New MIMO Transmission Schemes with Flexible Diversity-Multiplexing Tradeoff for Wireless Communications

Salim Abdelkareem Alkhawaldeh

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy at

Concordia University

Montreal, Quebec, Canada

September 2005

Your file Votre référence
ISBN: 0-494-09970-4
Our file Notre référence
ISBN: 0-494-09970-4

ABSTRACT

New MIMO Transmission Schemes with Flexible Diversity-Multiplexing Tradeoff for Wireless Communications

Salim Abdelkareem Alkhawaldeh, Ph. D.

Concordia University, 2005

The use of multiple transmit and receive antennas can lead to significant improvement of both the reliability and bandwidth efficiency in wireless communication systems. This has spurred a notable recent thrust area of research and development, namely, the so-called Multiple-Input Multiple-Output (MIMO) technology, that is aimed at making better use of multiple antennas in wireless communications.

One of the important discoveries in MIMO communications is the existence of an optimal tradeoff curve between data rate (bandwidth efficiency) and performance (reliability). This and the growing diversification in data rates and quality of services of wireless communications call for MIMO transmission schemes that can provide various rate-performance tradeoffs. This dissertation is aimed at developing new MIMO transmission schemes with flexible rate-performance tradeoff.

For flat fading channels, a formal approach for the design of a class of general space-time block codes, i.e., linear dispersion codes, is developed. This results in a new multi-layered linear coding scheme. With a common encoding structure, various rate-

performance tradeoffs can be achieved simply by choosing the number of data layers that enter the encoder. Once the data rate is chosen, phase shifts among input symbols can be applied to optimize the performance without loss of mutual information.

For frequency-selective MIMO channels, a new space-time orthogonal frequency division multiplexing (OFDM) modulation scheme is introduced to avoid the complicated code design. This is a significant departure from the conventional MIMO OFDM design. The benefit is that a large number of existing codes optimized for single-input fading channels are also optimal for MIMO channels. In addition, it allows an easy spatial multiplexing scheme to support multiple data layers and hence various data rates.

Last, a pilot-symbol-aided channel estimator is developed for the proposed MIMO OFDM transmission scheme. It is shown that, in terms of performance, the proposed scheme is equivalent to the conventional MMSE channel estimator but with much reduced computational complexity.

This dissertation is dedicated

to my parents,

my brothers and my sisters.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude and appreciation to my supervisors Dr. Yousef R. Shayan and Dr. Xiaofeng Wang for their great guidance, advice, and support. During the period of my study, they always follow up my research and encourage me for more progress and contribution.

I would like to thank the examining committee of this dissertation for their invaluable comments and suggestions.

I wish to express my profound appreciation to my parents, brothers and sisters for their continuous help, support and encouragement.

Finally, I would like to thank all my friends for their encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

1-D	One Dimensional
2-D	Two Dimensional
AWGN	Additive White Gaussian Noise
BLER	Block Error Rate
CP	Cyclic Prefix
D-BLAST	Diagonal Bell Labs Layered Space-Time Architecture
FDFR	Full-Diversity Full-Rate
FER	Frame Error Rate
FFT	Fast Fourier Transform
IFFT	Inverse Fast Fourier Transform
i.i.d	independently identically distributed
IR	Impulse Response
ISI	Inter-Symbol Interference
LD	Linear Dispersion
MIMO	Multiple-Input Multiple-Output
ML	Maximum Likelihood
MMSE	Minimum Mean-Square Error
MSE	Mean-Square Error
OFDM	Orthogonal Frequency Division Multiplexing

SIMO	Single-Input Multiple-Output
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
S/P	Serial-to-Parallel
STBC	Space-Time Block Codes
STC	Space-Time Coding
STC-OFDM	Space-Time-Coded OFDM
ST-OFDM	Space-Time OFDM
STTC	Space-Time Trellis Codes
V-BLAST	Vertical Bell Labs Layered Space–Time Architecture

Chapter 1

Introduction

1.1 Literature Review

A remarkable recent advance in wireless communications is the so-called Multiple-Input Multiple-Output (MIMO) technology. It makes use of multiple transmit and receive antennas to achieve higher data rate and reliability. Since the discovery of the significant gain in capacity of MIMO channels [1, 2], tremendous research and development efforts have been invested in MIMO technology. As a result, numerous MIMO transmission schemes have been developed. Early examples include Space-Time Trellis Codes (STTC) [3] and Space-Time Block Codes (STBC) [4, 5] aimed at improved error performance (reliability) and spatial multiplexing schemes [6, 7] aimed at higher data rate. A common drawback of early MIMO transmission schemes is that they are not flexible in rate-performance tradeoff.

Interestingly, Zheng and Tse have shown that there exists an optimal tradeoff curve

between diversity and multiplexing [8]. Since the diversity gain characterizes the error performance as a function of Signal-to-Noise Ratio (SNR) and multiplexing gain determines the data rate, the multiplexing-diversity tradeoff is a tradeoff between rate and performance. Although it was shown in [8] that Diagonal Bell Labs Layered Space-Time Architecture (D-BLAST) [6] allows optimal multiplexing-diversity tradeoff, it is not clear how to achieve this with structured codes instead of Gaussian random codes. In addition, space-time codes carved from properly constructed lattices achieve optimal multiplexing-diversity tradeoff for flat fading channels [9]. However, the decoding complexity is still prohibitively high even for a moderate number of transmit and receive antennas.

Recently, the need to support multi-rate wireless services has attracted a great deal of research attention to the design of MIMO transmission schemes that provide flexible tradeoff between performance and data rate. Among these, Linear Dispersion (LD) coding [10, 11, 12] has notable advantages.

The advantages of the LD codes, however, come at the cost of difficulties in design. This is primarily because a simultaneous optimization of error performance and rate is intractable. In [10], only the mutual information between transmit and receive signals is maximized. The resultant codes do not necessarily yield optimal performance. In [11], diversity gain is included in the optimization objectives in addition to the mutual information. The resulting designs only yield limited performance improvement unless SNR is very high. This is because the diversity gain as a performance index is misleading, particularly for block codes. The diversity gain alone does not guarantee satisfactory performance due to the existence of very small eigenvalues and hence small

determinants of the difference codeword matrices [12]. To remedy this problem, coding gain was also taken into account in [12] but at the cost of loss in channel capacity. Another problem with the above improved LD codes [11, 12] is that they are optimized only for full-rate transmission and, hence, are not optimal for applications that require high performance but less data rate. There also exist designs based on number theory but only for very limited applications [13, 14, 15].

For frequency-selective MIMO channels, the design of space-time codes becomes even more challenging. A popular design for wideband MIMO channels is to employ several Orthogonal Frequency Division Multiplex (OFDM) modulators, one for each transmit antenna, so as to allow the applications of many existing schemes developed for flat fading channels. The use of OFDM transforms a dispersive wideband channel into a set of flat narrowband channels to avoid the Inter-Symbol Interference (ISI) that otherwise exists in single-carrier transmission over wideband channels. To date, OFDM in conjunction with Space-Time Coding (STC) is widely regarded as the enabling technology for future wireless communications [16-19].

However, it was shown in [20-22] that the space-time codes designed for both slow and fast flat fading channels, in general, cannot achieve the full diversity gain available in frequency-selective channels. Although design criteria have been established for MIMO OFDM systems [20], they are difficult to apply in systematic code design. It is worthy of noting that an approach for the design of trellis codes with full diversity was given in [23]. However, the resulting codes usually have significantly smaller coding gain than that of the corresponding block codes.

In [24], a linear block precoder is introduced before OFDM modulation for

frequency-selective MIMO channels. It was shown that through the design of linear precoding, a certain frequency diversity gain can be ensured for an arbitrary space-time code at the cost of data rate. Although this scheme allows flexible tradeoff between diversity and rate, the tradeoff can be far from optimum because of the limitations in the precoder.

Besides the transmission schemes, another important task in MIMO OFDM systems is channel estimation. However, channel estimation for OFDM systems is not trivial. In fact, it can be a difficult task for frequency-selective MIMO channels because potentially a large number of channel taps are to be estimated.

There are two types of channel parameter estimators: decision-directed and pilot-symbol-aided. The main idea of the decision directed estimation is to use decision feedback of information data to track the channel variation after the initial training [25-31]. These decision-directed schemes suffer from significant degradation in performance when the channel varies fast. To overcome this problem, different one dimensional (1-D) and two dimensional (2-D) interpolation filters have been proposed for pilot-symbol-aided channel parameter estimation for OFDM systems [32-35]. Due to the frequent appearance of pilot symbols in time, pilot-symbol-aided schemes are more capable of tracking the channel time variation compared to the decision-directed parameter estimation techniques. In practice, channel statistics are unknown, therefore, a robust channel estimator based on 2-D filtering was proposed in [35]. A prominent drawback of these filter-based schemes is that the filtering, particularly 2-D filtering, involves tremendous amount of computation.

1.2 Problem Statements and Objectives

As mentioned in the previous section, MIMO diversity and spatial multiplexing approaches lack the flexibility of tradeoff between bandwidth efficiency and reliability. Although there exist schemes that provide various rate-performance tradeoffs such as [10, 11, 12] for flat fading channels and [11] for frequency-selective channels, the tradeoffs are not optimal and are only within a limited range.

For channel estimation, all the existing pilot-symbol-aided schemes are based on 1-D or 2-D interpolation filters. A prominent drawback of these schemes is their high computational complexity, which will severely limit their applications in MIMO channels.

Observing the above problems, our goal in this research is to develop new MIMO transmission schemes with flexible rate-performance tradeoff. Both flat and frequency-selective MIMO channels will be considered. For frequency-selective channels, OFDM will be the choice of primary modulation. Three specific objectives are listed below:

- To develop a formal approach for the design of linear dispersion space-time codes for flat fading MIMO channels. Using this formal approach, linear dispersion codes that are flexible in rate-performance tradeoff are to be designed.
- To investigate the interaction between OFDM modulation and space-frequency coding and to develop a new transmit diversity scheme for frequency-selective MIMO channels. The new scheme shall allow easy code design. In addition, a spatial multiplexing scheme will be developed to take the advantages of the new MIMO OFDM transmission scheme. The diversity and multiplexing schemes together form a

multi-layered MIMO OFDM transmission scheme that allows flexible rate-performance tradeoff.

- To design a robust pilot-symbol-aided channel estimator for the proposed MIMO OFDM transmission scheme with much reduced computational complexity as compared to the conventional filter-based estimators.

1.3 Contributions

In this dissertation, two multi-layered coding schemes are developed for flat and frequency-selective MIMO channels, respectively. In addition, a pilot-symbol-aided channel estimation scheme is proposed for frequency-selective fading channels.

For flat fading MIMO channels, we investigate the design of linear dispersion codes, aiming at flexible encoding schemes that allow various rate-performance tradeoffs under a common coding structure. First, the capacity of linear dispersion codes is studied. It is shown that the maximum multiplexing gain of a linear dispersion code is the number of symbols per channel use of the code (i.e., coding rate in symbols). In addition, conditions for the linear dispersion matrices to achieve various multiplexing gains are established. Based on this, we develop a general multi-layered linear dispersion coding scheme that allows various multiplexing gains simply by choosing a subset of the dispersion matrices from a larger set. When all the dispersion matrices are used, the channel capacity is preserved. Furthermore, phase shifting among input symbols is proposed to maximize both diversity and coding gains without loss of mutual information (i.e., multiplexing gain). This leads to an optimization problem which is simple compared to conventional

schemes.

For frequency-selective MIMO channels with OFDM modulation, we introduce a general multi-layered space-frequency coding scheme. In the heart of this approach is a novel space-time OFDM modulator. The proposed space-frequency OFDM modulator translates a MIMO channel into a Single-Input Multiple-Output (SIMO) channel without the loss of system freedom (the available diversity gain). This translation simplifies code design as compared to that in the conventional MIMO OFDM approach. Instead of more complicated space-time codes, those designed for single-input fading channels can be used with the proposed space-time modulation. For bandwidth-efficient applications, a channel multiplexing scheme is developed to work with the proposed space-time modulator. Unlike the conventional spatial multiplexing schemes, an arbitrary number of data streams can be created and each layer occupies all the transmit antennas all the time. As a result, all the available degrees of freedom are preserved for each layer and a full range of optimal tradeoff between data rate and reliability is possible.

Last, we introduce a pilot-symbol-aided channel estimator for Single-Input Single-Output SISO or SIMO OFDM systems. Since in the proposed MIMO OFDM transmission scheme, a MIMO channel is transformed into a SIMO channel, the proposed channel estimator can be directly used with the proposed MIMO OFDM transmission scheme. The new channel estimator is highly robust to time variation of wireless channels. In addition, it is equivalent to the conventional 2-D Minimum Mean-Square Error (MMSE) channel estimator but with much reduced computational complexity. The reduction in complexity is achieved by employing the 2-D Inverse Fast Fourier Transform (IFFT), 2-D Fast Fourier Transform (FFT) and a 2-D weighting function

instead of a 2-D filter. The weighting function is derived based on the Mean-Square Error (MSE) criterion and is simple to implement. For cases where channel statistics are not available, we propose a robust estimator based on a simple 2-D windowing function. Furthermore, we propose an enhanced channel estimator that can further improve the performance of the robust estimator.

1.4 Organization

This dissertation is organized as follows. In Chapter 2, we discuss preliminaries that will be useful for the subsequent development of the dissertation. In Chapter 3, we propose a general multi-layered linear dispersion coding scheme for flat fading MIMO channels which provides flexible rate-performance tradeoff. In Chapter 4, we introduce a new transmit diversity scheme and discuss the design of code and interleaver for MIMO OFDM systems. A spatial multiplexing scheme is also developed, which, in conjunction with the proposed diversity scheme, provide various rate-performance tradeoffs over frequency-selective fading MIMO channels. In Chapter 5, we propose a pilot-symbol-aided channel estimator for OFDM wireless communication systems with much reduced complexity as compared to the conventional filter based approaches. Conclusions and future works are presented in Chapter 6.

Chapter 2

Preliminaries

In this chapter, some preliminaries pertinent to the subsequent development of the dissertation are introduced. These include MIMO channel capacity, transmit diversity via space-time coding, spatial multiplexing, and multiplexing-diversity tradeoff.

2.1 Capacity of MIMO Channels

Consider a MIMO fading channel between N_t transmit and N_r receive antennas as shown in Figure 2.1. The path gain of the l^{th} resolvable path from transmit antenna n to receive antenna m is denoted by $h_{m,n}(l)$ with $0 \le l \le L-1$ and L as the number of resolvable paths. Unless mentioned otherwise, a block Rayleigh fading channel model will be assumed in this dissertation where the path gains remain constant over the duration of a transmission block and change from block to block [3, 5, 36, 37, 38]. Also, it is assumed

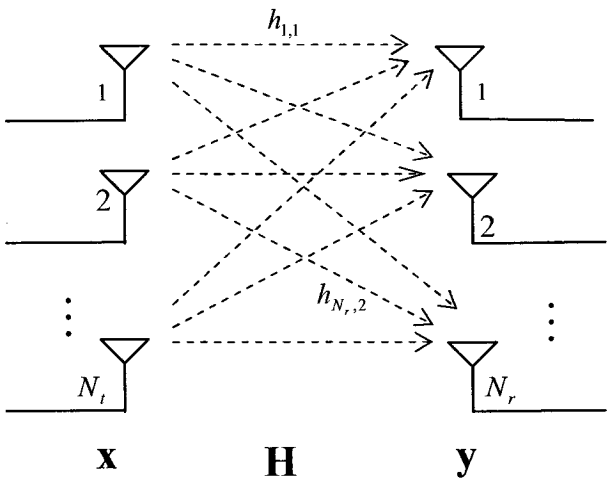

Figure 2.1. A wireless fading channel with N_t transmit and N_r receive antennas.

that these gains are samples of independent complex Gaussian random variables with zero mean and variance 1, *i.e.,*

$$E\left[h_{m,n}(k)h_{s,t}(j)\right]=\begin{cases}0 & (m,n,k)\neq(s,t,j)\\ 1 & (m,n,k)=(s,t,j)\end{cases}.$$

If the number of resolvable paths, L, is equal to 1, then the channel is said to be *non-frequency-selective (flat) fading* or *narrowband,* otherwise, it is called *frequency-selective* or *wideband.*

Denote $s_n(t)$ as the signal transmitted over the nth transmit antenna and $y_m(t)$ as the signal received at the mth receive antenna at time t, then the received vector can be written as

$$\mathbf{y}(t) = \sqrt{\frac{P}{N_t}} \sum_{l=0}^{L-1} \mathbf{H}(l)\mathbf{s}(t-l) + \mathbf{v}(t) \tag{2.1}$$

where $\mathbf{y}(t) = [y_1(t)\, y_2(t) \cdots y_{N_r}(t)]^T$, $\mathbf{s}(t) = [s_1(t)\, s_2(t) \cdots s_{N_t}(t)]^T$, $\mathbf{v}$ is the complex Additive White Gaussian Noise (AWGN) vector with independently identically distributed (i.i.d.) entries of zero mean and variance N_0, i.e., $v_m(t) \sim CN(0, N_0)$, P is the transmit power, and $\mathbf{H}(l)$ is the channel impulse response associated with delay l and is given by

$$\mathbf{H}(l) = \begin{bmatrix} h_{1,1}(l) & \cdots & h_{1,N_t}(l) \\ \vdots & & \vdots \\ h_{N_r,1}(l) & \cdots & h_{N_r,N_t}(l) \end{bmatrix}. \tag{2.2}$$

When the channel is known at the transmitter, its capacity can be readily found [39]. Of particular interest in this work is the case where the channel is unknown at the transmitter. In such a case, the ergodic channel capacity sets an upper bound for a practically achievable data rate.

The ergodic capacity of a flat fading channel with N_t transmit and N_r receive antennas is given by [1, 2]

$$C = \left(\max_{\mathbf{R}_s} E \log_2 \det\left(\mathbf{I}_{N_r} + \frac{P}{N_t N_0} \mathbf{H}\mathbf{R}_s\mathbf{H}^H \right) \right) \tag{2.3}$$

where $\mathbf{R}_s = E\{\mathbf{s}\mathbf{s}^H\}$ is the covariance matrix of the complex Gaussian input vector $\mathbf{s}$ and the expectation (E) is taken with respect to the channel realization $\mathbf{H}$. The above maximization is achieved when the covariance matrix is identity matrix, as such $\mathbf{R}_s = \mathbf{I}_{N_t}$ and the ergodic capacity can be written as

$$C = E\log_2 \det\left(\mathbf{I}_{N_r} + \frac{P}{N_t N_0}\mathbf{H}\mathbf{H}^H\right) \tag{2.4}$$

At high SNR, the above equation can be simplified to

$$C \approx \min(N_t, N_r)\log_2 \frac{P}{N_0}. \tag{2.5}$$

That is, the channel capacity is the minimum of N_t and N_r times of the capacity of a SISO AWGN channel.

In frequency-selective MIMO OFDM systems, the impulse response of the channel between transmit antenna n and receive antenna m can be written as [39]

$$h_{m,n}(t,\tau) = \sum_{l=0}^{L-1} \alpha_{m,n}(t,l)\delta[\tau - \tau_l(t)] \tag{2.6}$$

where $\delta(\tau)$ is the Dirac Delta function, $\alpha_{m,n}(t,l)$ and $\tau_l(t)$ are the complex path gain and time delay of the lth path at time t, respectively. Under the assumption of slow fading, $\alpha_{m,n}(t,l)$ can be treated as constant over the duration of an OFDM frame, but changes independently from frame to frame. The frequency response of the channel between transmit antenna n and receive antenna m at tone k and time t can be expressed as [39]

$$H_{m,n}[t,k] \overset{\Delta}{=} H_{m,n}(tT_s, k\Delta f) = \sum_{l=0}^{L-1} \alpha_{m,n}(t,l)e^{-j2\pi k\Delta f \tau_l} \tag{2.7}$$

where T_s and Δf are the frame duration and the tone spacing, respectively.

The ergodic capacity of each tone is achieved with Gaussian inputs with covariance matrix being identity. Hence the ergodic channel capacity can be written as [39, 40]

$$C = \frac{1}{K}\sum_{k=1}^{K} E_{\Phi_k} \log_2 \det\left(\mathbf{I}_{N_r} + \frac{P}{N_t N_0} \Phi_k \Phi_k^H \right) \tag{2.8}$$

where K is the length of OFDM frame and Φ_k is an $N_r \times N_t$ channel matrix with $H_{m,n}[t,k]$ as its (m, n)-th entry.

2.2 Diversity and Design Criteria for Fixed-Rate Space-Time Codes

Time-varying multipath fading is a fundamental problem in wireless communication systems. To combat this problem, an effective way is to achieve diversity gain by using multiple transmit and receive antennas. When the same signal is transmitted through independent paths between transmit and receive antennas, multiple independently faded replicas of this signal will be provided at the receiver. This refers to so called *spatial diversity* [3-5, 41-44] in which more reliable reception can be achieved. The diversity gain is the slope of the probability of error curve at high SNR. Figure 2.2 represents the block diagram of a MIMO diversity communication system.

One way to exploit spatial diversity is to apply coding across time and space. Such a technique of creating redundancy in the transmitted signal is called space-time coding. Unlike in the 1-D communication systems, space-time modulator maps a sequence of bits $b_1\ b_2\ \cdots\ b_{N_b}$ to a sequence of $N_t \times 1$ vector symbols $\mathbf{s}(t) = [s_1(t)\, s_2(t) \cdots s_{N_t}(t)]^T$, $t = 1, 2, \cdots, T$. Then, the transmission rate is $R = \frac{N_b}{T}$ bits/s/Hz .

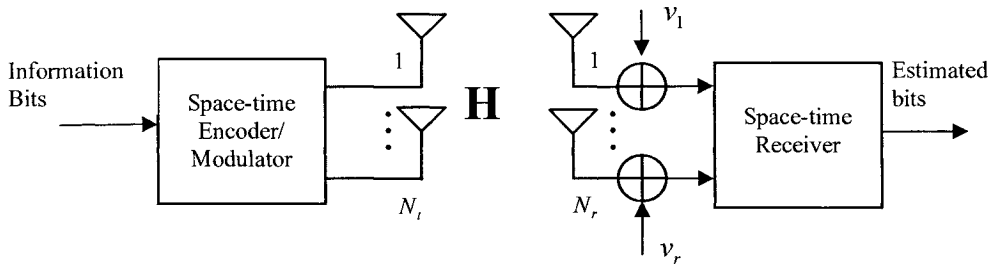

Figure 2.2. Block diagram of a MIMO diversity communication system.

Design criteria of space-time codes over flat fading channels can be obtained by considering the pair-wise error probability. Assuming Maximum Likelihood (ML) decoding and a constant channel during a transmission block (i.e., a block fading channel as considered in this dissertation), the pair-wise error probability of codeword $\mathbf{X}$ being transmitted but the decision being erroneously made in favor of $\hat{\mathbf{X}}$ is [3]

$$P(\mathbf{X} \to \hat{\mathbf{X}}) \le \left(\prod_{n=1}^{N_t} \left(1 + \lambda_n \frac{P}{N_t N_0} \right) \right)^{-N_r} \tag{2.9}$$

where λ_n is the nth eigenvalue of the matrix $(\mathbf{X} - \hat{\mathbf{X}})(\mathbf{X} - \hat{\mathbf{X}})^H$. When $\frac{P}{N_0} >> 1$, the pair-wise error probability can be approximated as

$$P(\mathbf{X} \to \hat{\mathbf{X}}) \le \left(\prod_{n=1}^{r} \lambda_n \right)^{-N_r} \left(\frac{P}{N_t N_0} \right)^{-rN_r} \tag{2.10}$$

where r is the rank of matrix $(\mathbf{X} - \hat{\mathbf{X}})(\mathbf{X} - \hat{\mathbf{X}})^H$. In SISO uncoded system, the pair-wise error probability can be approximated as

$$P(\mathbf{X} \to \hat{\mathbf{X}}) \le \left(\frac{P}{N_t N_0} \right)^{-1} . \tag{2.11}$$

Comparing (2.10) and (2.11), it can be seen that a diversity gain of rN_r and coding gain of $\left(\prod_{n=1}^{r} \lambda_n \right)^{\frac{1}{r}}$ are achieved. Hence, two design criteria for block flat fading channels with independent gains were derived in [3]:

- *The Rank Criterion:* In order to achieve the maximum diversity gain $N_t N_r$, the matrix $(\mathbf{X} - \hat{\mathbf{X}})(\mathbf{X} - \hat{\mathbf{X}})^H$ has to be full rank for any codewords' pair. If $(\mathbf{X} - \hat{\mathbf{X}})(\mathbf{X} - \hat{\mathbf{X}})^H$ has minimum rank *r*, then a diversity of *rm* is achieved.

- *The Determinant Criterion:* For diversity gain of *rm*, the coding advantage is the minimum of the *r*th roots of the sum of determinants of all $r \times r$ principal cofactors of $(\mathbf{X} - \hat{\mathbf{X}})(\mathbf{X} - \hat{\mathbf{X}})^H$ taken over all codewords pairs. For diversity gain of $N_t N_r$, the minimum of the determinant of $(\mathbf{X} - \hat{\mathbf{X}})(\mathbf{X} - \hat{\mathbf{X}})^H$ taken over all codewords pairs must be maximized.

It is worthy of noting that the rank and determinant criteria developed for channels with independent fade coefficients are also applicable for spatially correlated channels.

The above design criteria were for block flat fading channels. Another representitive extreme case is when the fading is sufficiently fast such that the channel coefficients over two successive symbol durations are mutually independent. For such a fast fading channel, denote $V(\mathbf{X}, \mathbf{X}')$ as the set of time instances $1 \le t \le T$ such that $\mathbf{x}(t)$ and $\mathbf{x}'(t)$ are different where $\mathbf{x}(t)$ denotes the transmitted symbols at time *t*, i.e., $\mathbf{x}(t) = [x_1(t)\, x_2(t) \cdots x_{N_t}(t)]^T$

and let $|V(\mathbf{X},\mathbf{X}')|$ denote the number of elements of $V(\mathbf{X},\mathbf{X}')$. Then, the pair-wise error probability is given by [3]

$$P(\mathbf{X}\to\hat{\mathbf{X}}) \le \prod_{t\in v(\mathbf{X},\mathbf{X}')}\left(\left\|\mathbf{x}(t)-\mathbf{x}'(t)\right\|^2 \frac{P}{N_t N_0}\right)^{-N_r}. \tag{2.12}$$

Hence, two design criteria for fast fading channels were derived in [3]:

- *The Distance Criterion:* To achieve the diversity advantage *vm* for any two codewords $\mathbf{X}$ and $\mathbf{X}'$, $\mathbf{x}(t)$ and $\mathbf{x}'(t)$ must be different at least for v values during T symbol durations.
- *The Product Criterion:* To maximize the coding advantage, we need to maximize the minimum products $\prod_{t\in v(\mathbf{X},\mathbf{X}')}\left\|\mathbf{x}(t)-\mathbf{x}'(t)\right\|^2$ taken over codewords $\mathbf{X}$ and $\mathbf{X}'$.

The above criteria can be readily used to derive systematic design procedure for trellis space-time codes with regular structure. However, they are not easy to use in the design of block codes. This is because the evaluation of diversity gain and coding gain over a large number of codewords pairs for a reasonable block length is computationally prohibitive. It is also worthy of noting that the design criteria derived using the pair-wise error probability are for applications with a fixed data rate. These criteria do not tell how to efficiently use bandwidth by adapting the data rate in multi-rate wireless services.

2.3 Spatial Multiplexing

In addition to spatial diversity, it is also possible to simultaneously transmit several independent data streams for higher bandwidth efficiency in MIMO channels. The

increase in data rate by the transmission of parallel independent data streams is called spatial multiplexing gain and the corresponding signal processing is called *spatial multiplexing* [6, 7]. Traditional spatial multiplexing schemes were designed for flat fading channels. Recently, spatial multiplexing is combined with OFDM modulation for frequency-selective channels [45, 46]. Figure 2.3 represents the block diagram of a spatial multiplexing communication system.

In spatial multiplexing, independent data streams are spatially multiplexed and subsequently demultiplexed at the receiver. During every symbol period, the space-time mapper/multiplexer maps N encoded data streams onto N_t output streams, one for each transmit antenna. If the output of the space-time mapper at time t is $\mathbf{s} = [s_1\ s_2 \cdots s_{N_t}]^T$, then the received $N_r \times 1$ vector can be written as

$$\mathbf{y} = \sqrt{\frac{P}{N_t}}\mathbf{H}\mathbf{s} + \mathbf{v} \tag{2.13}$$

where $\mathbf{v}$ is an $N_r \times 1$ vector realization of AWGN with i.i.d entries $v_m \sim CN(0, N_0)$.

Bellow we describe D-BLAST [6] as a typical spatial multiplexing scheme. The basic idea in D-BLAST is to create layers (logical channels) in such a way that signals from different layers can be effectively separated from each other. To do so, the input data stream is demultiplexed into N_t streams (layers) with equal rate and each layer is divided into substreams. Each layer is transmitted from different antennas at different time in diagonal fashion as shown in Figure 2.4. As can be seen, the channel is partitioned into N_t logical channels one for each layer. Each layer occupies the two-dimensional channel

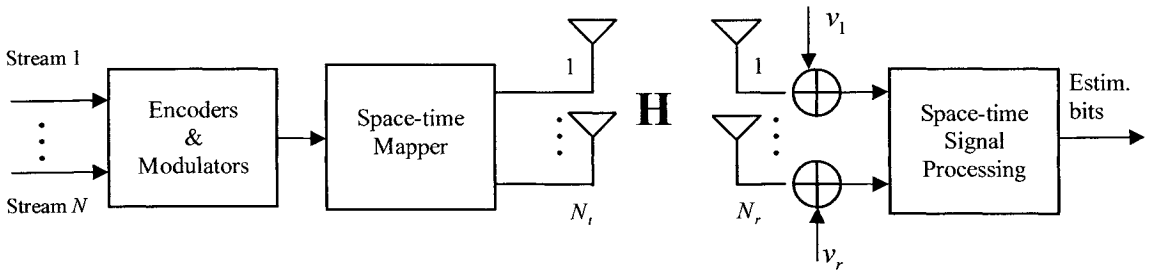

Figure 2.3. Block diagram of a spatial multiplexing communication system.

in a round-robin fashion: if at time slot t, a layer occupies transmit antenna n, then, at time slot t+1, the same layer will occupy transmit antenna (n+1) mod N_t. Outer codes could be used in each layer independently and each diagonal constitutes a complete codeword.

At the receiver, the layers can be separated by using the decorrelating or MMSE receiver and decoding is performed diagonal-by-diagonal. For each diagonal layer, interference from lower diagonals is cancelled using the decoded information bits; interference from the upper diagonal is eliminated by projecting the received signal onto the null space of the upper diagonals. Since D-BLAST effectively creates N_t logical channels with equal capacity, uniform energy allocation is optimal.

Although D-BLAST achieves high bandwidth efficiency, it often results in loss of diversity gain and, hence, loss of performance. Another drawback of this signal processing based approach is that it often requires equal or more receive than transmit antennas.

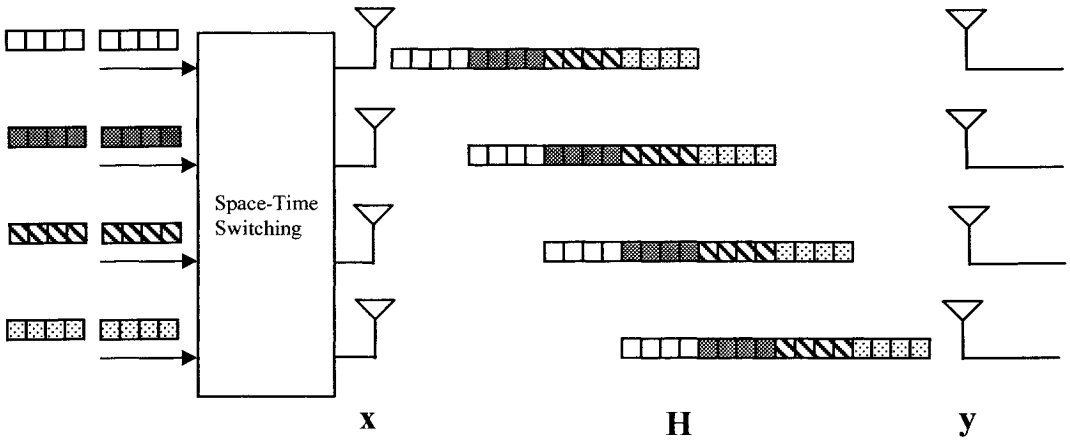

Figure 2.4. Block diagram of D-BLAST.

2.4 Diversity and Multiplexing Tradeoff

As mentioned in the introduction chapter, the maximal available multiplexing and diversity gains cannot be achieved simultaneously and there exists an optimal tradeoff. A MIMO transmission scheme is said to achieve a spatial multiplexing gain G_m if, at high SNR, its data rate is [8]

$$R(\text{SNR}) \approx G_m \log(\text{SNR}) \quad (bps/Hz). \tag{2.14}$$

In other words, the multiplexing gain is defined as

$$G_m = \lim_{\text{SNR}\to\infty} \frac{R(\text{SNR})}{\log(\text{SNR})}. \tag{2.15}$$

It is noted that the data rate R is G_m times the data rate of a SISO channel with the same SNR. This MIMO channel is equivalent to G_m parallel spatial channels where

information symbols can be transmitted independently in parallel through the spatial channels. This result suggests that the number of degrees of freedom is equal to G_m [8].

On the other hand, a MIMO transmission scheme is said to achieve a spatial diversity gain G_d if, at high SNR, the average error probability [8]

$$P_e \approx \mathrm{SNR}^{-G_d}. \tag{2.16}$$

In other words, the diversity gain is defined as

$$G_d = \lim_{\mathrm{SNR}\to\infty} \frac{\log(P_e)}{\log(\mathrm{SNR})}. \tag{2.17}$$

It is important to note that if a MIMO transmission scheme has a nonzero multiplexing gain G_m, the data rate R must increase as SNR increases. Hence, the error probabilities P_e in (2.16) and (2.17) are evaluated at different data rates. It is shown in [8] that there exists an optimal tradeoff between multiplexing and diversity, or, equivalently between data rate and performance. The optimal tradeoff between multiplexing and diversity is a piece-wise line shown in Figure 2.5. Specifically, the maximum diversity gain G_d of a scheme with multiplexing gain $G_m \le \min(N_t, N_r)$ is [8]

$$G_d = (N_t - G_m)(N_r - G_m). \tag{2.18}$$

As can be seen in Figure 2.5, the total number of degrees of freedom provided by the channel is equal to $\min(N_t, N_r)$ and this happens when the optimal tradeoff curve intersects the spatial multiplexing advantage axis, i.e., $G_{m\max} = \min(N_t, N_r)$ and $G_d = 0$. In contrast, the diversity advantage is maximum when the optimal tradeoff curve intersects diversity advantage-axis, i.e., $G_{d\max} = N_t N_r$ and $G_m = 0$.

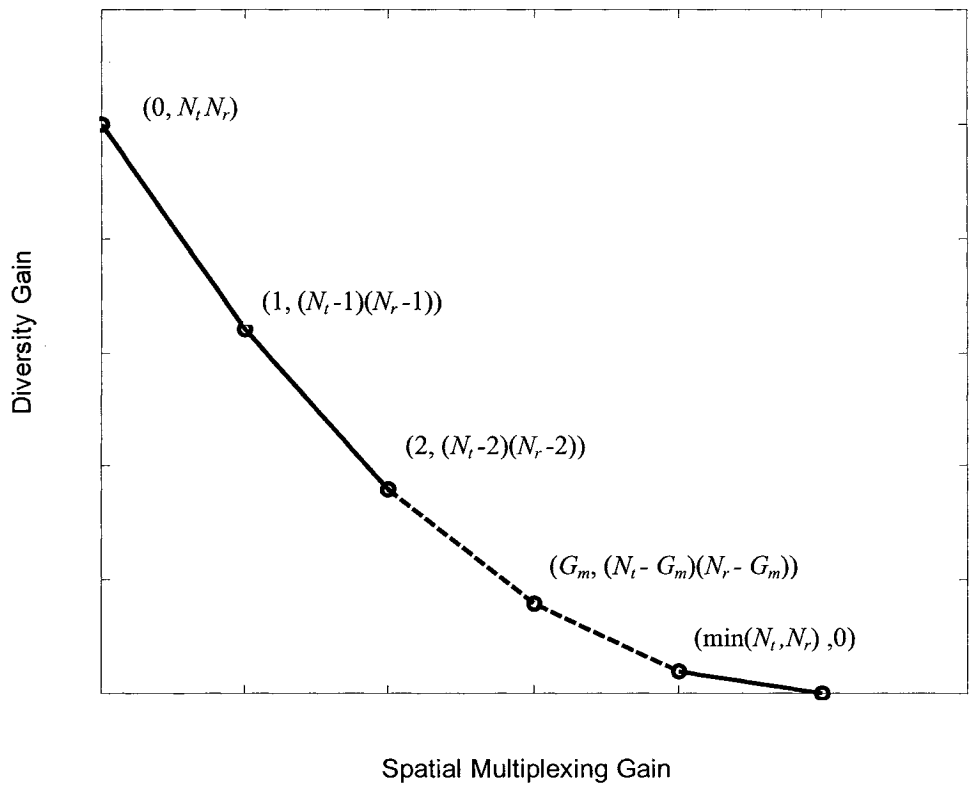

Figure 2.5. Optimal tradeoff curve between the diversity and multiplexing gains [8].

Chapter 3

Multi-Layered Linear Dispersion Codes for Flat Fading MIMO Channels

3.1 Introduction

Recently, MIMO transmission schemes that provide tradeoff between performance and data rate have attracted a great deal of research attention. One of those is LD coding [10]. LD codes have significant advantages compared to conventional space-time block codes. In addition to multiplexing-diversity tradeoff, LD codes subsume Vertical Bell Labs Layered Space–Time architecture V-BLAST [7] and many existing space-time block codes [5] as special cases, and allow suboptimal linear receivers with greatly reduced complexity [10]. Also, they were designed for any number of transmit and receive antennas and have simple encoding process. However, in the design of LD code,

simultaneous optimization of error performance and rate is intractable. In [10], linear dispersion codes were designed to maximize the mutual information between transmitted and received signals. However, they cannot guarantee diversity and coding gains, hence, optimal performance is not guaranteed. In contrast, Full-Diversity Full-Rate (FDFR) codes have been proposed [11] via constellation rotation. Although these codes achieve full diversity, a large portion of their codeword pairs have very small determinants which results in small pair-wise coding gain. As a result, some of the diversity orders cannot be achieved unless the SNR is extremely high. In addition to full diversity, coding gain was considered in [12] but at the cost of loss in channel capacity. Another drawback of the above improved LD codes in [11] and [12] is that they are optimized only for full-rate transmission. Hence, they are not optimal for applications that require high performance but less data rate.

To allow flexible and efficient multiplexing-diversity tradeoff, it is desirable to have a combined spatial multiplexing and coding approach. In this chapter, the design of linear dispersion codes is investigated, aiming at flexible encoding schemes that allow various rate-performance tradeoffs. Based on the mutual information between transmitted and received signals, conditions for the linear dispersion matrices to achieve various multiplexing gains are established. This led us to develop a general multi-layer linear dispersion coding scheme to provide various multiplexing gains simply by choosing a subset of the dispersion matrices from a larger set. Furthermore, without loss of mutual information, both diversity and coding gains are maximized by applying phase shifting among input symbols. Because of the regular structure of the proposed coding scheme (i.e., structured dispersion matrices), the optimization of phase shifts is simple and can be

efficiently carried out by a computer search. This substantially simplifies the selection of constellation and the number of layers (multiplexing gain) in multi-rate applications.

The rest of the chapter is organized as follows. In Section 3.2, the system model is described and a formal approach for the design of linear dispersion codes is developed. Based on this formal approach, a general multi-layered linear dispersion coding scheme that allows various rate-performance tradeoffs is proposed in Section 3.3. In Section 3.4, we compare the proposed multi-layered codes with some existing LD codes. Simulation results are presented in Section 3.5 and conclusions are drawn in Section 3.6.

3.2 Design of Linear Dispersion Codes

In this section, we present the system model and describe our approach to the design of linear dispersion space-time codes.

A MIMO communication system consisting of N_t transmit and N_r receive antennas over a flat fading channel will be considered in this study. The complex gain of the channel between transmit antenna n and receive antenna m is denoted by $h_{m,n}$. It is assumed that the channel gains are samples of independent circularly symmetric complex Gaussian random variables with zero mean and variance 1 (i.e., $h_{m,n} \sim CN(0,1)$), and are perfectly known to the receiver but unknown to the transmitter. The channel is assumed to undergo block fading, i.e., it keeps constant over the duration of a block and varies independently from block to block.

Linear dispersion codes are space-time block codes. With a block of K input symbols, $\mathbf{s} = [s_1\ s_2 \cdots s_K]^T$, an $N_t \times T$ codeword matrix is constructed as [10]

$$\mathbf{S} = \sum_{k=1}^{K} \mathbf{M}_k s_k + \sum_{k=1}^{K} \mathbf{N}_k s_k^* \tag{3.1}$$

where $\mathbf{M}_k, \mathbf{N}_k \in C^{N_t \times T}$ are the dispersion matrices for the kth symbol. Since the use of dispersion matrices for the conjugates of input symbols does not provide evident improvement [10], only the dispersion matrices for the original input symbols will be considered in this study, i.e.,

$$\mathbf{S} = \sum_{k=1}^{K} \mathbf{M}_k s_k \,. \tag{3.2}$$

Within one block, the received signal can be written as

$$\mathbf{Y} = \sqrt{\frac{P}{N_t}} \mathbf{H}\mathbf{S} + \mathbf{V} \tag{3.3}$$

where $\mathbf{H} \in C^{N_r \times N_t}$ is the channel matrix whose (m,n)th entry is $h_{m,n}$, $\mathbf{V} \in C^{N_r \times T}$ is the additive complex Gaussian noise matrix with i.i.d entries, i.e., $v_{m,t} \sim CN(0, N_0)$, and P is the total transmitted power with entries in $\mathbf{S}$ having unit variance.

Let $\mathbf{y}$ be the vector of length $N_r T$ formed by concatenating the T columns of $\mathbf{Y}$, i.e., $\mathbf{y} = \text{vec}(\mathbf{Y})$, and similarly let $\mathbf{v} = \text{vec}(\mathbf{V})$. Then (3.3) can be rewritten as

$$\mathbf{y} = \sqrt{\frac{P}{N_t}} \overline{\mathbf{H}} \mathbf{M}\mathbf{s} + \mathbf{v} = \sqrt{\frac{P}{N_t}} \hat{\mathbf{H}}\mathbf{s} + \mathbf{v} \tag{3.4}$$

where $\mathbf{M} = [\text{vec}(\mathbf{M}_1)\, \text{vec}(\mathbf{M}_2) \cdots \text{vec}(\mathbf{M}_K)]$ with $\text{tr}(\mathbf{M}^H \mathbf{M}) = N_t T$, $\hat{\mathbf{H}} = \overline{\mathbf{H}}\mathbf{M}$, and $\overline{\mathbf{H}} = \mathbf{I}_T \otimes \mathbf{H}$ with $\otimes$ as the kronecker product operator.

Evidently, with the definition in (3.2), a linear dispersion code is completely characterized by its K dispersion matrices, $\{\mathbf{M}_k\}_{k=1}^{K}$. Below, we will refer to a linear dispersion code with K inputs and output matrix of size $N_t \times T$ as a (T, K) LD code. If

the input symbols are drawn from a finite alphabet of size Q, the coding rate of a (T, K) LD code is $\frac{K \log_2 Q}{T}$ bits per channel use. Another useful definition is the *coding rate in symbols*, which is $R = \frac{K}{T}$ for a (T, K) LD code. As can be seen, there exist salient design choices to achieve different coding rate. A well-designed LD encoding scheme shall support a variety of multiplexing-diversity tradeoffs by varying the constellation size Q and/or the coding rate in symbols R. This makes LD codes particularly attractive for future multi-rate wireless communications.

The performance of an LD code depends not only on the set of dispersion matrices but also on the alphabet of the input symbols. This with the large number of possible codewords even for a reasonable value of K makes the design of an LD code difficult. In particular, a direct application of Tarokh's well-known criteria to the design of LD codes is impractical and also inappropriate [10]. Instead, maximizing the mutual information has been used as the primary criterion in the design of LD codes [10]. In this design method, Gaussian input was often assumed in order to make the optimization problem tractable. The resulting LD codes do not necessarily lead to optimal performance for symbols drawn from practical finite alphabets. To alleviate this problem, error probability measures have been used as the design objective. A typical error performance measure is the diversity gain [11]. However, diversity gain can be misleading. Small eigenvalues in the difference matrix of a pair of codewords may not contribute to the error exponent unless the signal-to-noise ratio is very large. Other performance measures including coding gain [12] and union bound [47-49] have been considered. The resulting optimization problems are highly nonlinear and symbol alphabet dependent. The

solutions, therefore, are symbol alphabet and rate dependent and do not share a common structure, which is undesirable for applications that require multiple or adaptive data rates.

In order to design an LD encoding scheme that provides variable data rates and flexible multiplexing-diversity tradeoff, a hybrid approach is needed. First, a *good* encoding structure must be adopted that is able to achieve full channel capacity and, with minimum change in encoding structure, can provide different encoding rates. Second, for every data rate, error performance must be optimized [50].

Below, we establish a theorem that will be useful for the design of a good encoding structure.

Theorem 3.1: For a given coding rate in symbols $R \le N_t$, the expected mutual information between transmitted and received signal $E[I(\mathbf{s};\mathbf{y})]$ is maximized when the dispersion matrices satisfy

$$\mathbf{G}_t \mathbf{G}_j^H = \begin{cases} \mathbf{D} & \text{when } t = j \\ \mathbf{0}_{N_t \times N_t} & \text{otherwise} \end{cases} \tag{3.5}$$

where the expectation E is taken with respect to channel $\mathbf{H}$, $\mathbf{G}_t$ of size $N_t \times RT$, $t = 1, \cdots, T$, consists of rows from $(t-1)N_t + 1$ to tN_t of $\mathbf{M}$, and $\mathbf{D}$ is any diagonal matrix with only R nonzero diagonal entries of value $\frac{N_t}{R}$.

Proof: Consider the maximization of $E[I(\mathbf{s};\mathbf{y})]$ in terms of $\mathbf{M}$. Denote $\mathbf{D}_t$ as the matrix consisting of rows and columns from $(t-1)N_t + 1$ to tN_t of $\mathbf{M}\mathbf{M}^H$, i.e., $\mathbf{D}_t$ is the tth $N_t \times N_t$ diagonal block of $\mathbf{M}\mathbf{M}^H$. Since $\mathbf{M}\mathbf{M}^H$ is positive semi-definite, then from (3.4) the mutual information with Gaussian inputs $\mathbf{s}$ is given by [1, 2]

$$E[I(\mathbf{s};\mathbf{y})] = E\log_2 \det\left(\mathbf{I}_{N_r T} + \frac{P}{N_t N_0}\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right)$$

$$\leq E\log_2 \prod_{t=1}^{T} \det\left(\mathbf{I}_{N_r} + \frac{P}{N_t N_0}\mathbf{H}\mathbf{D}_t\mathbf{H}^H\right)$$

$$= \sum_{t=1}^{T} E\log_2 \det\left(\mathbf{I}_{N_r} + \frac{P}{N_t N_0}\mathbf{H}\mathbf{D}_t\mathbf{H}^H\right) \tag{3.6}$$

The equality in (3.6) holds when all other entries in $\mathbf{M}\mathbf{M}^H$ that are excluded in $\mathbf{D}_t$ for $t = 1, 2, \cdots, T$ are zeros, i.e.,

$$\mathbf{G}_t\mathbf{G}_j^H = \mathbf{0}_{N_t \times N_t} \quad \forall t \neq j. \tag{3.7}$$

We now consider the optimization of each term in the summation of the last equation of (3.6). First, note that the maximum rank of $\mathbf{M}\mathbf{M}^H$ is RT, hence we assume that $\text{rank}(\mathbf{D}_t) = R$. Let $\text{trace}(\mathbf{D}_t) = C_t$. Apparently, $\sum_{t=1}^{T} C_t = \text{trace}(\mathbf{M}\mathbf{M}^H) = N_t T$. Since $\mathbf{D}_t$ is symmetric and positive semi-definite, one can write $\mathbf{D}_t = \mathbf{U}\mathbf{\Lambda}_t\mathbf{U}^H$ with $\mathbf{\Lambda}_t$ being diagonal and $\mathbf{U}$ unitary. Furthermore, $\mathbf{\Lambda}_t$ has R nonzero entries, say they are the i_1th, i_2th, $\cdots$ and i_Rth diagonal entries. Let $\tilde{\mathbf{\Lambda}}_t$ be the $R \times R$ matrix obtained by collecting the rows and columns $i_1, i_2, \cdots, i_R$ of $\mathbf{\Lambda}_t$. Similarly, let $\tilde{\mathbf{U}}_t$ be the $N_t \times R$ matrix that collects the columns $i_1, i_2, \cdots, i_R$ of $\mathbf{U}$. Then, it is clear that

$$E\log_2 \det\left(\mathbf{I}_{N_r} + \frac{P}{N_t N_0}\mathbf{H}\mathbf{D}_t\mathbf{H}^H\right) = E\log_2 \det\left(\mathbf{I}_{N_r} + \frac{P}{N_t N_0}\tilde{\mathbf{H}}\tilde{\mathbf{\Lambda}}_t\tilde{\mathbf{H}}^H\right) \tag{3.8}$$

where $\tilde{\mathbf{H}} = \mathbf{H}\tilde{\mathbf{U}}$. It can be easily checked that the entries of $\tilde{\mathbf{H}}$ are i.i.d complex Gaussian random variables with zero mean and variance 1. Hence, (3.8) is maximized when the

diagonal entries in $\tilde{\mathbf{\Lambda}}_t$ are equal and $\tilde{\mathbf{\Lambda}}_t = \frac{C_t}{R}\mathbf{I}_R$ [1]. Then (3.6) can be rewritten as

$$E[I(\mathbf{s};\mathbf{y})] \le \sum_{t=1}^{T} E \log_2 \det\left(\mathbf{I}_{N_r} + \frac{PC_t}{N_t N_0 R}\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H\right). \tag{3.9}$$

Note that $\log_2 \det\left(\mathbf{I}_{N_r} + x\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H\right)$ is a concave function of x and so is $E \log_2 \det\left(\mathbf{I}_{N_r} + x\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H\right)$. Further, by the fact that $\sum_{t=1}^{T} C_t = N_t T$, (3.9) is maximized when $C_t = N_t$, for $t = 1, 2, \cdots, T$, i.e., $\mathbf{G}_t\mathbf{G}_t^H = \mathbf{D}$ where $\mathbf{D}$ is only diagonal matrix of R nonzero entries of value $\frac{N_t}{R}$. This with (3.7), has proved Theorem 3.1. □

Two corollaries follow from Theorem 3.1.

Corollary 3.1: The maximum achievable multiplexing gain of LD codes with coding rate in symbols $R \le N_t$ is $\min(R, N_r)$.

A coding scheme is said to preserve the channel capacity if for any realization of the channel, say $\mathbf{H}$, the instantaneous channel capacity is $\log_2 \det\left(\mathbf{I}_{N_r} + \frac{P}{N_t N_0}\mathbf{H}\mathbf{H}^H\right)$.

Corollary 3.2: A linear dispersion code preserves the channel capacity if and only if the coding is at least full rate (i.e., $R \ge N_t$ symbols per channel use) and the dispersion matrices satisfy

$$\mathrm{tr}(\mathbf{M}_k{}^H\mathbf{M}_j) = \begin{cases} 1 & \text{when } k = j \\ 0 & \text{otherwise} \end{cases}. \tag{3.10}$$

Proof: Let $R = N_t$ in (3.5). Hence, from Theorem 3.1, the mutual information is maximized when $\mathbf{D} = \mathbf{I}_{N_t}$ and $\mathbf{M}\mathbf{M}^H = \mathbf{I}_{N_t T}$. As such, the instantaneous channel capacity is

$$I(\mathbf{s};\mathbf{y}) = \log_2 \det\left(\mathbf{I}_{N_r} + \frac{P}{N_t N_0} \mathbf{H}\mathbf{H}^H \right). \tag{3.11}$$

This has shown the sufficiency of the conditions. To establish the necessity, assume $R < N_t$ or (3.10) does not hold, then it is easy to show that $I(\mathbf{s};\mathbf{y})$ is smaller than that given by (3.11). Hence, the corollary has been proved. □

We are now ready to make some observations regarding to the construction of linear dispersion matrices.

1. Recall that the nth row of $\mathbf{G}_t$ collects the dispersion weights of all symbols at transmit antenna n and time t, i.e., $\mathbf{G}_t(n,:) = [M_1(n, t)\, M_2(n, t)\, \cdots\, M_K(n, t)]$ with $M_k(n,t)$ as the (n, t)th entry of $\mathbf{M}_k$. Hence we can call the nth row of $\mathbf{G}_t$ as the dispersion vector associated with transmit antenna n at time t. Equation (3.5) specifies that all the dispersion vectors are either zero or mutually orthogonal to each other. If one choose a multiplexing gain $R \le N_t$, then there are exactly R nonzero dispersion vectors at a given time. In other words, only R transmit antennas are occupied at any time.
2. If $\mathbf{M}_k$ and $\mathbf{M}_j$ have nonzero entries at different locations, i.e., $M_k(n,t)M_j(n,t) = 0,\ \forall(n, t)$, then $\mathrm{tr}(\mathbf{M}_k{}^H \mathbf{M}_j) = 0$.
3. For full diversity, each dispersion matrix must have at least N_t nonzero entries at different rows and different columns.

Our objective is to design a set of $N_t T$ dispersion matrices for a coding block of size T. This set of dispersion matrices must satisfy (3.10) to ensure full multiplexing gain $R = N_t$. If the multiplexing gain R is chosen to be less than N_t, one can simply take

RT dispersion matrices from the set that satisfy (3.5). With this objective in mind and based on the above observations, we may construct the dispersion matrices as follows.

Rule 1: The N_tT dispersion matrices are divided into N_t groups, each of which has T matrices. Any two matrices from different groups must have nonzero entries at different locations. This guarantees that the trace of the Hermitian product of any two matrices from different groups be zero, i.e., satisfy (3.10).

Rule 2: Assign each dispersion matrix N_t nonzero entries at different columns and rows. This ensures that a) full diversity is possible, b) any two dispersion vectors associated either with the same transmit antenna or the same transmit time are orthogonal to each other.

Rule 3: Within a group, apply spatial weighting to make sure any two dispersion vectors associated with different times and different antennas are orthogonal. This can be readily done by applying spatial weightings as will be shown in the next section.

3.3 Proposed Multi-Layered Linear Dispersion Codes for Flat Fading MIMO Channels

In this section, the design rules established in the last section will be used to develop a new multi-layered linear dispersion coding scheme.

In the proposed scheme the encoding is performed block by block. Each encoding block takes NT symbols as input, (i.e., $K = NT$), T symbols per layer (group), and output a codeword matrix of size $N_t \times T$, where T is an integer multiple of N_t. Hence, the transmission rate is N symbols per channel use if N layers are used. The dispersion

matrix for the tth symbol of layer i is given by

$$\mathbf{M}_{t,i} = \frac{e^{j\phi_{t \bmod(N_t)}} e^{j\theta_i} \mathbf{W}_t \mathbf{P}_i \mathbf{A}^{\lfloor (t-1)/N_t \rfloor}}{\sqrt{N}}, \quad t = 1, \cdots, T, \quad i = 1, \cdots, N, \tag{3.12}$$

where $\{\phi_t\}_{t=1}^{N_t}$ is a set of phase shifts among symbols in each layer, $\{\theta_i\}_{i=1}^{N}$ is a set of phase shifts among layers, $\mathbf{W}_t$ is an $N_t \times N_t$ diagonal matrix whose n^{th} diagonal entry is $\exp\left(\frac{-j2\pi(n-1)(t-1)}{N_t}\right)$, $\mathbf{P}_i$ is an $N_t \times T$ circulant matrix whose first row has 1 as its ith entry and zeros elsewhere, i.e,

$$\mathbf{P}_i = \begin{bmatrix} \overbrace{0 \ \cdots \ 0}^{i-1} \ 1 \ 0 & & \cdots & & 0 \\ 0 \quad \cdots \quad 0 \ 1 & & & & \\ \vdots & & \ddots \ \ddots & & \vdots \\ 0 \quad \cdots \quad 0 & & & 1 \ 0 \ \cdots \ 0 \end{bmatrix}, \tag{3.13}$$

A is a $T \times T$ circulant matrix given by

$$\mathbf{A} = \begin{bmatrix} \overbrace{0 \ \cdots \ 0}^{N_t} \ 1 & 0 & \cdots & 0 \\ \vdots \ \ddots & \ddots & \ddots & \ddots \ \vdots \\ & & & 0 \\ 0 & 0 & & 1 \\ 1 \ \ddots & & \ddots & 0 \\ 0 \ \ddots & & & \vdots \\ \vdots \ \ddots & & & \\ 0 \ \cdots \ 0 \ 1 & 0 & \cdots & 0 \end{bmatrix}, \tag{3.14}$$

and $\lfloor x \rfloor$ is the greatest integer smaller than or equal to x.

The construction of dispersion matrices specified in (3.12) is done according to the design rules and satisfies (3.5) and (3.10) regardless of the values of phase shifts ϕ_t and

θ_i. Therefore, this will allow a free choice of ϕ_t and θ_i without loss of mutual information. By Corollary 1, for a given number of layers $1 \le N \le N_t$, a multiplexing gain of $\min(N, N_r)$ is possible. In addition, by Corollary 2, if $N = N_t$ layers are used, the proposed construction in (3.12) preserves channel capacity.

It is interesting to note that the proposed scheme is a generalization of the scheme proposed in [11]. The proposed scheme here is developed rigorously based on Theorem 3.1 and applies to any block length of T. Following this development, it is now clear that the number of layers directly determines the multiplexing gain and, hence, data rate. Furthermore, in [11] the phase shifts were selected to ensure full rank of the difference matrix between any two codeword pairs. The resultant codes may not have satisfactory performance due to the existence of near-zero eigenvalues of the difference matrices associated with a large portion of the codeword pairs. In comparison, we present below a different strategy to optimize the performance through the design of phase shifts. Before we proceed to optimize the phase shifts, we note that the codeword matrix for one block is constructed as

$$\mathbf{S} = \sum_{i=1}^{N} \sum_{t=1}^{T} \mathbf{M}_{t,i} s_{t,i} \tag{3.15}$$

where $s_{t,i}$ is the tth symbol of the ith layer. As can be seen from (3.12) and (3.15), the kth symbol of the ith layer will be spatially spread using the spreading vector $\text{diag}(\mathbf{W}_t)$ and phase shifted before it is combined with other $N_t - 1$ symbols of the same layer. Denoting $x_{t,l}$ as the lth entry of the tth diagonal of $\mathbf{S}$, i.e.,

$$\mathbf{S}=\begin{bmatrix} x_{1,1} & x_{2,1} & \cdots & x_{T,1} \\ x_{T,2} & x_{1,2} & \cdots & x_{T-1,2} \\ \vdots & & & \\ x_{T-N_t+2,N_t} & \cdots & x_{T,N_t} & x_{1,N_t} \cdots x_{T-N_t+1,N_t} \end{bmatrix}. \tag{3.16}$$

Then from (3.15) we have

$$x_{i+cN_t,l} = \begin{cases} \dfrac{e^{j\theta_i}(e^{j\phi_1}q_{l1}s_{1+cN_t,i}+e^{j\phi_2}q_{l2}s_{2+cN_t,i}+\cdots+e^{j\phi_{N_t}}q_{lN_t}s_{N_t+cN_t,i})}{\sqrt{N}} & \text{for } 1\le i\le N \\ 0 & \text{otherwise} \end{cases} \tag{3.17}$$

where $l=1,\cdots,N_t$ and $c=0,1,\cdots,\dfrac{T}{N_t}-1$.

We now take two steps to optimize the performance. First, we consider the codeword pairs that are most vulnerable to error. To do this, we note that the average Euclidean distance of a pair of codewords $\mathbf{S}$ and $\mathbf{S}'$ is

$$\begin{aligned} d(\mathbf{S},\mathbf{S}') &= E\left[\text{tr}(\mathbf{H}(\mathbf{S}-\mathbf{S}')(\mathbf{S}-\mathbf{S}')^H\mathbf{H}^H\right] \\ &= N_r\sum_{t=1}^{T}\sum_{l=1}^{N_t}\left|x_{t,l}-x'_{t,l}\right|^2. \end{aligned} \tag{3.18}$$

Hence, we consider two codewords that differ by just 1 entry in $\mathbf{S}$, say $x_{1,1}$ in (3.17). Since $x_{1,1}$ is a linear combination of $s_{t,1}, t=1,\ldots,N_t$ this implies that the codeword pair will also differ at symbols $x_{1,2},x_{1,3},\ldots,x_{1,N_t}$. Then, the root of the determinant of $(\mathbf{S}-\mathbf{S}')(\mathbf{S}-\mathbf{S}')^H$ is the product of N_t distances and it is given by

$$D=\left|\prod_{n=1}^{N_t}d_i\right| \tag{3.19}$$

where

$$d_l = x_{1,l} - x'_{1,l} = \frac{e^{j\theta_i}\left(e^{j\phi_1} q_{l1}\left(s_{1,1} - s'_{1,1}\right) + e^{j\phi_2} q_{l2}\left(s_{2,1} - s'_{2,1}\right) + \cdots + e^{j\phi_{N_t}} q_{lN_t}\left(s_{N_t,1} - s'_{N_t,1}\right)\right)}{\sqrt{N}}. \quad (3.20)$$

Note that, for the codeword pairs that differ at only one diagonal in the codeword matrices, the pairwise coding gain is

$$G_c = D^{\frac{2}{N_t}} \Big/ N_t$$

It is clear that the term $e^{j\theta_i}$ has no effect on (3.19) because $\left|e^{j\theta_i}\right| = 1$ for any value of θ_i. Hence, we only consider the optimization of $\{\phi_t\}_{t=1}^{N_t}$ for this purpose. Note that $\phi_t = \phi_{t+cN_t}$. To optimize the performance of those vulnerable codeword pairs, and hence the associated diversity and coding gains, the phase shifts among symbols $\{\phi_t\}_{t=1}^{N_t}$ are chosen to maximize the minimum determinant D taken over all possible pairs $\hat{\mathbf{s}} = \left(s_{1,1}\ s_{2,1} \cdots s_{N_t,1}\right)$ and $\hat{\mathbf{s}}' = \left(s'_{1,1}\ s'_{2,1} \cdots s'_{N_t,1}\right)$, i.e,

$$\{\phi_t\}_{t=1}^{N_t} = \arg\ \max \min_{\{\hat{\mathbf{s}}, \hat{\mathbf{s}}'\}} D. \quad (3.21)$$

Since the symbols are taken from finite alphabets, the optimal D must be positive. In fact, when the constellation is a complex integer ring, one can readily find phase shifts to ensure a positive D [11]. If only one layer is used, the phase shifts obtained according to (3.21) maximize the diversity and coding gains of the code. In general, they maximize the diversity and coding gains of the codeword pairs that have small Euclidean distance. In addition, we establish the following theorem that shows the bound of the diversity gain of the code.

Theorem 3.2: In the proposed scheme, for a given number of layers $1 \le N \le N_t$ and regardless of the values of $\{\theta_i\}_{i=1}^{N}$, the diversity gain is bounded as

$$G_d \geq (N_t - N + 1)N_r. \tag{3.22}$$

Proof: First, we note the following facts:

a) With phase shifts $\{\phi_{t)}\}_{t=1}^{N_t}$ as defined in (3.21), if there is a nonzero entry in a diagonal of a difference matrix $\mathbf{E} = \mathbf{S} - \mathbf{S}'$, then all the entries at that diagonal must be nonzero.

b) Within any contiguous N_t diagonals of a difference matrix **E**, there are at most N contiguous nonzero diagonals and at least $N_t - N$ contiguous zero diagonals.

By the above two facts, in any difference matrix **E**, one can always find a square upper triangular submatrix of size $N_t - N + 1$ with nonzero diagonal entries. The rank of this submatrix is full. Hence, the rank of **E** is at least $N_t - N + 1$. Then the diversity gain $G_d \geq (N_t - N + 1)N_r$. This has proved Theorem 3.2. □

Note that the maximum multiplexing gain of the MIMO channel is $\min(N_t, N_r)$. Hence, we only consider the number of layers $N \leq \min(N_t, N_r)$. In such a case, Theorem 3.2 shows that the diversity gain of the proposed scheme is usually large except when N_t and N_r are small but N is close to $\min(N_t, N_r)$. It is worthy of noting that the improvement in performance quickly diminishes as the diversity gain increases beyond 4 [3], and in such a case, the Euclidean distance can be a good performance measure [51]. From (3.18) to (3.21), it is clear that the optimal phase shifts from (3.21) will also closely maximize the Euclidean distance. When N_t and N_r are relatively small and N is close to $\min(N_t, N_r)$, the diversity gain is important. In such a case, one can apply phase shifts $\{\theta_i\}_{i=1}^{N}$ to ensure full diversity such as that in [11]. However, the coding gain can be close

to zero, particularly, when the constellation is large. Therefore, we apply phase shifts $\{\theta_i\}_{i=1}^{N}$ among layers to maximize diversity and coding gains for all the codeword pairs when the constellation size is small. This can be done by maximizing the minimum coding gain G_c over all possible codeword pairs, i.e,

$$\{\theta_i\}_{i=1}^{N} = \arg \max \min_{\{\mathbf{s},\mathbf{s}'\}} G_c. \tag{3.23}$$

Equations (3.21) and (3.23) can be solved easily and efficiently by using a computer search and we show some results in Tables 1, 2 and 3. In these tables, it is assumed that $\phi_1 = \theta_1 = 0$ as reference. As shown in Tables 1, 2 and 3, for the same transmitted power and data rate the multi-layer codes provide better minimum determinant D_{min} than the single-layer codes. For example, the two-layer code with QPSK modulation has much larger D_{min} than the one-layer code with 16PSK or 16QAM modulation. Also by using 3 transmit antennas, the three-layer code with BPSK modulation has much larger D_{min} than the one-layer code with 8PSK modulation. This shows that one shall always try to use a larger number layers for a given data rate. Note $D_{\min}^{2/N_t} / N_t$ is the minimum coding gain of the codeword pairs that differ at only one diagonal. Form the tables, it can be seen that the coding gain of the code G_c becomes significantly smaller than $D_{\min}^{2/N_t} / N_t$ when the constellation size is relatively large. In such cases, the effect of phase shifts among layers θ is negligible. Hence, for large constellations, only phase shifts within layers $\{\phi_t\}_{t=1}^{N_t}$ are needed. However, it is worthy of mentioning that a small pair-wise coding gain does not necessarily lead to a large pairwise error probability when the diversity gain is relatively large. Instead, the pairwise Euclidean distance is a better indicator, which is maximized by phase shifts $\{\phi_t\}_{t=1}^{N_t}$.

Table 1. Optimum phase shifts and corresponding D_{min} for one-layer codes.

Modulation	N_t	N	ϕ_2 (degree)	$D_{min}/2$ (G_c)
BPSK	2	1	90	2
QPSK	2	1	45	1
8PSK	2	1	22.5	0.2929
16QAM	2	1	45	0.2

Table 2. Optimum phase shifts and corresponding D_{min} for two-layer codes.

Modulation	N_t	N	ϕ_2(degree)	θ_2(degree)	$D_{min}/2$	G_c
BPSK	2	2	90	90	1	1
QPSK	2	2	45	30,60,120,150	0.5	0.2588
8PSK	2	2	22.5	13, 32, 58, 77	0.1121	0.0312
8QAM	2	2	33.5	38	0.368	0.0534
16QAM	2	2	45	18.5, 26.5, 63.5, 71.5	0.1	0.0416

Table 3. Optimum phase shifts and corresponding D_{min} for one-layer and three-layer codes.

Modulation	N_t	N	ϕ_2	ϕ_3	θ_2	θ_3	$D_{\min}^{3/2}/3$
8PSK	3	1	21	63	N/A	N/A	0.0031
BPSK	3	3	4	128	20	160	0.0370

3.4 Comparison of Linear Dispersion Codes

In this section, we compare the proposed multi-layered codes with some existing LD codes.

3.4.1 Alamouti's Scheme

Consider channels with two transmit antennas, for which the Alamouti's scheme is designed for. If the number of receive antennas is more than 1, then a multiplexing gain of 2 is possible and two layers can be used in our scheme. Suppose $T = 2$ and denote the four input symbols by $s_{1,1}, s_{2,1}, s_{1,2}, s_{2,2}$. If the input symbols are taken from a real constellation (e.g., PAM), then the optimal phase shifts will be $\phi_1 = 0$, $\phi_2 = \frac{\pi}{2}$, $\theta_1 = 0$, $\theta_2 = \frac{\pi}{2}$, and the transmitted code word is

$$\mathbf{S} = \begin{bmatrix} s_{1,1} + js_{2,1} & j(s_{1,2} + js_{2,2}) \\ j(s_{1,2} - js_{2,2}) & s_{1,1} - js_{2,1} \end{bmatrix} \tag{3.24}$$

Define $x_1 = s_{1,1} + js_{2,1}$ and $x_2 = j(s_{1,2} - js_{2,2}) = js_{1,2} + s_{2,2}$, then

$$\mathbf{S} = \begin{bmatrix} x_1 & -x_2^* \\ x_2 & x_1^* \end{bmatrix} \tag{3.25}$$

which is exactly the Alamouti's scheme. Thus, the proposed scheme subsumes the Alamouti's scheme as a special case. For instance, the Alamouti's scheme with 16QAM is identical of the proposed 2-layered scheme with 4PAM and $N_t = T = 2$. Although it is a two-layered scheme, the symbols are limited to be real. As such, it looses half the channel capacity at high SNR when the number of receive antennas is two. On the other

hand, the proposed scheme is possible to achieve full channel capacity regardless the number of receive antennas.

3.4.2 Damen's Scheme

Damen et.al. proposed LD codes for two-transmit antennas based on number theory [14]. The codeword is constructed as

$$\mathbf{S} = \begin{bmatrix} s_{1,1} + e^{j\phi} s_{2,1} & e^{j\theta}(s_{1,2} + e^{j\phi} s_{2,2}) \\ e^{j\theta}(s_{1,2} - e^{j\phi} s_{2,2}) & s_{1,1} - e^{j\phi} s_{2,1} \end{bmatrix} \tag{3.26}$$

with $\theta = \frac{\phi}{2}$. Comparing (3.26) with (3.12) when $N = N_t = T = 2$, the two schemes has identical construction except that θ is constrained to be $\frac{\phi}{2}$ in Damen's scheme. This constraint simplifies the search of phase shifts but results loss in coding gain. On the other hand, in the proposed scheme, ϕ is chosen to maximize the coding gain of the codeword pairs that differ only in one layer instead of the coding gain of the code. As such, the complexity of the search of ϕ is reduced from $O(g^N)$ to $O(g)$ where g is a function of the constellation and N is the number of layers. Although θ has also to be searched in the proposed scheme, the search is only needed when the constellation size is small. For instance, for QPSK constellation, Damen's scheme yields the optimal coding gain 0.2369 at $\phi = 2\theta = 0.5$, while the proposed scheme yields the optimal coding gain 0.2588 at $\phi = \frac{\pi}{4}$ and $\theta = \frac{\pi}{6}$. Note that Damen's scheme only applies for channels with two transmit antennas.

3.4.3 Ma's Full-Diversity Full-Rate Codes

As mentioned before, the proposed scheme can be seen as a generalization of Ma's scheme [11], although developed with different approaches. One of the main differences is the optimization of phase shifts. In Ma's scheme, the phase shifts are chosen only to ensure full rank of the difference matrix without optimizing the coding gain. This can lead to solutions with unsatisfactory performance because the existence of a large portion of codeword pairs with small *effective* diversity gains [12]. While a direct optimization of the error performance is intractable, a better strategy, which is taken in this chapter, is to ensure large coding gain and diversity gain of the codeword pairs that have small Euclidean distance. Further performance comparison of the two schemes will be presented in the next section.

3.5 Simulation Results

In this section, we provide simulation results to compare the proposed scheme with several existing LD codes. The channel model described in Section 3.2 was assumed and maximum likelihood decoding was performed for all the schemes.

Figure 3.1 compares the proposed scheme with Hassibi's scheme [10] and Ma's scheme [11] over a channel with two transmit and two receive antennas. The code with the best performance in [10] was chosen. In all the schemes, two layers are employed with BPSK modulation. The modulation block is chosen to be 4 channel uses, i.e., 8 bits per block. As can be observed from the figure, at Block Error Rate (BLER) = $2*10^{-4}$, a performance gain of approximately 3 dB is achieved for the proposed scheme over

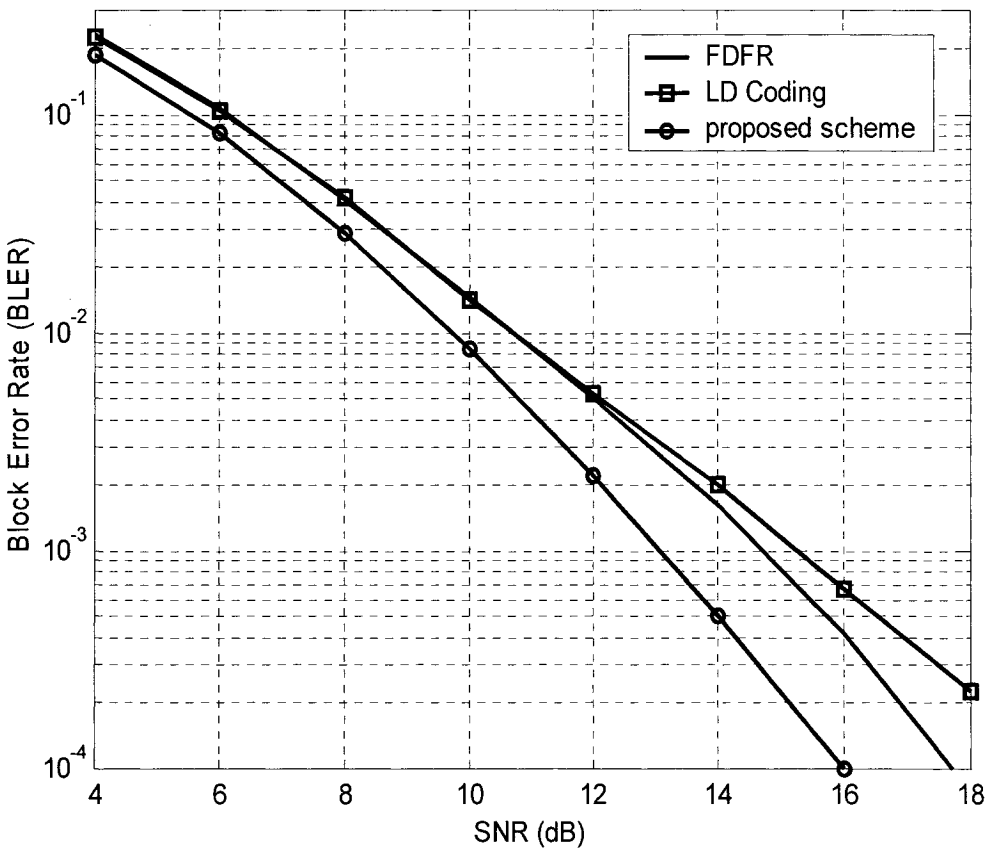

Figure 3.1. BLER performance comparison of FDFR, LD coding and the proposed scheme: 2 layers, 2 transmit, 2 receive antennas and BPSK modulation.

Hassibi's scheme and 1.8 dB over Ma's scheme. This is because the proposed scheme maximizes the diversity and coding gains without loss of mutual information. Hassibi's code preserves channel capacity but does not guarantee good performance since the diversity and coding gains were not explicitly optimized. Ma's code maximizes the diversity gain and preserves channel capacity but does not necessarily yield high coding gain which affects the performance of the system as discussed before.

In Figure 3.2, the performance of several two-layer codes with QPSK constellation are provided. Again, two transmit and two receive antennas are assumed. As expected, similar behavior as in the case of Figure 3.1 is observed. Specifically, the proposed two-layer code performs the best, followed by Damen's, Ma's, and Hassibi's codes. In addition, performance curves of Alamout's code and the proposed one-layer code, both with 16QAM, are provided. As mentioned before, Alamouti's code with 16QAM is identical with the proposed two-layer code with 4PAM. Hence, its performance is the worst among the two-layer codes. To examine the effects of the number of layers or symbol coding rate on performance, one can compare the proposed two-layer and one-layer codes. As can be seen, with the same data rate, the proposed two-layer code has approximately 2.5 dB gain over the proposed one-layer code at BLER=0.01. This shows that, for a given data rate, one shall choose a larger number of layers against constellation size to achieve better performance as predicted by Theorem 3.1.

Now, we examine the effects of the number of layers on a channel that consists of multiple transmit antennas but only one receive antenna. In such a case, the available multiplexing gain is 1. However, codes with more layers (i.e., higher coding rate in symbols) are expected to perform better at a practical SNR. Specifically, three transmit

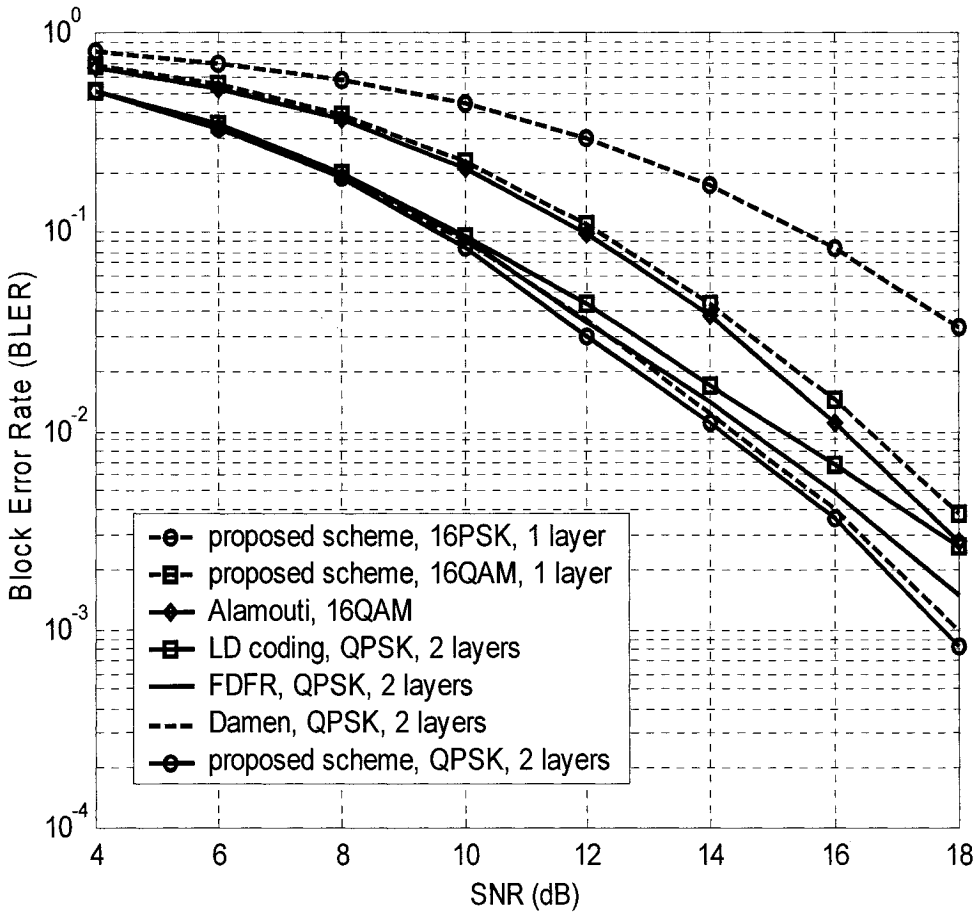

Figure 3.2. BLER performance comparison of the proposed and conventional schemes: 2 transmit and 2 receive antennas.

antennas were assumed. Performance curves of three full-rate codes: the proposed, Hassibi's, and Ma's codes, are illustrated in Figure 3.3. With BPSK and $T = 3$, each modulation block consists of 9 bits. In addition, the performance of the proposed one-layer code with 8PSK is also provided. As can be seen, the proposed 3-layer code performs significantly better than others. At BLER $2*10^{-3}$, the proposed 3-layer code has a performance gain about 2 dB over the other two full-rate codes. It is interesting to note in this case, the proposed one-layer code outperforms the other two full-rate codes when SNR is greater than 20dB. This is because the proposed one-layer code ensures full-diversity gain and optimal coding gain. Although Ma's code also enjoys full diversity gain, some of its diversity orders may not contribute to the decay of error performance till SNR is extremely high. It is also interesting to note that the two curves of the proposed 3-layer and one-layer codes consistently have the same slopes, with the former outperforms the latter. This, again, demonstrate that one shall maximize the number of layers for a given target data rate. However, the effects of the number of layers is less significant than the case shown in Figure 3.2 because there is only one receive antenna.

Last, let us compare the proposed scheme and Alamouti's space-time block code [4] with two transmit and one receive antennas as shown in Figure 3.4. For Alamouti's STBC, 16QAM modulation was used. For the proposed scheme, 2 layers and 4AM modulation were used. This makes the data rate for both of them is equal. As can be observed, the BLER of the proposed scheme overlaps exactly with BLER of the STBC. This is because, the space-time block code preserves full channel capacity when the number of transmit and receive antennas is equal to two and one, respectively. However, the Alamouti's STBC code only works for channels with two transmit antennas.

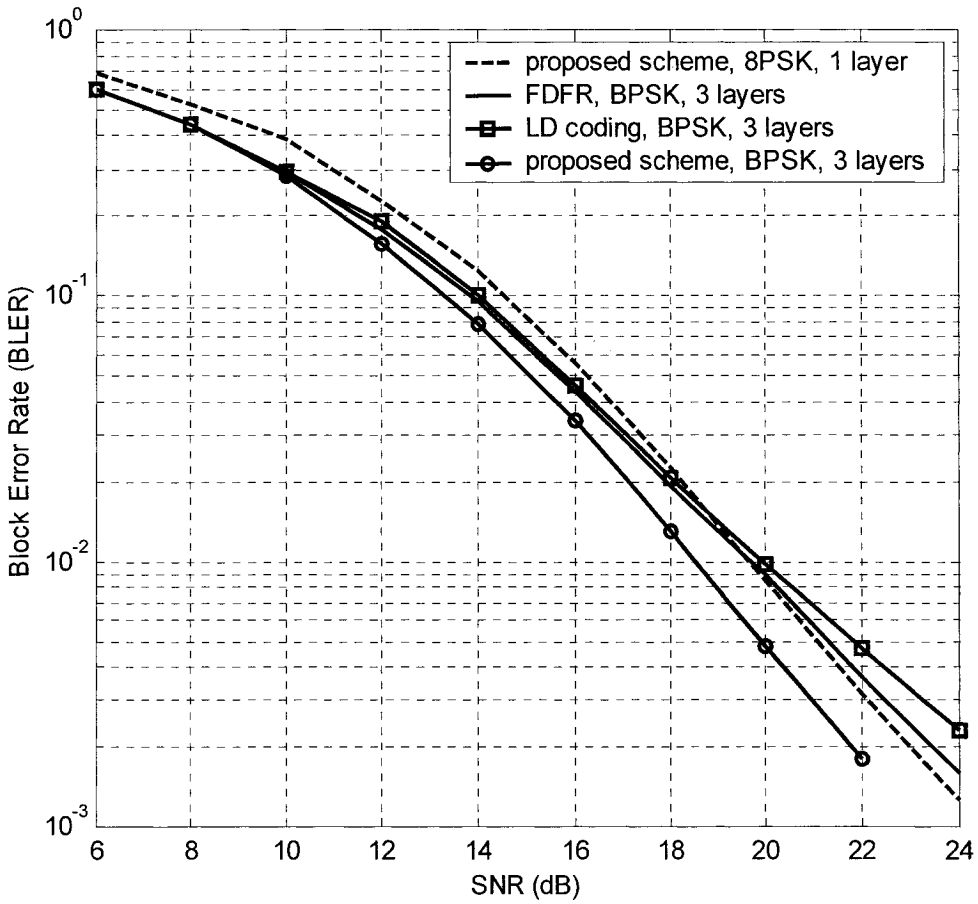

Figure 3.3. BLER performance comparison of FDFR, LD coding and the proposed scheme: 3 transmit and 1 receive antennas.

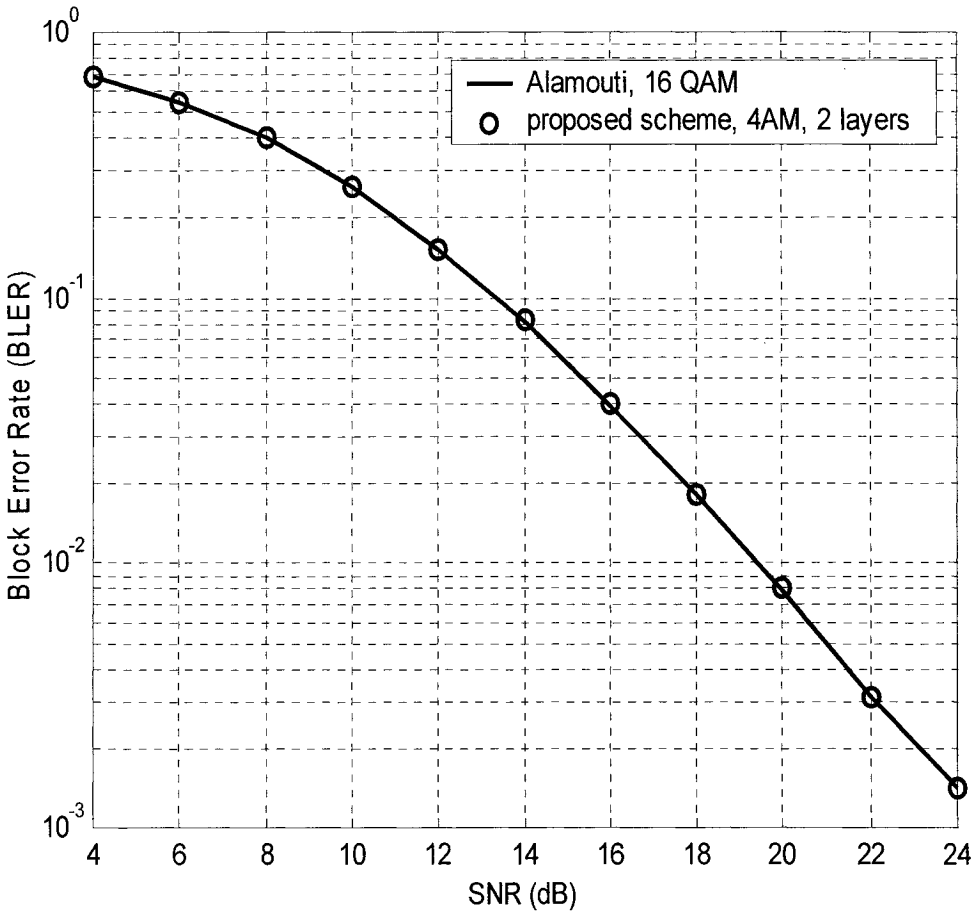

Figure 3.4. BLER performance comparison of the proposed scheme and space-time block code: 2 transmit and 1 receive antennas.

In summary, the above results show that the proposed scheme provides flexible tradeoff between performance and data rate. Furthermore, it shows that the proposed scheme provides significant performance gain as compared to the conventional schemes.

3.6 Conclusions

In this chapter, we have investigated the design of LD codes with flexible rate-performance tradeoff. A study on the capacity of LD coded channel was carried out. It shows that the maximum multiplexing gain of a linear dispersion code is exactly the coding rate defined as the number of symbols per channel use of the code. This also implies that full-rate LD codes have to be used to preserve the channel capacity. Conditions on the dispersion matrices for various multiplexing gains were also established. With these results, a general multi-layered LD coding scheme has been proposed. With the new coding scheme, various multiplexing gains can be obtained simply by increasing the coding rate and augmenting the set of existing dispersion matrices. Furthermore, the error performance is optimized by applying phase shifts among input symbols without the loss of mutual information. Simulation results demonstrated that the proposed coding scheme significantly outperforms existing space-time block codes under various data rates.

Chapter 4

Multi-Layered Space-Frequency Coding for Frequency-Selective MIMO Channels

4.1 Introduction

In frequency-selective channels, the transmissions of adjacent symbols will overlap each other, giving rise to ISI. One way to overcome the problem of ISI is to use OFDM to transfer a frequency-selective channel into a set of flat or narrowband channels.

In the conventional MIMO OFDM [16-20], N_t OFDM modulators are employed, one for each transmit antenna. A block diagram of this approach is illustrated in Figure 4.1 where the space-time process could be a space-time encoder (e.g., [3]) and/or a spatial multiplexer (e.g.,[6]). The data is transmitted frame by frame. The channel matrix is made block circulant by adding a Cyclic Prefix (CP) to each frame of the signal at the

transmitter and then removing it at the receiver. The noise-free received signal at the kth tone after the FFT can be expressed as

$$\begin{aligned}\mathbf{y}(k) &= [y_1(k)\ y_2(k)\ \cdots\ y_{N_r}(k)]^T \\ &= \sqrt{P/N_t}\,\mathbf{\Lambda}(k)[x_1(k)\ x_2(k)\ \cdots\ x_{N_t}(k)]^T \\ &= \sqrt{P/N_t}\,\mathbf{\Lambda}(k)\mathbf{x}(k) \quad \text{for } k = 1, \cdots, K \end{aligned} \tag{4.1}$$

where K is the FFT length, and $\mathbf{\Lambda}(k)$ is an $N_r \times N_t$ complex matrix. Denoting the channel matrix associated with the lth resolvable path by $\mathbf{H}(l) = [h_{m,n}(l)]_{m,n}$, $\mathbf{\Lambda}(k)$ can be written as

$$\mathbf{\Lambda}(k) = \sum_{l=0}^{L-1} \mathbf{H}(l) \exp\left\{\frac{-j2\pi(k-1)l}{K}\right\} \tag{4.2}$$

where L is the number of resolvable paths. As can be seen from (4.2), OFDM has transformed the frequency-selective fading channel into a set of MIMO flat subchannels. Or from the point of view of the signal before the IFFT, the resultant channel appears as a fast fading narrowband channel. Therefore, many of the MIMO techniques developed for narrowband fast fading channels are directly applicable.

When space-time coding is applied before OFDM modulation, it is often called Space-Time-Coded OFDM (STC-OFDM). In STC-OFDM, the coding is actually performed across transmit antennas and frequency tones and, thus, is also called *space-frequency coding*. Both space-time trellis codes [3], [21] and block codes [4], [46] can be applied. However, it was shown in [20] that the conventional space-time codes in general cannot achieve the full space and frequency diversity and suggested that new code design

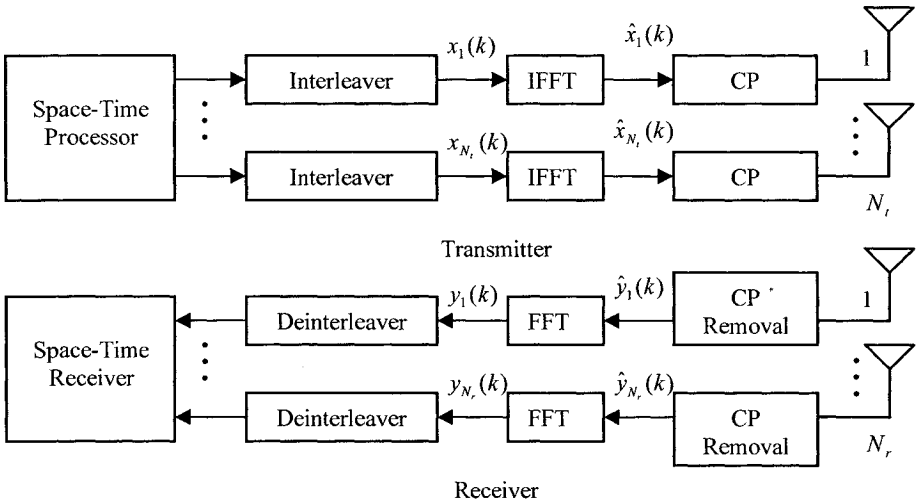

Figure 4.1. Block Diagram of a conventional MIMO OFDM transceiver.

procedures are needed. Unfortunately, code design in STC-OFDM shown in Figure 4.1 is not trivial, owing to the facts that the encoding is a two-dimensional process and the channel gains are correlated among different tones.

Spatial multiplexing gain can also be realized in MIMO OFDM systems. A MIMO transmission scheme is said to achieve a spatial multiplexing gain G_m if, at high SNR, it achieves a data rate

$$R = G_m \log(1 + \text{SNR}) \quad (bps/Hz) \tag{4.3}$$

which is G_m times the capacity of a SISO channel with the same SNR. Along this line, much of the recent research attention has been focused on the combined use of D-BLAST and V-BLAST with OFDM. In these schemes, N_t independently encoded data streams are spatially rearranged before they are sent to the N_t IFFT's as shown in Figure 4.1.

Although encoding for each data stream is a one-dimensional process, the diversity gain depends on the spatial arrangement of the independently encoded data streams. As a result, code design has to consider the spatial arrangement and becomes a difficult task.

A recent study in [8] demonstrates that there exists a fundamental tradeoff between diversity and multiplexing: The maximal available multiplexing and diversity gains cannot be achieved simultaneously and there exists an optimal tradeoff curve. Unfortunately, none of the aforementioned schemes are flexible in providing this tradeoff.

To this end, we emphasize that, in the conventional MIMO OFDM approach, the separation of the degrees of freedom in space and frequency gives rise to the difficulties in code design as well as in providing tradeoff flexibility between reliability versus data rate. This has motivated us to develop a new multi-layered space-frequency coding scheme to provide various rate-performance tradeoffs. The proposed scheme is based on a new space-time OFDM (ST-OFDM) modulation scheme that translates a MIMO channel into a SIMO channel of the same degrees of freedom. As a result, code design becomes much simpler and conventional codes for single-input channels can be used. In fact, the criteria for code design coincide with those codes developed for single-input fast fading channels. In addition, a new multiplexing scheme is developed to be used with the proposed ST-OFDM. In the proposed multiplexing scheme, each layer sees all the transmit antennas all the time and hence maintains all the degrees of freedom. It is further shown that the proposed multiplexing scheme achieves a multiplexing gain that equals the number of layers created. This, together with coding, allows a full range of optimal tradeoffs between data rate and reliability. Several examples are given to demonstrate the

advantages of the proposed approach over the conventional MIMO OFDM approach.

The remainder of the chapter is organized as follows. In Section 4.2, we introduce a new ST-OFDM modulation scheme. In Section 4.3, we discuss the design of codes and interleavers to be used with the proposed modulation scheme. In Section 4.4, we describe the multiplexing scheme and discuss the corresponding channel capacity and outage probability. Last, simulation results are presented in Section 4.5 and conclusions are drawn in Section 4.6.

4.2 A New Space-Time OFDM Modulator

In this section, we introduce a ST-OFDM modulator that effectively transfers a MIMO channel into a SIMO channel without the loss of the degrees of freedom. We consider communications over wideband MIMO channels with N_t transmit and N_r receive antennas. The gain of the lth resolvable path from transmit antenna n to receive antenna m is denoted by $h_{m,n}(l)$ with $0 \le l \le L-1$ and L as the number of resolvable paths. A block Rayleigh fading channel model will be assumed where the path gains remain constant over the duration of a transmission block and change from block to block. Furthermore, the path gains are assumed to be symmetric complex Gaussian random variables.

In Figure 4.2, a MIMO transceiver equipped with the proposed space-time modulation scheme is illustrated. In such a transceiver, data is transmitted frame by frame and each frame of data is first encoded, mapped to symbols, and interleaved to form K samples $\mathbf{x} = [x_1\ x_2 \cdots x_K]^T$. Here the encoder has only one output stream instead of N_t

streams as in STC-OFDM. The encoded data stream $\mathbf{x}$ is passed through a K-point IFFT and the resulting signal is given by

$$\hat{\mathbf{x}} = \frac{1}{\sqrt{K}}\mathbf{F}^H\mathbf{x} \tag{4.4}$$

where $\mathbf{F}$ is the FFT matrix whose (k_1,k_2) th entry is $f_{k_1,k_2} = \exp\left\{-\frac{j2\pi(k_1-1)(k_2-1)}{K}\right\}$.

The output of the IFFT, $\hat{\mathbf{x}}$, is used to form N_t blocks with the nth block as its cyclic shift to the top by n-1, i.e.,

$$\hat{\mathbf{x}}_n = [\hat{x}_n\ \hat{x}_{n+1}\cdots\hat{x}_K\ \hat{x}_1\cdots\hat{x}_{n-1}]^T = \mathbf{P}_n\hat{\mathbf{x}} \quad \text{for } n = 1,2,\cdots,N_t \tag{4.5}$$

where $\mathbf{P}_n$ is an $K\times K$ circulant permutation matrix whose first row has 1 as its nth entry and zeros elsewhere, i.e,

$$\mathbf{P}_n = \begin{bmatrix} \overbrace{0\ \ldots\ 0}^{n-1} & 1 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ & & & & 0 \\ 0 & & 0 & & 1 \\ 1 & \ddots & & \ddots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & \ddots & & & \\ 0\ \ldots\ 0 & 1 & 0 & \ldots & 0 \end{bmatrix}. \tag{4.6}$$

Clearly, $\mathbf{P}_1$ is the identity matrix and $\hat{\mathbf{x}}_1 = \mathbf{x}$.

The N_t signal blocks will be used to form N_t data matrices and then transmitted block by block. First, for each of the N_t blocks the last $N_t(L-1)$ samples are inserted at the beginning of the block as the cyclic prefix. The resultant signals are passed through a

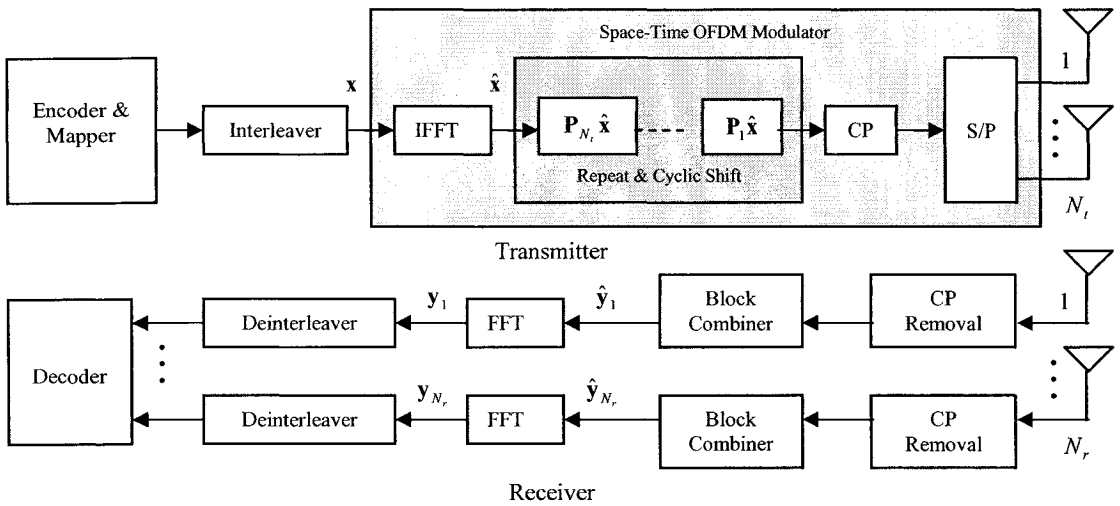

Figure 4.2. Block Diagram of the proposed MIMO transceiver using space-time OFDM transceiver.

$1:N_t$ Serial-to-Parallel (S/P) converter to form a data matrix with N_t rows, which is then transmitted over the N_t transmit antennas. Assuming that N_t divides K, a total of $(L-1)+\frac{K}{N_t}$ transmission durations are needed for each of the N_t blocks and the loss in bandwidth efficiency due to cyclic prefix is $\frac{(L-1)N_t}{K+(L-1)N_t}$. For instance, if $K=4, N_t=L=2$, the 4 IFFT output samples will be used to form 2 blocks and transmitted from right to left over the two antennas as

$$\begin{bmatrix} x_4 & x_2 & x_4 & | & x_3 & x_1 & x_3 \\ x_1 & x_3 & \underbrace{x_1}_{prefex} & | & x_4 & x_2 & \underbrace{x_4}_{prefex} \end{bmatrix}.$$

The case where N_t does not divide K will be addressed shortly. Since IFFT outputs

are transmitted over both space and time, the IFFT together with the subsequent framing process will be referred to as the *ST-OFDM modulation.*

At the receiver, after removing the first L-1 samples corresponding to the CP, the signal of the nth block received by the mth antenna can be written as

$$\begin{aligned}\hat{\mathbf{y}}_{m,n} &= \sqrt{P/N_t}\mathbf{H}_m\mathbf{P}_n\hat{\mathbf{x}} + \mathbf{n}_{m,n} \\ &= \sqrt{P/N_t}\mathbf{H}_{m,n}\hat{\mathbf{x}} + \mathbf{n}_{m,n}\end{aligned} \tag{4.7}$$

where $\mathbf{n}_{m,n}$ is complex Gaussian noise with zero mean and variance $N_0\mathbf{I}$, and $\mathbf{H}_m$ is an $\frac{K}{N_t} \times K$ block circulant matrix given by

$$\mathbf{H}_m = \begin{bmatrix} \mathbf{h}_m^T(0) & 0_{1\times N_t} & \cdots & 0_{1\times N_t} & \mathbf{h}_m^T(L-1) & \cdots & \mathbf{h}_m^T(2) & \mathbf{h}_m^T(1) \\ \mathbf{h}_m^T(1) & \mathbf{h}_m^T(0) & 0_{1\times N_t} & \cdots & 0_{1\times N_t} & \mathbf{h}_m^T(L-1) & \cdots & \mathbf{h}_m^T(2) \\ & \ddots & & \ddots & & \ddots & & \ddots \\ 0_{1\times N_t} & \cdots & 0_{1\times N_t} & & \mathbf{h}_m^T(L-1) & \cdots & \mathbf{h}_m^T(1) & \mathbf{h}_m^T(0) \end{bmatrix}. \tag{4.8}$$

In (4.8), $\mathbf{h}_m^T(l) = [h_{m,1}^T(l)\ h_{m,2}^T(l)\ \cdots\ h_{m,N_t}^T(l)]$ collects all the channel gains from N_t transmit antennas to the mth receive antenna associated with the lth delay path. Since $\mathbf{P}_n$ is a permutation matrix, the rows in $\mathbf{H}_{m,n} = \mathbf{H}_m\mathbf{P}_n$ are cyclic shifts of the corresponding rows in $\mathbf{H}_m$ to the right by n-1.

The block combiner at the receiver collects the signals from all of the N_t blocks and interleaves them sample by sample. From (4.7) and (4.8), the resulting output is given by

$$\hat{\mathbf{y}}_m = \begin{bmatrix} \hat{y}_{m,1}(1) \\ \vdots \\ \hat{y}_{m,N_t}(1) \\ \hat{y}_{m,1}(2) \\ \vdots \\ \hat{y}_{m,N_t}(K/N_t) \end{bmatrix} = \sqrt{\frac{P}{N_t}} \begin{bmatrix} \mathbf{H}_{m,1}(1,\ :) \\ \vdots \\ \mathbf{H}_{m,N_t}(1,\ :) \\ \mathbf{H}_{m,1}(2,\ :) \\ \vdots \\ \mathbf{H}_{m,N_t}(K/N_t,\ :) \end{bmatrix} \hat{\mathbf{x}} + \mathbf{n}_m$$

$$= \sqrt{P/N_t}\,\tilde{\mathbf{H}}_m \hat{\mathbf{x}} + \mathbf{n}_m \tag{4.9}$$

where $\mathbf{H}_{m,n}(k,\ :)$ stands for the kth row of matrix $\mathbf{H}_{m,n}$. Note that $\mathbf{H}_{m,n}(k,\ :)$ is the cyclic shift of $\mathbf{H}_{m,1}(k,\ :)$ to the right by n-1 and $\mathbf{H}_{m,1} = \mathbf{H}_m$ is block ciculant as given in (4.8). Therefore, $\tilde{\mathbf{H}}_m$ is an $K \times K$ circulant matrix whose first column collects all the $N_t L$ path gains associated with receive antenna m and is given by

$$\begin{aligned} \tilde{\mathbf{h}}_m = [&h_{m,1}(0)\ h_{m,N_t}(1)\ h_{m,N_t-1}(1) \cdots h_{m,1}(1) \cdots h_{m,N_t}(L-1) \cdots \\ &h_{m,1}(L-1)\ \underbrace{0 \cdots 0}_{N_t}\ h_{m,N_t}(0) \cdots h_{m,2}(0)]^T \end{aligned} \tag{4.10}$$

From (4.4), (4.9), (4.10), and by the fact that an FFT matrix diagonalizes a circulant matrix of the same size, the output of the FFT at receive antenna m can be written as

$$\mathbf{y}_m = \frac{\sqrt{P}}{K\sqrt{N_t}} \mathbf{F}\tilde{\mathbf{H}}_m \mathbf{F}^H \mathbf{x} + \frac{1}{\sqrt{K}} \mathbf{F}\mathbf{n}_m$$

$$= \sqrt{\frac{P}{N_t}} \mathbf{\Lambda}_m \mathbf{x} + \mathbf{z}_m \tag{4.11}$$

where complex Gaussian noise $\mathbf{z}_m$ has zero mean and variance $N_0\mathbf{I}$ and

$$\mathbf{\Lambda}_m = \mathrm{diag}(\mathbf{F}\tilde{\mathbf{h}}_m). \tag{4.12}$$

The above equality still holds after removing all the zeros in $\tilde{\mathbf{h}}_m$ and the corresponding columns in $\mathbf{F}$, i.e.,

$$\mathbf{\Lambda}_m = \mathrm{diag}(\hat{\mathbf{F}}\hat{\mathbf{h}}_m)\,. \tag{4.13}$$

By (4.10), $\hat{\mathbf{h}}_m$ consists of the first $N_tL - N_t + 1$ and the last $N_t - 1$ entries in $\tilde{\mathbf{h}}_m$ and, correspondingly, $\hat{\mathbf{F}}$ consists of the first $N_tL - N_t + 1$ and the last $N_t - 1$ columns of $\mathbf{F}$.

As can be seen from (4.11), the original frequency-selective multiple-input channel has now been transformed into K independent flat single-input channels. Or looking at the time domain, (4.9) shows as if the signal $\hat{\mathbf{x}}$ arrived at antenna m through a single-input channel with Impulse Response (IR), $\tilde{\mathbf{h}}_m$, given by (4.10), although it was actually transmitted over N_t antennas. As a result of this channel transformation, conventional codes designed for single-input fast fading channels can be applied. In addition, this transformation maintains the degrees of freedom (i.e., available diversity order), which is apparent by the fact that there are N_tL independent random variables in $\tilde{\mathbf{h}}_m$. Unlike STC-OFDM where the available system freedom is exploited by the two-dimensional encoding process, the ST-OFDM modulation *projects* all the degrees of freedom onto one virtual domain so that they can be capitalized by one-dimensional encoding. In the rest of the chapter, $\tilde{\mathbf{h}}_m$ in (4.10) and $\hat{\mathbf{h}}_m$ in (4.13) will be referred to as the IR and the *effective* IR of the equivalent single-input channel, respectively.

In the above, it has been assumed that N_t divides K. In case where N_t does not divide K, the same structure shown in Figure 4.2 can be used with a slightly different framing process. The framing process now generates N_t' blocks as if there were N_t'

transmit antennas, where N_t' is the smallest integer that is greater than N_t and divides K. However, for each data block, the last $N_t' - N_t$ streams at the output of the S/P converter are not transmitted. By doing this, the impulse response $\tilde{\mathbf{h}}$ will be different from that given in (4.10) but still has exactly $N_t L$ nonzero entries. Since this case will not change the subsequent results, we will only consider the case where N_t divides K in the rest of the chapter.

4.3 Code and Interleaver Design

In this section, we discuss the design of codes and interleavers to be used with the ST-OFDM modulation for applications with reasonably high SNR. For cases where SNR is low but the available diversity is larger (e.g., there exists a large number of antennas), conventional designs based on maximizing Euclidean distance can be used [51].

For the simplicity of analysis, we assume that channels seen by different receive antennas are independent but those associated with one receive antenna are correlated. That is,

$$E\left[\hat{\mathbf{h}}_{m_1}\hat{\mathbf{h}}_{m_2}^{H}\right] = \begin{cases} 0 & \text{if } m_1 \neq m_2 \\ \mathbf{R}_h & \text{if } m_1 = m_2 \end{cases}.$$

This may happen when the transmit antennas are closely spaced but receive antennas are sufficiently far apart. We further assume that $\mathbf{R}_h$ has full rank and $\mathrm{tr}(\mathbf{R}_h) = N_t$. For a complete treatment of more general cases, one may follow the development in [52].

As in [3], we begin with the analysis of pair-wise error probability by assuming ML

decoding. An upper bound of the probability of codeword $\mathbf{x}$ being transmitted but the decision being erroneously made in favor of $\mathbf{x}'$ can be found as

$$P(\mathbf{x} \to \mathbf{x}') \le \exp\left(\frac{-d^2(\mathbf{x}, \mathbf{x}')}{4N_0}\right) \tag{4.14}$$

where $d^2(\mathbf{x}, \mathbf{x}')$ is the Euclidean distance between the two codewords. By (4.11) and (4.13), we have

$$d^2(\mathbf{x}, \mathbf{x}') = \frac{P}{N_t}\sum_{m=1}^{N_r}\left\|\Lambda_m(\mathbf{x}-\mathbf{x}')\right\|^2$$

$$= \frac{P}{N_t}\sum_{m=1}^{N_r}\left\|\mathrm{diag}(\hat{\mathbf{F}}\hat{\mathbf{h}}_m)(\mathbf{x}-\mathbf{x}')\right\|^2. \tag{4.15}$$

Decompose the channel correlation matrix as $\mathbf{R}_h = \mathbf{B}\mathbf{B}^H$ and define $\mathbf{g}_m = \mathbf{B}^{-1}\hat{\mathbf{h}}_m$. Apparently, $\mathbf{g}_m$ consists of independent identically distributed (i.i.d.) symmetric Gaussian random variables with zero mean and variance $1/L$. We can then rewrite (4.15) as

$$d^2(\mathbf{x}, \mathbf{x}') = \frac{P}{N_t}\sum_{m=1}^{N_r}\left\|\mathrm{diag}(\hat{\mathbf{F}}\mathbf{B}\mathbf{g}_m)(\mathbf{x}-\mathbf{x}')\right\|^2. \tag{4.16}$$

Denoting V as the set collecting all the indexes corresponding to nonzero entries in $\mathbf{x}-\mathbf{x}'$, i.e., $V = \{k \mid x(k) - x'(k) \neq 0\}$, $\hat{\mathbf{F}}(V,:)$ as the matrix consisting of those rows of $\hat{\mathbf{F}}$ whose indexes are in V, and $\mathbf{E}$ as the diagonal matrix whose kth diagonal entry is the kth nonzero entry of $\mathbf{x}-\mathbf{x}'$, (4.16) can be rewritten as

$$d^2(\mathbf{x}, \mathbf{x}') = \frac{P}{N_t}\sum_{m=1}^{N_r}\mathbf{g}_m^H\mathbf{B}^H\hat{\mathbf{F}}(V,:)^H\mathbf{E}^H\mathbf{E}\hat{\mathbf{F}}(V,:)\mathbf{B}\mathbf{g}_m. \tag{4.17}$$

Substituting (4.17) into (4.14) and averaging (4.14) with respect to the distribution of $|\hat{\mathbf{g}}_m(k)|$, the average pair-wise error probability at high SNR region can be found as

$$P(\mathbf{x} \to \mathbf{x}') \le \left[\hat{\det}\left(\mathbf{B}^H \hat{\mathbf{F}}(V,:)^H \mathbf{E}^H \mathbf{E}\hat{\mathbf{F}}(V,:)\mathbf{B}\right)\right]^{-N_r} \exp\left(\frac{P}{4N_t N_0}\right)^{-rN_r} \tag{4.18}$$

where r is the rank of $\mathbf{E}\hat{\mathbf{F}}(V,:)\mathbf{B}$ or, equivalently, the rank of $\mathbf{E}\hat{\mathbf{F}}(V,:)$, and $\hat{\det}(\mathbf{A})$ stands for the product of all the nonzero eigenvalues of $\mathbf{A}$. From the above equation, it can be observed that a diversity gain $G_d(\mathbf{x},\mathbf{x}') = rN_r$ and a coding gain $G_c(\mathbf{x},\mathbf{x}') = \frac{1}{N_t}\left[\hat{\det}\left(\mathbf{B}^H \hat{\mathbf{F}}(V,:)^H \mathbf{E}^H \mathbf{E}\hat{\mathbf{F}}(V,:)\mathbf{B}\right)\right]^{\frac{1}{r}}$ have been achieved for the codeword pair.

Recall that $\hat{\mathbf{F}}(V,:)$ is of size $|V| \times N_t L$ with $|V|$ being the number of entries in set V. Furthermore, since $\hat{\mathbf{F}}(V,:)$ is a submatrix of the Vandermonder matrix $\mathbf{F}$, $\hat{\mathbf{F}}(V,:)$ is full rank, i.e.,

$$\text{rank}\left(\hat{\mathbf{F}}(V,:)\right) = \min\left(|V|, N_t L\right).$$

This with the fact that $\mathbf{E}$ has a rank $|V|$ leads to

$$r = \min\left(|V|, N_t L\right). \tag{4.19}$$

Following the definition in [53], we refer to $|V|$ as *the effective length* of the codeword pair and the minimal $|V|$ taking over all distinct codeword pairs as the effective length of the code. Since we are primarily interested in codeword pairs that have the worst pair-wise error probability, we may assume $|V| \le N_t L$ in the following analysis. Under this assumption, $r = |V|$. Furthermore, since matrices $\mathbf{A}\mathbf{A}^H$ and $\mathbf{A}^H\mathbf{A}$ have the same set of

nonzero eigenvalues, we can write

$$G_c(\mathbf{x},\mathbf{x}') = \frac{1}{N_t}\left[\det\left(\mathbf{E}\hat{\mathbf{F}}(V,:)\mathbf{B}\mathbf{B}^H\hat{\mathbf{F}}(V,:)^H\mathbf{E}^H\right)\right]^{\frac{1}{r}}$$

$$= \frac{1}{N_t}\left[\left|\det(\mathbf{E})\right|^2 \det\left(\hat{\mathbf{F}}(V,:)\mathbf{R}_h\hat{\mathbf{F}}(V,:)^H\right)\right]^{\frac{1}{r}}$$

$$= \frac{1}{N_t}\left[\prod_{i\in V}|e_i|^2 \det\left(\hat{\mathbf{F}}(V,:)\mathbf{R}_h\hat{\mathbf{F}}(V,:)^H\right)\right]^{\frac{1}{r}}. \qquad (4.20)$$

Note that the product term $\prod_{i\in V}|e_i|^2$ in the above equation is *the squared product distance* of the codeword pair [53], which is independent of the channel. On the other hand, the term $\det\left(\hat{\mathbf{F}}(V,:)\mathbf{R}_h\hat{\mathbf{F}}(V,:)^H\right)$ is related to the channel statistics and locations on which a pair of codewords differ. Since the channel is unknown to the transmitter, it is desirable to neglect this term in code design and leave its maximization by the interleaver as will be discussed shortly.

Define the diversity and coding gains of a code as the minimal pair-wise diversity gain and coding gain taking over all distinct codeword pairs, respectively. From the above analysis, the diversity gain of a code is the product of N_r and the effective length of the code, and the coding gain is directly proportional to the minimal squared product distance of the code. To summarize, we have the following design criteria:

- *The Effective-Length Criterion:* If a diversity gain of rN_r is the design target $\left(r \le N_t L\right)$, the effective length of the code must be r.

- *The Squared-Product-Distance Criterion:* The minimal squared product distance of

the code must be maximized.

It is evident that the above code design criteria are exactly those well-known criteria for single-input fast fading channels. Therefore, the codes designed for single-input fading channels such as those in [53] and [54] can be directly used here.

As shown in (4.20), in addition to the squared product distance, the coding gain also depends on channel statistics and the locations of symbols in which the pair of codewords differ. In general, maximizing the term $\det\left(\hat{\mathbf{F}}(V,:)\mathbf{R}_h\hat{\mathbf{F}}(V,:)^H\right)$ in (4.20) is not feasible when channel statistics are unknown to the transmitter. However, for channels with i.i.d. path gains (i.e., $\mathbf{R}_h = \frac{1}{L}\mathbf{I}$), it was shown in [55] that $\det\left(\hat{\mathbf{F}}(V,:)\hat{\mathbf{F}}(V,:)^H\right)$ is maximized if the pair of codewords differ at locations that are separate by an integer multiple of $\frac{K}{LN_t}$. Therefore, a block interleaver with an interleaving depth $\frac{K}{LN_t}$ is nearly optimal for trellis codes over channels with i.i.d path gains [55]. Interestingly, this is also a "good" design for trellis codes in general. First of all, if the effective length of the code is $|V| = N_t L$, $\hat{\mathbf{F}}(V,:)$ has full rank and

$$\det\left(\hat{\mathbf{F}}(V,:)\mathbf{R}_h\hat{\mathbf{F}}(V,:)^H\right) = \det\left(\hat{\mathbf{F}}(V,:)\hat{\mathbf{F}}(V,:)^H\right)\det(\mathbf{R}_h).$$

Hence if the two codewords differ at symbols that are $\frac{K}{LN_t}$ tones apart, $\det\left(\hat{\mathbf{F}}(V,:)\hat{\mathbf{F}}(V,:)^H\right)$ is maximized [55]. In general when the channel is unknown to the transmitter, we can only seek to maximize the following lower bound

$$\det\left(\hat{\mathbf{F}}(V,:)\mathbf{R}_h\hat{\mathbf{F}}(V,:)^H\right) \geq \det\left(\hat{\mathbf{F}}(V,:)\hat{\mathbf{F}}(V,:)^H\right)\prod_{i=1}^{|V|}\lambda_{\mathbf{R}}(i)$$

where $\lambda_{\mathbf{R}}(i)$ is the ith smallest eigenvalue of $\mathbf{R}_h$. Again, the above lower bound is maximized when the two codewords differ at symbols that are $\frac{K}{LN_t}$ tones apart.

4.4 Channel Multiplexing

In the coded ST-OFDM described above as well as in the conventional STC-OFDM, higher data rate can be achieved by using larger constellations. However, this is not desirable because the constellation size and, hence, the computational complexity, grow exponentially with the data rate. More importantly, there exists a fundamental tradeoff between multiplexing and diversity [8]. As will be shown later in this section, the coded ST-OFDM alone only provides a multiplexing gain 1. One approach to achieving high data rate with reasonable complexity is to transmit multiple independent data streams in parallel as in D-BLAST. The technique of creating a number of independent data streams (or layers) that simultaneously share the channel is called *channel multiplexing*.

The ST-OFDM modulation allows a simple channel-multiplexing mechanism as shown in Figure 4.3. As can be seen in the figure, the K encoded data streams are first independently modulated using the ST-OFDM modulator, the N_t outputs from each of the layers are then weighted and summed together before they are transmitted.

The use of transmit weights effectively creates different channels for different layers. Denote $w_{i,n}$ as the weight associated with layer i and transmit antenna n, then the lth

path gain of the channel from transmit antenna n to receive antenna m associated with layer i can be written as $w_{i,n}h_{m,n}(l)$. Denoting the weight vector of layer i by $\mathbf{w}_i = [w_{i,N_t} \quad w_{i,N_t-1} \cdots w_{i,1}]^T$ and collecting the weights according to the ordering of channel gains in $\hat{\mathbf{h}}_m$ into a diagonal matrix as

$$\mathbf{W}_i = \text{diag}\left([w_{i,1} \underbrace{\mathbf{w}_i^T \cdots \mathbf{w}_i^T}_{L-1 \text{ terms}} w_{i,N_t} \cdots w_{i,2}] \right) \tag{4.21}$$

then the effective IR of the equivalent single-input channel seen by layer i at receive antenna m can be written as

$$\hat{\mathbf{h}}_{i,m} = \mathbf{W}_i \hat{\mathbf{h}}_m . \tag{4.22}$$

It follows from (4.11) and (4.22) that the FFT output at receive antenna m is

$$\mathbf{y}_m = \sum_{i=1}^{N} \mathbf{y}_{m,i} + \mathbf{z}_m = \sqrt{\frac{P}{N_t N}} \sum_{i=1}^{N} \mathbf{\Lambda}_{m,i} \mathbf{x}_i + \mathbf{z}_m \tag{4.23}$$

where P is the total transmit power shared by all the N layers, $\mathbf{x}_i$ denotes the signal of the ith layer before IFFT, $\mathbf{y}_{m,i}$ is the contribution of the ith layer to the FFT output at receive antenna m, and $\mathbf{\Lambda}_{m,i}$ is a diagonal matrix given by

$$\mathbf{\Lambda}_{m,i} = \text{diag}(\hat{\mathbf{F}}\hat{\mathbf{h}}_{i,m}) = \text{diag}(\hat{\mathbf{F}}\mathbf{W}_i\hat{\mathbf{h}}_m) . \tag{4.24}$$

With the above channel multiplexing, the channel seen by the N coded signal streams is an N-input N_r-output channel. The channel matrix, $\mathbf{\Phi}_k$, seen by the kth tone of all the N layers is of size $N_r \times N$ and its (m, i)th entry can be found from (4.23) and (4.24) as

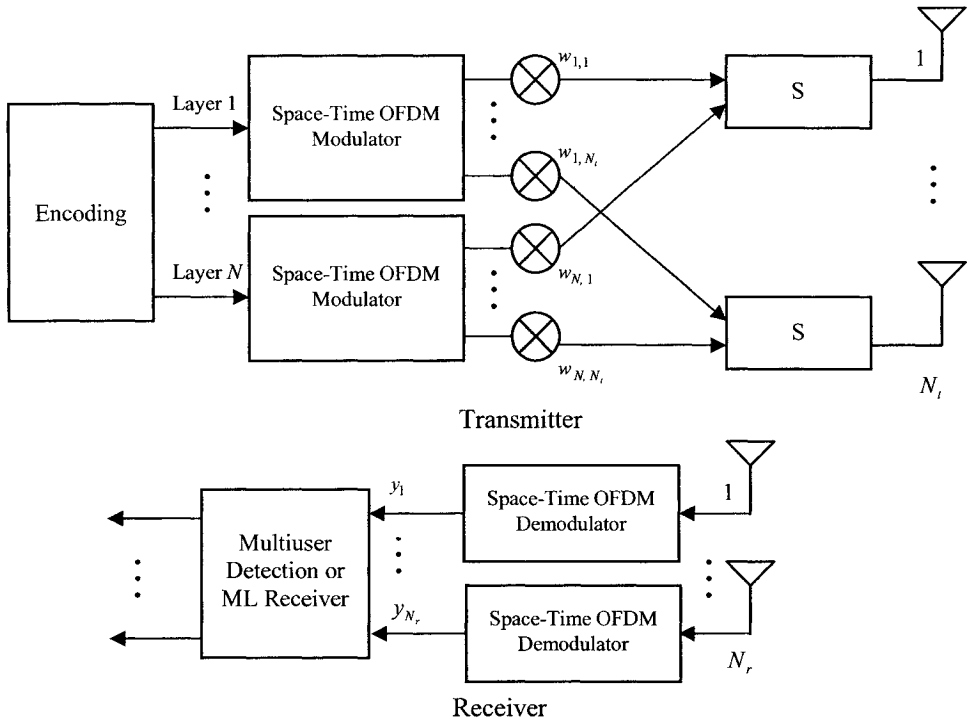

Figure 4.3. Block diagram of the proposed multiplexing scheme using space-time OFDM modulation.

$$\phi_k(m,i) = \lambda_{m,i}(k,k) = \hat{\mathbf{F}}(k,:)\mathbf{W}_i\hat{\mathbf{h}}_m. \tag{4.25}$$

It is interesting to note that the proposed multiplexing subsumes D-BLAST and V-BLAST multiplexing as special cases. If $\mathbf{w}_i$ is taken as the ith standard coordinate vector, it is vertical multiplexing; while if a time-variant weight vector is used with the $(\text{mod}(i+k, N_t)+1)$ th standard coordinate vector as the weight vector for the kth block of layer i (recall that there are N_t blocks in each frame), then it is diagonal multiplexing.

However, in conventional schemes that employ N_t OFDM's, the multiplexing is applied before OFDM whereas the proposed multiplexing is applied after OFDM.

4.4.1 Weight Design

A set of good multiplexing weights shall preserve channel capacity when a maximal number of layers is used and, in the mean time, facilitate harvesting of diversity gain through coding. When the channel, $\mathbf{\Phi}$, is unknown to the transmitter, the channel capacity is [1, 2]

$$\log_2 \det\left(\mathbf{I}_{N_r} + \frac{P}{N_t N_0}\mathbf{\Phi}\mathbf{\Phi}^H\right).$$

From the above, it is clear that any set of $N = \min(N_t, N_r)$ orthogonal multiplexing weight vectors of equal L_2-norm preserves the capacity. To allow easy harvesting of diversity gain, the multiplexing weights must not be zeros for any layer at any transmit antenna to preserve the degrees of system freedom. Apparently, both V-BLAST and D-BLAST have zero weights. To summarize, the following two conditions are to be satisfied:

$$\begin{cases} \mathbf{w}_i^H \mathbf{w}_j = 0 & \forall\, i \neq j \\ |w_{i,n}| = 1 & \forall (i,n) \end{cases}. \tag{4.26}$$

In addition to nonzero weights, equal transmit power among antennas is also enforced in the second condition. It will be useful to note that the first condition ensures independent gains in the equivalent N-input N_r-output channel when the original channel is spatially uncorrelated, i.e.,

$$\begin{cases} \phi_k(i,m) \sim CN(0,N_t) & \forall (i,k,m) \\ E[\phi_k(i,m)\phi_k(j,u)] = 0 & \forall (i,m) \neq (j,u) \end{cases}. \tag{4.27}$$

A set of weight vectors satisfying the conditions in (4.26) can be readily found as

$$w_{i,n} = \exp\left(\frac{-j2\pi(i-1)(n-1)}{N_t}\right) \quad \forall\, i = 1,\ldots,N \text{ and } n = 1,\ldots,N_t. \tag{4.28}$$

We conjecture that the set of multiplexing weight vectors defined in (4.28) is optimal in the sense that it provides the best performance when random encoding is performed.

4.4.2 Code and Receiver Design

Since the weighting does not change the channel statistics, if the N layers are independently encoded, the criteria of code design developed in the last section also applies here. In contrast, in the conventional multiplexing approach to wideband MIMO channels, the channel seen by each layer during a frame is either a constant channel with $N_r L$ degrees of freedom as in V-BLAST or a time varying channel cycling among transmit antennas with each providing $N_r L$ degrees of freedom as in D-BLAST. As a result, the available diversity gain for each layer is reduced or difficult to obtain. To formalize this, we consider the communication over a frequency-selective channel with i.i.d. path gains (i.e., $h_{m,n}(l) \sim CN(0,1/L)$) and we establish the following theorems.

Theorem 4.1: For a conventional spatial multiplexing scheme where layers are separate in space at any time, let a codeword matrix of size $N_t \times K$ collect the symbols from all layers according to their spatial and frequency locations (i.e., whose (n, k)th entry is the symbol to be transmitted over antenna n at tone k). Denote V_n as the set collecting the

indexes of symbols in which the two codewords $\mathbf{X}$ and $\mathbf{X}'$ differ at transmit antenna n. Then the pair-wise diversity gain is bounded as

$$G_d^c(\mathbf{X},\mathbf{X}') \geq N_r \times \max_{\{\tilde{V}_1,\tilde{V}_2,\cdots,\tilde{V}_{N_t}\}} \left| \bigcup_{1\leq n\leq N_t} \tilde{V}_n \right| \tag{4.29}$$

where $\tilde{V}_n \in V_n$ and $\left|\tilde{V}_n\right| \leq L$ and the maximization is taken over all possible selections of $\tilde{V}_n$.

Proof: Denote $\mathbf{Q}$ as the first L columns of a $K \times K$ FFT matrix and $\mathbf{G} = [\mathbf{Q}\,\mathbf{Q}\cdots\mathbf{Q}]$ as a complex matrix of size $KN_t \times LN_t$. Also, denote $\mathbf{E} = [\mathbf{E}_1\,\mathbf{E}_2\cdots\mathbf{E}_{N_t}]$ with $\mathbf{E}_n$ being a diagonal matrix of size K whose kth diagonal entry, $e_n(k)$, is $x_n(k) - x_n'(k)$. Then, the lower bound of the error probability between two codewords $\mathbf{X}$ and $\mathbf{X}'$ can be

$$P(\mathbf{X}\to\mathbf{X}') = \frac{1}{\hat{\det}(\mathbf{G}^H\mathbf{E}^H\mathbf{E}\mathbf{G})}\left(\frac{P}{4N_tN_0}\right)^{-r}$$

where $\hat{\det}(\mathbf{A})$ denotes the product of the nonzero eigenvalues of $\mathbf{A}$. Define $\tilde{V}$ as the union of N_t sets $\left\{\tilde{V}_n : \left|\tilde{V}_n\right| \leq L\right\}$ with the maximum size among all possible choices of $\tilde{V}_n$ with sizes less than or equal to L. There must exist a permutation matrix $\mathbf{P}$ such that $\mathbf{PE} = [\mathbf{E}_a^T\mathbf{E}_b^T]^T$ with $\mathbf{E}_a$ collecting rows of $\mathbf{E}$ with indexes in $\tilde{V}$. Denote $\hat{\mathbf{E}}_a$ as the matrix obtained by removing columns of $\mathbf{E}_a$ with indexes not in $\tilde{V}$ and $\hat{\mathbf{G}}$ as the matrix by removing rows of $\mathbf{G}$ with indexes $cK + i$ with any $c \geq 0$ and $i \notin \tilde{V}$.

Then we have

$$\mathbf{A} = \mathbf{E}_a\mathbf{G}\mathbf{G}^H\mathbf{E}_a^H = \hat{\mathbf{E}}_a\hat{\mathbf{G}}\hat{\mathbf{G}}^H\hat{\mathbf{E}}_a^H. \tag{4.30}$$

Note that each of the N_t diagonal blocks in $\hat{\mathbf{G}}$ has at most L rows taken from $\mathbf{Q}$, it is either a generalized Vandermonde matrix of size L or a submatrix of it and, thus, $\hat{\mathbf{G}}\hat{\mathbf{G}}^H$ is full rank and positive definite. Since $\hat{\mathbf{E}}_a$ is full-row rank, $\mathbf{A}$ in (4.30) is full rank, i.e., $\text{rank}(\mathbf{A}) = |\widetilde{V}|$. This with the facts that $\mathbf{A}$ is a submatrix of matrix $\mathbf{PEGG}^H\mathbf{E}^H\mathbf{P}^T$ and $\text{rank}(\mathbf{EGG}^H\mathbf{E}^H) = \text{rank}(\mathbf{PEGG}^H\mathbf{E}^H\mathbf{P}^T)$ has proved Theorem 4.1. □

Theorem 4.2: For the proposed multiplexing scheme, let a codeword matrix of size $N \times K$ collect symbols according to their associated layers and frequency locations (i.e., whose (*i, k*)th entry is the symbol of layer *i* to be transmitted at tone *k*). Denote U_i as the set collecting indexes of symbols in which the two codewords differ at layer *i*. Then the pair-wise diversity gain is bounded as

$$\begin{aligned} &G_d^n(\mathbf{X},\mathbf{X}') \geq N_r \times \max_{S:S\in\bigcup_{i=1}^{N}U_i} |S| \\ &\text{subject to} \sum_{i=1}^{N} |S \cup U_i| \leq N_t L \end{aligned} \tag{4.31}$$

where set $S \in \bigcup_{i=1}^{N} U_i$ and the maximization is taken over all possible selections S satisfying $\sum_{i=1}^{N} |S \cup U_i| \leq N_t L$.

Proof: We first assume that there is only one receive antenna. When N_r receive antennas are used, the diversity gain is just N_r times of the diversity gain when one receive antenna is used.

From (4.23), the Euclidean distance between two codewords $\mathbf{X}$ and $\mathbf{X}'$ is given by

$$d^2(\mathbf{X},\mathbf{X}') = \frac{P}{N_t N}\sum_{i=1}^{N}\left\|\mathbf{\Lambda}_{m,i}(\mathbf{x}_i - \mathbf{x}_i')\right\|^2 \tag{4.32}$$

where $\mathbf{x}_i$ and $\mathbf{x}_i'$ are the ith columns of $\mathbf{X}$ and $\mathbf{X}'$, respectively. Let U be the union of the N difference sets, i.e.,

$$U \in \bigcup_{i=1}^{N} U_i \tag{4.33}$$

and assume that the elements, $\{i_u, u = 1,2,\ldots,|U|\}$, in U are ordered in an ascending order. Then N diagonal matrices of size $|U|$ can be defined as

$$\mathbf{E}_i = \mathrm{diag}\left([x_i(i_1) - x_i'(i_1)\ \ x_i(i_2) - x_i'(i_2)\ \cdots\ x_i(i_{|U|}) - x_i'(i_{|U|})]\right) \tag{4.34}$$

where $x_i(k)$ and $x_i'(k)$ are the kth entries in $\mathbf{x}_i$ and $\mathbf{x}_i'$, respectively. It follows from (4.24) and (4.32) that

$$\begin{aligned} d^2(\mathbf{X},\mathbf{X}') &= \frac{P}{N_t N}\sum_{i=1}^{N}\left\|\mathbf{E}_i\hat{\mathbf{F}}(U,:)\mathbf{W}_i\hat{\mathbf{h}}\right\|^2 \\ &= \frac{P}{N_t N}\hat{\mathbf{h}}^H\mathbf{G}^H\mathbf{E}^H\mathbf{E}\mathbf{G}\hat{\mathbf{h}} \end{aligned} \tag{4.35}$$

where $\hat{\mathbf{F}}(U,:)$ is the matrix collecting rows of $\hat{\mathbf{F}}$ with indexes in U,

$$\mathbf{G} = \left[\mathbf{W}_1^H\hat{\mathbf{F}}(U,:)^H\ \ \mathbf{W}_2^H\hat{\mathbf{F}}(U,:)^H\cdots\ \mathbf{W}_N^H\hat{\mathbf{F}}(U,:)^H\right]^H \tag{4.36}$$

and

$$\mathbf{E} = [\mathbf{E}_1\ \mathbf{E}_2\cdots\ \mathbf{E}_N]. \tag{4.37}$$

Therefore, the pair-wise diversity gain is given as

$$G_d(\mathbf{X},\mathbf{X}') = \min\left(N_t L, \mathrm{rank}(\mathbf{G}^H\mathbf{E}^H\mathbf{E}\mathbf{G})\right). \tag{4.38}$$

Let S be a subset of U satisfying

$$\sum_{i=1}^{N} |S \cup U_i| \le N_t L . \tag{4.39}$$

Then there exists a permutation matrix $\mathbf{P}$ such that

$$\mathbf{PE} = \left[\mathbf{E}_a^H \mathbf{E}_b^H\right]^H \tag{4.40}$$

where $\mathbf{E}_a$ collects rows of $\mathbf{E}$ associated with symbol indexes $t \in S$. By (4.39), $\mathbf{E}_a$ is of size $|S| \times |U| N$. Let $q = \exp(-j2\pi)$, then it follows from (4.28) that the (u,k)th entry in $\hat{\mathbf{F}}(U,:)\mathbf{W}_i$ can be written as

$$g_i(u,k) = \begin{cases} q^{(i_u-1)(k-1)/K - i(k-1)/N_t} & \text{for } k \le N_t L - N_t + 1 \\ q^{(i_u-1)(k+K-N_t L-1)/K - i(k-1)/N_t} & \text{for } k > N_t L - N_t + 1 \end{cases} . \tag{4.41}$$

The above equation shows that any $N_t L$ rows from matrix $\mathbf{G}$ make up a generalized Vandermonde matrix and, thus, are mutually linear independent of each other. Now for an arbitrary vector $\mathbf{a} \in C^{|S| \times 1}$, it follows from (4.39) that $\mathbf{a}^H \mathbf{E}_a$ has at most $N_t L$ nonzero entries. Consequently

$$\mathbf{a}^H \mathbf{E}_a \mathbf{G} \neq 0 . \tag{4.42}$$

This shows that $\text{rank}(\mathbf{E}_a \mathbf{G}) = |S|$. By Equation (4.35) and the fact that $\text{rank}(\mathbf{G}^H \mathbf{E}^H \mathbf{E} \mathbf{G}) \ge \text{rank}(\mathbf{G}^H \mathbf{E}_a^H \mathbf{E}_a \mathbf{G})$, Theorem 4.2 has been proved. □

If the same code is used for both the conventional spatial multiplexing and the proposed multiplexing schemes and a pair of codewords may differ at arbitrary locations due to a random interleaving, it can be shown that the lower bound of the proposed

scheme is always larger than that of the conventional schemes. For codes with an effective length greater than or equal to $N_t L$, the lower bound of the diversity gain of the conventional spatial multiplexing schemes is $N_r L$ and that of the proposed scheme is $\frac{N_r L N_t}{N}$. Note that the above lower bounds are obtained without considering the finite alphabet of transmitted symbols. Practical alphabet often leads to larger diversity gain. For instance, we have found that the proposed scheme always achieves full diversity, $N_t N_r L$, for BPSK and QPSK modulations. In fact, full diversity gain for the proposed scheme can be guaranteed for any constellations carved from the Gaussian integer ring by appropriate constellation rotations similar to that in [11]. On the other hand, the above lower bounds are indicative in that extra nonzero eigenvalues are often significantly smaller than the first group of eigenvalues indicated by the bounds. Because of this, constellation rotation is not pursued. In addition, the diversity loss only occurs when the number of layers, N, is greater than 1, which implies $N_t > 1$ and $N_r > 1$. When $\frac{N_r L N_t}{N}$ is reasonably large, say greater than 4, the benefit of extra diversity gain diminishes and a more appropriate performance index is the Euclidean distance [51, 56, 57]. Furthermore, if the data rate is fixed against the channel capacity as SNR increases, full diversity is impossible [8].

Since one of the main motivations of using multiple layers is to provide high data rate transmission with reasonable complexity. Therefore, suboptimal detection schemes rather than the optimal ML detection schemes are of primary interest. Various suboptimal multi-user detection schemes, such as decorrelating detection, MMSE detection, and combined successive cancellation and linear detection, can be used to separate the layers

such that independent decoding can be applied for each layer [6, 7]. Or iterative joint detection and decoding schemes can be used for better performance [58].

It is worthy of noting that the N layers do not have to be independently encoded. In fact, the proposed OFDM modulation and multiplexing schemes together create an N-input N_r-output wideband channel. Any MIMO technique developed to be used in conjunction with OFDM such as that in [23] can also be used here.

4.4.3 Channel Capacity and Outage Probability

For the above channel multiplexing scheme, we establish the following theorem that shows the channel capacity.

Theorem 4.3: At high SNR, the ergodic capacity (in bps/Hz) of the N-input N_r-output channel created by using a set of N orthogonal weight vectors satisfying (4.26) is given by

$$C(N) = \min(N, N_r)\log\left(\frac{P}{N_0}\right) + \sum_{i=|N-N_r|+1}^{\max(N,N_r)} E\left[\log \chi_{2i}^2\right] + o(1) \tag{4.43}$$

where χ_{2i}^2 has a Chi-square distribution with $2i$ degrees of freedom.

Proof: Note that the channel capacity is the average capacity of the K orthogonal subchannels. For each subchannel, the channel matrix consists of entries as samples of independent zero-mean Gaussian random variables with variance N_t as shown in (4.27). The ergodic capacity of such an N-input N_r-output channel is well-known and given by (4.43) [1] [2]. Thus, the capacity of the MIMO OFDM channel is also given by (4.43).

This has proved Theorem 4.3. □

As can be seen, an N-layer system with $N \le N_r$ achieves a multiplexing gain N. Note that, in the above, the loss in bandwidth efficiency due to cyclic prefixes has been neglected. When cyclic prefix is considered, the loss in bandwidth efficiency is $\frac{(L-1)N_t}{K+(L-1)N_t}$.

To study the outage probability, we took a Monte-Carlo approach and randomly generated ten thousand channel realizations. With $N_t = N_r = L = 4$, the outage probabilities with the conventional MIMO OFDM scheme [6] and the proposed multiplexing scheme are compared in Figure 4.4. For both of them, the optimal receiver was assumed. As expected, when the number of layers equals 4, the outage probability curves of the two schemes coincide with each other. When the number of layers is less than 4, only a subset of the transmit antennas were used in the conventional approach. As a result, the outage probability of the conventional approach is consistently larger than that of the proposed multiplexing scheme. It can also be observed from the figure that the slope of the outage curve becomes steeper as the number of layers decreases. This shows that the available diversity gain increases as the multiplexing gain decreases [8].

4.5 Simulation Results

In this section, we provide simulation results to compare the proposed approach with the conventional approach including both diversity and multiplexing schemes. In all simulations below, channels with i.i.d path gains were assumed and the OFDM frame length was 64. A data frame is considered as the collection of all the information bits

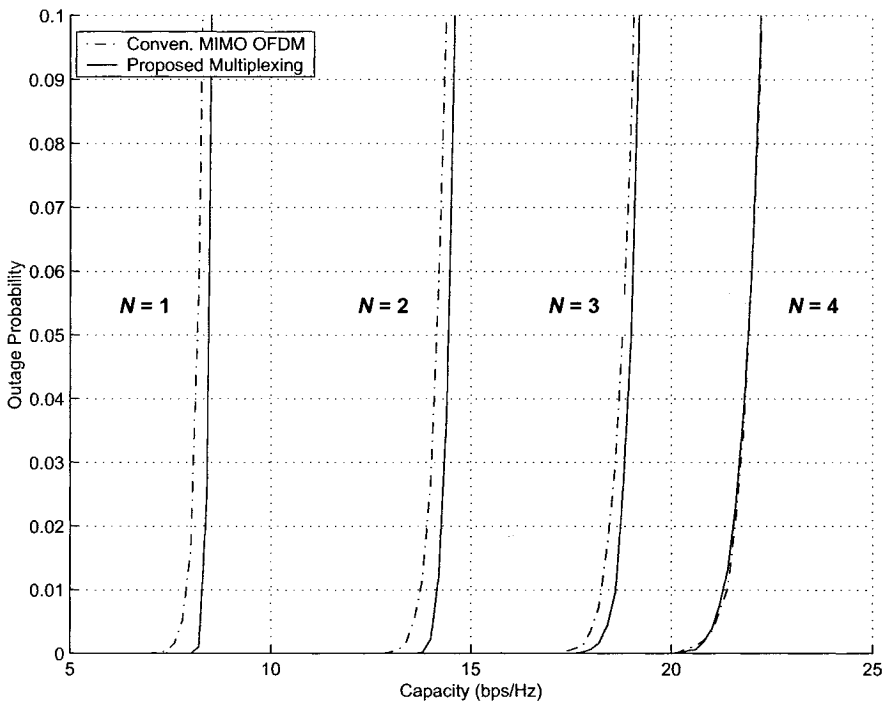

Figure 4.4. Outage probabilities of the conventional MIMO OFDM schemes and the proposed multiplexing scheme.

transmitted over one OFDM frame. In addition, block interleavers with appropriate interleaving depths were employed to maximize the coding gain unless mentioned otherwise.

As a first example, the proposed ST-OFDM in conjunction with conventional trellis codes for single-input channels is compared with STC-OFDM. Two transmit and one receive antennas were assumed. The channel has 4 equal-gain resolvable paths. For STC-OFDM, two space-time trellis codes with 8 and 32 states and 4-PSK constellation given in [3] were considered. It was conjectured that these two codes are optimal 2-ST trellis codes for flat quasi-static fading channels in the sense that they have the maximal coding gains [3]. For the given frequency-selective channel, the two codes provide diversity gain 2 and 3, respectively. To ensure roughly equal data rate as that of the STC-OFDM scheme, we used 2/3 trellis codes with 8-PSK constellation [53] for the proposed scheme. Note that the decoding complexity of a 4-PSK 2-ST trellis codes will be about twice that of a conventional 2/3-rate 8-PSK trellis code with the same number of states. To ensure fairness, two trellis codes with 16 and 64 states were chosen to compare with the ST codes with 8 and 32 states, respectively. The effective lengths and, hence, the diversity gains of the two conventional codes are, respectively, 3 and 4. In Figure 4.5, the resulting Frame Error Rate (FER) is plotted against the SNR at each receive antenna. As can be observed, the 8-state ST code yields the worst performance while the proposed ST-OFDM with the 64-state trellis code provides the best performance. In addition, the performance gap becomes larger as SNR increases due to different diversity gains.

As a second example, we compare the proposed multiplexing scheme with D-BLAST [6] and Ma's scheme [11]. Note that D-BLAST was shown to provide the optimal

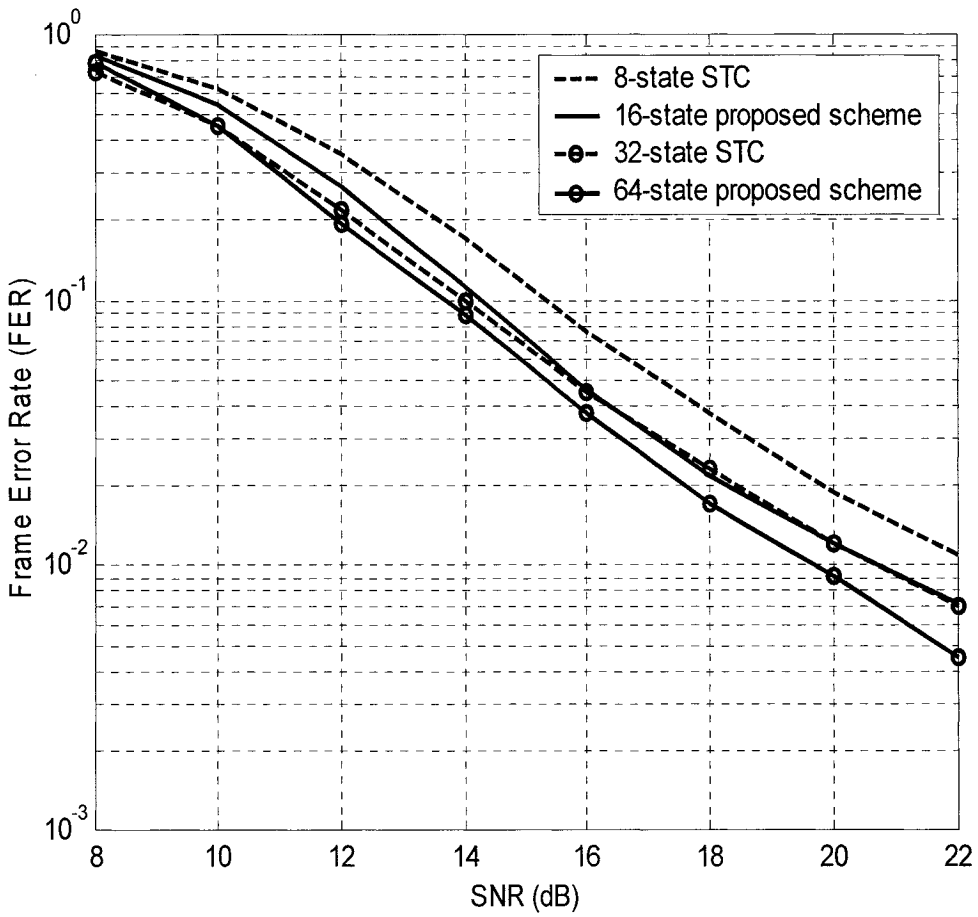

Figure 4.5. Performance comparison of STC-OFDM and proposed coded ST-OFDM: 2 transmit, 1 receive antennas, and 4 resolvable paths.

multiplexing-diversity tradeoff as long as the outer encoder is powerful enough [8]. Two layers were created over a channel with two transmit and two receive antennas and $L = 2$ equal-gain paths. A data frame of 128 bit is assigned to two layers and transmitted over one OFDM symbol. For the proposed scheme and D-BLAST, an 8-state QPSK trellis code with an effective length 4 was used for both layers. The minimum squared product distance and the minimum Euclidean distance are 64 and 12 respectively. No outer code was employed for Ma's scheme because excessively high complexity of decoding. Interleaving was not applied and ML decoding was employed for all the three schemes. The FER of the three schemes is plotted against SNR in Figure 4.6. As can be observed, the proposed scheme and Ma's scheme have the same diversity gain while D-BLAST has a smaller diversity gain. However, within the SNR range of concern, both the proposed scheme and D-BLAST significantly outperform Ma's scheme because of their larger coding gain. The performance gap between the proposed scheme and Ma's scheme at $\mathrm{FER} = 4 * 10^{-3}$ is about 3.5 dB. This demonstrates the advantage of trellis codes over block codes in terms of coding gain.

As a third example, we compare the proposed multiplexing scheme with D-BLAST using two transmit and two receive antennas with a suboptimal receiver. At the receiver, an MMSE detector is first employed to separate the signals from different layers and then independent decoding is performed for each layer. The conventional trellis code with 16 states with an effective length 3 [53] was used for both multiplexing schemes. The resulting FER curves are given in Figure 4.7. As can be seen, increasing the number of resolvable paths L from 2 to 4 has less effect on the performance of the proposed scheme. This is because the diversity gain in the proposed scheme is 3×2 if ML detection is used,

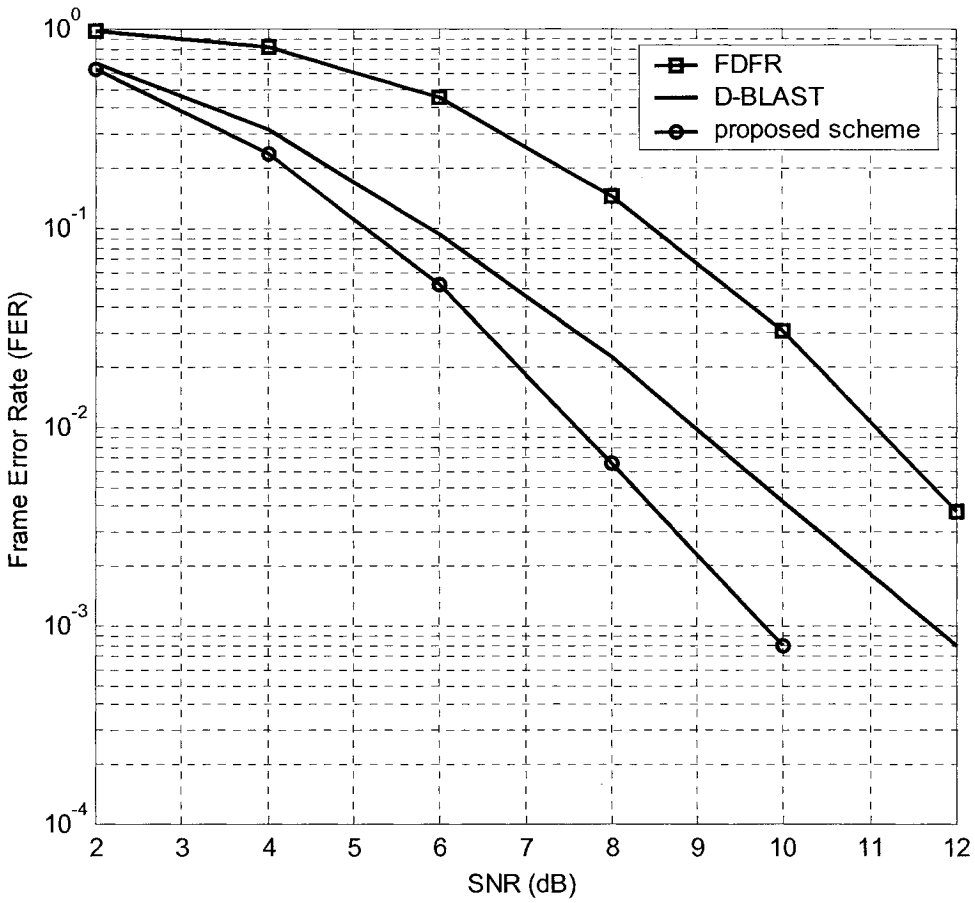

Figure 4.6. Performance comparison of the proposed multiplexing scheme, D-BLAST, and MA's scheme: 2 transmit and 2 receive antennas, 2 resolvable paths, 2 layers, and in total 2 bits per channel use, ML decoding.

regardless of the number of resolvable paths. However, in D-BLAST, the diversity gain will be only 2×2 when the number of resolvable paths is 2. Since MMSE was used, it is expected that the actual diversity gains will be less than those provided by ML detection. However, we believe that the diversity gain of the proposed scheme will still be larger than that of D-BLAST when $L = 2$. Comparing the proposed scheme with D-BLAST when $L = 4$, we can see that the two curves are in parallel. This can be explained as follows: although both schemes have the same diversity gain, the proposed scheme has less low-diversity codeword pairs (see Theorems 4.1 and 4.2).

To show the advantage of multiplexing over diversity in high data-rate applications, the FER curve of a 16-PSK 16-state ST trellis code is also plotted in Figure 4.7. The code is a delay-diversity code, i.e., the signal transmitted over antenna 2 will be the signal over antenna 1 with a delay of one symbol duration. Apparently, this ST code has a diversity gain of 2×2 when $L = 2$ and provides the same data rate as the above multiplexing schemes. The decoding complexity is also comparable to that of the multiplexing schemes. However, its performance is significantly worse than that of the multiplexing schemes including both D-BLAST and the proposed multiplexing scheme. Note that the FER curve of the 16-PSK ST code has a steeper slope because a diversity gain exactly 4 has been achieved.

Last, we provide some more simulation results of multiplexing schemes in Figure 4.8. It was assumed that there were 4 transmit and 4 receive antennas and 4 resolvable equal-gain paths. Using the proposed scheme and D-BLAST with the 16-state trellis code, $N = 2$ and $N = 4$ layers are created. As expected, the proposed scheme outperforms D-BLAST in both cases. It can also be observed that the performance of 2-layer cases is

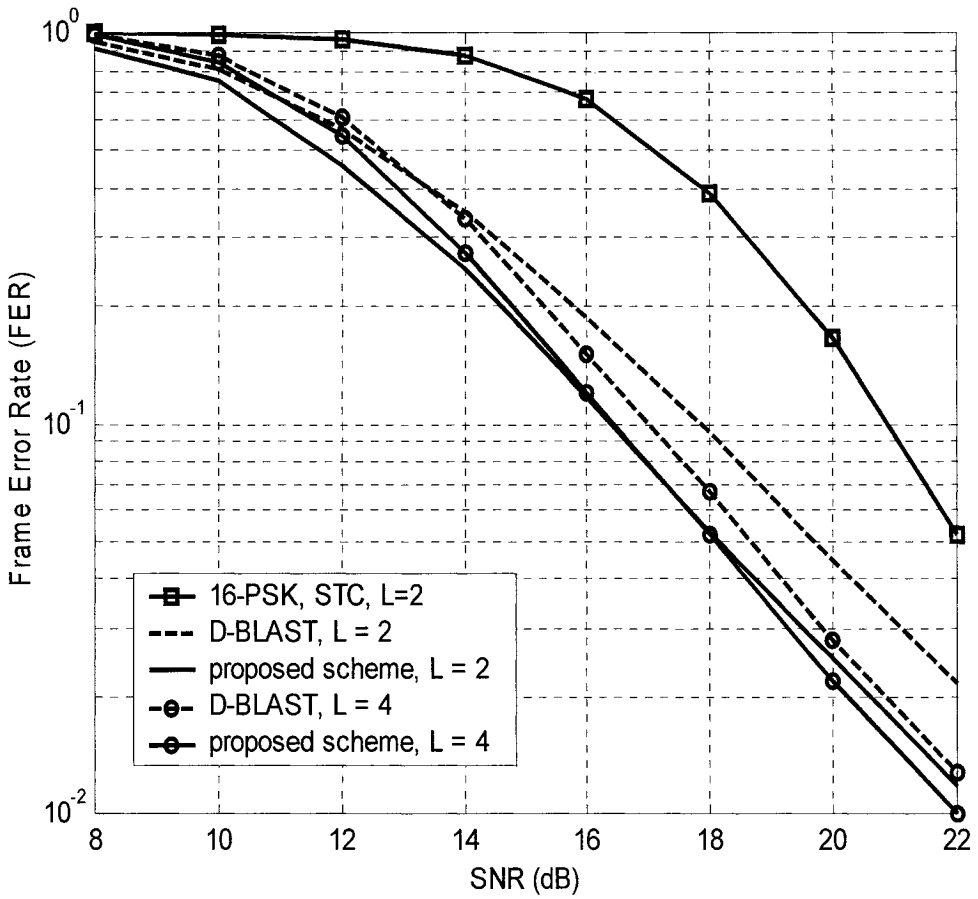

Figure 4.7. Performance comparison of D-BLAST, proposed multiplexing scheme, and STC-OFDM: 2 transmit and 2 receive antennas, MMSE receiver for multiplexing schemes.

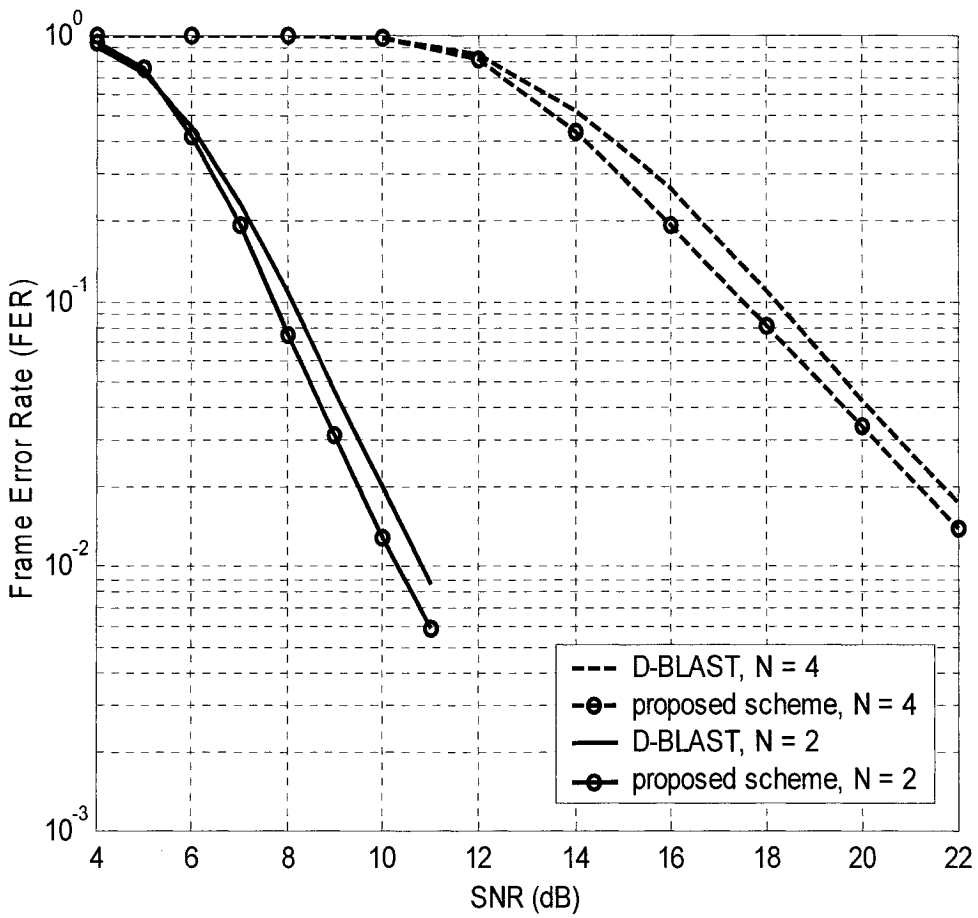

Figure 4.8. Performance comparison of D-BLAST and proposed multiplexing scheme: 4 transmit and 4 receive antennas, MMSE receiver.

significantly better than that of 4-layer cases and the gap becomes larger as SNR increases. At an FER of 0.02, the performance gap is more than 10 dB. Of course, this performance gain is obtained at the cost of half the data rate. This and the second example demonstrate the importance of tradeoff designs.

4.6 Conclusions

In this chapter, we have introduced an integrated approach to achieving both diversity and multiplexing gains for MIMO frequency-selective channels. The proposed approach is based on a new ST-OFDM modulation scheme that translates a MIMO channel into a SIMO channel. As a result, the design of codes to achieve diversity gain becomes easier and conventional codes designed for single-input fast-fading channels can be used. In addition, we developed a new multiplexing scheme to be used with the proposed ST-OFDM. In the proposed multiplexing scheme, each layer sees all the transmit antennas all the time. This ensures that all the available diversity gain is preserved for each layer. It was shown that the proposed multiplexing scheme achieves a multiplexing gain that equals the number of layers created. This, together with coding, allows a full range of optimal tradeoffs between data rate and reliability. Simulation results have demonstrated the flexibility in tradeoff and the superior performance of the proposed approach compared to the traditional MIMO OFDM approach.

Chapter 5

A Pilot-Symbol-Aided Channel Estimator for OFDM Wireless Communications

5.1 Introduction

OFDM is an important technique for high data-rate wireless communications [59-66] and finds applications in many communication standards such as ETSI, IEEE 802.11a and HIPPERLANII, DAB, DVB, BRAN, ARIB and MMAC. The use of OFDM transfers a dispersive wideband channel into a set of flat narrowband channels, so as to avoid the ISI that otherwise exists in single-carrier transmission over wideband channels. Recently, OFDM in conjunction with space-time coding [16-19] has also attracted a great deal of attention as a suitable technique for realizing the high channel capacity provided by the MIMO fading channels.

In OFDM systems, channel estimation is an important task. The knowledge about a

channel at the receiver allows coherent demodulation, which provides a 3 dB performance gain as compared to differential demodulation [67]. However, channel estimation for OFDM systems is not trivial. In fact, it can be a difficult task for frequency-selective MIMO channels because potentially a large number of channel taps are to be estimated.

One way to estimate the channel for OFDM systems is to use decision feedback of information data to track the channel variation after initial training [25-31]. Decision-directed schemes of this kind suffer from significant degradation in performance when the channel varies fast. To overcome this problem, pilot symbols can be inserted at frequency (tone) and time in a regular fashion. At the receiver, channel coefficients at pilot locations can be readily estimated based on known pilot symbols. Then, these estimates of coefficients are interpolated to obtain channel coefficients at other locations. If the 2-D Nyquist sampling condition is satisfied [68], channel coefficients at other locations can be recovered without error from those at pilot locations. Because of the frequent appearance of pilot symbols in time, pilot-symbol-aided schemes are more capable of tracking the channel time variation compared to the decision-directed parameter estimation techniques. Various 1-D interpolation filters have been proposed [32, 33]. By contrast, 2-D filtering was proposed [34, 35] to further exploit the correlation of channel coefficients associated with different tones at different time. In [34, 69] it was shown that 2-D channel estimation significantly outperforms 1-D channel estimation with respect to overhead (pilot density), MSE performance, and latency. A prominent drawback of these schemes is that they require channel statistics that are usually unknown. For this reason, a robust channel interpolator that requires no channel statistics

was proposed in [35], at the cost of notable performance degradation due to noise enhancement. Another drawback of these filter-based schemes is that the filtering, particularly 2-D filtering, involves large amount of computation.

In this chapter, our goal is to develop a computationally efficient channel estimator for OFDM wireless communication systems, particularly for the use of the new OFDM MIMO scheme proposed in Chapter 4. To reduce the complexity in the aforementioned conventional schemes based on 2-D filtering, we propose a pilot-symbol-aided channel estimator for OFDM wireless communication systems based on 2-D weighting. The proposed scheme is highly robust when the channel varies fast. Although, the performance of the proposed estimator is equivalent to the 2-D interpolation filter presented in [35], the proposed scheme is much simpler in terms of complexity. The reduction in complexity is achieved by employing the 2-D IFFT, FFT and a 2-D weighting function instead of a 2-D interpolating filter. For cases where channel statistics are not available, we propose a robust estimator that involves a simple 2-D windowing function. In addition, we propose an enhanced channel estimator that can further improve the performance of the robust estimator. Several examples show that the proposed simple windowing function is highly effective for realistic situations when the channel statistics are not available.

5.2 Preliminaries

Figure 5.1 depicts the block diagram of an OFDM wireless communication system. The baseband impulse response of the multi-path fading channel can be expressed

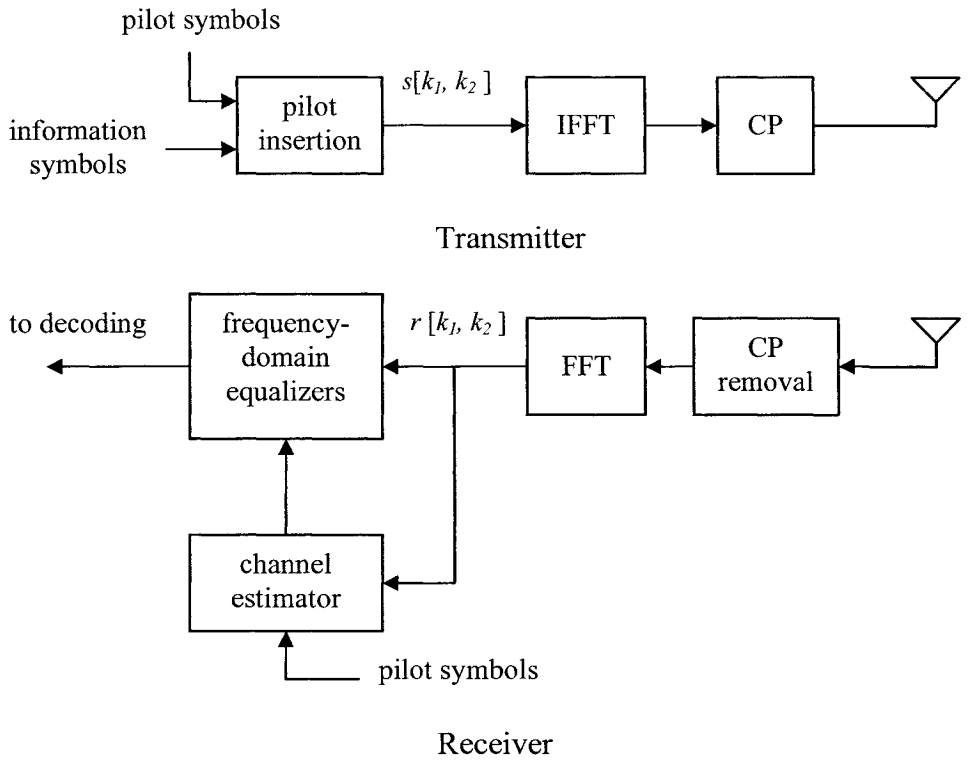

Figure 5.1. The baseband model of an OFDM wireless communication system with pilot-symbol-aided channel estimator.

as [70, 71]

$$h(t,\tau) = \sum_l \alpha_l(t) \exp\{j2\pi f_c \tau_l(t) + j2\pi f_{D,l}(t)[t - \tau_l(t)]\}\delta[\tau - \tau_l(t)] \qquad (5.1)$$

where f_c is the carrier frequency, and $\alpha_l(t), \tau_l(t)$ and $f_{D,l}(t)$ are the real gain, time delay and the Doppler shift of the lth multipath component at time t, respectively.

From (5.1), the impulse response of the channel can be rewritten as

$$h(t,\tau) = \sum_l \overline{\alpha}_l(t)\delta[\tau - \tau_l(t)] \tag{5.2}$$

where $\overline{\alpha}_l(t)$ is the complex path gain at time t. Under the assumption of slow fading, $\overline{\alpha}_l(t)$ can be treated as constant over the duration of a frame, but changes from frame to frame depending on the Doppler shift. Furthermore, the path gains are assumed to be samples of narrow-band zero-mean complex Gaussian random variables that are mutually independent. At time t, the frequency response of the wireless channel presented in (5.2) is defined as

$$H(t,f) = \int_{-\infty}^{\infty} h(t,\tau)e^{-j2\pi f\tau}d\tau = \sum_l \overline{\alpha}_l(t)e^{-j2\pi f\tau_l} \ . \tag{5.3}$$

From (5.3), the frequency response of the wireless channel at tone k_2 of OFDM frame k_1 can be defined as

$$H[k_1,k_2] = H(k_1T_s, k_2\Delta f) = \sum_l \overline{\alpha}_l(k_1T_s)e^{-j2\pi k_2\Delta f\tau_l} \tag{5.4}$$

where T_s and Δf are the frame duration and the tone spacing, respectively.

In OFDM systems, information bits are transmitted frame by frame. For each frame, the K-point IFFT is first applied to the information and pilot symbols and then a cyclic prefix (CP) of length L_c-1 is inserted before the signal being transmitted. At the receiver, the CP is removed and the remaining K samples are sent to the FFT. If L_c is greater than or equal to the length of the channel impulse response, the channel matrix becomes block circulant and there is no interference among tones. As such, the received signal at the output of the FFT associated with tone k_2 of frame k_1 can be written as

$$r[k_1,k_2] = H[k_1,k_2]s[k_1,k_2] + z[k_1,k_2] \tag{5.5}$$

where $s[k_1,k_2]$ is the transmitted symbol at tone k_2 of frame k_1, and $z[k_1,k_2]$ is the FFT of the channel noise at tone k_2 of frame k_1. If the original channel noise is white Gaussian with zero-mean and variance σ_z^2, it can be shown that $z[k_1,k_2]$ is zero-mean complex white Gaussian noise with variance σ_z^2 and $E\{z^*[k_1,k_2]z[k_1',k_2']\}=0$ for any $(k_1,k_2)\neq(k_1',k_2')$ where superscript * denotes conjugate operation.

If pilot symbols are inserted periodically, temporal estimates of the dispersive channel parameters at pilot locations can be obtained from (5.5) as

$$\tilde{H}[k_{1p},k_{2p}]=r[k_{1p},k_{2p}]s^*[k_{1p},k_{2p}]=H[k_{1p},k_{2p}]+z[k_{1p},k_{2p}]s^*[k_{1p},k_{2p}]$$

$$=H[k_{1p},k_{2p}]+\tilde{z}[k_{1p},k_{2p}] \tag{5.6}$$

where k_{1p} and k_{2p} represent the frame and tone indices of pilot locations, respectively. With these temporal estimates, channel coefficients at other locations may be obtained by using a 2-D upsampling filter [35]. A prominent drawback of this approach is that the 2-D filtering involves intensive computation.

5.3 Proposed Channel Estimation Scheme

In this section, we describe the proposed pilot-symbol-aided channel estimator, which uses a weighting function in the delay and frequency domain.

5.3.1 Channel Estimation Based on MMSE Weighting

Figure 5.2. shows the block diagram of the proposed pilot-symbol-aided channel

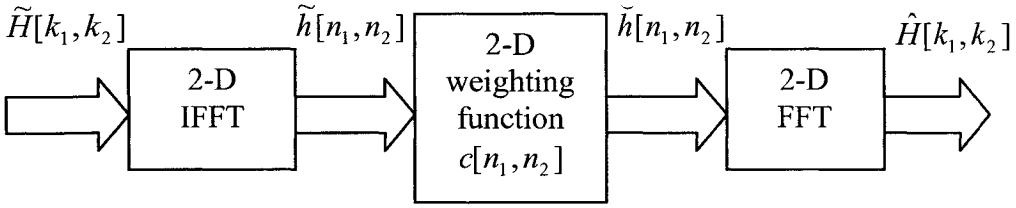

Figure 5.2. Block diagram of the proposed pilot-symbol-aided channel estimator for OFDM wireless communication systems.

estimator. The proposed estimator is based on the observation that a 2-D filtering is equivalent to a 2-D weighting function at the transformed domains. The proposed scheme works block by block with each block covering a number of OFDM frames. For each processing block, temporal estimates of channel parameters at pilot locations are first obtained as in (5.6), and zeros are inserted at other locations to form an up-sampled signal $\tilde{H}[k_1,k_2]$. This signal is sent to the 2-D IFFT as shown in Figure 5.2. The output signal of the 2-D IFFT is therefore at the delay and frequency domains corresponding to tones and frames, respectively. Then, a 2-D weighting function is applied and the weighted 2-D signal is sent to the 2-D FFT to provide the channel estimation at every tone and frame $\hat{H}[k_1,k_2]$.

Let us denote the two-dimensional signal $\tilde{H}[k_1,k_2]$ by $\tilde{H}[\mathbf{k}]$, where the two-element vector is given by $\mathbf{k}=[k_1,k_2]^T$. Now let us define $\mathbf{n}=[n_1,n_2]^T$ and the sampling matrix

$$\mathbf{V}=\begin{bmatrix} N_f & 0 \\ 0 & K \end{bmatrix} \tag{5.7}$$

with K and N_f denoting OFDM frame length and number of frames in each processing block, respectively. Then, the output of the 2-D IFFT can be written as

$$\tilde{h}[\mathbf{n}] = \frac{1}{|\det \mathbf{V}|}\sum_{\mathbf{k}} \tilde{H}[\mathbf{k}]\exp(j2\pi \mathbf{k}^T \mathbf{V}^{-1}\mathbf{n}) = h_p[\mathbf{n}] + \tilde{z}[\mathbf{n}] \tag{5.8}$$

where

$$h_p[\mathbf{n}] = \left\{ \frac{1}{|\det \mathbf{V}|}\sum_{\mathbf{k}} H_p[\mathbf{k}]\exp(j2\pi \mathbf{k}^T \mathbf{V}^{-1}\mathbf{n}) \right\} \tag{5.9}$$

with

$$H_p[\mathbf{k}] = \begin{Bmatrix} H[\mathbf{k}] & \text{when } \mathbf{k} = \mathbf{k}_p \text{ (at pilot locations)} \\ 0 & \text{otherwise} \end{Bmatrix}. \tag{5.10}$$

The 2-D signal at the output of weighting function is given by

$$\breve{h}(\mathbf{n}) = c[\mathbf{n}]\tilde{h}[\mathbf{n}] \tag{5.11}$$

where $c[\mathbf{n}]$ is the 2-D weighting function. The optimal MMSE 2-D weighting function can be obtained by minimizing

$$\text{MSE} = \det(\mathbf{V}) E\left\{ \left| h[\mathbf{n}] - c[\mathbf{n}]\tilde{h}[\mathbf{n}] \right|^2 \right\}. \tag{5.12}$$

By letting the derivative of the MSE in (5.12) be zero and solving it with respect to $c[\mathbf{n}]$, we have the optimal weighting function as

$$c[\mathbf{n}] = \frac{E\{h[\mathbf{n}]\tilde{h}^*[\mathbf{n}]\}}{E\{|\tilde{h}[\mathbf{n}]|^2\}}. \tag{5.13}$$

Using (5.8), (5.13) can be written as

$$c[\mathbf{n}] = \frac{E\{h[\mathbf{n}]h_p^*[\mathbf{n}]\}}{E\{|h_p[\mathbf{n}]|^2\} + \rho} \tag{5.14}$$

where $\rho = E\{\tilde{z}^2[\mathbf{n}]\}$.

In (5.14),

$$E\{h[\mathbf{n}]h_p^*[\mathbf{n}]\} = E\left\{\frac{1}{|\det \mathbf{V}|^2}\sum_{\mathbf{k}} H[\mathbf{k}]\exp(2\pi\mathbf{k}^T\mathbf{V}^{-1}\mathbf{n}).\sum_{\mathbf{k}} H_p^*[\mathbf{k}]\exp(-j2\pi\mathbf{k}^T\mathbf{V}^{-1}\mathbf{n})\right\}$$

$$= E\left\{\frac{1}{|\det \mathbf{V}|^2}\sum_{\mathbf{k}}\sum_{\mathbf{k}'} H[\mathbf{k}]H_p^*[\mathbf{k}']\exp(2\pi[\mathbf{k}-\mathbf{k}']^T\mathbf{V}^{-1}\mathbf{n})\right\}. \tag{5.15}$$

By substituting $\mathbf{m} = \mathbf{k} - \mathbf{k}'$,

$$E\{h[\mathbf{n}]h_p^*[\mathbf{n}]\} = \left\{\frac{1}{|\det \mathbf{V}|^2}\sum_{\mathbf{k}}\sum_{\mathbf{m}} E\{H[\mathbf{k}]H_p^*[\mathbf{k}-\mathbf{m}]\}\exp(2\pi\mathbf{m}^T\mathbf{V}^{-1}\mathbf{n})\right\}$$

$$= \left\{\frac{1}{|\det \mathbf{V}|}\sum_{\mathbf{m}} R_{\mathbf{m}}\exp(2\pi\mathbf{m}^T\mathbf{V}^{-1}\mathbf{n})\right\} \tag{5.16}$$

where $R_{\mathbf{m}} = E\{H[\mathbf{k}]H_p^*[\mathbf{k}-\mathbf{m}]\}$ is the correlation function between $H[\mathbf{k}]$ and $H_p[\mathbf{k}]$.

In similar manner,

$$E\{|h_p[\mathbf{n}]|^2\} = \left\{\frac{1}{|\det \mathbf{V}|}\sum_{\mathbf{m}} R_{p\mathbf{m}}\exp(2\pi\mathbf{m}^T\mathbf{V}^{-1}\mathbf{n})\right\} \tag{5.17}$$

where $R_{p\mathbf{m}} = E\{H_p[\mathbf{k}]H_p^*[\mathbf{k}-\mathbf{m}]\}$ is the autocorrelation function of $H_p[\mathbf{k}]$.

The MSE for the proposed MMSE estimator can be expressed as

$$\text{MSE} = \det(\mathbf{V}) E\left\{\left|h[\mathbf{n}] - c[\mathbf{n}]\tilde{h}[\mathbf{n}]\right|^2\right\}. \tag{5.18}$$

By substituting (5.8) and using the orthogonality principle [72],

$$\begin{aligned}\text{MSE} &= \det(\mathbf{V}) E\left\{\left|h[\mathbf{n}]\right|^2 - c[\mathbf{n}]h_p[\mathbf{n}]h^*[\mathbf{n}]\right\} \\ &= 1 - \det(\mathbf{V}) E\{c[\mathbf{n}]h_p[\mathbf{n}]h^*[\mathbf{n}]\}.\end{aligned} \tag{5.19}$$

From (5.14),

$$\begin{aligned}\text{MSE} &= 1 - \det(\mathbf{V}) \frac{E\{h[\mathbf{n}]h_p^*[\mathbf{n}]\}}{E\{\left|h_p[\mathbf{n}]\right|^2\} + \rho} E\{h_p[\mathbf{n}]h^*[\mathbf{n}]\} \\ &= 1 - \det(\mathbf{V}) \frac{\left|E\{h[\mathbf{n}]h_p^*[\mathbf{n}]\}\right|^2}{E\{\left|h_p[\mathbf{n}]\right|^2\} + \rho}.\end{aligned} \tag{5.20}$$

Note, in the above, the channel has been assumed to be wide-sense stationary so that an MMSE estimator based on Wiener theory is optimal in the sense of MSE. More robust estimation schemes can be developed based on the Kalman filtering [73] for non stationary channels, however, at the cost of degraded performance for realistic channels that are closely stationary.

It is interesting at this point to compare the 2-D filtering and the proposed scheme in terms of complexity. If the channel statistics are known, a 2-D interpolating filter can be derived. However, this 2-D interpolating filter will need KN_f multiplications for each output sample. Of course, simplifications can be made, for instance, by shortening the filter length, but this is at the cost of performance degradation. Note that the 2-D FFT of size KN_f requires $\frac{1}{2}\log_2 KN_f$ multiplications per sample. Hence, the proposed scheme

(2-D IFFT, 2-D weighting, and 2-D FFT) require $\log_2(KN_f)+1$ multiplications per sample. This means the proposed scheme has much reduced computational complexity compared to a 2-D interpolating filter.

5.3.2 Robust (Window) Estimator

In the previous subsection, we define the MMSE estimator based on the channel statistics in (5.16) and (5.17). However channel statistics are often not available in practice. For such cases a robust channel estimator may be obtained by taking note of the following observations. First, $h[\mathbf{n}]$ will take zero values when n_2 is greater than or equal to the number of resolvable paths L, which is in fact often far smaller than the FFT length K. Furthermore, if the pilot sampling is dense enough, $h[\mathbf{n}] = M_f h_p[\mathbf{n}]$ for $n_2 \le L$ where M_f is the spacing between pilot symbols in the frequency domain. By ignoring the noise in (5.14), the weighting function $c[\mathbf{n}]$ becomes flat window function. Therefore, we propose a robust estimator with the windowing function given by

$$\tilde{c}[\mathbf{n}] = \begin{Bmatrix} M_f, & n_2 \le L \\ 0, & \text{otherwise} \end{Bmatrix}. \tag{5.21}$$

If the number of resolvable paths is unknown, the length of cyclic prefix can be used.

The MSE for the proposed robust (window) estimator can be expressed as

$$\text{MSE} = \det(\mathbf{V}) E\left\{ \left| h[\mathbf{n}] - \tilde{c}[\mathbf{n}]\tilde{h}[\mathbf{n}] \right|^2 \right\}. \tag{5.22}$$

By using (5.8) and (5.21), we obtain

$$\begin{aligned}\mathrm{MSE} = \det(\mathbf{V})\{E\{|h[\mathbf{n}]|^2\} - M_f E\{h[\mathbf{n}]h_p^*[\mathbf{n}]\} - M_f E\{h^*[\mathbf{n}]h_p[\mathbf{n}]\} \\ + M_f^2 E\{|h_p[\mathbf{n}]|^2\} - M_f^2 E\{\tilde{z}^2[\mathbf{n}]\}\},\end{aligned} \tag{5.23}$$

for $n_1 \le N_f$ and $n_2 \le L$.

In the above expression, for $n_1 \le N_f$ and $n_2 \le L$,

$$E\{\tilde{z}^2[\mathbf{n}]\} = \frac{\rho L}{K} \tag{5.24}$$

and

$$E\{|h[\mathbf{n}]|^2\} = M_f E\{h[\mathbf{n}]h_p^*[\mathbf{n}]\} \tag{5.25}$$

with

$$E\{h[\mathbf{n}]h_p^*[\mathbf{n}]\} = E\{h^*[\mathbf{n}]h_p[\mathbf{n}]\} = M_f E\{|h_p[\mathbf{n}]|^2\}. \tag{5.26}$$

Then,

$$\mathrm{MSE} = \frac{\det(\mathbf{V}) M_f^2 \rho L}{K} = \frac{\sigma_z^2 L M_f}{K} = \frac{L M_f}{K * \mathrm{SNR}}. \tag{5.27}$$

5.3.3 Enhanced Estimator

As mentioned in Section 5.3.1., the proposed scheme works block by block with each block covering a number of OFDM frames. An enhanced estimator can be implemented as in Figure 5.3. Suppose that the data in the previous block have been detected. For the current processing block, temporal parameter estimates of the channel at every location ($k_1 = 1, 2, \cdots, N_f$) can be obtained by applying the proposed simple

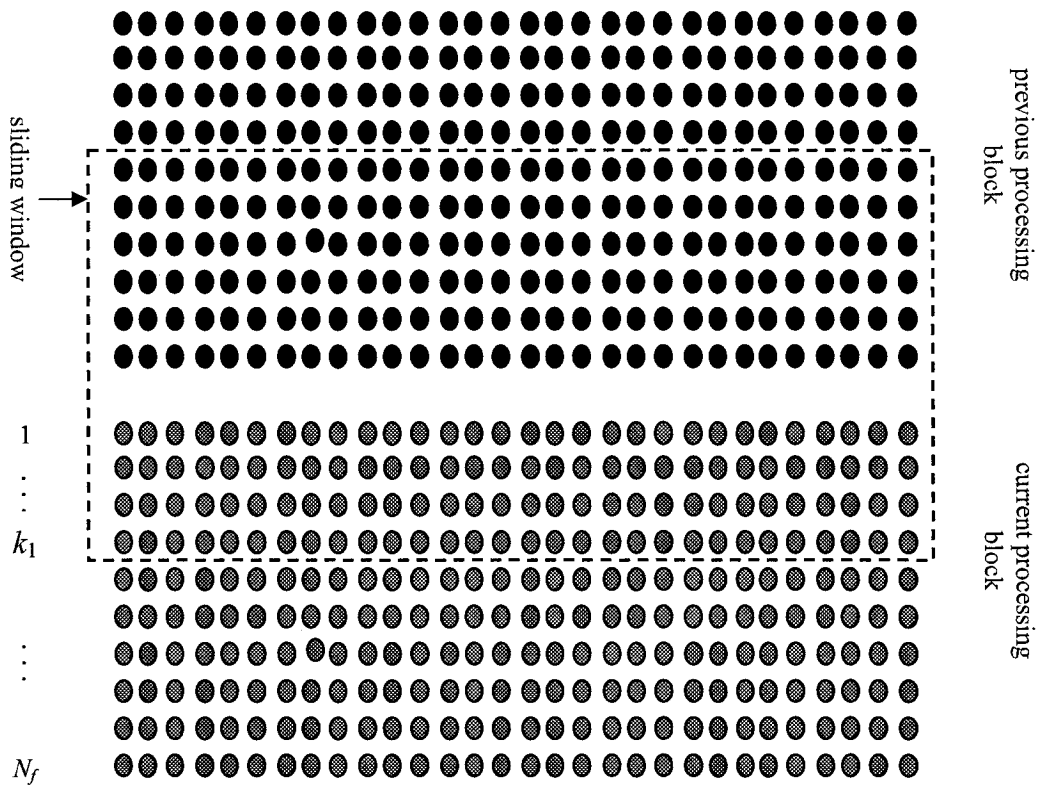

Figure 5.3. Enhanced channel parameter estimation.

windowing function presented in Section 5.3.2, using the pilot symbols. These temporal channel estimates of the current processing block can be further improved by using a window sliding in between previous and current processing blocks. Within each window, the channel parameters associated with the first $N_f - k_1$ frames are first refined by using the decision feedbacks of detected symbols. Then, these improved channel estimates together with the channel estimates associated with the last k_1 OFDM frames are used to replace the channel statistics in (5.14) in order to calculate the 2-D weighting function. Better channel parameter estimation can be obtained by using the resulting weighting

function. This estimator is referred to as *the enhanced estimator*. For slowly fading channels, accurate estimates of the channel can be obtained, and, hence, performance improvement over that of the robust estimator can be achieved. However, for channels with a large Doppler shift, the performance improvement might be less than the case of slow fading because the channel changes quickly and is less correlated from frame to frame.

The computational complexity of the enhanced estimator can be greatly reduced by observing that the length of the channel impulse response, L, is usually far smaller than the OFDM frame length, K. From (5.14), the weighting function of the MMSE estimator will have zeros for delay indexes greater than L. Therefore, the IFFT needs only generate outputs with delay indexes less than or equal to L, and only $L \times N_f$ weighting coefficients need to be calculated. Even though L may not be known initially at the receiver, it can be readily estimated by counting the number of IFFT outputs with absolute values significantly larger than others after a few processing windows.

5.3.4 Grid Design

In this section, we discuss the pilot symbol grid selection of the proposed scheme. To recover channel parameters, the 2-D sampling theorem [68] demands that

$$\begin{cases} \tau_{\max} \Delta f M_f \le 1/2 \\ f_{D,\max} T_s M_t \le 1/2 \end{cases}$$

where $\tau_{\max}$ is the one sided maximum echo delay and $f_{D,\max}$ is the maximum Doppler shift. For the proposed scheme, the pilot grid is designed as $M_t = 1$ and $M_f \le K/L_{\max}$,

where $L_{\max}$ is the maximum number of resolvable paths of the channel. For example, in Figure 5.4, for temporal channel estimation, 10 % pilot symbols are inserted in OFDM frames. In this case, $M_f = 10$ and $M_t = 1$. In Figure 5.5, 5 % pilot symbols are inserted with $M_f = 20$ and $M_t = 1$.

5.4 Simulation Results and Discussion

In this section, we provide simulation results of the proposed pilot-symbol-aided channel estimator for OFDM in wireless communication systems. For the channel model described in Section 5.2., the available bandwidth was chosen to be 1MHz and divided into 128 tones. This led to a symbol duration of 128 μs. In each OFDM frame, the first and the last four tones were used as guard tones. To avoid inter-symbol interference, an interval 32 μs was added as a guard interval. Therefore, the frame length was equal to 160 μs and the subchannel symbol rate was 6.250 k symbols/s [74]. In all the simulations, the estimation was carried out window by window, with each processing window covering 10 OFDM frames. In Figure 5.4, within each processing window, 10 % of the tones are reserved for pilot symbols. This 10 % grid is considered in Figures 5.6, 5.7, and 5.8. In Figure 5.5, 5 % grid is considered and used in Figure 5.9. For the enhanced estimator that uses a sliding window approach as shown in Figure 5.3, it was assumed that the two consecutive windows overlap by $k_1 = 1$ frame. For all simulations, BPSK modulation was assumed.

In Figure 5.6, we compare the proposed MMSE, robust, and enhanced estimators in terms of performance for an equal-gain channel with a Doppler shift $f_D = 40$ Hz. The

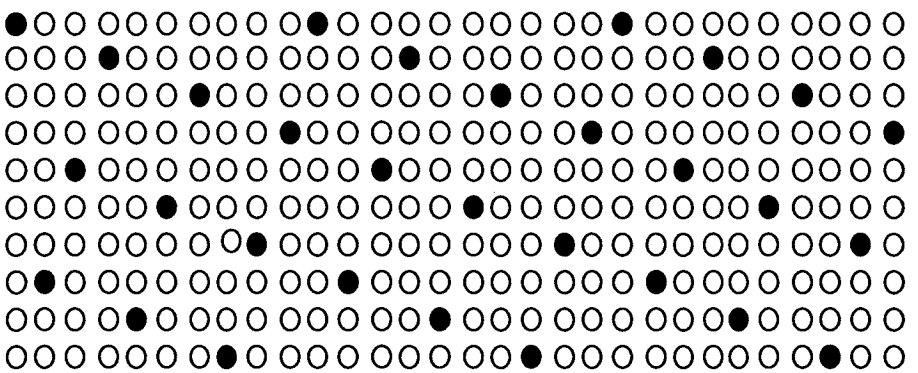

Figure 5.4. Sampling grid for pilot symbols with $M_f = 10$ and $M_t = 1$.

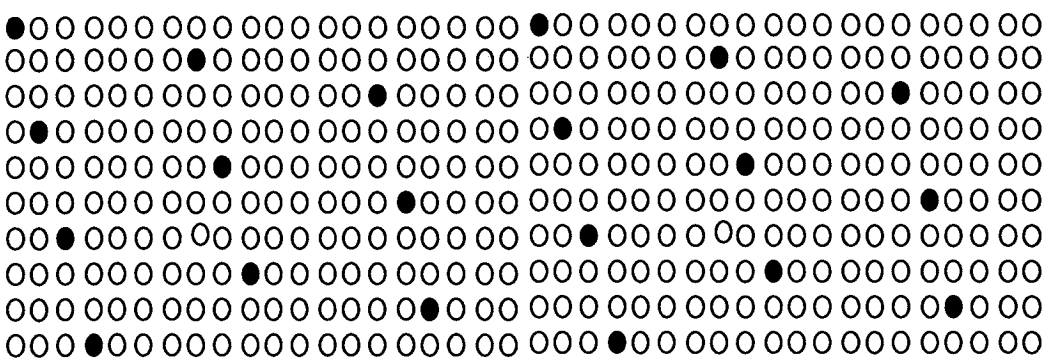

Figure 5.5. Sampling grid for pilot symbols with $M_f = 20$ and $M_t = 1$.

number of resolvable paths was set to be 4 and 8. As expected, a performance gain of approximately 3 dB is achieved for the 4-tap channel compared to the 8-tap channel. Furthermore, in both cases, the performance gain of the enhanced estimator with respect to the robust estimator ranges from about 3.5 dB at low SNR to 1 dB at high SNR.

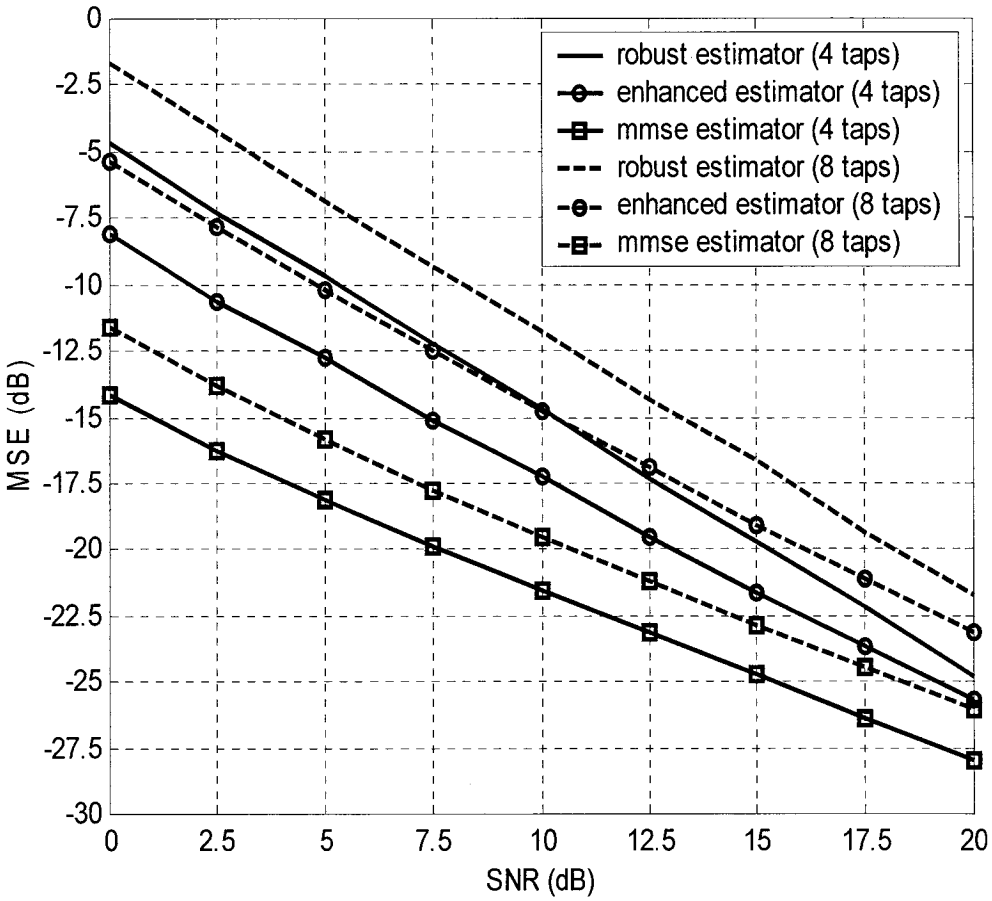

Figure 5.6. MSE of the robust, enhanced and MMSE estimators using 10% grid density for an equal-gain channel having $f_D = 40Hz$.

In Figure 5.7, the MSE curves of the optimal MMSE, robust, and enhanced estimators for an equal-gain channel with a Doppler shift $f_D = 200\,\text{Hz}$ are given. The number of resolvable paths was set to be 4 and 8. As can be again observed, a performance gain of approximately 3 dB is obtained for the 4-tap channel compared to the 8-tap channel. Comparing the case of $f_D = 200\,\text{Hz}$ to the case of $f_D = 40\,\text{Hz}$, we observe that for the same number of resolvable paths, the robust estimator yields approximately the same performance in both cases. The performance gain of the enhanced estimator with respect to the robust estimator when $f_D = 200\,\text{Hz}$ ranges from 2.3 dB at low SNR to 0 dB at high SNR, which is less significant than in the case of $f_D = 40\,\text{Hz}$. This result is due to the fact that for a large Doppler shift, the channel changes quickly and is less correlated from frame to frame and, thus, the information feedback from previous frames is less helpful.

Figure 5.8 compares the optimal MMSE, robust, and enhanced estimators in terms of performance for the ITU channel (Vehicular A) [75] with a Doppler shift $f_D = 40\,\text{Hz}$ and $f_D = 200\,\text{Hz}$. As expected, behavior similar to that in the case of the equal-gain channel is observed. In addition, all three estimators yield better performance in the current case because the tap energy is more focused on the first one or two taps.

Figure 5.9 shows the MSE performance of the optimal MMSE, robust, and enhanced estimators for an equal-gain 4-tap channel with Doppler shift $f_D = 40$ Hz and $f_D = 200$ Hz. In this case a grid density of 5% is used to show the effect of the grid density on the performance. As expected, a performance loss of approximately 3 dB is experienced in the case of 5% grid density compared to the case of 10% grid density.

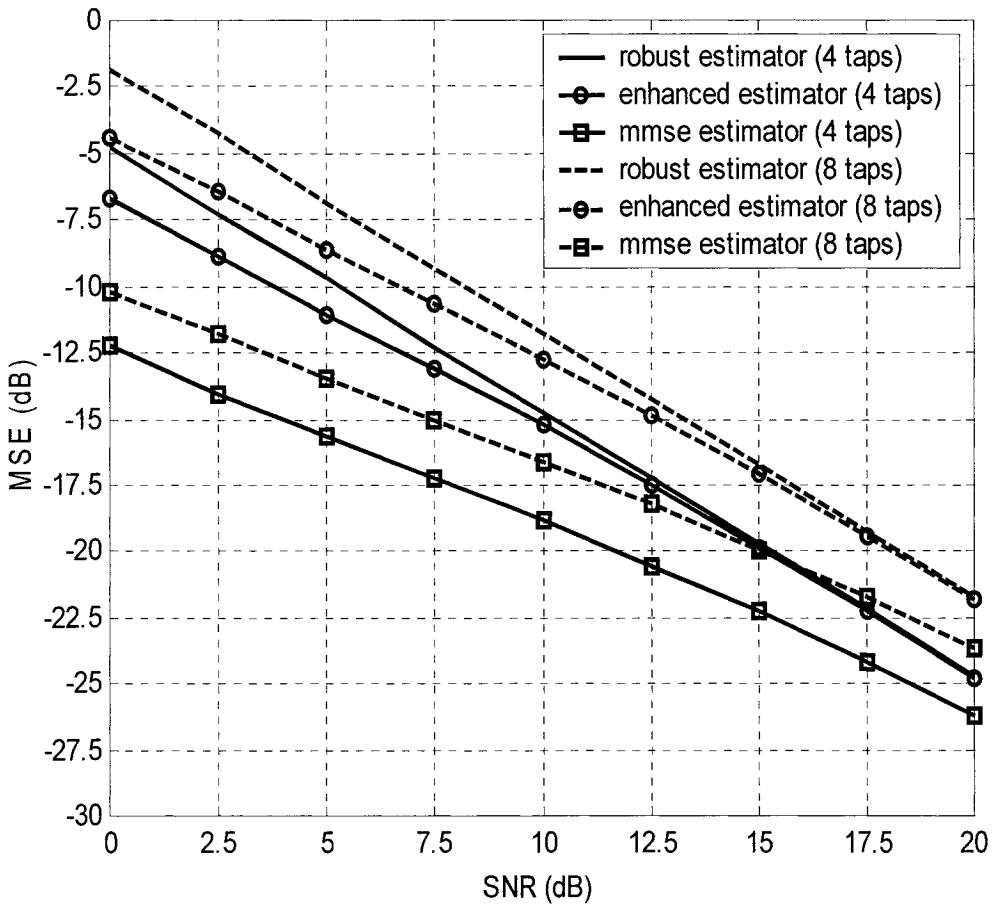

Figure 5.7. MSE of the robust, enhanced and MMSE estimators using 10% grid density for an equal-gain channel having $f_D = 200Hz$.

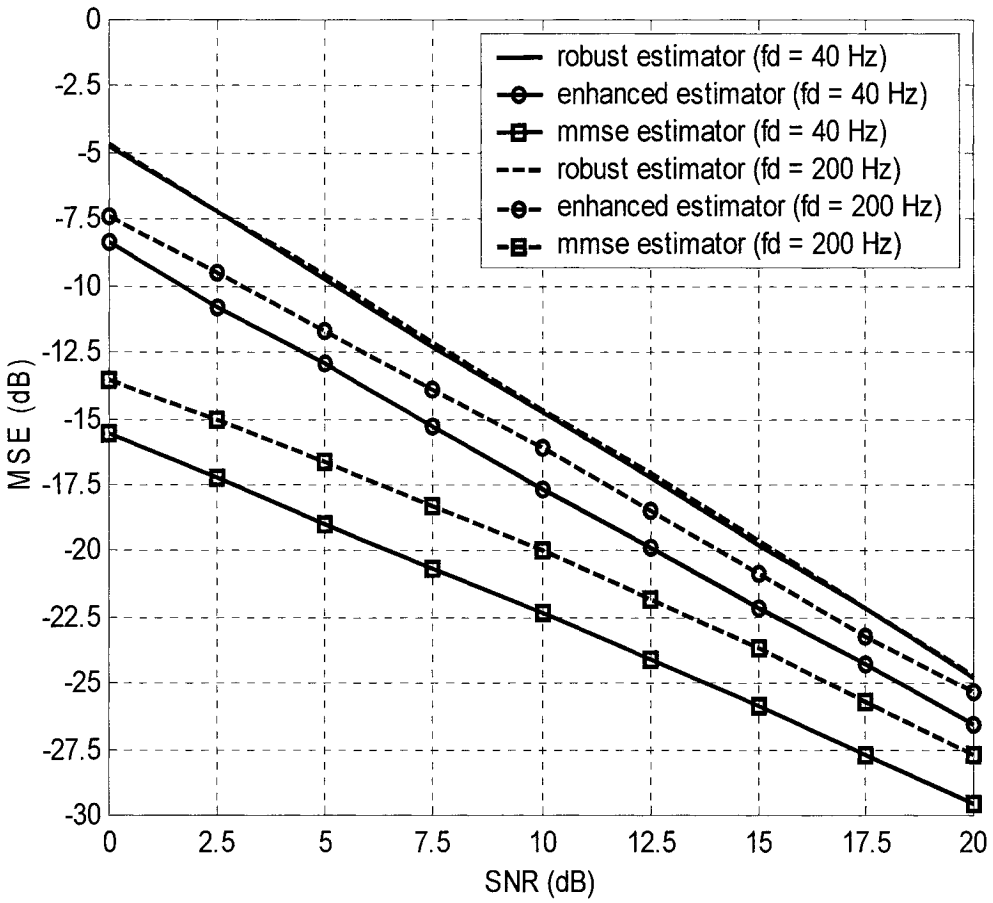

Figure 5.8. MSE of the robust, enhanced and MMSE estimators using 10% grid density for the ITU channel.

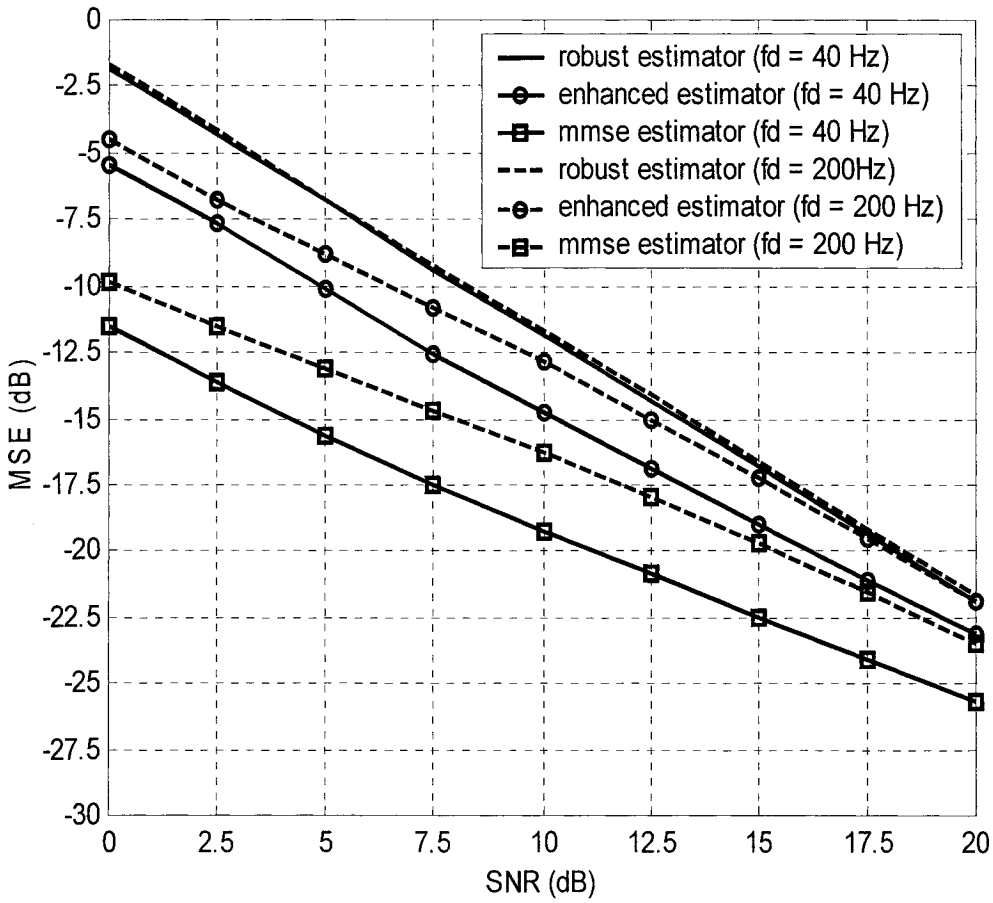

Figure 5.9. MSE of the robust, enhanced and MMSE estimators using 5% grid density for an equal-gain channel.

In summary, the above results show that the proposed robust estimator is capable of tracking channel variation and is also robust in dealing with different pilot sampling grids. At very-high SNR region, the performance of the robust estimator will converge to that of the optimal MMSE. When the SNR is low and the Doppler shift is relatively small, the performance gap between the robust and the optimal estimators is notable. In this case, the proposed enhanced estimator is an effective way to reduce the gap.

5.5 Conclusions

In this chapter, we have investigated a pilot-symbol-aided channel estimator for OFDM wireless systems using 2-D IFFT, 2-D FFT and a 2-D weighting function. Compared to the decision-directed channel parameter estimator, the proposed pilot-symbol-aided estimator is highly robust to Doppler shift. Also, the proposed channel estimator has much reduced complexity compared to conventional 2-D interpolation filters. The reduction in complexity is achieved by employing the 2-D IFFT, 2-D FFT and a 2-D weighting function instead of a 2-D filter. However, channel statistics are often not available in practice. Therefore, we proposed a robust channel estimator using simple windowing function. Moreover, we introduced an enhanced estimator that can improve the performance of the robust estimator. This can be obtained by using decision feedback of decoded information symbols and sliding window approach. Simulation results demonstrated that the robust and enhanced estimators are highly effective and robust to channel time variation when the channel statistics are unknown.

Chapter 6

Conclusions and Future Works

6.1 Conclusions

In this dissertation, we have developed two MIMO transmission schemes capable of various rate-performance tradeoffs for flat and frequency-selective channels, respectively.

For flat fading MIMO channels, we developed a formal approach for the design of LD codes. Based on this, we proposed a general multi-layer LD coding scheme that allows various multiplexing gains simply by increasing the coding rate and augmenting the set of existing dispersion matrices. To optimize the performance, we proposed to maximize both the diversity and coding gains by applying phase shifts among input symbols without the loss of mutual information. As a result, significant performance gain can be achieved compared to the existing LD codes. Simulation results were provided to demonstrate the effectiveness of the proposed scheme.

For frequency-selective MIMO channels, we introduced a general multi-layered space-frequency coding scheme that provides flexible tradeoff between data rate and performance. In the proposed scheme, a new space-time OFDM modulation scheme is introduced that translates a MIMO frequency-selective channel into a mathematically equivalent SIMO channel. Consequently, the conventional codes designed for fast fading single-input channels can be used to exploit diversity in both spatial and frequency domains. It was shown that the diversity gain of the proposed scheme equals the effective length of the conventional trellis codes used. This is in contrary to the case of STC-OFDM where no systematic methods exist for the design of space-time trellis codes with a given target of diversity gain.

Using the proposed space-time modulator, we also developed a channel multiplexing scheme, which subsumes V-BLAST [7] and D-BLAST [6] as suboptimal designs. Unlike in the conventional spatial multiplexing schemes where a data stream occupies only one transmit antenna at a time, each of the data streams in the proposed scheme sees all the transmit antennas at any time. This significantly simplifies the design of codes to achieve diversity gain. Another advantage of the proposed scheme is that an arbitrary number of data streams (layers) can be created. When the number of data streams, N, is less than or equal to the smaller of the numbers of transmit and receive antennas, a multiplexing gain N is achieved. Therefore, a full range of tradeoff between diversity and multiplexing gains is possible. Simulations were carried out and results were given to show the merits of the proposed scheme.

Last, we introduced a pilot-symbol-aided channel estimator for SIMO OFDM systems. The proposed estimator can be directly used with the proposed ST-OFDM

modulation scheme presented in Chapter 4. The new channel estimator is capable to track the time variation of wireless channels compared to the decision directed schemes. Also, it is based on 2-D IFFT, 2-D FFT and a 2-D weighting function instead of a 2-D interpolation filter. This makes the computational complexity very low compared to the conventional 2-D interpolation filter. In practice, channel statistics are unknown, therefore, a robust estimator based on a simple 2-D windowing function was proposed. To further improve the performance of the robust estimator, an enhanced channel estimator was proposed with greatly improved performance. Simulation results demonstrated that the proposed robust and enhanced estimators are highly effective for realistic situations when channel statistics are not available.

6.2 Future Research Directions

As a continuation of the work in this dissertation, the following research topics can be identified:

1. *Channel Layering*

The proposed multi-layered MIMO OFDM transmission scheme in chapter 4 allows channel layering in order to achieve higher data rate. This is because the channel seen by the information source is a one-dimensional channel, which is an interleaved version of channels associated with all transmit antennas. Apparently, different sets of weights applied to transmit antennas will result in different one-dimensional channels seen by the information source. Consequently, multiple information streams can be transmitted over different one-dimensional channels. However, two important issues remain to be

addressed:

- Joint iterative (turbo) detection and decoding

The creation of multiple layers allows higher data rate for a given bandwidth. However, as the number of layers increases, ML decoding is too complicated to implement. To save the computation, with minimum loss in performance one can borrow the turbo principles [58, 77] to perform joint iterative layer separation and decoding.

- Antenna weight design

So far, we have found that orthogonal codes are optimal for equal-gain multipath fading channels. However, practical wireless channels often have largely unbalanced gains among resolvable paths. In this case, orthogonal codes are no longer the optimum. More effort is required to design antenna weights based on the knowledge of channel delay profile.

2. *Channel Estimation*

We have proposed a robust pilot-symbol-aided channel estimator for SIMO OFDM systems. Next step would be to generalize our proposed estimator to the case of OFDM MIMO channels, then optimize it for general space-time coding schemes.

3. *Code Design with Asymmetrical Constellations*

For SISO channels, trellis codes with asymmetrical constellations have been shown to provide higher coding gain than the traditional trellis codes based on symmetrical constellations [76]. We expect that the improvement in coding gain with asymmetrical constellations will be even more significant for MIMO channels with appropriate constellation design. This will be another research direction.

References

[1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecom.*, vol. 10, pp. 585-595, Nov. 1999.

[2] G. J. Foshini and M. J. Gans, "On the limits of wireless communications in a fading environment when using multiple antennas, *Wireless Personal Communications*, vol. 6, no. 3, pp. 311-335, Mar. 1998.

[3] V. Tarokh, N. Seshadri, and R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744-765, March 1998.

[4] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun*, vol. 16, pp. 1451-1458, Oct. 1998.

[5] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456-1467, July 1999.

[6] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp 41-59, 1996.

[7] G. J. Foschini, G. Golden, R. Valenzuela, and P. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1841-1852, Nov. 1999.

[8] Lizhong Zheng and David N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. on Info.*, vol. 49 , no. 5, pp. 1073 - 1096, May 2003.

[9] H. ElGamal, G. Caire, and M. O. Damen, "Lattice coding and decoding achieve the optimal diversity–multiplexing tradeoff of MIMO channels," *IEEE Trans. on Information Theory*, vol. 50 , no. 6, pp. 968-985, June 2004.

[10] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 1804-1824, July 2002.

[11] X. Ma and G. B. Giannakis, "Full-diversity full-rate complex-field space-time coding," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2917-2930, Nov. 2003.

[12] R. W. Heath Jr. and A. J. Paulraj, "Linear dispersion codes for MIMO systems based on frame theory," IEEE Trans. Signal Processing, vol. 50, no. 10, pp. 2429-2441, Oct. 2002.

[13] M. O. Damen, N. C. Beaulieu, and J.-C. Belfiore, "A number theory based dual transmit antennas space-time code," in *Proc. GLOBECOM Conf.*, 1, San Antonio, TX, pp. 485-489, Nov. 25-29, 2001.

[14] M. O. Damen, A. Tewfik, and J.-C. Belfiore, "A construction of a space-time code based on number theory," *IEEE Trans. Infor. Theory*, vol. 48, pp. 753-760,

March 2002.

[15] H. El Gamal, and M. O. Damen, "An Algebraic number theoretic framework for space-time coding," in *Proc. IEEE ISIT*, vol. 32, pp. 132, Lausanne, Switzerland, June 30-July 5, 2002.

[16] D. Agrawal, V. Tarokh, A. Naguib, and N. Sheshadri, "Space-time coded OFDM for high data-rate wireless communication over wideband channels," in *Proc. 48th IEEE Vehicular Technology Conf.*, vol. 3, pp. 2232-2236, Ottawa, Canada, May 18-21, 1998.

[17] S. Mudulodu, and A. Paulraj, "A transmit diversity scheme for frequency-selective fading channels," in *Proc. 2000 IEEE Global Telecommunications Conf.,* vol. 2, pp. 1089-1093, San Francisco, CA, Nov. 27 - Dec.1, 2000.

[18] R. Blum, Y.(Geoffrey) Li, J. Winters and Q. Yan, "Improved Space-time coding for MIMO-OFDM wireless communications," *IEEE Trans. Commun.*, vol. 49, no. 11, pp. 1873-18, Nov. 2001, 78.

[19] Salim Alkhawaldeh and Y. R. Shayan, "An extended space-time transmit diversity scheme for frequency selective fading channels," *IEEE' CCECE conf., 2003*, Montreal, pp. 1663-1666, May 2003.

[20] H. Bolcskei and A. J. Paulraj, "Space-frequency coded broadband OFDM Systems," *Proc. IEEE Wireless Communications and Networking (WCNC) 2002*, vol. 1, pp. 1-6, 2002.

[21] B. Lu and X. Wang, "Space-time code design in OFDM Systems," *Proc. IEEE Globecom Conference*, vol. 2, pp. 1000-1004, San Francisco, CA, Nov. 2000.

[22] H. El Gamal, A. R. Hammons Jr., Youjian Liu, M.P. Fitz, O.Y. Takeshita, "On

the design of space-time and space-frequency codes for MIMO frequency-selective fading channels," *IEEE Trans. on Info.*, vol. 49, no. 9, pp. 2277 - 2292, Sept. 2003.

[23] Z. Liu, Y. Xin, and G. B. Giannakis, "Space-time-frequency coded OFDM over frequency-selective fading channels," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2465-2476, Oct. 2002.

[24] H. Bolcskei, M. Borgmann, and A. J. Paulraj, "Space-frequency coded MIMO-OFDM with variable multiplexing-diversity tradeoff," in *Proc. IEEE International Conference on Communications (ICC) 2003*, vol. 4, pp. 11-15, May 2003.

[25] J. Beek, O. Edfors, M. Sandell, S. Wilson, and P.O. Borjesson, "On channel estimation in OFDM systems," *in Proc. 45th IEEE Vehicular Technology Conf.*, vol. 2, pp. 815-819, Chicago, IL, July 25-28, 1995.

[26] O. Edfors, M. Sandell, J. van de Beek, S. Wilson, and P.O. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, pp. 931-939, July 1998.

[27] Ye (Geoffrey) Li, L.J. Cimini, and N.R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channel," *IEEE Trans. Commun.*, vol. 46, pp. 902-915, July 1998.

[28] Y. Li and N. Sollenberger, "Adaptive antenna arrays for OFDM systems with co-channel interference," *IEEE Trans. Commun.*, vol. 47, pp. 217-229, Feb. 1999.

[29] Ye (Geoffrey) Li and N. Seshadri, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE J. Select. Areas Comm.*,

vol. 17, pp. 461-471, Mar. 1999.

[30] Ye (Geoffrey) Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE trans. Wireless Commun.*, vol. 1, no. 1, pp. 67-75, Jan. 2002.

[31] H. Minn, D.I. Kim and V.K. Bhargava, "A reduced complexity channel estimation for OFDM systems with transmit diversity in mobile wireless channels," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 799-807, May 2002.

[32] A. Aghamohammadi, H. Meyr, and G. Ascheid, "A new method for phase synchronization and automatic gain control of linearly modulated signals on frequency-flat fading channels," *IEEE Trans. Commun.* , vol. COM-39, pp. 25-29, Jan. 1991.

[33] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686-693, Nov. 1991.

[34] P. Hoeher, S. Kaiser, and P. Robertson, "Two-Dimensional pilot-symbol-aided channel estimation by Wiener filtering," *in Proc. 1997 IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 3, pp. 1845-1848, Munich, Germany, Apr. 21-24, 1997.

[35] Ye (Geoffrey) Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 1207-1215, July 2000.

[36] V. Tarokh, N. Seshadri, and R. Calderbank, "Combined array processing and space-time coding," *IEEE Trans. on Inform. Theory*, vol. 45, no. 4, pp. 1121-1128, May 1999.

[37] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Y. Kuo, "Signal design for transmitter

diversity wireless communication systems over Rayleigh fading channels," *IEEE Trans. Comm.*, vol. 47, no. 4, pp. 527-537, April 1999.

[38] Salim Alkhawaldeh, X. Wang and Y. R. Shayan, "Phase-shift-based layered linear space-time codes," *10th IEEE sympo. on Comp. and Comm. ISCC 2005*, pp. 281-286, Cartagena, Murcia, Spain, July, 2005.

[39] Branka Vucetic and Jinhong Yuan, "Space-time Coding," John Wiley and Sons Ltd, England, 2003.

[40] H. Bolckei, D. Gesbert, A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Transactions on Communications*, Vol. 50, pp. 225 – 234, Feb. 2002.

[41] A. Wittneben, "Base station modulation diversity for digital simulcast," in *Proc. IEEE VTC*, vol. 3, pp. 505-511, May 1993.

[42] N. Seshadri and J. H. Winters, "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity," *Int. Journal. Wireless Inf. Networks*, vol. 1, no. 1, pp. 49-60, Jan. 1994.

[43] R. W. Heath Jr. and A. Paulraj, "Transmit diversity using decision-directed antenna hopping," in *Proc. of ICC Comm. Theory Mini-Conf.*, pp. 141-145, June 1999.

[44] Salim Alkhawaldeh, X. Wang, and Y. R. Shayan, "A New Divesrity Scheme for Frequency Selective MIMO Channels," presented at the 14th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PRIMC2003), Beijing, China, Sept. 7-10, 2003.

[45] R. J. Piechocki, P. N. Flectcher, A. R. Nix, C. N. Canagarajah, and J. P. McGeehan, Performance evaluation of BLAST-OFDM enhanced Hiperlan/2 using simulated and measured channel data," *Electron. Lett.*, vol. 37, no. 18, pp. 1137–1139, Aug. 2001.

[46] Y. Xin, and G.B. Giannakis, "High-Rate Space-Time Layered OFDM," *IEEE Communications Letters*, vol. 6, no. 5, pp. 187-189, May 2002.

[47] J. Geng, M. Vajapeyam and U. Mitra, "Union bound of space-time block codes and decomposable error patterns," Proc. *IEEE International Symposium on Information Theory, pp. 62, 29 June-4 July, 2003.*

[48] S. Sandhu and A. Paulraj, "Union bound on error probability of linear space-time block codes," Acoustics, Speech, and Signal Processing, 2001. In *Proc. IEEE ICASSP 2001*, vol.4, pp. 2473 – 2476, 7-11 May 2001 .

[49] M. Vajapeyam, J. Geng and U. Mitra, "Low SNR design of space-time block codes based on union bound and indecomposable error patterns," *IEEE International Conference on Communications*, vol. 2, pp. 707 – 711, 20-24 June 2004.

[50] Salim Alkhawaldeh, X. Wang, and Y. R. Shayan, "Multi-layered linear dispersion codes for flat fading MIMO Channels," submitted to *IEEE Transaction on Signal Processing.*

[51] J. Wang, E. B, and K. Y, "Asymptotic performance of space-frequency codes over broadband channels," *IEEE Communications letters*, vol. 6, pp. 523-525, Dec. 2002.

[52] B. Lu, X. Wang, and K. R. Narayanan, "LDPC-based space-time coded OFDM

systems over correlated fading channels: performance analysis and receiver design," *IEEE Trans. on Commun.*, vol. 50, no. 1, pp. 74-88, Jan. 2002.

[53] C. Schlegel and D. J. Costello, "Bandwidth efficient coding for fading channels: code construction and performance analysis," *IEEE J. Select. Areas in Commun.*, vol. 7, no. 9, pp. 1356-1368, Dec. 1989.

[54] S. Hamidreza Jamali and Tho Le-Ngoc, "*Coded-modulation techniques for fading channels*,: Boston: Kluwer Academic Publishers, 1994.

[55] X. Wang, Y. R. Shayan, and M. Zeng, "On the code and interleaver design of broadband OFDM systems," *IEEE Commun. Letters*, vol. 8, no. 11, pp. 653-655, Nov. 2004.

[56] J. Yuan, Z. Chen, B. Vucetic, and W. Firmanto, "Performance and design of space-time coding in fading channels," *IEEE Trans. on Commun.*, vol. 51, pp. 1991-1996, Dec. 2003.

[57] R. Heath Jr. and A. Paulraj, "Switching between multiplexing and diversity based on constellation distance," *Proc. of Alterton Conf. on Commun. Control and Comp.*, Oct. 2000.

[58] X. Wang and H.V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. on Comm.*, vol. 46, pp.1046-1061, July 1999.

[59] S.B. Weinstein and P.M. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Trans. Comm. Tech.*, vol. COM-19, pp. 628-634, Oct. 1971.

[60] L.J. Cimini, Jr., "Analysis and simulation of a digital mobile channel using

orthogonal frequency division multiplexing," *IEEE Trans. Comm.*, vol. COMM-33, pp 665-675, July 1985.

[61] I. Kalet, "The multitone channel," *IEEE Trans. Comm.*, vol. 37, pp. 119-124, Feb. 1989.

[62] J.A.C. Bingham, "Multicarrier modulation for data transmission: an idea whose time has come," *IEEE Comm. Mag.*, vol. 28, pp. 5-14, May 1990.

[63] L.J. Cimini, Jr., B. Daneshrad, and N.R. Sollenberger, "Clustered OFDM with transmitter diversity and coding," *IEEE Global Telecomm. Conf.*, pp. 5-14, Nov. 1996.

[64] L. J. Cimini and N.R. Sollenberger, "OFDM with diversity and coding for high-bit-rate mobile data applications," *IEEE Global Telecomm. Conf.*, Phoenix, AZ, pp. 305-309, Nov. 1997.

[65] V. Mignone and A. Morello, "CD3-OFDM: a novel demodulation scheme for fixed and mobile receivers," *IEEE Trans. Comm.*, vol. 44, pp. 1144-1151, Sep. 1996.

[66] X. Wang, Salim Alkhawaldeh, and Y. R. Shayan, "A new approach to diversity and multiplexing gains for wideband MIMO channels," accepted for publication in *IEEE Transaction on Wireless Communications.*

[67] J. Broakis, *Digital communications*, McGraw-Hill, 1989.

[68] R.M. Mersereau and Theresa C. Speake, "The periodically sampled multidimensional signals," *IEEE Trans. Signal Processing*, vol. ASSP-31, Feb. 1983, pp. 188-194.

[69] P. Hoeher, S. Kaiser, and P. Robertson, "Pilot-symbol-aided channel estimation in

time and frequency," *in Proc. 1997 IEEE Global Telecommunications Conf., the mini-conf.,* Phoenix, AZ, Nov. 1997, pp. 90-96.

[70] G.L. Stuber, *Principles of mobile communication*, KAP, USA, 1996.

[71] Salim Alkhawaldeh, X. Wang, and Y. R. Shayan, "A robust pilot-symbol-aided channel estimator for OFDM wireless communications," *Canadian Journal of Electrical and Computer Engineering*, vol. 30, no.1, Winter 2005.

[72] A. Papoulis, *Probability, Random Variables, and stochastic Processes*, New York: McGraw-Hill, 1991.

[73] Harold W. Sorenson, "*Kalman filtering: theory and application*," IEEE Press, NewYork, 1985.

[74] Salim Alkhawaldeh, X. Wang, and Y. R. Shayan, "A robust pilot-symbol-aided channel estimator for OFDM wireless communications," *6th World Wireless Congress*, pp. 34-39, Palo Alto, May, 2005.

[75] Recommendation ITU-R M.1225, "Guidelines for evaluation of radio transmission technologies for IMT-2000," 1997.

[76] M. Isaka, M. Fossorier, R. Morelos-Zaragoza, S. Lin and H. Imai, "Multilevel coded modulation for unequal error protection and multistage decoding, II. Asymmetric constellations," *IEEE Trans. Comm.,* vol. 48, pp. 774-786, May 2000.

[77] P. D. Alexander, M. C. Reed, J. A. Asenstorfer and C. B. Schlegel, "Iterative multiuser interference reduction: turbo CDMA," *IEEE Trans. Comm.*, vol. 47, pp. 1008-1014, July. 1999.