information gives us probabilities of the other users' signals, which can be used to estimate their interference contributions. Hence, we will perform *soft interference cancellation* on the observed signal $\mathbf{y}$ based on the available a priori probabilities of the other users symbols before applying the MMSE filter. At the detector, soft estimates can be formed for each of the $k \in Q$ users as follows:

$$\tilde{x}_k = \sum_{\forall a_k \in \Omega_k} a_k\, p^{(a)}(x_k = a_k), \qquad \text{for all } k = 1, \dots, N_u. \tag{5.8}$$

When performing detection for user k, we want to cancel out the other users' interference components from the total received signal $\mathbf{y}$, i.e. soft symbol estimates of all symbols *except* for user k's symbol. We will denote the vector of all soft (estimated) symbols except for user k's symbol as $\tilde{\mathbf{x}}_{\bar{k}}$. If for example $k = 3$ and $Q = 4$, $\tilde{\mathbf{x}}_{\bar{3}} = [\tilde{x}_1, \tilde{x}_2, 0, \tilde{x}_4]^T$.

Using the received signal model $\mathbf{y} = \mathbf{Hx} + \mathbf{n}$, we denote the received signal for user k after soft interference cancellation as $\tilde{\mathbf{y}}_k = \mathbf{H}(\mathbf{x} - \tilde{\mathbf{x}}_{\bar{k}}) + \mathbf{n}$. The MMSE filter will be applied on the post-cancellation signal to further suppress the residual interference and hence should be based on the statistical properties of $\tilde{\mathbf{y}}_k$. Similar to how we

derived (4.5), (4.7), and (4.10) we can write the filter components as the following:

$$\mathbb{E}\{\tilde{\mathbf{y}}_k\tilde{\mathbf{y}}_k^H\} = \mathbb{E}\left\{(\mathbf{H}(\mathbf{x}-\tilde{\mathbf{x}}_{\bar{k}})+\mathbf{n})\left((\mathbf{x}-\tilde{\mathbf{x}}_{\bar{k}})^H\mathbf{H}^H+\mathbf{n}^H\right)\right\} \tag{5.9}$$

$$= \mathbf{H}\,\mathrm{Cov}(|\mathbf{x}-\tilde{\mathbf{x}}_{\bar{k}}|)\mathbf{H}^H+\sigma^2\mathbf{I} \tag{5.10}$$

$$\mathbb{E}\{x_k^*\tilde{\mathbf{y}}_k\} = \mathbb{E}\left\{x_k^*\left(\mathbf{H}(\mathbf{x}-\tilde{\mathbf{x}}_{\bar{k}})+\mathbf{n}\right)\right\} \tag{5.11}$$

$$= \mathbf{H}\,\mathbb{E}\left\{x_k^*(\mathbf{x}-\tilde{\mathbf{x}}_{\bar{k}})\right\}+\mathbb{E}\{\mathbf{n}\}\mathbb{E}\{x_k\} \tag{5.12}$$

$$= E_s\mathbf{H}\mathbf{e}_k. \tag{5.13}$$

The constant E_s is the energy of the transmitted symbol (i.e. $\mathbb{E}|x_k|^2$), and is often considered as normalized to 1. In the above equations, we will denote $\Lambda_k = \mathrm{Cov}(\mathbf{x}-\tilde{\mathbf{x}}_{\bar{k}})$ for notational ease. Λ_k is a diagonal matrix such that $[\Lambda_k]_{kk} = 1$ and $[\Lambda_k]_{jj} = \mathrm{Var}(u_j)$ for all $j \neq k$ where u_j is the uncancelled interference of user j. In other words, the diagonal entries of Λ_k represent the energy of user k's signal and the remaining interference energy components of the other users j. The expression for the variance of the uncancelled interference can be simplified as follows:

$$\mathbb{E}\left\{|x_j-\tilde{x}_j|^2\right\} = \mathbb{E}\left\{|x_j|^2\right\}-|\tilde{x}_j|^2. \tag{5.14}$$

where $\mathrm{Cov}(|\mathbf{x}-\tilde{\mathbf{x}}_{\bar{k}}|)$ is equivalently written as $\mathbb{E}\left\{|x_j-\tilde{x}_j|^2\right\}$. On successive iterations through the turbo loop, these variances should tend towards zero, as the soft interference estimates become more accurate with increasingly refined a priori inputs.

The filter $\tilde{\mathbf{w}}_k$ is of the form $\tilde{\mathbf{w}}_k = E\{\tilde{\mathbf{y}}_k\tilde{\mathbf{y}}_k^H\}^{-1}E\{x_k^*\tilde{\mathbf{y}}_k\}$, which involves the inversion of an $N_r \times N_r$ matrix. This can be a computationally intensive task but we can use the fact that $\mathbf{H}$ is a $N_r \times Q$ matrix and use the *matrix inversion lemma* to instead get by with inverting a $Q \times Q$ matrix. For the purposes of reliable multiuser detection, we normally limit the system such that $N_r > Q$. This leads to a smaller matrix inversion calculation which can give a savings in complexity. We use the known matrix identity found in [35]

$$(\mathbf{I}+\mathbf{AB})^{-1}\mathbf{A} = \mathbf{A}(\mathbf{I}+\mathbf{BA})^{-1} \tag{5.15}$$

to write $\tilde{\mathbf{w}}_k$, after substitution of (5.10) and (5.13), in the following form

$$
\begin{align}
\tilde{\mathbf{w}}_k &= (\mathbf{H}\mathbf{\Lambda}_k\mathbf{H}^H + \sigma^2\mathbf{I})^{-1}\mathbf{H}\mathbf{e}_k \tag{5.16} \\
&= \sigma^2(\mathbf{I} + \frac{1}{\sigma^2}\mathbf{H}\mathbf{\Lambda}_k\mathbf{H}^H)^{-1}\mathbf{H}\mathbf{e}_k \tag{5.17} \\
&= \sigma^2\mathbf{H}(\mathbf{I} + \frac{1}{\sigma^2}\mathbf{\Lambda}_k\mathbf{H}^H\mathbf{H})^{-1}\mathbf{e}_k \tag{5.18} \\
&= \mathbf{H}(\sigma^2\mathbf{I} + \mathbf{\Lambda}_k\mathbf{H}^H\mathbf{H})^{-1}\mathbf{e}_k. \tag{5.19}
\end{align}
$$

The output of the MMSE filter after soft interference cancellation for the k-th is then $\tilde{z}_k = \tilde{\mathbf{w}}_k^H\tilde{\mathbf{y}}_k$. We can use (5.19) to expand the output of the filter for user k as

$$
\begin{align}
\tilde{z}_k &= \tilde{\mathbf{w}}_k^H\tilde{\mathbf{y}}_k \tag{5.20} \\
&= \mathbf{e}_k^T(\sigma^2\mathbf{I} + \mathbf{H}^H\mathbf{H}\mathbf{\Lambda}_k)^{-1}\mathbf{H}^H(\mathbf{y} - \mathbf{H}\tilde{\mathbf{x}}_{\bar{k}}) \tag{5.21} \\
&= \mathbf{e}_k^T(\sigma^2\mathbf{I} + \mathbf{H}^H\mathbf{H}\mathbf{\Lambda}_k)^{-1}(\mathbf{H}^H\mathbf{y} - \mathbf{H}^H\mathbf{H}\tilde{\mathbf{x}}_{\bar{k}}). \tag{5.22}
\end{align}
$$

If this were a non-iterative system with hard decisions on the symbols, the filter output $\tilde{z}_k$ would be thresholded to produce the closest constellation point. However, as shown in Figure 5-2, the multiuser detector block is expected to produce extrinsic probabilities so we must convert the filtered output to a probability.

It is known that the output $\tilde{z}_k$ of the MMSE filter has a Gaussian distribution and that it can be represented as $\tilde{z}_k = \mu_k x_k + m_k$ where $m_k \sim \mathcal{CN}(0, \nu_k^2)$ [30, 12]. The distribution of $\tilde{z}_k$ can then be parameterized by μ_k and ν_k^2, which are be derived as follows:

$$
\begin{align}
\mu_k &= \mathbb{E}\{\tilde{z}_k x_k^*\} \tag{5.23} \\
&= \mathbb{E}\{\tilde{\mathbf{w}}_k^H\tilde{\mathbf{y}}_k x_k^*\} \tag{5.24} \\
&= \tilde{\mathbf{w}}_k^H\mathbb{E}\{(\mathbf{H}(\mathbf{x} - \tilde{\mathbf{x}}_{\bar{k}}) + \mathbf{n})x_k^*\} \tag{5.25} \\
&= \tilde{\mathbf{w}}_k^H\mathbf{H}\mathbf{e}_k \tag{5.26}
\end{align}
$$

where $\mathbf{H}\mathbf{e}_k$ can also be written as $\mathbf{h}_k$, the k-th column of $\mathbf{H}$ and

$$\nu_k^2 \;=\; \mathbb{E}\{|\tilde{z}_k|^2\} - (\mathbb{E}\{|\tilde{z}_k|\})^2 \tag{5.27}$$

$$\;=\; \mathbb{E}\{\tilde{\mathbf{w}}_k^H \tilde{\mathbf{y}}_k \tilde{\mathbf{y}}_k^H \tilde{\mathbf{w}}_k\} - \mu_k^2 \tag{5.28}$$

$$\;=\; \tilde{\mathbf{w}}_k^H \mathbb{E}\{\tilde{\mathbf{y}}_k \tilde{\mathbf{y}}_k^H\} \underbrace{\mathbb{E}\{\tilde{\mathbf{y}}_k \tilde{\mathbf{y}}_k^H\}^{-1} \mathbf{H}\mathbf{e}_k}_{\tilde{\mathbf{w}}_k} - \mu_k^2 \tag{5.29}$$

$$\;=\; \tilde{\mathbf{w}}_k^H \mathbf{H}\mathbf{e}_k - \mu_k^2 \tag{5.30}$$

$$\;=\; \mu_k - \mu_k^2. \tag{5.31}$$

Using these parameters, it is now possible to construct expressions for the extrinsic probabilities that are to be passed to the soft decoder. We know that the MMSE filter output has a Gaussian distribution $\tilde{z}_k \sim \mathcal{CN}(\mu_k a_k, \nu_k^2)$ when $\tilde{z}_k$ is conditioned such that $x_k = a_k$ where $a_k \in \Omega_k$. This allows us to write each extrinsic probability of the symbol set as:

$$p^{(e)}(x_k = a_k) \;\propto\; p(\tilde{z}_k | x_k = a_k) \tag{5.32}$$

$$\;=\; \alpha \exp\left(-\frac{1}{\nu_k^2} |\tilde{z}_k - \mu_k a_k|^2\right), \tag{5.33}$$

where α in the final expression is some constant normalization factor. The multiuser detector will calculate extrinsic probabilities $p^{(e)}$ for all Q users on the subcarrier tile and for each symbol in each user's constellation set Ω, i.e. $\{p^{(e)}(x_k = a_k)\}_{\forall a_k \in \Omega_k} \forall k \in \{Q\}$. These extrinsic probabilities for user k are a functions of $\tilde{z}_k$ and hence implicitly conditioned on the received observation $\mathbf{y}$ and the a priori probabilities $\{p^{(a)}(x_i = a_i)\}_{\forall a_i \in \Omega_i, i \neq k}$.

5.4   SISO Decoding for Iterative Framework

When decoding a data frame of Q overlapping users, the receiver will have a bank of N_s (one for each subcarrier tile) total multiuser detector blocks (of the form seen in Figure 5-2) implementing either the optimal MAP detection algorithm or the reduced-complexity MMSE algorithm with soft interference cancellation. The SISO decoding

Figure 5-3: Schematic of the the single user SISO decoding of the turbo loop from received messages of MUD block. One note is that most SISO decoders are designed to produce APP LLRs on the code bits. Our turbo loop is designed to pass only extrinsic messages so the input message is subtracted from the output message to remove the a priori contribution.

of each of the Q users is done separately and in parallel of one another. A SISO decoder block for a given user will need to receive the set of extrinsic probabilities of all the possible candidate symbols $a_k \in \Omega_k$ for each of the individual tiles in the data frame being decoded. These sets for each user k can be gathered from the bank of multiuser detectors and passed to each k-th user's SISO decoding block, as shown in Figure 5-3.

For a given user k, the N_s sets of extrinsic probabilities must be de-mapped to a $N_s \cdot \log_2 M$ sequence of code bit likelihoods, where M is the size of user k's constellation. We know that each of the N_s transmitted constellation symbols (i.e. x_k) is mapped to a code bit pattern vector $\phi(x_k) \rightarrow [c_1, \ldots, c_{log_2 M}]^T$ via some known mapping function. The de-mapper first derives extrinsic probabilities on each of the code bits $c_{k,i}$ in the codeword for the data frame by the following calculation:

$$p^{(e)}(c_{k,i} = 1) = \beta_1 \sum_{\forall a_k : c_{k,i}=1} p^{(e)}(x_k = a_k) \prod_{n \neq m} p^{(a)}(c_{k,n}) \qquad (5.34)$$

$$p^{(e)}(c_{k,i} = 0) = \beta_0 \sum_{\forall a_k : c_{k,i}=0} p^{(e)}(x_k = a_k) \prod_{n \neq m} p^{(a)}(c_{k,n}) \qquad (5.35)$$

We assume every possible candidate transmitted symbol a_k in the M-QAM constel-

52

lation is mapped to a known unique bit pattern $\phi(a_k) \to [c_1, \ldots, c_{log_2 M}]^T$ of length $\log_2 M$. In (5.34) and (5.35), the summation is over the candidate symbols (of the transmitted symbol x_k from which it was mapped) that can be mapped from the code bit pattern vector containing $c_{k,i}$. The product is over the prior probabilities for the mapped bit values of the symbols a_k chosen for the summation, excluding the the bit $c_{k,i}$. The constants β_1 and β_0 are the appropriate normalization factors. These extrinsic probabilities on code bits are calculated in turbo fashion with the similar marginalization approach used to find extrinsic probabilities on the candidate transmitted symbols a_k and take into account the prior code bit probabilities (i.e. LLRs).

The extrinsic probabilities for each of the code bits i of user k can be equivalently represented as a single log-likelihood ratio:

$$\lambda_k^{(e)}(c_{k,i}) = \log \frac{p^{(e)}(c_{k,i} = 1)}{p^{(e)}(c_{k,i} = 0)}. \tag{5.36}$$

A sequence of length $N_s \cdot \log_2 M$ of log-likelihood ratios (LLRs) is assembled for the code bit sequence and passed to the de-interleaver to undo the interleaving effects and to re-order the bits to the order expected by the SISO decoder. The SISO decoder receives this sequence of LLRs on the code bits and produces a sequence of APPs on the information bits and a sequence of the updated APPs of the bits in the codeword. Because we are working with LLRs, the input to the SISO decoder (which is essentially the a priori information of the code bits) can be subtracted from the output code bit APPs to produce the extrinsic LLRs of the code bits, as seen in Figure 5-3.

To continue the iterative process, the extrinsic code bit LLRs are re-interleaved and re-mapped to produce the sets of a priori probabilities on the symbols in the data frame. The re-mapping process converts the sequence of code bit probabilities $p^{(a)}(c_{k,i})$ to prior symbol probabilities that are passed to the detector block in the form $p^{(a)}(x_k = a_k)$. These are calculated as follows:

$$p^{(a)}(x_k = a_k) = \gamma \prod_{i=1}^{\log_2 M} p^{(a)}(c_{k,i}) \qquad (5.37)$$

where the product is over the $\log_2 M$ code bit priors of the mapped bits of the particular symbol $x_k = a_k$ and γ is another normalization constant. These sets of probabilities $\{p^{(a)}(x_k = a_k)\}$ of candidate symbols $a_k \in \Omega_k$ on each tile get gathered and passed to the appropriate multiuser detector blocks to start the next iteration of the turbo loop.

Chapter 6

Simulation Framework

6.1 Goals of Simulation

The purpose of Monte Carlo numerical simulation will be to understand and charac-
terize the general performance of the multiple antenna base station receiver imple-
menting the iterative detection and decoding methods described in Chapter 5. We
will avoid focusing on overly specific system scenarios and instead apply more general
system parameters that remain realistic and extensible to a wide range of setups.
The main focus of the simulation will be the error-rate performance of the iterative
detection and decoding algorithms and hence the most precision will be afforded to
the implementation of those blocks. When possible, other aspects of the simulation
framework will be abstracted away or generalized into concise forms.

Because of the OFDM modulation being used, each channel resource (i.e. block
of subcarrier tiles that hold one data frame of modulation symbols) can be processed
and treated separately. The OFDMA system can be viewed as a set of parallel com-
munication links, where each link consists of a channel resource and the Q overloaded
users. Rather than simulate with the intent of examining overall system-wide per-
formance, we will rather focus on performance at the individual physical layer link
level, namely the decoding of overlapping data frames of Q users on a single channel
resource. This allows us to abstract away many of the extraneous details of higher
layer implementation.

System Paramter	Value
Carrier Frequency	1.9 GHz
System Bandwidth	~10 MHz
Chip period	100 ns
FFT/IFFT size	1024 samples
Subcarrier spacing	~10 kHz

Table 6.1: Summary of system simulation parameters

The prime issues of interest in the simulation are the gains of iterative processing with respect to: the tradeoffs of performance with computational complexity, the effects of the overloading factor Q on individual single user performance, and the impacts of spatial diversity on performance. The choices of system parameters and decisions in the framework design are made to most conveniently probe these specific system aspects. The choices for non-essential parameters are made arbitrarily and can often be substituted with other realistic values.

As a note, the simulations are implemented in Matlab code for convenience and because of the availability of built-in toolboxes for common communication system simulation functions.

6.2 System Parameters, and Basic Assumptions

We are interested in simulating the transmitter and receiver sides of a synchronous uplink OFDMA system operating over a multiple antenna, multipath channel. We choose the system to operate at a carrier frequency of 1.9 GHz with a system bandwidth of approximately 10 MHz and a corresponding time sample (i.e. "chip") period of 100 ns. The size of the DFT and IDFT used by the transmitters and the receivers, respectively, will be set to 1024 points for a total of 1024 available subcarriers in the frequency domain. In practice, a portion of these subcarriers are set to be "guard" carriers and are hence unused but we will not model this in our system. The main parameters for the simulation are summarized in Table 6.1.

6.2.1 Synchronization

It is first important to define the meaning of "synchronous" for our multiuser system setup. As described in the previous chapters, the base station receiver will receive a superposition of the time-domain signals of all users within the system. The receiver will then matched filter (with the square root Nyquist shaping pulse) and sample the continuous time signal get the equivalent discrete time samples. The receiver will then take a length N segment of the samples (corresponding to one OFDM symbol) and take the DFT to get observations in the discrete frequency domain (i.e. subcarriers).

For this process to work properly and to allow for the receiver to recover the users' signals on the subcarriers they were originally transmitted on, the users passband modulators must be synchronized to the carrier frequency of the base station demodulator. Additionally, the users must transmit in a synchronous fashion such that their OFDM symbols (in the time-domain) arrive within the appropriate OFDM symbol period windows to the base station's time reference. This is a complex task because of propagations delays caused by the physical distances the signals must travel. However, the synchronization does not need to be achieved at the level of individual time samples (i.e. chips) but rather to within the interval of the cyclic prefix. Users' transmitted OFDM symbols that arrive within the cyclic prefix time interval of the base station's OFDM symbol period reference time can be adequately resolved (when coupled with proper channel estimation) in the frequency domain after the application of the DFT. This is because of the cyclic properties of the transmitted symbols with the addition of the cyclic prefix and the fact that time shifts and delays correspond to complex phase rotations in frequency. Hence, slightly misaligned users can be tolerated, as the channel estimation at the receiver can be used to correct for the phase rotations of the received observations. Synchronization of individual users with the base station is achieved through training sequences, pilots, and dedicated timing estimation algorithms and will be abstracted away from the simulation. Detailed analysis of OFDMA synchronization issues can be found in [18].

6.2.2 System Resources

Using the previous assumption that users within the system are synchronized to the central base station, the users' notion of the available channel resources will correspond to that of the base station's. As mentioned in the previous chapter, a single user's channel assignment will consist of a block of frequency subcarrier tiles on which it will modulate its frame of data symbols. For the purposes of our system simulation, this channel resource unit will consist of 16 contiguous subcarriers for 8 OFDM symbol periods for a total block of 128 subcarrier tiles, or exactly one data frame of transmitted symbols. Each user in the system will be assigned a single channel resource unit and overlapping users will overlap on all the tiles within their assigned block.

Additionally, for purposes of simplicity, users are assumed to be perfectly power-controlled such that the averaged received signal energies for each user seen the base station are equal.

6.2.3 Channel Estimation

Since processing in OFDM is done in the frequency domain, we will be interested in the frequency responses of each of the subcarrier tiles within a channel resource block during detection and decoding. These are obtained through pilot-aided channel estimation algorithms. The pilots are generally known symbols embedded into set tiles of the block. Estimation is performed by interpolating the observed pilot responses over the entire data frame to capture temporal and frequency correlation. Channel Estimation algorithms for adaptive antenna array systems are developed in detail in [17]. However, we will assume perfect channel estimation algorithms and that the channel response on each tile is known at the receiver.

Parameter	Value
Code Rate (Convolutional), R	1/2
Encoder Constraint Length, ν	7
Information Frame Size	122 bits
Codeword Size	256 bits
M-QAM Constellation Size	4
Channel Assignment Width	16 subcarriers
Channel Assignment Duration	8 OFDM symbol periods
Data Frame Size	128 symbols

Table 6.2: Summary of parameters for transmission and reception.

6.3 Implementation Parameters and Guidelines

The framework is designed to simulate the transmission and reception of data frames with the focus on a single channel resource within the 10 MHz frequency band. For every data frame period, there will be a set of Q overlapping users that will be detected and decoded for.

6.3.1 Transmitter

Each of the Q users will have an independently generated source of information to transmit over the channel. We will assume the information is efficiently source-coded such that the sequence of bits being generated behaves as a Bernoulli random process (i.e. memoryless, equal probability of 0's, 1's). The sequence of information bits is then to be passed to a convolutional encoder for channel coding. The convolutional encoder will apply a rate 1/2 non-recursive, non-systematic code (NSC) with a constraint length of 7 and the commonly used generator polynomial (in octal form) of [171, 133]. The incoming information sequence will be packaged into segments of 122 bits for the encoder. Each segment of bits corresponds to the total information bits to be transmitted in one data frame. The 122 bits will be combined with the termination bits for the encoder to produce information words of size 128 bits and codewords of 256 bits. The summary of parameters for data transmission (and reception) are summarized in Table 6.1.

The resulting sequence of coded bits will be mapped and modulated into the

physical data frame for transmission. However, before mapping, the code bits of the codeword to modulation symbols, the code bits are first interleaved to approximately eliminate correlation between the eventual mapped symbols and to reduce the effect of bursty channel errors for the receiver. The interleaving pattern is unique for each user and pseudo-random such that it is known at the receiver and is reversible. This sequence of 256 interleaved coded bits is then mapped to modulation symbols chosen from each user's M-QAM constellation set. In keeping with the goal of being general, we will assign all users to be transmitting the same unit energy 4-QAM constellation. In reality, the choice of constellation set can differ for each user and be adaptive to maximize data rate for given channel conditions and error rate requirements. Because $M = 4$, each symbol will be Gray-mapped from a $\log_2 M$ segment of the coded bit sequence. The mapping block will therefore produce a sequence of 128 constellation symbols to be modulated onto the subcarrier tiles. This sequence will be placed onto the 16 subcarrier $\times$ 8 OFDM symbol period grid in some pre-determined pattern known to both the transmitter and receiver.

The physical unit for transmission of each user is an OFDM symbol and hence the data frame to be transmitted is buffered and transmitted over 8 OFDM symbol periods. Along with the 16 assigned subcarriers per OFDM symbol, the remaining subcarriers within the 1024-point frequency band are loaded with values of 0 such that the user does not cause interference to other channel resources. The data-embedded 1024-point set of discrete frequency points is passed through an IDFT unit, implemented with the IFFT algorithm, to produce a time-domain sequence of 1024 chips/samples. These time-domain samples are pulse interpolated with square root Nyquist pulses for carrier frequency modulation and transmission over the channel. For the purposes of our simulated system, we will not model the time-domain conversion and interpolation procedures and instead assume ideal functionality. As we have shown in Chapters 2 and 3, the OFDM modulation allows us to efficiently simulate and process data transmission and reception in the frequency-domain.
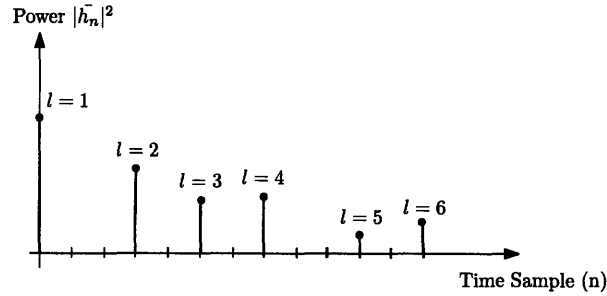
Power $|\bar{h}_n|^2$

$l=1$

$l=2$

$l=3$ $l=4$

$l=5$ $l=6$

Time Sample (n)

Figure 6-1: Example of a power delay profile for a single realization of a discrete-time multipath channel $\bar{h}_n$ with 6 non-zero taps.

6.3.2 Multipath Channel

The effects of the multipath channel can also be modeled in the frequency domain as a set of flat-fading complex channel gains, assuming the subcarrier widths are less than the coherence bandwidth of the channel and the channel is underspread (i.e. the delay spread is less than the coherence time of the channel) [1]. A data frame being simulated consists of 8 OFDM symbol periods. Normally, the mobility of the channel (i.e. speed of the user relative to the base station) determines the Doppler spectrum and the temporal correlation of the channel. However, we will consider a low mobility case such that the channel can be considered static for the duration of the data frame transmission. Since we are treating the transmitted data as frequency domain values, we are interested in obtaining frequency domain values of randomly generated (but realistic) channel realizations. The existing models for multipath channels and MIMO channels [36, 37] are based in the time domain however so we will first generate realizations of the multiple antenna channel in the time-domain.

We will use the ITU Pedestrian B channel with a non-uniform power delay profile (PDP) for our underlying channel model [38]. The Pedestrian B channel can be modeled as a tapped delay-line of discrete taps at predetermined locations in the time domain. We will assume a highly scattering environment such that each non-zero tap $\bar{h}_l$ of a channel realization can be modeled as a sample of a Rayleigh fading distribution $\bar{h}_l \sim \mathcal{CN}(0, \sigma_l^2)$ and that the total power gain of the channel is $\sum_{\forall l} \mathbb{E}\left\{|\bar{h}_l|^2\right\} = 1$. An example of the power delay profile for a single channel realization is illustrated in

Tap No.	Delay (ns)	Relative Power (σ_l^2)
1	0	0.406
2	2	0.330
3	8	0.131
4	12	0.064
5	23	0.067
6	37	0.002

Table 6.3: Parameters for tapped delay-line model of ITU Pedestrian B channel.

Figure 6 − 1.

The parameters used to model our Pedestrian B channel are shown in Table 6.3. For purposes of simplicity, the total power of the channel impulse response is generally normalized to 1 such that the channel has on average, unity power gain, as can be seen in Table 6.3. Our system simulations will use these channel parameters to generate random tapped delay-line.

The ITU Pedestrian B channel model provides the parameters to generate single realizations of single-input single-output (SISO) channels. Since we are considering a multiple antenna array at the receiving basestation we will need to extend the model for SISO channels to multiple-input multiple-output (MIMO) channels, or more technically in our case, single-input multiple-output (SIMO) channels. A SIMO channel can be viewed as a collection of SISO channels from the user's single transmitting antenna to each of the multiple receiver antennas at the base station. As an extension, the collection of SIMO channels for each of the users can be viewed as a single MIMO channel of geographically distributed transmit antennas. Due to possibilities of a less than ideal well-scattering environment, the MIMO/SIMO channels can suffer from the effects of transmitter and receiver correlation. Because of assumed geographical separation of users and independent local scattering environments, our simulation will make the assumption of no transmitter-side correlation.

We will start by generating a SISO channel for each user-receiver antenna pair where each tap is a sample of the appropriate Rayleigh processes. To model receiver-side correlation, we will color (i.e. apply correlation to) each randomly generated set of SIMO channel realizations on a tap-by-tap basis, since taps within a realiza-

tion are independent of one another [39]. A given user will have a discrete tapped-delay line channel to each of the N_r base station antennas. For each tap l, we will start with a randomly generated set of the complex gains to each receiving antenna: $\bar{\mathbf{h}}_l = [\bar{h}_l^{(1)}, \ldots, \bar{h}_l^{(N_r)}]^T$. We will produce a colored set of channel responses through a correlation matrix $\mathbf{R}$ to give us the new set of correlated complex gains $\check{\mathbf{h}}_l$ for the tap of interest:

$$\check{\mathbf{h}}_l = \mathbf{R}^{\frac{1}{2}} \bar{\mathbf{h}}_l \tag{6.1}$$

Correlation matrices for antenna arrays are commonly modeled with an exponential correlation model $\mathbf{R} = \left[\rho^{|i-j|}\right]_{i,j}$ such that the entry of the $N_r \times N_r$ matrix in the i-th row and j-th column is $\rho^{|i-j|}$ where $0 \leq \rho \leq 1$ and ρ is the correlation coefficient between the i-th and j-th antenna of the receiver array [40, 41]. For example, if $N_r = 4$, the colored tap gain would be calculated as follows:

$$\begin{bmatrix} \check{h}_l^{(1)} \\ \check{h}_l^{(2)} \\ \check{h}_l^{(3)} \\ \check{h}_l^{(4)} \end{bmatrix} = \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix} \begin{bmatrix} \bar{h}_l^{(1)} \\ \bar{h}_l^{(2)} \\ \bar{h}_l^{(3)} \\ \bar{h}_l^{(4)} \end{bmatrix}. \tag{6.2}$$

This model is intuitive in the physical sense that, assuming the antennas are indexed in spatial ordering from i and j, correlation decreases exponentially with physical antenna separation. Our simulation will explore the effects of different values of ρ on the receiver performance. The set of taps within each channel realization are colored through this approach to yield a set of colored time-domain channel responses to each of the correlated receiver antennas.

6.3.3 Multiple Antenna Receiver Front-End

We will not model the physical effects of antenna configurations, and to remain as general as possible, we will be considering uniform linear arrays of antennas at the base station with spacings of half wavelengths or greater. Additionally, we will be

considering arrays of 2 and 4 receiver antennas, which dictates that we will be dealing with 2×2 and 4×4 channel matrices $\mathbf{H}$ for each subcarrier tile in the data frame.

In a real system, channel estimation is performed to obtain frequency fading gains on the subcarrier tiles to each antenna from each of the given users. For example, in the scenario of 4 receiver antennas ($N_r = 4$) and 4 overlapping users ($Q = 4$), the receiver will calculate a total of 16 channel estimates per subcarrier tile and assemble them into the appropriate channel matrix $\mathbf{H}$ for use in detection. Since we are not modeling the details of channel estimation, we will assume the frequency fading gains to each of the antennas are perfect and use the channel realization generated in the previous section. We are interested in the frequency-domain representation of the channel and hence we will take the DFT of the colored time-domain realizations generated for the antennas to obtain correlated frequency-domain fading gains on the subcarriers. Again, as a note, we will assume the channel realizations are time-invariant in the duration of a data frame (i.e. 8 OFDM symbol periods).

Because of Parseval's relation [19], the average energy of the frequency responses (for each user) on the subcarriers across the entire frequency band is equal to the average energy of the time-domain channel response and hence unity in our case:

$$\sum_{n=0}^{N-1} |\bar{h}_{k,n}|^2 = \frac{1}{N} \sum_{f=0}^{N-1} |h_{k,f}|^2 = 1 \tag{6.3}$$

where the time-domain summation is over an N-sample window containing the 6 taps (in the Pedestrian B channel) of a user k and where the summation in frequency is over all N subcarrier responses for that user k. In our simulation, this relation will apply to each of the SISO links making up the total MIMO channel.

6.3.4 Iterative Receiver and Performance Measurement

The simulation will examine the performance of the two iterative detection algorithms in the context of the iterative receiver design described and elaborated in Chapter 5. The detection block will implement either the optimal MAP symbol detection algorithm of Section 5.2 or the MMSE detection algorithm with soft interference

cancellation described in Section 5.3. In the decoding block, the mapping and interleaving functions (in both directions of the loop) operate with knowledge of the users' constellation types (i.e. 4-QAM for all users) and their pseudorandom interleaving patterns, respectively. Likewise, the SISO decoder is implemented using the BCJR algorithm with knowledge of the generator polynomial for the convolutionally coded bits.

We will be testing the scenarios of 2 receiver antennas with 2 overloaded users ($N_r = 2$, $Q = 2$), 4 receiver antennas with 2 overloaded users ($N_r = 4$, $Q = 2$), and 4 antennas with 4 overloaded users ($N_r = 4$, $Q = 4$). The other principle issue of interest in our simulations will be the effect of the receiver antenna correlation at the base station and hence we will simulate cases of low ($\rho = 0.25$), moderate ($\rho = 0.50$), and high ($\rho = 0.75$) antenna correlations with the goal of determining the robustness of the optimal and reduced complexity MMSE iterative algorithms in the different receiver antenna and overloading factor scenarios. We will measure performance for different levels of iteration (1-3 iterations) with the single iteration case being equivalent to measuring the non-iterative receiver performance. For a point of performance comparison, and to judge the interference suppression capabilities of the algorithms relative to one another, an additional set of simulations will be run for the single-user, non-iterative, non-overloaded OFDMA scenario.

The simulation is run using each set of complete parameters for a range of E_s/N_o (SNR) values to measure and plot bit-error rate (BER) performance. To compute BER performance at a given E_s/N_o value and iteration stage, hard decisions are made on the information bit LLR outputs of each users SISO decoder at the end of that iteration stage and compared to the original transmitted information bit words for the corresponding users. Any bit error within a transmitted frame constitutes a frame error for that entire transmitted frame. The total number of individual bit errors for each frame in error are recorded and the simulation is run at that SNR value until at least 100 frame errors occur per user during the third iteration stage.

The final bit error rate is calculated at the termination of the simulation by dividing the total observed bit errors by the total number of bits transmitted in

the total frames transmitted. This includes the frames with bit errors and the frames with no bit errors:

$$\text{BER} = \frac{\text{total bit errors}}{122 \times (\text{total frames tx to acheive 100 frame errors})} \qquad (6.4)$$

where the denominator is multiplied by 122, the number of information bits in one information word.

Chapter 7

Numerical Results and Discussion

7.1 Simulated Cases

This section will present a summary of the numerical results obtained from Monte Carlo simulation of the iterative detection and decoding algorithms implemented in the system framework described in Chapter 6. We are mainly interested in the information bit error performance of the turbo Optimal MAP symbol detector and the turbo MMSE filter with soft interference cancellation. The results will be presented in the form of bit error-rate (BER) curves plotted with respect to the signal-to-noise ratio (i.e. E_s/N_o) per receiver antenna. The users' transmit powers are scaled and normalized to transmit at unit energy and the multipath channel has also been normalized to have, on average, unity gain on the subcarrier frequency responses.

To observe the effects of co-channel (i.e. inter-user) interference, we have simulated cases of the overloading factor set to $Q = 2$ and $Q = 4$. To observe the effects of spatial diversity, we have simulated scenarios with the number of receiver antennas set to $N_r = 2$ and $N_r = 4$. The antenna scenarios have been simulated for high correlation ($\rho = 0.75$), moderate correlation ($\rho = 0.5$) and low correlation ($\rho = 0.25$) coefficient values.
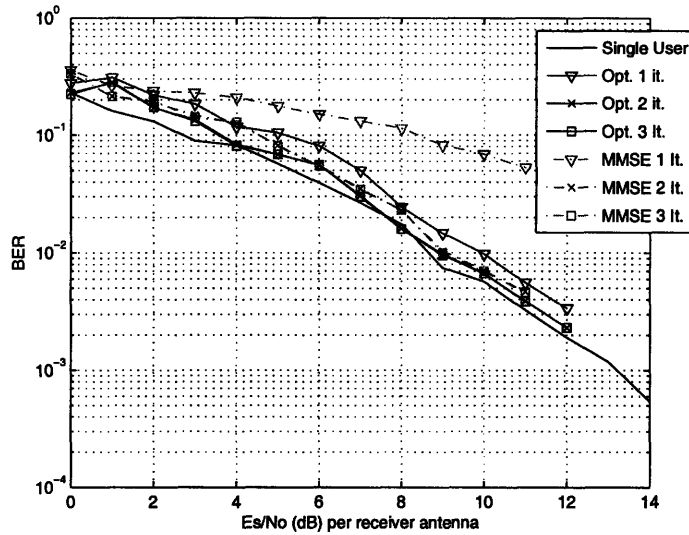
Figure 7-1: BER vs E_s/N_o curves for 2 Rx antennas, 2 users, $\rho = 0.25$.

7.1.1 Case 2x2 (2 Receiver Antennas, 2 overloaded users)

As a note for this simulation case and the following cases, the sample points of the collected results show a certain amount of variance especially at the lower ranges of E_s/N_o. This is a result of the simulation conditions we have imposed and the relatively small number of transmitted frames required to achieve the target frame error rate for termination. The high variance reflects the range of performance demonstrated by the iterative algorithms in randomly generated multipath MIMO channels at those E_s/N_o values.

Low Correlation ($\rho = 0.25$)

Figure 7-1 presents the simulated BER versus E_s/N_o performance comparison between the two iterative detection algorithms, optimal MAP symbol detection and MMSE filtering with soft interference cancellation, aggregated over the two users, with respect to single user performance in a 2 antenna receiver channel with a low correlation coefficient ($\rho = 0.25$).

The performance of the optimal MAP and MMSE algorithms after one iteration is essentially the performance of the respective non-iterative receivers. The MMSE
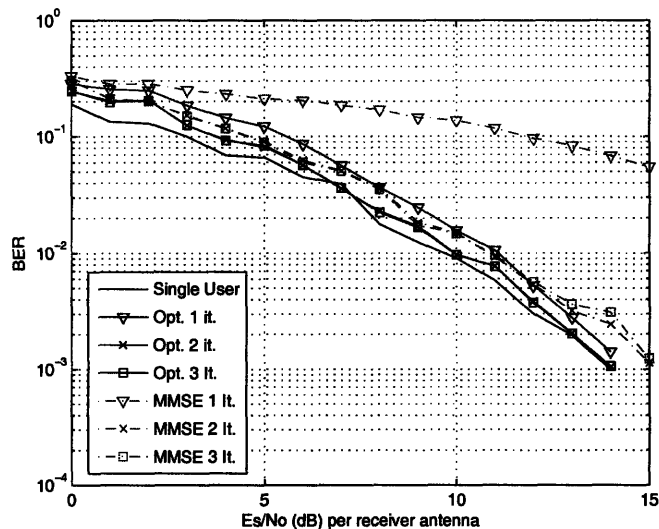
Figure 7-2: BER vs E_s/N_o curves for 2 Rx antennas, 2 users, $\rho = 0.5$.

algorithm performs poorly compared to the optimal MAP algorithm and its curve has a much slower rate of descent with increasing E_s/N_o. The optimal MAP algorithm performs much better in the first iteration and is on average, about 1 dB worse than the single user performance. After additional iterations however, both the optimal MAP and MMSE algorithms approach single user performance very closely. The soft-interference cancellation of MMSE algorithm provides larger gains with increasing E_s/N_o. In both scenarios, it can be seen that most of the iterative gain is exploited by the second iteration, with the third iteration providing only marginally small benefit. An important note is that the low-complexity MMSE filter approaches the performance of the much higher complexity optimal MAP algorithm with increasing of E_s/N_o even after only two iterations.

Moderate Correlation ($\rho = 0.5$)

Figure 7-2 presents the simulated BER versus E_s/N_o performance comparison between the two iterative detection algorithms, with respect to single user performance in a 2 antenna receiver channel with a moderate correlation coefficient ($\rho = 0.5$).

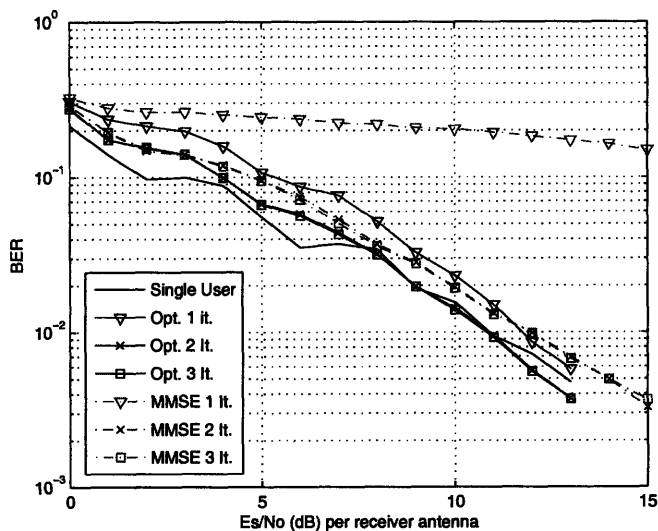The performance of the algorithms at $\rho = 0.50$ closely mirrors those of $\rho = 0.25$,

Figure 7-3: BER vs E_s/N_o curves for 2 Rx antennas, 2 users, $\rho = 0.75$.

the higher correlation of the former causing about a 1 dB loss in performance to each of the scenarios. The increased correlation between antennas leads to a decrease in the available spatial diversity of the channel and hence lower average performance.

High Correlation ($\rho = 0.75$)

Figure 7-3 presents the simulated BER versus E_s/N_o performance comparison between the two iterative detection algorithms, with respect to single user performance in a 2 antenna receiver channel with a moderate correlation coefficient ($\rho = 0.75$).

As expected, the increase in channel correlation and the decrease in spatial diversity causes an additional degradation in performance. Also, due to the high channel correlation, we see the slope of the MMSE BER curve start to taper off with higher E_s/N_o. However, the important point again for all three correlation levels is that performance disparities between the high complexity MAP detection algorithm and the low complexity MMSE algorithm diminish even with only one additional iteration of the detection/decoding loop.
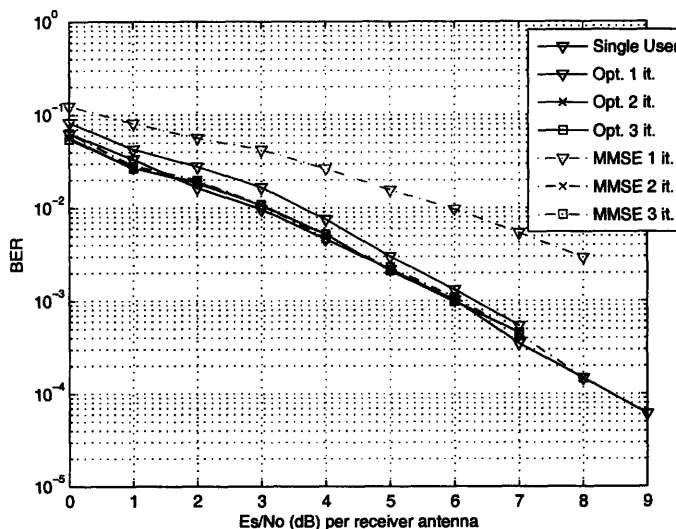
Figure 7-4: BER vs E_s/N_o curves for 4 Rx antennas, 2 users, $\rho = 0.25$.

7.1.2 Case 4x2 (4 Receiver Antennas, 2 overloaded users)

Low Correlation $(\rho = 0.25)$

Figure 7-4 presents the simulated BER versus E_s/N_o performance comparison between the two iterative detection algorithms, aggregated over the 2 users, with respect to single user performance in a 4 antenna receiver channel with a low correlation coefficient $(\rho = 0.25)$.

Compared to the 2x2 case, we have increased the degrees of freedom and spatial diversity within the channel with the two additional antennas without increasing the overloading factor. Understandably, performance for the $\rho = 0.25$ case is much improved over the corresponding cases in the 2 receiver antenna setup. The optimal MAP and MMSE algorithms approach the single user performance limit more closely with iterations than the 2 antenna case, as the effect of specific poor channel realizations can be averaged out by antenna diversity. We see again that most of the performance gains from the turbo loop are harnessed after 2 iterations.
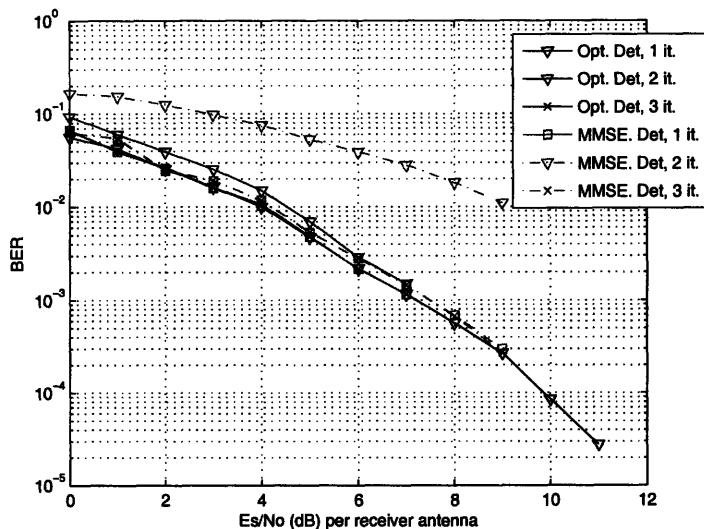
Figure 7-5: BER vs E_s/N_o curves for 4 Rx antennas, 2 users, $\rho = 0.5$.

Moderate Correlation ($\rho = 0.5$)

Figure 7-5 presents the simulated BER versus E_s/N_o performance comparison between the two iterative detection algorithms, with respect to single user performance in a 4 antenna receiver channel with a moderate correlation coefficient ($\rho = 0.5$).

The increase in correlation causes approximately a 1 dB loss in performance in all cases compared to the low correlation case in the 4 receiver antenna setup. However, the performance still approaches that of the single user scenario with increasing E_s/N_o.

High Correlation ($\rho = 0.75$)

Figure 7-6 presents the simulated BER versus E_s/N_o performance comparison between the two iterative detection algorithms, with respect to single user performance in a 4 antenna receiver channel with a high correlation coefficient ($\rho = 0.75$).

In addition to a shift in the curves (corresponding to a performance loss), the increased correlation causes both algorithms to fail to match the single user performance with successive iterations. The correlation of the receiver antennas is likely to be causing an amount of unresolvability in the interfering signals of the users that
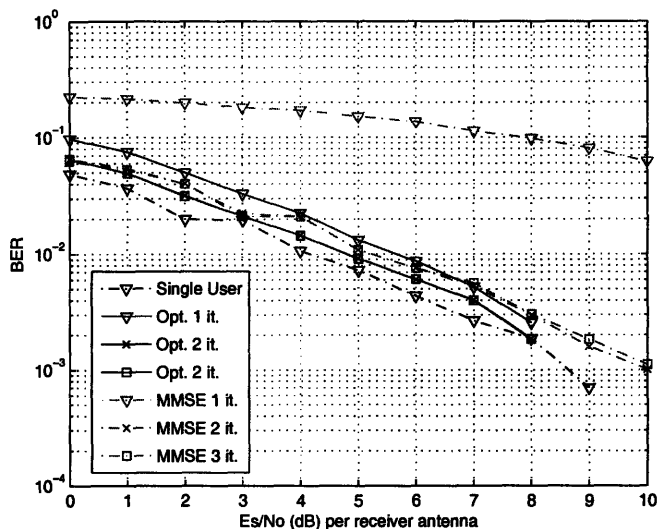
72

Figure 7-6: BER vs E_s/N_o curves for 4 Rx antennas, 2 users, $\rho = 0.75$.

limits the improvement from iteration of the turbo loop. Similar to the 2 receiver antenna case, the high correlation causes the MMSE algorithm curve to taper off from the optimal MAP and single user curves at higher E_s/N_o.

7.1.3   Case 4x4 (4 Receiver Antennas, 4 overloaded users)

Low Correlation $(\rho = 0.25)$

The optimal MAP algorithm for the 4 user scenario is exceedingly complex and is infeasible in realistic implementation because of the exponential growth in complexity with the number of users. Hence we will restrict our focus to the reduced complexity MMSE algorithm with soft interference cancellation. Figure 7-7 presents the simulated BER versus E_s/N_o performance of the MMSE filtering with soft interference cancellation algorithm, aggregated over the 4 users, with respect to single user performance in a 4 antenna receiver channel with a low correlation coefficient ($\rho = 0.25$).

We first notice that, in contrast to the 2 user cases, there is additional performance gain in the third iteration of the turbo loop. Because of the higher overloading factor, we are expecting more overall residual interference at every iteration stage
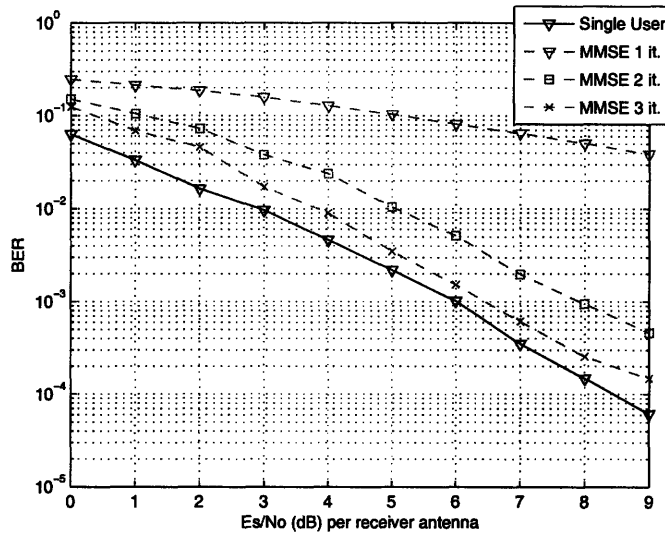
Figure 7-7: BER vs E_s/N_o curves for 4 Rx antennas, 4 users, $\rho = 0.25$.

when compared to the 2x2 or 4x2 cases. The third iteration will allow the receiver to additionally refine its symbol estimates from the second iteration and hence compensate for the additional two users (and cancel their interference contributions). We see that at on average, the low-complexity MMSE approach with soft interference cancellation will perform within 1 dB of the single user scenario after 3 iterations.

Moderate Correlation ($\rho = 0.5$)

Figure 7-8 presents the simulated BER versus E_s/N_o performance comparison between the two iterative detection algorithms, with respect to single user performance in a 4 antenna receiver channel with a moderate correlation coefficient ($\rho = 0.5$).

An increase in receiver antenna correlation in the 4 user case will significantly degrade the performance of the MMSE algorithm. The achievable BERs after multiple iterations remains, on average, 3 dB worse for low to moderate E_s/N_o. The performance loss will increase with higher E_s/N_o, as the curves can be seen to taper and diverge from the single user curve. Additionally, successive iterations will only yield marginal improvements at this level of moderate channel correlation.
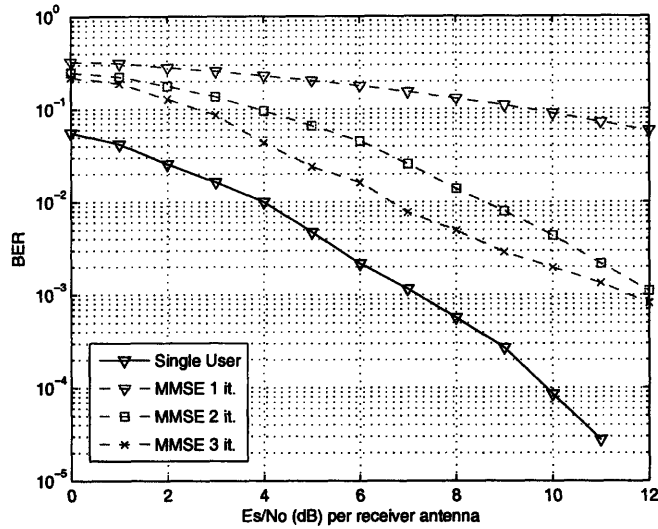
Figure 7-8: BER vs E_s/N_o curves for 4 Rx antennas, 4 users, $\rho = 0.5$.

High Correlation ($\rho = 0.75$)

Figure 7-9 presents the simulated BER versus E_s/N_o performance comparison between the two iterative detection algorithms, with respect to single user performance in a 4 antenna receiver channel with a moderate correlation coefficient ($\rho = 0.75$).

With high receiver antenna correlation, the performance degrades significantly from the moderate and low correlation scenarios. The benefits of additional iterations of the turbo loop also diminish more rapidly than the moderate and low correlation scenarios. Similar to the previous cases, the additional antenna correlation is most likely hindering the ability of the optimum combining MMSE filter to resolve and estimate the interfering users' contributions and hence its ability to improve performance through iterative interference cancellation.

7.2 Discussion

The results of numerical simulation clearly show the performance benefits of iterative detection and decoding in low to moderate receiver antenna correlation scenarios. The iterative processing does provide gain even in the highly correlated receiver antenna
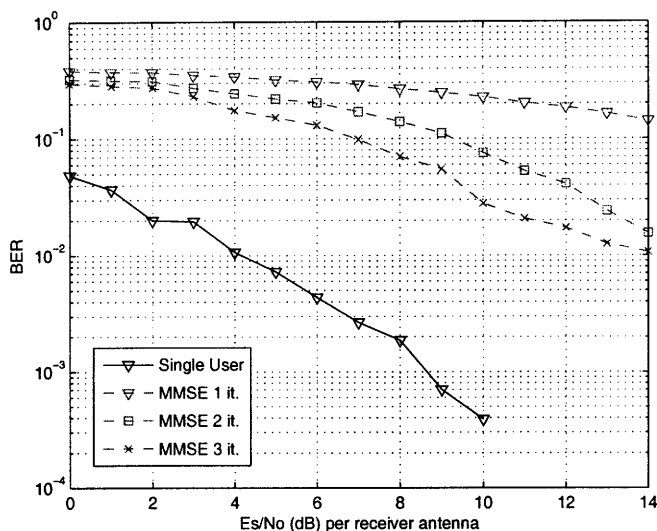
75

Figure 7-9: BER vs E_s/N_o curves for 4 Rx antennas, 4 users, $\rho = 0.75$.

cases but not in the drastic manner of the low and moderate cases. The iterative turbo loop provides a feasible approach to mitigating cochannel interference when the non-iterative algorithms are either too complex or cannot achieve the desired error performance.

From the results we can also see the tradeoff of complexity with performance between the optimal MAP symbol detector and the MMSE filtering approach. In the 2 user scenarios, the MMSE algorithms were able to approach non-iterative optimal MAP algorithm performance after only 2 iterations (for $\rho = 0.25$) while taking 3 iterations in the 4 case (for $\rho = 0.25$). Additionally, at higher E_s/N_o, the iterative optimal MAP and iterative MMSE algorithms were able to achieve very close to single user performance. The performance of the MMSE algorithm in the 4 user, 4 receiver antenna shows to be promising for low to moderate values of ρ. Implementation of the optimal MAP algorithm in this scenario is unrealistic and hence a reduced complexity algorithm such as the presented MMSE algorithm must be used if the channel is to support 4 overloaded users.

The original motivation behind developing these iterative detection and decoding algorithms was to provide a feasible receiver implementation for overloaded OFDMA

systems. As we can see from the results, additional receiver antennas provide the means (i.e. spatial diversity and degrees of freedom) to support multiple independent users on a MIMO channel when coupled with iterative receiver processing. The 4 antenna, 2 user case provides gain over the 2 antenna, 2 user case due to the additional spatial diversity provided by the 2 additional antennas. It is also clear that additional antennas provide better performance in highly correlated scenarios, as it provides more opportunities for antenna diversity in our exponential correlation model. One possible approach to mitigating the poor performance of the highly correlated 4 antenna, 4 user case would be to increase the number of receiver antennas and use an antenna selection algorithm for reduced-complexity optimum combining (i.e. MMSE filtering), which instead would come at the cost of additional hardware complexity [42].

It should be noted that the complexity in the iterative detection and decoding is shared between the detection algorithms we have explored and the remainder of the decoding block, including the mapping, interleaving, and SISO decoding processes. Hence the additional cost of each turbo loop iteration must be considered when deciding to implement either the non-iterative or iterative optimal MAP and MMSE algorithms in practice.

7.3   Topics for Future Study

7.3.1   Imperfect Channel Estimation

Our system framework and simulation parameters were designed with the assumption of perfect channel estimation of the users' channels at the base station receiver. Oftentimes, this is an overly optimistic assumption and scenarios of channel estimation error should be explored in depth. Because of the complex interdependencies of the constituent algorithms the iterative processing loop, it is difficult to generalize the effects of channel estimation error on the overall error performance. It could be likely that the performance and efficiency of the channel estimation algorithm could put

practical bounds or limitations on the marginal benefits of iteration.

7.3.2 Unequal Power Users for Soft Handoff Scenarios

The optimal MAP and MMSE algorithms currently assume equal received power from the overloaded users of interest. This is a reasonable assumption to make in a tightly power controlled cellular system. In this case, the users will receive commands from the base station for adaptively adjusting its transmitted power to the base station. We are operating under the design assumption that the base station will control the powers to receive equal power levels from all users within its cell.

However, there may be certain scenarios where this is not desirable or possible. One such case is for the implementation of uplink soft handoff between two or more base stations receivers in a multiple-cell network. The overloaded users within a data frame being processed at a given base station receiver may consist of users transmitting from other neighboring cells. These other-cell users will also be in active communication with the base station of its primary cell and hence will be power-controlled to the assigned power level of that cell. Due to path loss and shadowing of the transmitted signal through the wireless channel, there is likely to be unequal levels of received power from the interfering users seen at the given base station and the formulation of the algorithm parameters must be re-formulated to take into account the relative power levels between the users. Previous works have shown that unequal received power levels can actually be beneficial for overall aggregate performance [12].

7.3.3 Hard/Soft Combination Interference Cancellation

At the end of each iteration, the SISO decoder will produce APP LLRs on both the information bits and the code bits. The decoder can make temporary hard decisions on both the APP LLRs of the information bits and the code bits. If a frame error has not occurred, the hard estimate of the information bits should exactly produce the the hard estimate of the code bits when passed through convolutional encoding.

We can use the fact that codespaces of convolutional codes are designed for maxi-

mal Hamming distances between codeword sequences to make the approximation that if the re-encoded estimates of the information bits equal the current estimates of the code bits, the probability of the information bit estimate containing errors is small. This is because the probability of codeword mismatches are very small in properly designed decoders. We can be leverage this to trigger early termination of the detection and decoding processes of the users who meet this criterion in early iterations and save computational resources. Additionally, once a codeword match has been found through the described information, the prior symbol information being sent back to the detection block can be approximated as perfect "prior" symbol probabilities for that user. This procedure of hard interference cancellation in conjunction with the soft interference cancellation may or may not provide reliable performance while reducing the total iterations required by the detection and decoding loop.

7.4 Concluding Remarks

The scope of this thesis has been the formulation of an iterative multiple antenna receiver framework based on optimal MAP and reduced complexity MMSE multiuser detection algorithms. We explored the implementation issues associated with designing such a system in an overloaded cellular OFDMA framework and developed a simplified but realistic model for numerical simulation and performance verification. The benefits of OFDMA are that it eliminates the need for complex time-domain equalization procedures and that detection across subcarriers can be performed independently and in parallel. Through Monte Carlo simulation techniques, we have observed and analyzed the performance of the developed algorithms in the context of correlated multipath fading MIMO channels.

From the error performance results, we have drawn conclusions regarding the complexity tradeoffs of the iterative approach and the effectiveness of the iterative detection algorithms in supporting overloaded OFDMA systems for a range of different operating conditions. In SIMO/MIMO channels with low to moderate correlation between receiver antennas, the results show that the performance of the presented

iterative detection algorithms approach single user performance. In all cases, the iterative algorithms provide significant gain over the non-iterative approaches to detection and decoding. The low-complexity iterative MMSE algorithm with interference cancellation, when combined with adequate spatial diversity, provides a feasible solution to overloading standard OFDMA systems.

Bibliography

[1] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[2] J. H. Winters, "On the capacity of radio communication systems with diversity in a rayleigh fading environment," *IEEE Journal on Selected Areas in Communication*, vol. 5, pp. 871–876, June 1987.

[3] J. H. Winters, J. Salz, and R. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Transactions on Communications*, vol. 42, pp. 1740–1751, 1994.

[4] J. Salz and J. H. Winters, "Effect of fading correlation on adaptive arrays in digital mobile radio," *IEEE Transactions on Vehicular Technology*, vol. 43, no. 4, pp. 1049–1057, November 1994.

[5] P. Vandenameele, L. V. D. Perre, M. G. Engels, B. Gyselinckx, and H. J. D. Man, "A combined ofdm/sdma approach," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 11, pp. 2312–2321, November 2000.

[6] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," Bell Labs Technical Journal, Tech. Rep., 1996.

[7] X. Wang and H. V. Poor, *Wireless Communication Systems: Advanced Techniques for Signal Reception.* Prentice Hall, 2004.

[8] A. M. Tonello, "Asynchronous multicarrier multiple access: Optimal and suboptimal detection and decoding," Bell Labs Technical Journal, Tech. Rep. 7(3), 2003.

[9] J. Hagenauer, "The turbo principle: Tutorial introduction and state of the art," in *Proc. International Symposium on Turbo Codes*, September 1997, pp. pp. 1–11.

[10] H. V. Poor, "Turbo multiuser detection: An overview," in *6th International Symposium on Spread-Spectrum Technology and Application, IEEE*, September 2000, pp. 583–587.

[11] R. Koetter, A. C. Singer, and M. Tuchler, "Turbo equalization," *IEEE Signal Processing Magazine*, vol. 21, no. 1, pp. 67–80, January 2004.

[12] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded cdma," *IEEE Transactions on Communications*, vol. 47, no. 7, pp. 1046–1061, July 1999.

[13] A. Tarable, G. Montorosi, and S. Banedetto, "A linear front end for iterative soft interference cancellation and decoding in coded cdma," *IEEE Transactions on Wireless Communications*, vol. 4, no. No. 2, pp. 507–518, March 2005.

[14] H. E. Gamal and E. Geraniotis, "Iterative multiuser detection for coded cdma signals in awgn and fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. No. 1, pp. 30–41, January 2000.

[15] B. Farhang-Boroujeny, H. Zhu, and Z. Shi, "Markov chain monte carlo algorithms for cdma and mimo communications systems," *IEEE Transactions on Signal Processing*, vol. 54, no. 5, pp. 1896–1909, May 2006.

[16] M. Munster and L. Hanzo, "Co-channel interference cancellation techniques for antenna array assisted multiuser ofdm systems," in *First International Conference on 3G Mobile Communication Technologies*, London, UK, 2000, pp. 256–260.

[17] Y. Li and N. R. Sollenberger, "Adaptive antenna arrays for ofdm systems with cochannel interference," *IEEE Transactions on Communications*, vol. 47, no. 2, pp. 217–229, February 1999.

[18] M. Morelli, "Timing and frequency synchronization for the uplink of an ofdma system," *IEEE Transactions on Communications*, vol. 52, no. 2, pp. 296–306, February 2004.

[19] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, 2nd ed. Prentice Hall, 1999.

[20] G. Strang, *Introduction to Linear Algebra*, 3rd ed. Wellesley Cambridge Press, 2003.

[21] *IEEE C802.20-05/68 QFDD and QTDD: Technology Overview*, IEEE 802.20, October 2005.

[22] M. Chiani, M. Z. Win, A. Zanella, R. K. Mallik, and J. H. Winters, "Bounds and approximations for optimum combining of signals in the presence of multiple cochannel interferers and thermal noise," *IEEE Transactions on Communications*, vol. 51, no. 2, pp. 296–307, February 2003.

[23] M. C. Reed and P. D. Alexander, "Iterative multiuser detection using antenna arrays and fec on multipath channels," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 12, pp. 2082–2089, December 1999.

[24] J. G. Andrews, "Interference cancellation for cellular systems: A contemporary overview," *IEEE Wireless Communications Magazine*, vol. 12, no. 2, pp. 19–29, April 2005.

[25] S. Kaiser and J. Hagenauer, "Multi-carrier cdma with iterative decoding and soft-interference cancellation," in *Global Telecommunications Conference, GLOBECOM '97., IEEE*, vol. 1, 1997, pp. 6–10.

[26] W.-J. Choi, K.-W. Cheong, and J. M. Cioffi, "Iterative soft interference cancellation for multiple antenna systems," in *Wireless Communications and Networking Conference, 2000. WCNC, IEEE*, vol. 1, 2000, pp. 304–309.

[27] M. Witzke, S. Baro, F. Schreckenbach, and J. Hagenauer, "Iterative detection of mimo signals with linear detectors," in *Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers*, vol. 1, 2002, pp. 289–293.

[28] H. V. Trees, *Detection, Estimation, and Modulation Theory, Part I.* Wiley-Interscience, 2001.

[29] D. P. Bertsekas and J. N. Tsitsiklis, *Introduction to Probability.* Athena Scientific, 2002.

[30] S. Verdu, *Multiuser Detection.* Cambridge University Press, 1998.

[31] J. G. D. Forney, "The viterbi algorithm," *Proceedings of the IEEE*, vol. 61, pp. 268– 278, March 1973.

[32] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 284–287, March 1974.

[33] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo codes (1)," in *International Conference on Communications, ICC 93. Geneva. Technical Program, Conference Record, IEEE*, vol. 2, 1993, pp. 1064–1070.

[34] J. Boutros and G. Caire, "Iterative multiuser joint decoding: Unified framework and asymptotic analysis," *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 1772–1793, July 2002.

[35] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," online, February 16 2006.

[36] R. B. Ertel, P. Cardieri, K. W. Sowerby, T. S. Rappaport, and J. H. Reed, "Overview of spatial channel models for antenna array communications systems," *IEEE Personal Communications*, vol. 5, pp. 10–22, February 1998.

[37] K. I. Pedersen, J. B. Andersen, J. P. Kermoal, and P. Mogensen, "A stochastic multiple-input-multiple-output radio channel model for evaluation of space-time coding algorithms," in *IEEE VTS-Fall VTC 2000. 52nd*, vol. 2, 2000, pp. 893–897.

[38] M. Z. Win, G. Chrisikos, and A. Molisch, "Wideband diversity in multipath channels with non-uniform power delay profiles," Mitsubishi Electric Research Labratories, Inc., Tech. Rep., 2003.

[39] *802.20-PD-08. Channel Models Document for IEEE 802.20 MBWA System Simulations*, IEEE 802.20, September 2005.

[40] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatiall correlated mimo rayleigh-fading channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2363–2371, Octorber 2003.

[41] H. Shin, M. Z. Win, J. H. Lee, and M. Chiani, "On the capacity of doubly correlated mimo channels," *IEEE Transactions on Wireless Communications*, 2005, to appear.

[42] A. Molisch and M. Z. Win, "Mimo systems with antenna selection," *IEEE Microwave Magazine*, vol. 5, pp. 46– 56, March 2004.