WIDELY-LINEAR MMSE RECEIVERS FOR LINEAR DISPERSION
SPACE-TIME BLOCK-CODES

by

Amirhossein Shokouh Aghaei

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science
Graduate Department of Electrical and Computer Engineering
University of Toronto

Copyright © 2008 by Amirhossein Shokouh Aghaei
Abstract

Widely-Linear MMSE Receivers for Linear Dispersion Space-Time Block-Codes

Amirhossein Shokouh Aghaei
Master of Applied Science
Graduate Department of Electrical and Computer Engineering
University of Toronto
2008

Space-time coding techniques are widely used in multiple-input multiple-output communication systems to mitigate the effect of multipath fading in wireless channels. An important subset of space-time codes are linear dispersion (LD) codes, which are used for quasi-static Rayleigh flat fading channels when the channel state information (CSI) is only available at the receiver side. In this thesis, we propose a new receiver structure for LD codes. We suggest to use widely-linear minimum-mean-squared-error (WL-MMSE) estimates of the transmitted symbols in lieu of the sufficient statistics for maximum likelihood (ML) detection of these symbols. This structure offers both optimal and suboptimal operation modes. The structures of the proposed receivers in both modes are derived for general LD codes. As special cases, we study two important subsets of LD codes, namely orthogonal and quasi-orthogonal codes, and examine the performance of the proposed receivers for these codes.
Acknowledgements

I would like to express my sincere gratitude and appreciation to my supervisor Professor Konstantinos N. Plataniotis for his invaluable advice, guidance, and encouragement. Without his kind supports during my M.A.Sc. studies, this research would have been impossible. I am deeply grateful for all his helps during this period.

Furthermore, I would like to offer my sincere thanks to Professor Subbarayan Pasupathy for his insightful guidance and helpful suggestions throughout this work. His comments and guidelines have always been a source of inspiration for me. I would also like to thank other committee members Professor Teng Joon Lim and Professor Glenn Gulak for reviewing my work and offering their constructive comments.

I greatly appreciate the kind helps and supports of my colleagues and friends during these two years. Specially, I should thank my best friends: Ali Kalantarian, Houman Hosseinpour, Mahdi Ramezani, and Seyedhossein Seyedmahdi as well as my colleagues: Amin Alamdar, Haiping Lu, Mohammad Shahin Mahanta, Mohammad Sharif, Mohsen Heidarinejad, Sachin Kadloor, and Saeed Moradi. Moreover, I would like to thank Amir Ali Basri, Azadeh Kushki, and Payam Dehghani for hours of discussion and helpful feedbacks.

Last, but certainly not least, my infinite thanks go to my family for their never-ending support, encouragement, and unconditional love.
# Contents

## 1 Introduction

1.1 Improper Complex Signals in Communications ............................ 2  
1.2 Motivations and Problem Description ..................................... 4  
1.3 System Model ........................................................................... 7  
  1.3.1 Transmitter Structure ....................................................... 8  
  1.3.2 Channel Model ................................................................. 9  
  1.3.3 Receiver Structure ............................................................. 9  
1.4 Thesis Contributions ............................................................... 10  
1.5 Thesis Organization .................................................................. 11

## 2 Preliminaries

2.1 Notation Description .................................................................. 13  
2.2 Second Order Statistics of Complex Signals ................................. 14  
  2.2.1 Proper vs. Improper Complex Signals ................................... 16  
  2.2.2 Probability Density Function of Complex Gaussian Random Vectors 17  
2.3 * Noncircularity Matrix for Complex Random Vectors .................... 18  
2.4 Chapter Summary ....................................................................... 24

## 3 Widely-Linear MMSE Receivers for LD Codes  

3.1 Widely-Linear MMSE Receiver Structure ..................................... 26  
3.2 Analysis of WL-MMSE Receiver for Some LD Codes ....................... 28
3.2.1 WL-MMSE Receiver for an Orthogonal LD Code with $M_t = 2, M_r = 1, N = 2$ ................................................................. 29
3.2.2 Equivalency between WL-MMSE Estimation and Alamouti’s Combining Scheme for Orthogonal Codes .......................... 31
3.2.3 Optimal WL-MMSE Receiver for a Quasi-Orthogonal LD code with $M_t = 4, M_r = 1, N = 4$ ................................. 33
3.2.4 Suboptimal WL-MMSE Receiver for Quasi-Orthogonal Codes ............................. 36
3.3 Sufficiency of WL-MMSE Estimates for Improper Complex Gaussian Signals 38
3.4 Simulation Results .......................................................... 40
  3.4.1 Simulation Parameters ................................................ 41
  3.4.2 Performance Analysis for Orthogonal Codes ...................... 41
  3.4.3 Performance Analysis for Quasi-Orthogonal Codes ............. 42
3.5 Chapter Summary ......................................................... 46

4 ML Detection in The Presence of Improper Noise 48
  4.1 Need for ML Detection in The Presence of Improper Gaussian Noise .......... 49
  4.2 ML Decision Rule in The Presence of Improper Gaussian Noise ............... 51
  4.3 Circularizing Filter Followed by Conventional Detector ...................... 53
  4.4 Binary ML Detection in the Presence of Scalar Improper Noise ............. 56
    4.4.1 Detection Using Pseudo-Correlators and Correlators .................. 56
    4.4.2 Detection Using a Circularizing Filter Followed by Conventional Detector 57
    4.4.3 Decision Regions of The Detector ................................... 59
    4.4.4 Probability of Detection Error ...................................... 60
  4.5 Chapter Summary ......................................................... 64

5 Conclusions 66
  5.1 Research Summary ...................................................... 66
5.2 Future Work ............................... 67

A Proofs ........................................... 69
A.1 Proof of Theorem 2.1 ......................... 69
A.2 Proof of Theorem 2.2 ......................... 70
A.3 Proof of Theorem 4.1 ......................... 70
A.4 Proof of Lemma 4.1 ......................... 71
A.5 Proof of Proposition 4.1 ....................... 72
A.6 Derivation of Equation (4.40) ............... 72

B On widely-linearity of MMSE estimator ...... 75

Bibliography ........................................ 77
## List of Tables

2.1 Covariance matrices in the case of scalar complex random variable, in terms of $\sigma_z^2$, $\gamma_z^2$, $\alpha_z$, $\sigma_x^2$, $\sigma_y^2$, and $\sigma_{xy}$ . . . . . . . . . . . . . . . . . . . . . . . . . . . . 22

3.1 Orthogonal code of [1] and corresponding dispersion matrices . . . . . . . 29

3.2 Quasi-orthogonal code of [2] and corresponding dispersion matrices . . . 34

3.3 Summary of Simulation Parameters . . . . . . . . . . . . . . . . . . . . . 42
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Applications of improper signals in communication systems.</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td>Block diagram of the MIMO system model</td>
<td>7</td>
</tr>
<tr>
<td>2.1</td>
<td>Possible values of $\alpha_z$ in the complex plane.</td>
<td>21</td>
</tr>
<tr>
<td>2.2</td>
<td>Contours of pdf for a scalar zero mean complex valued Gaussian random variable</td>
<td>24</td>
</tr>
<tr>
<td>3.1</td>
<td>The WL-MMSE receiver structure for Alamouti orthogonal code</td>
<td>31</td>
</tr>
<tr>
<td>3.2</td>
<td>The WL-MMSE receiver structure for quasi-orthogonal code of Table 3.2</td>
<td>36</td>
</tr>
<tr>
<td>3.3</td>
<td>Suboptimal WL-MMSE receiver for quasi-orthogonal code of Table 3.2</td>
<td>37</td>
</tr>
<tr>
<td>3.4</td>
<td>Relationship between $y$, $\tilde{x}$, and $x$.</td>
<td>40</td>
</tr>
<tr>
<td>3.5</td>
<td>Bit-error probability versus SNR for quasi-orthogonal code of Table 3.2, using 4-PSK modulation and one receive antenna.</td>
<td>43</td>
</tr>
<tr>
<td>3.6</td>
<td>Bit-error probability versus SNR for quasi-orthogonal code of Table 3.2, using 8-PSK modulation and one receive antenna.</td>
<td>43</td>
</tr>
<tr>
<td>3.7</td>
<td>Bit-error probability versus SNR for quasi-orthogonal code of Table 3.2, using 16-PSK modulation and one receive antenna.</td>
<td>44</td>
</tr>
<tr>
<td>3.8</td>
<td>Bit-error probability versus SNR for quasi-orthogonal code of Table 3.2, using 16-QAM modulation and one receive antenna.</td>
<td>45</td>
</tr>
<tr>
<td>3.9</td>
<td>Bit-error probability versus SNR for quasi-orthogonal code of Table 3.2, using 64-QAM modulation and one receive antenna.</td>
<td>46</td>
</tr>
</tbody>
</table>
4.1 Binary ML detectors implementation as a combination of correlators and pseudo-correlators ................................................................. 57
4.2 Binary ML detectors implementation as a combination of circularizing filter followed by conventional detector. ................................. 58
4.3 Decision regions and decision boundary of the ML detector in (a) general case, (b) simple antipodal example. .......................................... 60
4.4 Contours of constant $P_e$ in $\alpha_n$-plane for fixed SNR $(\frac{|s_1-s_0|^2}{\sigma_n^2})$. .............................................................. 62
4.5 Error probability for the special example of $s^1 = -s^0 = 1$ with $\Re\{\alpha_n\} = 0$ and $\Im\{\alpha_n\} = \rho/2$. .................................................. 63
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASK</td>
<td>Amplitude Shift Keying</td>
</tr>
<tr>
<td>BPSK</td>
<td>Binary Phase Shift Keying</td>
</tr>
<tr>
<td>CSI</td>
<td>Channel State Information</td>
</tr>
<tr>
<td>DS-CDMA</td>
<td>Direct-Sequence Code-Division-Multiple-Access</td>
</tr>
<tr>
<td>LD</td>
<td>Linear Dispersion</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple-Input Multiple-Output</td>
</tr>
<tr>
<td>ML</td>
<td>Maximum Likelihood</td>
</tr>
<tr>
<td>MMSE</td>
<td>Minimum Mean Square Error</td>
</tr>
<tr>
<td>OFDM</td>
<td>Orthogonal Frequency Division Multiplexing</td>
</tr>
<tr>
<td>OQPSK</td>
<td>Offset Quadrature Phase Shift Keying</td>
</tr>
<tr>
<td>PDF</td>
<td>Probability Density Function</td>
</tr>
<tr>
<td>PSK</td>
<td>Phase Shift Keying</td>
</tr>
<tr>
<td>QAM</td>
<td>Quadrature Amplitude Modulation</td>
</tr>
<tr>
<td>SNER</td>
<td>Signal to Noise Ratio</td>
</tr>
<tr>
<td>ST-BC</td>
<td>Space-Time Block-Code</td>
</tr>
<tr>
<td>V-BLAST</td>
<td>Vertical-Bell Laboratories -Layered -Space-Time</td>
</tr>
<tr>
<td>WL</td>
<td>Widely Linear</td>
</tr>
<tr>
<td>ZF</td>
<td>Zero-Forcing</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Since the first studies of communication systems, there has been a great interest in analyzing the signals and/or systems using complex-valued variables and functions. This interest is due to the fact that the complex domain provides a useful framework which simplifies the analysis of communication systems. In response to this wide usage of complex entities, many definitions and solutions which had originally been developed for real-valued entities were modified in order to be applicable to complex-valued entities as well.

Estimation and detection of transmitted messages in communication systems are among those problems which had extensively been studied for real-valued signals and were generalized for complex signals in various scenarios. Although these problems seem to be exhaustively studied in the literature, there are still many cases which have not yet been studied in this context. This chapter provides a brief discussion on the need for further studying the problems of estimation and detection of complex-valued signals in the multiple-input multiple-output (MIMO) communication systems.
1.1 Improper Complex Signals in Communications

One of the major applications of the complex domain in communication systems is low-pass representation of real-valued passband signals/systems. The low-pass, also called complex envelope, representation of the signal is of particular interest in many applications, in that the energy of the complex envelope of the signal is concentrated in low-frequency components, which in turn simplifies the analysis and processing of these signals [3]. Furthermore, the process of converting a passband signal into a low-pass one is reversible; hence, the passband signal can be uniquely defined from its complex envelope.

The real and imaginary parts of the complex envelope of a passband signal, are usually referred to as the inphase and quadrature parts of that signal. In many practical applications, the inphase and quadrature parts of the signals satisfy a particular set of symmetry conditions, called circular-symmetry [4] or properness [5] conditions, which will be explained in detail in Section 2.2.1. A simple example where these conditions are satisfied is the case that inphase and quadrature parts of the signal are uncorrelated to each other and have equal powers.

Due to the great extent of the applications in which the properness conditions are satisfied, the complex envelop signals are commonly presumed to be proper. As a consequence, most of the transceiver structures designed in the literature have an implicit assumption of dealing with proper complex signals. However, during the past decade it has been shown that there exist certain applications in which this assumption does not hold true. In these applications the complex envelope of the signal-of-interest or interfering signals are not proper. As a case in point, it has been shown that the complex envelope in some modulation schemes such as M-ary amplitude shift keying (ASK), binary phase shift keying (BPSK), offset quadrature phase shift keying (OQPSK) and minimum-shift keying (MSK) are improper. Moreover, some space-time block codes may also result in improper complex signals.
Fig. 1.1 provides a quick overview of some important applications in which the signals-of-interest or the interfering signals have improper complex envelopes. As it is illustrated in this figure, previous works on this area can be divided into the following major categories:

**Direct-sequence code-division-multiple-access (DS-CDMA) systems:** The first studies of improper signals in communication systems were in the context of DS-CDMA systems after [6] showed that in DS-CDMA systems using BPSK modulation, the multiple access interference is improper. Based on this result, [7–16] have designed and analyzed new receiver structures to improve the performance of DS-CDMA systems in different scenarios. It is noteworthy that even when a complex modulation scheme is used in CDMA systems, the receiver might encounter an improper signal. This problem might be resulted from using an improper complex
constellation or from using iterative multiuser receiver structures. These two issues are studied in [17] and [18], respectively.

**Multicarrier systems:** In these systems, there might be two possible sources of improper interference. First, when an improper narrowband signal from another communication system is interfering with the multicarrier system. This might happen in both wireless multicarrier systems (e.g., overlay networks) and wired multicarrier systems (due to crosstalk or radio frequency interference) [19, 20]. Second, in baseband OFDM systems if a colored noise exists at the receiver input, it leads to an improper noise at the output of the discrete Fourier transformer [21–23].

**Multiple-input multiple-output (MIMO) systems:** Some space-time codes used for MIMO channels generate improper transmit signals. The impropriety of the transmitted signal can be due to the special structure of the space-time coding scheme (e.g. linear dispersion [24], orthogonal [25], and quasi-orthogonal [2] codes), or due to the usage of an improper modulation scheme (e.g ASK, BPSK, MSK, or OQPSK) before space-time coding. The former case has been studied in [26–28], while [29–32] have studied the latter case.

In this thesis, we focus on the MIMO systems dealing with improper signals. We will study the problem of detection and estimation of improper complex signals in MIMO transceivers using linear dispersion codes.

### 1.2 Motivations and Problem Description

Since the emergence of commercial wireless communication systems, there has been an ever-increasing demand for higher data rates for transmitting more information over wireless networks. One of the most restrictive limitations of wireless channels which has prevented wireless transceivers from easily achieving high data rates is time-varying
multipath-fading [33]. In a typical additive noise channel with QPSK modulation, the bit-error-rate of the system can be easily decreased from $10^{-3}$ to $10^{-4}$ by less than 2 dB increase in the transmission power. However, in a Rayleigh flat fading channel with the same modulation, this reduction in the bit-error-rate requires 10 dB increase in the transmission power.

In order to mitigate the effect of multipath fading in wireless channels, multiple-input multiple-output (MIMO) wireless systems are being used. MIMO systems deploy the concepts of transmit/receive diversity in order to provide the receiver with multiple copies of the data. MIMO transmission techniques can be categorized into the following two major groups: schemes that require channel information at the transmitter side, and schemes that do not need this information at the transmitter. The former schemes usually acquire the channel information through feedback from the receiver. Since wireless channels are time-varying, the feedback should be updated accordingly. This might impose a prohibitive overhead on the reverse channel. The latter schemes, however, do not require any information about the channel at the transmitter side; hence, they are more cost effective in many applications.

One of the most important existing techniques which does not need the channel information at the transmitter is the linear dispersion (LD) space-time block-coding. This scheme is used for transmission over quasi-static flat fading channels, and only requires the knowledge of channel state information (CSI) at the receiver side. In general, every space-time code whose codewords are constructed from linear combination of input symbols and their complex conjugates is called an LD code [24]. This scheme encompasses a wide range of space-time codes, such as V-BLAST [34], orthogonal [1, 25], and quasi-orthogonal [2] codes, each of which is designed based on a certain criterion. Orthogonal and quasi-orthogonal codes are designed to decrease the complexity of the decoding algorithm, while V-BLAST is designed to increase the throughput of the MIMO channel. It is also possible to design LD codes such that the mutual information between the
transmitted and received signals is maximized \cite{24}.

For decoding LD codes at the receiver side, there exist various strategies, including maximum likelihood (ML) decoding, sphere decoding \cite{35,36}, and successive canceling and nulling \cite{37,38}. The first method (ML detection), which is based on finding the most likely transmitted block, suffers from computational complexity of searching over all possible codewords, unless orthogonal or quasi-orthogonal codes are used. In order to decrease this complexity, the second algorithm (sphere decoding) tries to reduce the number of possible codewords over which the search should be performed. In sphere decoding, only those codewords which are located within a sphere centered around the received signal are considered for ML detection, and the remaining codewords are presumed unlikely to be transmitted. Finally, the main idea behind the third method (successive nulling and canceling) is to estimate the transmitted symbols in a sequential order. In this method, the symbols are detected one by one from the strongest SNR to the weakest one. In detection of each symbol, the effect of previously detected symbols is reduced from the received signals before detecting the new symbol (canceling), and the effect of all remaining undetected symbols will be treated as interference (interference nulling). The nulling stage can be accomplished by using different equalization concepts such as zero-forcing (ZF) or minimum mean square error (MMSE) equalization. Note that if we ignore the canceling stage, this method can be implemented in one step, namely equalization.

It has been shown in \cite{27} that when LD codes are used at the transmitter side, the transmitted symbols and received signals are jointly improper. However, most of the existing receiver structures have not taken into account this impropriety. In the literature, only few works \cite{26,27,39} have considered impropriety of the LD codes while designing the receiver. Nevertheless, these works have only considered some special members of LD codes as follows. The works in \cite{27} and \cite{39} have considered LD codes for transmission over the MIMO channel; however, they have assumed that a convolutional data encoder is used prior to the space-time coder and have proposed iterative detection strategies for
this case. In [26], a new receiver is designed to equalize the intersymbol interference of Alamouti code, which is a member of LD codes. In fact, there exists no general decoding scheme for LD codes which takes into account the impropriety of these codes.

This thesis proposes a general decoding scheme for LD codes, which takes into account the inherent impropriety of these codes. This decoding scheme is based on using the minimal sufficient statistics for ML detection of transmitted symbols. As it is shown in Fig. 1.2, the proposed receiver first estimates the transmitted symbols from the received signals using MMSE criterion. Then, it utilizes these estimates in lieu of the sufficient statistics for ML detection of the transmitted symbols. This framework is deployed in this thesis to design both optimal (ML) and suboptimal receivers. Due to the impropriety of LD codes, MMSE estimation of the transmitted symbols in our proposed structure requires linear processing of not only the received signals, but their complex conjugates as well [40,41]. This estimator is called a *widely-linear* (WL) estimator [40], as opposed to strictly-linear estimators which only perform linear processing on the received signals.

### 1.3 System Model

This thesis considers a MIMO communication system as shown in Fig 1.2. This transceiver utilizes $M_t$ transmit antennas and $M_r$ receive antennas to communicate over a frequency-flat quasi-static fading channel.

![Figure 1.2: Block diagram of the MIMO system model](image-url)
1.3.1 Transmitter Structure

At the first stage of Fig. 1.2, data bits are mapped into complex symbols \((x_i)\) using a set of complex-valued constellations, e.g. PSK or QAM. Each of the symbol mappers can use a different constellation for mapping the data bits; however, in practice they are usually selected to be identical. These transmitted symbols are assumed to be uncorrelated with each other, the channel coefficients, and the additive noise of the channel. Now, suppose the channel is constant for an interval of \(T\) symbols, as will be justified shortly. The transmitter uses this interval to transmit a block of \(N\) symbols, denoted by \(x = [x_1, x_2, \ldots, x_N]^T\), over \(M_t\) transmit antennas. The value of \(N\) has to satisfy the inequality \(N \leq TM_t\), where the value of \(M_t\) is determined depending on the application and the value of \(T\) is governed by the channel characteristics which will be discussed later on (refer to [24] for a discussion on implications of different choices of \(N\)).

In order to transmit this block of symbols \((x)\), an LD space-time block-code forms a matrix of size \(T \times M_t\), denoted by \(S\), as follows:

\[
S = \sum_{n=1}^{N} A^{(n)}x_n + B^{(n)}x_n^*,
\]

where \(A^{(n)}\) and \(B^{(n)}\), called dispersion matrices, are fixed \(T \times M_t\) complex matrices, and \(x_i\) is usually selected from PSK or QAM constellations. Dispersion matrices can be defined based on different criteria. As an example, in orthogonal and quasi-orthogonal codes shown in tables 3.1-3.2 (on pages 29 and 34), dispersion matrices are selected subject to the constraint that all or some of the columns of the LD code are orthogonal to each other [1, 2, 25]. They can also be selected subject to the constraint that mutual information between transmitted and received signals is maximized [24].

Finally, at the \(i\)th symbol time, the \(i\)th row of \(S\) (i.e., \(s_i = [s_{i1}, s_{i2}, \ldots, s_{iM_t}]\)) is transmitted from \(M_t\) transmit antennas.
1.3.2 Channel Model

Let $h_{ij}$ denote the channel coefficient from transmitter $i$ to receiver $j$. Using Rayleigh fading model, these coefficients can be characterized as i.i.d. circularly symmetric complex Gaussian random variables with zero-mean, unit variance, and the following temporal autocorrelation function: $R_h(\tau) = J_0(2\pi f_D T_s \tau)$, where $J_0(\cdot)$ is the zero order Bessel function, $f_D$ is the doppler shift, and $T_s$ is the symbol duration. This channel has a coherence time of $T_c \propto \frac{1}{f_D}$, during which the channel coefficients are highly correlated.

When the duration of transmission of LD code is much less than the coherence time of the channel (i.e., $T_s T \ll T_c$), it is reasonable to assume that the channel is changing slow enough so that during the transmission of each block code, the channel coefficients are fixed; hence, the term quasi-static. In this thesis, it is assumed that the CSI is available for the receiver in terms of matrix $H \in \mathbb{C}^{M_t \times M_r}$, whose elements are $h_{ij}$ for $i = 1, \ldots, M_t$ and $j = 1, \ldots, M_r$. Accordingly, the following baseband input/output relationship can be used for the channel to determine the output of the channel at the $i$th symbol time:

$$r_i = s_i H + n_i = \left(\sum_{n=1}^{N} (a_{i}^{(n)} x_n + b_{i}^{(n)} x_n^*) \right) H + n_i.$$  \hspace{1cm} (1.2)

In this equation, $r_i = [r_{i1}; r_{i2}; \ldots; r_{iM_r}] \in \mathbb{C}^{1 \times M_r}$ denotes the received vector, where $r_{ij}$ is the signal received by the $j$th antenna at time $i$. Also, $s_i$, $a_{i}^{(n)}$, and $b_{i}^{(n)}$ are the $i$th rows of $S$, $A^{(n)}$, and $B^{(n)}$, respectively. The additive noise component in (1.2), denoted by $n_i \in \mathbb{C}^{1 \times M_r}$, is assumed to be spatially and temporally white with circularly symmetric complex Gaussian distribution.

1.3.3 Receiver Structure

During transmission of all rows of $S$ in $T$ symbol times, the received vectors $r_i$, $i = 1, \ldots, T$ will be collected in the following observation vector

$$y = [r_1; r_2; \ldots; r_T]^T \in \mathbb{C}^{(TM_r) \times 1}.$$   \hspace{1cm} (1.3)
The proposed receiver in Fig. 1.2 uses ML criterion for detection of $\mathbf{x}$, but it does not apply ML criterion directly to the observation vector $\mathbf{y}$. This receiver first estimates the transmitted vector using a widely linear MMSE estimator. It is proved in Chapter 4 that this estimate, denoted by $\tilde{\mathbf{x}}$, provides the minimal sufficient statistics for $\mathbf{x}$. Thus, $\tilde{\mathbf{x}}$ provides the same information about $\mathbf{x}$ as $\mathbf{y}$ does, and the ML detection criterion is applied to $\tilde{\mathbf{x}}$ in the second stage to detect the transmitted vector $\mathbf{x}$. Finally, the outputs of the ML detector, denoted by $\hat{x}_i$, will be de-mapped to a bit sequence at the last stage.

1.4 Thesis Contributions

In brief, the main contributions of this thesis are as follows:\footnote{Some of these results are published in [42] and [43].}

- In this work, it is proved that when the transmitted symbols and received signals are \textit{jointly improper Gaussian} signals, widely-linear MMSE estimates of the transmitted symbols are the \textit{minimal sufficient statistics} for detection of these symbols. This result is the main rationale behind our proposed structure for decoding LD codes. Moreover, for non-Gaussian symbols derived from PSK or QAM constellations, we have shown with simulation results that the assumption of sufficiency is still valid. It should be noted that in the literature, the issue of sufficiency of strictly-linear MMSE estimates has already been proved for \textit{proper} Gaussian signals (see [44]), but the discussion for \textit{improper} signals is absent in the literature.

- A new ML receiver structure for decoding LD codes is proposed. This receiver applies the ML detection criterion to the WL-MMSE estimates of the transmitted symbols rather than the original received signals. This structure offers both optimal and suboptimal operation modes. In the optimal mode, the receiver exhibits ML performance with the same complexity as conventional ML detectors. In suboptimal mode, the complexity of detection will be reduced from $O(L^N)$ to $O(LN)$,
where $L$ and $N$ denote the constellation size and the number of transmitted symbols, respectively. Simulation results show that the suboptimal receiver performs reasonably close to the optimal case, less than 1 dB away, for QAM constellations.

- The problem of ML detection of $M$-ary deterministic but unknown signals in the presence of improper noise is studied in this thesis. This ML detector is used in our proposed scheme for detection of transmitted symbols from the output of the WL-MMSE estimator. To the best of our knowledge, the problem of ML detection in improper noise has not been addressed in the literature till now.\textsuperscript{2} Our work shows that the ML detector in this case performs pseudo correlation [4] as well as conventional correlation of the observation to the signals-of-interest. As an alternative solution, we propose a filter, called circularizing filter, for converting improper signals to proper ones. This filter can be used as a preprocessing step for conventional tools (conventional detectors in this thesis) to enable them to deal with improper signals.

- By studying the structure of proposed WL-MMSE receiver for Alamouti code and comparing it to the decoding strategy of [1], we provide new insight into Alamouti’s combining scheme in [1]. We show that the combining scheme proposed in [1] is a scaled version of WL-MMSE estimation if certain conditions are satisfied.

1.5 Thesis Organization

Chapter 2 includes explanation of the notation used throughout the thesis, as well as a more detailed description of improper complex signals and their second order characteristics. Furthermore, a new approach for characterizing improper signals is introduced in

\textsuperscript{2}Although [45] has studied the mean square estimation and detection of improper signals-of-interest that are observed in proper noise, the problem of ML detection in the presence of improper noise has not been studied in general yet; except few works which have studied some special systems with BPSK signaling (e.g. [12, 14, 46]).
this chapter, which will be used in analyzing these signals in subsequent chapters.

Chapter 3 studies the proposed WL-MMSE for decoding LD codes. Due to the importance of orthogonal and quasi-orthogonal codes as two subsets of LD codes, this chapter examines the structure of the proposed receiver for these codes in detail.

Chapter 4 studies maximum likelihood detection of transmitted signals from the sufficient statistics provided by WL-MMSE estimator of chapter 3. It will be shown that this new observation might include improper additive noise. As a result, the structure of maximum likelihood detector in the presence of improper noise is studied in this chapter. Two equivalent ML detector structures will be proposed and their performance will be analyzed.

Finally, Chapter 5 concludes this work and proposes possible areas of research for continuing this work.
Chapter 2

Preliminaries

This chapter gives a brief overview of the concept of impropriety, and presents the tools required for analyzing complex signals in the rest of this thesis. The issues discussed in Sections 2.2 are borrowed from the literature, whereas Section 2.3 includes new analysis and tools. In order to distinguish these two cases, the latter section is marked with an asterisk.

2.1 Notation Description

Throughout this thesis, the following notation is used: lowercase letters for scalar variables (e.g. $z$), boldface lowercase letters for vectors (e.g. $\mathbf{z}$), and boldface uppercase letters for matrices (e.g. $\mathbf{Z}$). $\mathbb{R}$ and $\mathbb{C}$ represent the real and complex domains. $\Re\{z\}$ and $\Im\{z\}$ represent the real and imaginary parts of a complex variable $z$. The complex conjugate, transpose, and complex conjugate transpose of a vector $\mathbf{z}$ will be denoted by $\mathbf{z}^*$, $\mathbf{z}^T$, and $\mathbf{z}^H$, respectively. For an invertible matrix $\mathbf{Z}$, $\mathbf{Z}^{-1}$ represents the inverse of $\mathbf{Z}$ and $\mathbf{Z}^{-*}$ represents the complex conjugate of $\mathbf{Z}^{-1}$. The identity matrix of size $k \times k$ is denoted by $\mathbf{I}_k$.

Let $z = x + jy$ be a complex random variable. The variance and pseudo-variance [4] of $z$ will be respectively represented by $\sigma_z^2 \triangleq \mathbb{E}\{|z - \bar{z}|^2\} = \sigma_x^2 + \sigma_y^2$ and $\gamma_z^2 \triangleq \mathbb{E}\{(z - \bar{z})^2\} =$
\[(\sigma^2_x - \sigma^2_y) + j(2\sigma_{xy}), \text{ where } \sigma_{xy} \text{ denotes the covariance between } x \text{ and } y.\]
Similarly, the conventional covariance matrix and pseudo-covariance matrix between two complex valued random vectors \(z_1\) and \(z_2\) are denoted by 
\[
C_{z_1z_2} \triangleq \mathbb{E}\{(z_1 - \bar{z}_1)(z_2 - \bar{z}_2)^H\} \quad \text{and} \quad C_{z_1z_2^*} \triangleq \mathbb{E}\{(z_1 - \bar{z}_1)(z_2 - \bar{z}_2)^T\}.
\]

### 2.2 Second Order Statistics of Complex Signals

Let \(z = [z_1, z_2, \ldots, z_K]^T\) be a random vector each element of which is a complex-valued random variable with the following definition: \(z_i = x_i + jy_i\) for \(i = 1, \ldots, K\). Using this representation, we can decompose \(z\) into its real and imaginary parts: \(z = x + jy\), where \(x = \Re\{z\} = [x_1, x_2, \ldots, x_K]^T\) and \(y = \Im\{z\} = [y_1, y_2, \ldots, y_K]^T\). This leads us to a mapping from any complex vector \(z \in \mathbb{C}^K\) to a real vector \(w(z) \in \mathbb{R}^{2K}\), or simply \(w_z\), using the following definition:

\[
w_z = \begin{bmatrix} x \\ y \end{bmatrix} = [x_1, \ldots, x_K, y_1, \ldots, y_K]^T. \tag{2.1}
\]

This transformation is an isomorphism from \(\mathbb{C}^K\) onto \(\mathbb{R}^{2K}\), and there exists a one-to-one relationship between the members of these two vector spaces. In order to have a linear transformation matrix for generating \(w_z\) not only the vector \(z\) but also \(z^*\) is required. This motivates us to use an augmented vector \(v(z) \in \mathbb{C}^{2K}\), or simply \(v_z\), including both \(z\) and \(z^*\) as its elements:\footnote{The factor \(\frac{1}{\sqrt{2}}\) assures us to have the same power in augmented vector \(v_z\) compared to the vector \(z\).}

\[
v_z = \frac{1}{\sqrt{2}}[z^T, z^H]^T = \frac{1}{\sqrt{2}}[z_1, z_2, \ldots, z_K, z_1^*, z_2^*, \ldots, z_K^*]^T. \tag{2.2}
\]

It follows that the linear transformation from \(v_z\) to \(w_z\) is in the form of

\[
w_z = T_K v_z, \quad \text{where} \quad T_K = \frac{1}{\sqrt{2}} \begin{bmatrix} I_K & I_K \\ -jI_K & jI_K \end{bmatrix}. \tag{2.3}
\]
\( T_K \) is a unitary transformation which preserves the power of the random vector during the transformation.

Now consider the problem of finding first and second order statistics of complex-valued random vector \( z \). Owing to the fact that these statistics are well defined for real-valued random vectors in the literature, one can make use of the aforementioned mapping to find the first and second order statistics of \( z \) based on the statistics of \( w_z \).

Using this approach, it can be easily shown that the first order statistics of \( z \) requires the knowledge of \( \mathbb{E}\{w_z\} = [\bar{x}^T, \bar{y}^T]^T \), where \( \bar{x} = \mathbb{E}\{x\} \) and \( \bar{y} = \mathbb{E}\{y\} \). This results in a simple generalization of the first order statistics for complex vectors as follows: \( \bar{z} = \mathbb{E}\{z\} = \bar{x} + j\bar{y} \). Similarly, second order characterization of \( z \) requires the knowledge of the covariance matrix of \( w_z \) defined as follows:

\[
C_{w_zw_z^H} = \mathbb{E}\{(w_z - w_{\bar{z}})(w_z - w_{\bar{z}})^H\} = \begin{bmatrix}
C_{xx^T} & C_{xy^T} \\
C_{yx^T} & C_{yy^T}
\end{bmatrix}.
\tag{2.4}
\]

This is equivalent to the knowledge of \( C_{v_zv_z^H} = T_K^H C_{w_zw_z^H} T_K \), which is given by

\[
C_{v_zv_z^H} = \mathbb{E}\{(v_z - v_{\bar{z}})(v_z - v_{\bar{z}})^H\} = \frac{1}{2} \begin{bmatrix}
C_{zz^H} & C_{zz^T} \\
C_{zz^T}^* & C_{zz^{H*}}
\end{bmatrix}.
\tag{2.5}
\]

Equation (2.5) reveals that, unlike real-valued random vectors, second order characterization of a complex-valued random vector \( z \), requires knowledge of both the following matrices [4]:

\[
\text{covariance of } z : \quad C_{zz^H} = \mathbb{E}\{(z - \bar{z})(z - \bar{z})^H\}, \tag{2.6}
\]

\[
\text{pseudo-covariance of } z : \quad C_{zz^T} = \mathbb{E}\{(z - \bar{z})(z - \bar{z})^T\}. \tag{2.7}
\]

The following equations show the relationship between covariance/pseudo-covariance matrices of \( z \) and the covariance matrices of real vectors \( x \) and \( y \):

\[
C_{zz^H} = C_{xx^T} + C_{yy^T} + j(C_{yx^T} - C_{xy^T}) \quad (2.8)
\]

\[
C_{zz^T} = (C_{xx^T} - C_{yy^T}) + j(C_{xy^T} + C_{yx^T}) \quad (2.9)
\]
Note that $C_{zz}^H$ is a Hermitian matrix, while $C_{zz}^T$ is symmetric. Equations (2.8) and (2.9) show that $C_{zz}^H$ is not sufficient to uniquely determine $C_{xx}^T$, $C_{yy}^T$, and $C_{xy}^T$. Consequently, knowledge of $C_{zz}^T$ as well as $C_{zz}^H$ is necessary for completely characterizing the random vector $z$ in second order.

### 2.2.1 Proper vs. Improper Complex Signals

In the above discussion, it was mentioned that complete second order characterization of complex-valued random vector $z$ requires both $C_{zz}^T$ and $C_{zz}^H$. By convention, however, when $C_{zz}^T = 0$, instead of explicitly mentioning the value of $C_{zz}^T$, it is verbally mentioned that $z$ is a proper or circularly symmetric random vector.

**Definition 2.1** Random vector $z$ is called proper [4] or circularly symmetric [5] if $C_{zz}^T = 0$; otherwise, it is called improper or non circularly symmetric.

From (2.9), it can be seen that a proper random vector $z$ has the following properties:

- Real and imaginary parts of $z$ have equal covariance matrices, i.e., $C_{xx}^T = C_{yy}^T$.

- The cross-covariance matrix between the real and imaginary parts of $z$ is an anti-symmetric matrix, i.e., $C_{xy}^T = -C_{yx}^T$.

The former condition requires all elements of this vector ($z_k = x_k + jy_k$, $k = 1, \ldots, K$) to have their power equally distributed between their real and imaginary parts (i.e., $\sigma_{z_k}^2 = \sigma_{y_k}^2$), and the latter condition requires each $z_k$ to have uncorrelated real and imaginary parts (i.e., $\sigma_{x_ky_k} = 0$). Note that these conditions on $z_k$ are necessary but not sufficient conditions to get $C_{xx}^T = C_{yy}^T$ and $C_{xy}^T = -C_{yx}^T$.

A simple example for proper vector $z = x + jy$ is the case where $x$ and $y$ are uncorrelated ($C_{xy}^T = 0$) and have the same covariance matrices. However, it should be noted that a proper $z$ can still have correlated real and imaginary parts (i.e., $C_{xy}^T \neq 0$), while satisfying the conditions above. As a case in point, consider $z = [x + jy, y - jx]^T$,
where $x$ and $y$ are uncorrelated and have equal powers (i.e., $\sigma_x^2 = \sigma_y^2 = \sigma_0$). This vector has correlated $x$ and $y$. However, since

$$C_{xx} = C_{yy} = \sigma_0 I_2, \quad C_{xy} = -C_{yx} = \sigma_0$$

we get $C_{zz} = 0$; hence, $z$ is proper.

In special case, when $K = 1$ we have a scalar random variable $z$, whose second order statistics are characterized by its variance and pseudo-variance as follows:

$$\sigma_z^2 \triangleq C_{zz} = E\{(z - \bar{z})(z - \bar{z})^*\} = E\{|z - \bar{z}|^2\} = \sigma_x^2 + \sigma_y^2 \in \mathbb{R}$$
$$\gamma_z^2 \triangleq C_{zz} = E\{(z - \bar{z})(z - \bar{z})\} = E\{(z - \bar{z})^2\} = (\sigma_x^2 - \sigma_y^2) + j(2\sigma_{xy}) \in \mathbb{C}$$

In this case, $z$ is improper if the power of $z$ is not equally distributed between its real and imaginary parts ($\sigma_x^2 \neq \sigma_y^2$) or there exists a correlation between these two parts ($\sigma_{xy} \neq 0$).

### 2.2.2 Probability Density Function of Complex Gaussian Random Vectors

By definition, $z$ is a complex Gaussian random vector if and only if the elements of $w_z$ are jointly Gaussian random variables [41]. In order to find the probability density function (pdf) of a Gaussian random vector $z$, we utilize the well known multivariate Gaussian distribution for $w_z$, given by:

$$f_{w_z}(w_z) = \frac{1}{\sqrt{2\pi C_{w_z}w_z^T}} \exp\left\{-\frac{1}{2}(w_z - w_{\bar{z}})^TC_{w_z}^{-1}(w_z - w_{\bar{z}})\right\}.$$  

Since there exists a unitary transformation between $v_z$ and $w_z$, it can be shown [41] that the pdf of $z$ is in the form of:

$$f_z(z) = f_{v_z}(v_z) = \frac{1}{\sqrt{2\pi C_{v_z}v_z^H}} \exp\left\{-\frac{1}{2}(v_z - v_{\bar{z}})^HC_{v_z}^{-1}(v_z - v_{\bar{z}})\right\}.$$
where $v_z = \frac{1}{\sqrt{2}}[z^T, z^H]^T$ is the augmented vector of $z$, and $C_{v_z v_z^H}$ is given in (2.5). Since $C_{v_z v_z^H}$ is a block matrix, its inverse can be expressed as follows [47]

$$C_{v_z v_z^H}^{-1} = 2 \begin{bmatrix} Q_z^{-*} & -Q_z^{-*}P_z^* \\ -Q_z^{-1}P_z & Q_z^{-1} \end{bmatrix} \quad (2.15)$$

where $Q_z = C_{zz}^* - C_{zz}^T C_{zz}^{-1} C_{zz}^T$ and $P_z = C_{zz}^* C_{zz}^{-1}$. By substituting (2.15) in (2.14), $f_z(z)$ can be decomposed as follows:

$$f_z(z) = \frac{1}{\sqrt{\pi |C_{zz}^H|}} \exp \left\{ -z^H Q_z^{-*}z + \Re \{z^T Q_z^{-1}P_z z\} \right\}$$

$$= \frac{1}{|\pi C_{zz}^H|} \exp \left\{ -z^H C_{zz}^H z \right\} \times \frac{1}{\sqrt{|C_{zz}^{-1} Q_z|}} \exp \left\{ \Re \{ -z^H P_z^H Q_z^{-1}P_z z + z^T Q_z^{-1}P_z z \} \right\}.$$  

(2.16)

The first term in (2.16) only depends on the covariance matrix of $z$, while the second term requires the knowledge of pseudo-covariance of $z$ as well.

As a special case, when $z$ is a proper Gaussian vector, $C_{zz}^T = 0$ and $Q_z = C_{zz}^H$. In this case, the pdf of $z$ can be simplified to

$$f_z(z) = \frac{1}{|\pi C_{zz}^H|} \exp \left\{ -z^H C_{zz}^H z \right\}, \quad (2.17)$$

which is the well-known expression for the pdf of a proper complex Gaussian vector.

### 2.3 Noncircularity Matrix for Complex Random Vectors

In Section 2.2.1, it was mentioned that the pseudo-covariance matrix $C_{zz}^T$ is usually used to distinguish improper signals from proper ones. However, this matrix is not suitable for comparing two improper signals, in that simple scaling of the elements of $z$ changes $C_{zz}^T$. 
In general, let $\tilde{z} = Dz$, where $z \in \mathbb{C}^K$ and $D = \text{diag}(d_1, \ldots, d_K) \in \mathbb{R}^{K \times K}$. Then, we get $C_{\tilde{z}\tilde{z}} = DC_{zz}D \neq C_{zz}$. This shows that, $C_{zz}$ depends on not only the impropriety of the $z$, but the power of elements of $z$ as well.

In order to mitigate this dependency on the power of $z$, the following definition, proposes a new matrix for measuring the noncircularity (or impropriety) of the $z$.

**Definition 2.2** Given a complex random vector $z$, the noncircularity matrix of $z$ is defined as $A_z \triangleq C_{zz}^{-\frac{1}{2}}C_{zzr}C_{zzr}^{-\frac{1}{2}}$, where $C_{zz}$ and $C_{zzr}$ are the covariance matrix and pseudo-covariance matrix of $z$ defined in (2.6) and (2.7).

The following Theorem shows that the noncircularity matrix $A_z$ is independent from the changes in the powers of elements of $z$, and only conveys information about the impropriety of $z$.

**Theorem 2.1** The noncircularity matrix $A_z$, defined in Definition 2.2, is invariant under real-valued scaling of the elements of $z$.

*Proof:* see Appendix A.

In particular, when $z$ is a scalar random variable ($K = 1$), we get a complex scalar value for $A_z$, which will be denoted by $\alpha_z$, as follows

$$\alpha_z = \frac{\gamma_z^2}{\sigma_z^2} = \left( \frac{\sigma_x^2 - \sigma_y^2}{\sigma_x^2 + \sigma_y^2} \right) + j \left( \frac{2\sigma_{xy}}{\sigma_x^2 + \sigma_y^2} \right). \quad (2.18)$$

In this case, we call $\alpha_z$ the *noncircularity coefficient* of $z$. Note that the real and imaginary parts of $\alpha_z$ are normalized measures of the power difference between $x$ and $y$ and the correlation between $x$ and $y$, respectively.

To exemplify, suppose that $z$ is randomly selected from a constellation of size $L$, denoted by $\Omega = \{z_1, z_2, \ldots, z_L\}$. The following are simple examples for possible values of $\alpha_z$ depending on the structure of this constellation:

- $\Omega = \{e^{j\frac{\pi}{4}}, e^{-j\frac{\pi}{4}}, e^{j(\pi+\frac{\pi}{4})}, e^{j(\pi-\frac{\pi}{4})}\}$: In this case, the real and imaginary parts of $z$ are uncorrelated and have equal powers. Therefore, we get $\alpha_z = 0$. 

• $\Omega = \{e^{j\pi/8}, e^{-j\pi/8}, e^{j(\pi + \pi/8)}, e^{j(\pi - \pi/8)}\}$: This constellation has a rectangular shape, which results in a power imbalance between the real and imaginary parts of $z$. Consequently, we get $\alpha_z = \frac{1}{\sqrt{2}}$.

• $\Omega = \{-3e^{j\pi/8}, -e^{j\pi/8}, e^{j\pi/8}, 3e^{j\pi/8}\}$: In this case, the real part and imaginary part of $z$ are completely correlated to each other ($\Re\{z\} = \Im\{z\}$). Consequently, we get $\alpha_z = j$.

• $\Omega = \{-3e^{j\pi/8}, -e^{j\pi/8}, e^{j\pi/8}, 3e^{j\pi/8}\}$: In this case, not only the real and imaginary parts of $z$ are correlated to each other, but also there exists a power imbalance between these parts. As a result, we get $\alpha_z = \frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$.

As it can be seen in the above examples, the real part of $\alpha_z$ depends on the power imbalance between the real and imaginary parts of $z$, whereas the imaginary part of $z$ depends on the correlation between real and imaginary parts of $z$.

Decomposition of the pseudo-variance as $\gamma_z^2 = \sigma_z^2 \alpha_z$ gives us this ability to distinguish between the changes in $\gamma_z^2$ caused by changing the power of $z$ as opposed to the changes caused by changing the structure of the $z$. The following theorem shows that the noncircularity coefficient $\alpha_z$ lies within the unit circle for all random variables.

**Theorem 2.2** The magnitude of the the noncircularity coefficient of a complex random variable is upper bounded by one (i.e. $0 \leq |\alpha_z| \leq 1$).

*Proof:* see Appendix A.

Fig. 2.1 illustrates different possible values of $\alpha_z$ and the corresponding values for $\sigma_z^2$, $\sigma_y^2$, and $\sigma_{xy}$. It can be seen that $\alpha_z$ takes its minimum magnitude ($\alpha_z = 0$) in the case of circularly symmetric random variable $z$, when $\gamma_z^2 = 0$, and becomes larger and larger as the value of $|\gamma_z^2|$ increases and finally takes its maximum magnitude ($|\alpha_z| = 1$) when $|\gamma_z^2|$ becomes equal to $\sigma_z^2$.

In Section 2.2.2, it was mentioned that second order characterization of $z$ requires knowledge of $C_{v_z v_z^T}$ or $C_{w_z w_z^T}$. Table 2.1 summarizes different equivalent representations
Chapter 2. Preliminaries

Figure 2.1: Possible values of $\alpha_z$ in the complex plane.

of these covariance matrices. The second column of this table reveals the fact for a given signal power, $\alpha_z$ can completely characterize both $C_{v_z v_z^T}$ and $C_{w_z w_z^T}$.

Accordingly, the following lemma shows the relationship between eigen values/vectors of $C_{v_z v_z^T}$ and $C_{w_z w_z^T}$ and the noncircularity coefficient $\alpha_z$.

**Lemma 2.1** Let $C_{v_z v_z^T} = QA_1Q^H$ and $C_{w_z w_z^T} = UA_2U^T$ be the eigen decomposition of covariance matrices defined in (2.5) and (2.4). $A_1$ and $A_2$ are diagonal matrices whose diagonal elements are, respectively, the eigen values of $C_{v_z v_z^T}$ and $C_{w_z w_z^T}$ in non-increasing order, and $Q$ and $U$ are unitary matrices whose columns are the eigenvectors of $C_{v_z v_z^T}$ and $C_{w_z w_z^T}$ corresponding to the eigenvalues in $A_1$ and $A_2$, respectively. Then

1. $A_1 = A_2 = \Lambda = \frac{1}{2} \begin{bmatrix} \sigma_z^2 + |\gamma_z|^2 & 0 \\ 0 & \sigma_z^2 - |\gamma_z|^2 \end{bmatrix} = \frac{\sigma_z^2}{2} \begin{bmatrix} 1 + |\alpha_z| & 0 \\ 0 & 1 - |\alpha_z| \end{bmatrix}$,
Table 2.1: Covariance matrices in the case of scalar complex random variable, in terms of $\sigma_z^2$, $\gamma_z^2$, $\alpha_z$, $\sigma_x^2$, $\sigma_y^2$, and $\sigma_{xy}$

<table>
<thead>
<tr>
<th></th>
<th>$\sigma_z^2, \gamma_z^2$</th>
<th>$\sigma_z^2, \alpha_z$</th>
<th>$\sigma_z^2, \sigma_x^2, \sigma_y^2, \sigma_{xy}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{v,v'}$</td>
<td>$\frac{1}{2} \begin{bmatrix} \sigma_z^2 &amp; \gamma_z^2 \ (\gamma_z^2)^* &amp; \sigma_z^2 \end{bmatrix}$</td>
<td>$\frac{1}{2} \begin{bmatrix} 1 &amp; \alpha_z \ \alpha_z^* &amp; 1 \end{bmatrix}$</td>
<td>$\frac{1}{2} \begin{bmatrix} \sigma_z^2 + \sigma_y^2 &amp; (\sigma_z^2 - \sigma_y^2) + j(2\sigma_{xy}) \ (\sigma_z^2 - \sigma_y^2) - j(2\sigma_{xy}) &amp; \sigma_z^2 + \sigma_y^2 \end{bmatrix}$</td>
</tr>
<tr>
<td>$C_{w,v'}^{-1}$</td>
<td>$\frac{2}{\sigma_z^2 -</td>
<td>\gamma_z</td>
<td>^2} \begin{bmatrix} \sigma_z^2 &amp; -\gamma_z^2 \ -(\gamma_z^2)^* &amp; \sigma_z^2 \end{bmatrix}$</td>
</tr>
<tr>
<td>$C_{w,w'}$</td>
<td>$\frac{1}{2} \begin{bmatrix} \sigma_z^2 + \Re(\gamma_z^2) &amp; \Im(\gamma_z^2) \ \Im(\gamma_z^2) &amp; \sigma_z^2 - \Re(\gamma_z^2) \end{bmatrix}$</td>
<td>$\frac{1}{2} \begin{bmatrix} 1 + \Re(\alpha_z) &amp; \Im(\alpha_z) \ \Im(\alpha_z) &amp; 1 - \Re(\alpha_z) \end{bmatrix}$</td>
<td>$\begin{bmatrix} \sigma_z^2 &amp; \sigma_{xy} \ \sigma_{xy} &amp; \sigma_y^2 \end{bmatrix}$</td>
</tr>
<tr>
<td>$C_{w,w'}^{-1}$</td>
<td>$\frac{2}{\sigma_z^2 -</td>
<td>\gamma_z</td>
<td>^2} \begin{bmatrix} \sigma_z^2 - \Re(\gamma_z^2) &amp; -\Im(\gamma_z^2) \ -\Im(\gamma_z^2) &amp; \sigma_z^2 + \Re(\gamma_z^2) \end{bmatrix}$</td>
</tr>
</tbody>
</table>

$|C_{v,v'}| = \frac{\sigma_z^4 - |\gamma_z|^4}{4}$

$|C_{w,w'}| = \frac{\sigma_z^4(1 - |\alpha_z|^2)}{4}$
2. \( \mathbf{U} = \begin{bmatrix} \cos \frac{\varphi}{2} & -\sin \frac{\varphi}{2} \\ \sin \frac{\varphi}{2} & \cos \frac{\varphi}{2} \end{bmatrix} = \mathbf{R}_w(\varphi_2) \),

3. \( \mathbf{Q} = \mathbf{T}^H \mathbf{U} = \begin{bmatrix} e^{j\frac{\varphi}{2}} & e^{j\frac{\varphi + \pi}{2}} \\ e^{-j\frac{\varphi}{2}} & e^{-j\frac{\varphi + \pi}{2}} \end{bmatrix} = \begin{bmatrix} e^{j\frac{\varphi}{2}} & 0 \\ 0 & e^{-j\frac{\varphi}{2}} \end{bmatrix} \mathbf{T}^H = \mathbf{R}_v(\varphi_2) \mathbf{T}^H \)

where \( \varphi \) denotes the phase of noncircularity coefficient \( \alpha_z \), and \( \mathbf{R}_w(\varphi_2) \) and \( \mathbf{R}_v(\varphi_2) \) are unitary transformations on \( \mathbf{w}_z \) and \( \mathbf{v}_z \), respectively, both of which result in the rotation of \( \mathbf{z} \) in the complex plane by phase \( \varphi_2 \).

In order to get a better insight to the information conveyed by \( \alpha_z \), let us study the pdf of a Gaussian random variable \( \mathbf{z} \), which is given by

\[
f_{\mathbf{z}}(\mathbf{z}) = \frac{1}{\sqrt{2\pi \mathbf{C}_{\mathbf{w}_z \mathbf{w}_z}^T}} \exp \left\{ -\frac{1}{2} (\mathbf{w}_z - \mathbf{w}_z)^T \mathbf{C}_{\mathbf{w}_z \mathbf{w}_z}^{-1} (\mathbf{w}_z - \mathbf{w}_z) \right\}
\]

(2.19)

\[
f_{\mathbf{z}}(\mathbf{z}) = \frac{1}{\sqrt{2\pi \mathbf{C}_{\mathbf{v}_z \mathbf{v}_z}^T}} \exp \left\{ -\frac{1}{2} (\mathbf{v}_z - \mathbf{v}_z)^H \mathbf{C}_{\mathbf{v}_z \mathbf{v}_z}^{-1} (\mathbf{v}_z - \mathbf{v}_z) \right\}.
\]

(2.20)

Without loss of generality, assume \( \mathbf{z} \) is zero mean. According to (2.19), the pdf of \( \mathbf{z} \) is exponentially related to a quadratic term of the form \( \mathbf{w}_z^T \mathbf{C}_{\mathbf{w}_z \mathbf{w}_z}^{-1} \mathbf{w}_z \). Therefore, the equiprobable points in the \( \mathbf{z} \)-plane are the solutions of the equation \( \mathbf{w}_z^T \mathbf{C}_{\mathbf{w}_z \mathbf{w}_z}^{-1} \mathbf{w}_z = c_0 \), where \( c_0 \) is a constant value. This leads to a well known elliptic equation of the form

\[
c_0' = \frac{x^2}{\sigma_x^2} - \frac{(2\sigma_{xy})^2}{\sigma_y^2} x y + \frac{y^2}{\sigma_y^2} = 2 \frac{\sigma_{xy}^2}{\sigma_y^2}
\]

(2.21)

where \( c_0' = (1 - \frac{\sigma_{xy}^2}{\sigma_x^2 \sigma_y^2})c_0 \), implying that the equiprobable points in the \( \mathbf{z} \)-plane are located on rotated elliptical contours. The angle of rotation, denoted by \( \phi \), for these ellipses can be calculated as follows

\[
\tan 2\phi = \frac{-2\sigma_{xy}}{\sigma_x^2 - \sigma_y^2} = \frac{2\sigma_{xy}}{\sigma_x^2 - \sigma_y^2} = \tan \varphi \quad \Rightarrow \quad \phi = \frac{\varphi}{2}
\]

(2.22)

where \( \varphi = \angle \alpha_z \). Furthermore, it can be shown that the ratio between the major and minor axes of each ellipse is equal to \( \frac{1 + |\alpha_z|}{1 - |\alpha_z|} \). Therefore, the magnitude and phase of \( \alpha_z \) are sufficient to characterize the structure of pdf of a Gaussian random variable.
Chapter 2. Preliminaries

Figure 2.2: Contours of pdf for a scalar zero mean complex valued Gaussian random variable

2.4 Chapter Summary

In this chapter, we presented a brief overview of the definition and properties of improper complex signals. Since impropriety is a second-order characteristic of the signal, we focused on the second order properties of complex signals. It was shown that second-order characterization of complex random vectors requires the knowledge of their pseudo-covariance matrix as well as conventional covariance matrix. This fact was particularly studied for complex Gaussian vectors, and it was shown that the pdf of these vectors depends on both covariance and pseudo-covariance matrices.

Moreover, it was shown that pseudo-covariance matrix is not suitable for characterizing the impropriety of a vector, in that it is sensitive to simple scaling of the vector. As a result, we proposed to use a new matrix, called noncircularity matrix, for characterizing the improper structure of complex vectors. The properties of this matrix was then studied for a simple case of scalar complex random variable. It was shown that the noncircularity coefficient of a complex random variable conveys all the required information for characterizing its impropriety. This new measure will be utilized in Chapter 4 for analyzing the structure of ML detectors in the presence of improper noise.
Chapter 3

Widely-Linear MMSE Receivers for LD Codes

This chapter considers the MIMO system model of Section 1.3, and proposes a new receiver structure for this system. As mentioned in Section 1.2, existing methods for decoding LD codes do not take into account the inherent improper structure of these codes. Our proposed method, however, exploits this characteristic of LD codes to develop a new decoding strategy. Due to the fact that the transmitted and received signals in our system model are jointly improper, we suggest using WL-MMSE estimates of the transmitted symbols in lieu of the sufficient statistics for ML detection of these symbols, as follows.

Assume that the transmitter uses LD code of (1.1) to transmit the vector \( \mathbf{x} = [x_1, x_2, \ldots, x_N]^T \) through the MIMO channel. The proposed receiver in Fig. 1.2 uses the ML criterion for detection of \( \mathbf{x} \), but it does not apply the ML criterion directly to the observation vector \( \mathbf{y} \). This receiver first estimates the transmitted vector using a widely-linear MMSE estimator. This estimate, denoted by \( \tilde{\mathbf{x}} \), provides the minimal sufficient statistics for \( \mathbf{x} \). Thus, \( \tilde{\mathbf{x}} \) provides the same information about \( \mathbf{x} \) as \( \mathbf{y} \) does, and the ML detection criterion is applied to \( \tilde{\mathbf{x}} \) at the second stage to detect the transmitted vector.
\section{3.1 Widely-Linear MMSE Receiver Structure}

It has been shown in \cite{40} that when a random variable \( x \) is to be estimated from a complex-valued observation vector \( y \), the widely-linear estimator of the form

\[ \tilde{x} = f y + g y^* \]

(3.1)

gives the minimum mean squared error if

\begin{align*}
    f &= [C_{xy}^{-1} - C_{xy} C_{yy}^{-1} C_{yy}^{-1} C_{yy}^{-1}]^{-1}, \\
    g &= [C_{xy}^{-1} - C_{xy} C_{yy}^{-1} C_{yy}^{-1} C_{yy}^{-1}]^{-1}. 
\end{align*}

(3.2) (3.3)

This estimator is called widely linear with respect to \( y \) since it linearly processes both \( y \) and \( y^* \) \cite{40}. By applying the results of \cite{40} to the system model of Section 1.3, we propose to estimate the transmitted symbols \( x_i \), \( i = 1, \ldots, N \), from observation vector \( y \) in (1.3) using \( \tilde{x}_i = f_i y + g_i y^* \), where \( f_i \) and \( g_i \) can be determined from (3.2) and (3.3) by substituting \( x \) with \( x_i \) in these equations. It can be shown that when the LD code given in (1.1) is used at the receiver side, we get

\begin{align*}
    C_{x_i y} &= \sigma^2 x_i \left[ a_1^{(i)} H, a_2^{(i)} H, \ldots, a_T^{(i)} H \right]^* + \gamma^2 x_i \left[ b_1^{(i)} H, b_2^{(i)} H, \ldots, b_T^{(i)} H \right]^*, \\
    C_{x_i y^T} &= \sigma^2 x_i \left[ b_1^{(i)} H, b_2^{(i)} H, \ldots, b_T^{(i)} H \right] + \gamma^2 x_i \left[ a_1^{(i)} H, a_2^{(i)} H, \ldots, a_T^{(i)} H \right], \\
    C_{yy} &= C_{nn} + \sum_{n=1}^{N} \left[ \sigma^2_{x_n} \left( A^{(n)H} A^{(n)H} + B^{(n)H} B^{(n)H} \right) \right. \\
    &\quad \left. + \gamma^2_{x_n} A^{(n)H} B^{(n)H} + (\gamma^2_{x_n})^* B^{(n)H} A^{(n)H} \right], \\
    C_{yy^T} &= \sum_{n=1}^{N} \left[ \sigma^2_{x_n} \left( A^{(n)H} B^{(n)T} + B^{(n)H} A^{(n)T} \right) \right. \\
    &\quad \left. + \gamma^2_{x_n} A^{(n)H} A^{(n)T} + (\gamma^2_{x_n})^* B^{(n)H} B^{(n)T} \right].
\end{align*}

(3.4) (3.5) (3.6) (3.7)
where $a_j^{(i)}$ and $b_j^{(i)}$ are the $j^{th}$ rows of the dispersion matrices $A^{(i)}$ and $B^{(i)}$ in (1.1), $C_{nn'}$ is the covariance matrix of the noise, and $\sigma^2_{x_i}$ and $\gamma^2_{x_i}$ are the variance and pseudo-variance of $x_i$.

A simple inspection of the equations (3.2)-(3.7) reveals that there are two transmission scenarios for which the receiver has to use the widely-linear estimator of (3.1) and the conventional strictly-linear estimator ($\tilde{x}_i = C_{x_i y'} C_{yy'}^{-1} y$) is not optimal:

1. The LD code transmits complex conjugate of the symbol ($x^*_i$), i.e., the dispersion matrix $B^{(i)}$ is nonzero. This condition holds true for orthogonal and quasi-orthogonal codes proposed by [1, 2, 25] and also for LD codes proposed in [24].

2. The modulation scheme used for $x_i$ generates improper symbols, i.e., $\gamma^2_{x_i} \neq 0$. As mentioned in Section 2.2.1, $x_i$ is improper if the real and imaginary parts of $x_i$ are correlated to each other or have nonequal powers. The $\gamma^2_{x_i} \neq 0$ condition holds true for BPSK modulation, as studied in [32] for layered ST-BCs, as well as some complex-valued modulation schemes that are proposed for quasi-orthogonal codes in [48].

In both cases, $C_{x_i y'}$ becomes nonzero, and (3.1) provides the minimum estimation error [40].

After estimating the vector $\tilde{x} = [\tilde{x}_1, \ldots, \tilde{x}_N]^T$, the receiver applies the ML criterion to $\tilde{x}$ in order to find the most likely transmitted vector. The following theorem, which is proved in Section 3.3, justifies why the receiver is allowed to detect the transmitted symbols using these widely-linear MMSE estimates instead of the original observation $y$.

**Theorem 3.1** For jointly Gaussian\(^1\) complex vectors $x$ and $y$, the WL-MMSE estimate of $x$, denoted by $\tilde{x} = [\tilde{x}_1, \ldots, \tilde{x}_N]$, is a minimal sufficient statistic for detection of $x$ from $y$.

\(^1\)In this section and Section 3.2, we assume that the vectors $x$ and $y$ are jointly Gaussian. In Section 3.4, the validity of this assumption for different cases will be examined.
Proof: see Section 3.3

This theorem implies that vector $\tilde{x} = [\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_N]^T$ conveys the same information about $x$ as the observation vector $y$. Thus, the ML criterion can be applied to $\tilde{x}$ instead of $y$ as follows:

$$\hat{x} = \arg \max_x f_{y|x}(y|x) = \arg \max_x f_{\tilde{x}|x}(\tilde{x}|x).$$

(3.8)

Note that ML detection of $x$ both from $y$ and from $\tilde{x}$ requires us to search over all possible combinations of $x_1, \ldots, x_N$ to find the most likely transmitted vector $\hat{x} = [\hat{x}_1, \ldots, \hat{x}_N]^T$. The only exception is when the code has a specific orthogonal or quasi-orthogonal structure which results in reduction of the joint ML decision rule to separate decision rules [1, 2, 25]. Thus, for nonorthogonal LD codes, the complexity of detecting $\hat{x}$ is $O(L^N)$, where $L$ denotes the constellation size. However, the main advantage of using $\tilde{x}$ instead of $y$ is that the resulting receiver can work in two operational modes: optimal and suboptimal. Wherever the complexity of the receiver is of major importance, the joint detector of Fig. 1.2 can be replaced by separate detectors of the form $\hat{x}_i = \arg \max_{x_i} f_{\tilde{x}_i|x_i}(\tilde{x}_i|x_i)$. This will reduce the complexity of searching from $O(L^N)$ to $O(LN)$. Although the resulting receiver is no longer optimal in ML sense; the simulation results show that it exhibits only a slight loss in the performance and performs reasonably close to the optimal receiver. An example of this suboptimal receiver will be discussed in Section 3.2.4.

### 3.2 Analysis of WL-MMSE Receiver for Some LD Codes

In the previous section, it was mentioned that there are two transmission scenarios which require utilization of WL-MMSE receiver, namely, $B^{(i)} \neq 0$ or $\gamma^2_{x_i} \neq 0$. This section
Chapter 3. Widely-Linear MMSE Receivers for LD Codes

\[ N = T = M_t = 2 \]

\[
\begin{bmatrix}
  x_1 \\
  x_2 \\
  -x_2^* \\
  x_1^*
\end{bmatrix}
\]

\[
\begin{bmatrix}
  1 & 0 \\
  0 & 0 \\
  0 & 1 \\
  0 & 0
\end{bmatrix}
\]

\[
\begin{bmatrix}
  0 & 0 \\
  0 & 1 \\
  0 & 0 \\
  -1 & 0
\end{bmatrix}
\]

Table 3.1: Orthogonal code of \([1]\) and corresponding dispersion matrices

focuses on the former scenario and studies the structure of the proposed WL-MMSE receiver for two important cases of LD codes which satisfy \(B^{(i)} \neq 0\).

3.2.1 WL-MMSE Receiver for an Orthogonal LD Code with \(M_t = 2, M_r = 1, N = 2\)

The space-time code shown in Table 3.1, which was originally presented by Alamouti in \([1]\), is a simple example of an LD code with two transmit antennas. This code is studied in this section as a representative of orthogonal codes, which are a subset of LD codes. Orthogonal codes are of particular interest in MIMO systems due to the fact that orthogonality of the columns of transmitted block \(S\) (i.e., \(S^H S = (\sum_{i=1}^{N} |x_i|^2)I_N\)) results in a simple detector at the receiver \([25]\).

Since Alamouti code has nonzero \(B^{(i)}\) for \(i = 1, 2\), it can be concluded from the discussion of previous section that a widely-linear MMSE estimator is required to estimate \(x_1\) and \(x_2\) at the receiver as follows. The observation vector \(y\) for this code can be written as

\[
y = \begin{bmatrix} r_{11} \\ r_{21} \end{bmatrix} = \begin{bmatrix} h_{11}x_1 + h_{21}x_2 + n_1 \\ -h_{11}x_2^* + h_{21}x_1^* + n_2 \end{bmatrix}.
\] (3.9)
Using (3.4)-(3.7), it can be shown that

\[
C_{yy} = \begin{bmatrix}
|h_{11}|^2 \sigma_{x_1}^2 + |h_{21}|^2 \sigma_{x_2}^2 + \sigma_n^2 & h_{11} h_{21}^* \gamma_{x_1} - h_{11} h_{21} \gamma_{x_2}^*

h_{11} h_{21} (\gamma_{x_2}^*)^* - h_{11} h_{21}^* (\gamma_{x_1})^* & |h_{11}|^2 \sigma_{x_1}^2 + |h_{21}|^2 \sigma_{x_2}^2 + \sigma_n^2
\end{bmatrix},
\]

(3.10)

\[
C_{yy} = \begin{bmatrix}
h_{11} \gamma_{x_1}^2 - h_{21} \gamma_{x_2}^2 & h_{11} h_{21} (\sigma_{x_1}^2 - \sigma_{x_2}^2)

h_{11} h_{21}^* (\sigma_{x_1}^2 - \sigma_{x_2}^2) & h_{11} h_{21}^* (\gamma_{x_1})^* + h_{21} (\gamma_{x_2}^*)^*
\end{bmatrix},
\]

(3.11)

\[
C_{x_1 y} = \begin{bmatrix}
h_{11} \gamma_{x_1} & h_{21} \gamma_{x_1}^2
\end{bmatrix}, \quad C_{x_2 y} = \begin{bmatrix}
h_{11} \gamma_{x_1}^2 & h_{21} \gamma_{x_1}^2
\end{bmatrix},
\]

(3.12)

\[
C_{x_1 y} = \begin{bmatrix}
h_{11} \gamma_{x_1}^2 & h_{21} \gamma_{x_1}^2
\end{bmatrix}, \quad C_{x_2 y} = \begin{bmatrix}
h_{21} \gamma_{x_2}^2 - h_{11} \gamma_{x_2}^2
\end{bmatrix}.
\]

(3.13)

By substituting these values in (3.2)-(3.3), the WL estimators for \(x_1\) and \(x_2\) can be easily determined.

Note that the above expressions place no constraints on the constellations used for \(x_1\) and \(x_2\). As a special case, if \(x_1\) and \(x_2\) are selected from proper constellations, such as QAM or PSK constellations for which \(\gamma_{x_1}^2 = \gamma_{x_2}^2 = 0\), and if \(x_1\) and \(x_2\) have the same power (i.e., \(\sigma_{x_1}^2 = \sigma_{x_2}^2 = \sigma_x^2\)), then the received signal will be proper (\(C_{yy} = 0\)). Nevertheless, the special structure of the orthogonal code (\(B^{(i)} \neq 0\)) causes non-zero pseudo-covariances between \(x_i\) and \(y\), i.e., \(C_{x_i y} = [0 \ h_{21} \sigma_x^2\] and \(C_{x_2 y} = [0 \ -h_{21} \sigma_x^2]\). This, in turn, results in non-zero vectors \(g_i\) in (3.1), which necessitate using of both the observation vector \(y\) and its complex conjugate \(y^*\) in the estimation process, as follows:

\[
\tilde{x}_1 = c (h_{11}^* r_{11} + h_{21} r_{21}^*),
\]

(3.14)

\[
\tilde{x}_2 = c (h_{21}^* r_{11} - h_{11} r_{21}^*).
\]

(3.15)

By substituting \(r_{11}\) and \(r_{21}\) from (3.9) in (3.14) and (3.15), the following expressions are obtained

\[
\tilde{x}_1 = \frac{\sigma_x^2 \left(|h_{11}|^2 + |h_{21}|^2\right)}{\sigma_n^2 \left(|h_{11}|^2 + |h_{21}|^2\right) + \sigma_n^2} x_1 + \frac{\sigma_x^2 \left(h_{21}^* n_1^* + h_{11} n_1\right)}{\sigma_n^2 \left(|h_{11}|^2 + |h_{21}|^2\right) + \sigma_n^2} n_1 = c' x_1 + n_1'(3.16)
\]

\[
\tilde{x}_2 = \frac{\sigma_x^2 \left(|h_{11}|^2 + |h_{21}|^2\right)}{\sigma_n^2 \left(|h_{11}|^2 + |h_{21}|^2\right) + \sigma_n^2} x_2 + \frac{\sigma_x^2 \left(h_{21}^* n_1 - h_{11} n_2\right)}{\sigma_n^2 \left(|h_{11}|^2 + |h_{21}|^2\right) + \sigma_n^2} n_2 = c' x_2 + n_2'(3.17)
\]

where \(c' = \frac{\sigma_x^2 \left(|h_{11}|^2 + |h_{21}|^2\right)}{\sigma_n^2 \left(|h_{11}|^2 + |h_{21}|^2\right) + \sigma_n^2}\). Since the transmitted symbols (\(x_1, x_2\)) and additive noises (\(n_1, n_2\)) are uncorrelated to each other, there exists no correlation between \(\tilde{x}_1\) and \(\tilde{x}_2\) in...
(3.16) and (3.17). As a result, two separate ML detectors can be used for detection of each \(x_i\) from \(\tilde{x}_i\), using the following set of decision rules:

\[
\hat{x}_i = \arg \max_{x_i} \Re \{ \tilde{x}_i (c'x_i)^* \} - \frac{1}{2} |c'x_i|^2 \quad \text{for} \quad i = 1, 2. \tag{3.18}
\]

The resulting receiver is illustrated in Fig. 3.1. At the first stage, the WL-MMSE estimator calculates \(\tilde{x}_1\) and \(\tilde{x}_2\). At the next stage, ML detection of \(x_1\) and \(x_2\) is performed using (3.18).

\[
\begin{align*}
\tilde{x}_1 &= c(h_{11}^*r_{i1} + h_{21}^*r_{2i}^*) & \text{ML Detector} & \hat{x}_1 \\
\tilde{x}_2 &= c(h_{21}^*r_{i1} - h_{11}^*r_{2i}^*) & \text{ML Detector} & \hat{x}_2
\end{align*}
\]

Figure 3.1: The WL-MMSE receiver structure for Alamouti orthogonal code

### 3.2.2 Equivalency between WL-MMSE Estimation and Alamouti’s Combining Scheme for Orthogonal Codes

A commonly used technique for detecting \(x_i\) for orthogonal codes calls for applying the ML criterion on the observation vector \(y\) [25], which results in the following decision rule:

\[
\hat{x} = \arg \max_{x} \Re \{ (SH)^H y - \frac{1}{2} |SH|^2 \}. \tag{3.21}
\]

By applying this decision rule to Alamouti code and after simple manipulations, the argument of this decision rule reduces to the summation of two separate terms each of which depends on only one of the unknown symbols \(x_i\); hence, we get

\[
\begin{align*}
\hat{x}_1 &= \arg \max_{x_1} \Re \{ (h_{11}^*r_{11} + h_{21}^*r_{21}^*)x_1^* \} - \frac{1}{2} (|h_{11}|^2 + |h_{21}|^2)|x_1|^2, \tag{3.19} \\
\hat{x}_2 &= \arg \max_{x_2} \Re \{ (h_{21}^*r_{11} - h_{11}^*r_{21}^*)x_2^* \} - \frac{1}{2} (|h_{11}|^2 + |h_{21}|^2)|x_2|^2. \tag{3.20}
\end{align*}
\]

\(^3\)For \(M_r > 1\), the term \(SH\) in this decision rule will be replaced by the vectorized version of \(SH\); i.e., \([Sh_1^T, \ldots, (Sh_{M_r})^T]^T\), where \(h_j\) denotes the \(j\)th column of matrix \(H\).
In other words, the ML detection of \( x_1 \) and \( x_2 \) can be accomplished separately. As described in [1], this receiver can be implemented in two steps. At the first step, the received signals are passed through a combiner to generate

\[
\hat{x}_1 = h_{11}^* r_{11} + h_{21}^* r_{21}, \tag{3.21}
\]

\[
\hat{x}_2 = h_{21}^* r_{11} - h_{11}^* r_{21}. \tag{3.22}
\]

Then, these combined signals are deployed in separate ML detectors to detect \( \hat{x}_1 \) and \( \hat{x}_2 \).

If we compare the decoding scheme of (3.18) with (3.19) and (3.20), it can be seen that by substituting the value of \( \tilde{x}_i \) from (3.14) and (3.15), the decision rule of (3.18) reduces to (3.19) and (3.20). As a result, these two decoders are equivalent, and the WL-MMSE receiver proposed in this thesis for Alamouti code performs the same as the ML receiver of [1].

Indeed, the outputs of combiners in (3.21) and (3.22) in Alamouti’s decoding method are scaled versions of WL-MMSE estimations of \( x_1 \) and \( x_2 \) given by (3.14) and (3.15), with an scaling factor of \( (|h_{11}|^2 + |h_{21}|^2) + \frac{\sigma_n^2}{\sigma_x^2} \). Similarly, it can be shown that the generalization of combining scheme of [1] for any other orthogonal code is equivalent, up to a scale, to WL-MMSE estimation of transmitted symbols. However, there are two important differences between these methods:

- Alamouti’s combining scheme is a direct result of orthogonality of the code and cannot be generalized to other non-orthogonal LD codes, whereas the WL-MMSE estimator in (3.1) is derived for general LD codes and does not require any further restriction on the structure of LD code. Consequently, WL-MMSE estimation is the only existing approach which provides a general combining scheme for non-orthogonal LD codes.

- The combining schemes of (3.14) and (3.15) are derived for proper \( x_1 \) and \( x_2 \) with equal powers. Consequently, whenever one of these conditions is not satisfied, these
expressions will be changed and Alamouti’s combining scheme no longer yields WL-MMSE estimation of $x_i$.

### 3.2.3 Optimal WL-MMSE Receiver for a Quasi-Orthogonal LD code with $M_t = 4, M_r = 1, N = 4$

Quasi-orthogonal codes are another important subset of LD codes which can provide full coding rate [2], and also can achieve full diversity gain [48–50] by simple rotation of the input symbols. Since these codes are not orthogonal, Alamouti’s combining scheme is not applicable to these codes; however, the general WL-MMSE estimation method of (3.1) can be applied to these codes similar to other LD codes.

Consider the quasi-orthogonal code given in Table 3.2. The observation vector at the receiver can be expressed as follows:

$$
y = \begin{bmatrix}
  r_{11} \\
  r_{21} \\
  r_{31} \\
  r_{41}
\end{bmatrix} = \begin{bmatrix}
  h_{11}x_1 + h_{21}x_2 + h_{31}x_3 + h_{41}x_4 + n_1 \\
  -h_{11}x_2^* + h_{21}x_1^* - h_{31}x_4^* + h_{41}x_3^* + n_2 \\
  -h_{11}x_3^* - h_{21}x_4^* + h_{31}x_1^* + h_{41}x_2^* + n_3 \\
  h_{11}x_4 - h_{21}x_3 - h_{31}x_2 + h_{41}x_1 + n_4
\end{bmatrix}.
$$  
(3.23)

For mathematical tractability, we have restricted our study to the case of $M_r = 1$; however, the results can be easily generalized to $M_r > 1$. Moreover, it will be assumed that all symbols are selected from proper modulation schemes with equal powers, i.e., $\sigma^2_{x_i} = \sigma^2_x$ and $\gamma^2_{x_i} = \gamma^2_x = 0$ for $i = 1, \ldots, 4$. Under these assumptions, by using (3.4)-(3.7) and substituting the corresponding values from Table 3.2, the WL-MMSE estimation of $\mathbf{x}$, can be expressed as

$$
\tilde{\mathbf{x}} = \begin{bmatrix}
  c_1 & 0 & 0 & -c_2 \\
  0 & c_1 & c_2 & 0 \\
  0 & c_2 & c_1 & 0 \\
  -c_2 & 0 & 0 & c_1
\end{bmatrix} \begin{bmatrix}
  h_{11}^* & h_{21} & h_{31} & h_{41}^* \\
  h_{21}^* & -h_{11} & h_{41} & -h_{31}^* \\
  h_{31}^* & h_{41} & -h_{11} & -h_{21}^* \\
  h_{41}^* & -h_{31} & -h_{21} & h_{11}^*
\end{bmatrix} \begin{bmatrix}
  r_{11} \\
  r_{21} \\
  r_{31} \\
  r_{41}
\end{bmatrix}.
$$  
(3.24)
where \( c_1 = \frac{\sigma_x^2}{\sigma_z^2 - \sigma_b^2} \), \( c_2 = \frac{\sigma_x^2 b}{\sigma_z^2 - \sigma_b^2} \), \( a = \sigma_x^2 (\sum_{i=1}^{4} |h_{i1}|^2) + \sigma_n^2 \), and \( b = 2 \sigma_x^2 \Re\{h_{11}^* h_{41} - h_{21} h_{31}^*\} \).

By substituting (3.23) in (3.24), it can be shown that

\[
\begin{bmatrix}
\tilde{x}_1 \\
\tilde{x}_4
\end{bmatrix} =
\begin{bmatrix}
1 - \frac{\sigma_x^2}{\sigma_z^2} c_1 & \frac{\sigma_x^2}{\sigma_z^2} c_2 \\
\frac{\sigma_x^2}{\sigma_z^2} c_2 & 1 - \frac{\sigma_n^2}{\sigma_z^2} c_1
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_4
\end{bmatrix} +
\begin{bmatrix}
n_1' \\
n_4'
\end{bmatrix}
\]  

(3.25)
where \( \sigma_{n_i}^2 = (c_i^2 + c_3^2)(\sum_{i=1}^{4} |h_{i1}|^2) - 4c_1c_2\Re\{h_{i1}^*h_{41} - h_{21}h_{31}^*\}, \) for \( i = 1, \ldots, 4, \) and \( \sigma_{n_i'n_i'} = -\sigma_{n_i'n_i'} = 2(c_1^2 + c_3^2)\Re\{h_{i1}^*h_{41} - h_{21}h_{31}^*\} - 2c_1c_2(\sum_{i=1}^{4} |h_{i1}|^2) \). Thus, the ML detection of \( \hat{x} \) requires joint detection of the pair \( (x_1, x_4) \) from \( (\tilde{x}_1, \tilde{x}_4) \) and joint detection of the pair \( (x_1, x_4) \) from \( (\tilde{x}_2, \tilde{x}_3) \) using the following decision rules, as illustrated in Fig. 3.2.

\[
(\hat{x}_1, \hat{x}_4) = \arg \max_{(x_1,x_4)} \Re \left\{ \left| \tilde{x}_1 - (1 - \frac{\sigma_n^2}{\sigma_x^2} c_1)x_1 - \frac{\sigma_n^2}{\sigma_x^2} c_2 x_4 \right|^2 + \left| \tilde{x}_4 - (1 - \frac{\sigma_n^2}{\sigma_x^2} c_1)x_4 - \frac{\sigma_n^2}{\sigma_x^2} c_2 x_4 \right|^2 \\
- \frac{2\sigma_{n_1'n_1'}}{\sigma_{n_1'}^2} (\tilde{x}_1 - (1 - \frac{\sigma_n^2}{\sigma_x^2} c_1)x_1 - \frac{\sigma_n^2}{\sigma_x^2} c_2 x_4) (\tilde{x}_4 - (1 - \frac{\sigma_n^2}{\sigma_x^2} c_1)x_4 - \frac{\sigma_n^2}{\sigma_x^2} c_2 x_4) \right\} \tag{3.27}
\]

\[
(\hat{x}_2, \hat{x}_3) = \arg \max_{(x_2,x_3)} \Re \left\{ \left| \tilde{x}_2 - (1 - \frac{\sigma_n^2}{\sigma_x^2} c_1)x_2 + \frac{\sigma_n^2}{\sigma_x^2} c_2 x_3 \right|^2 + \left| \tilde{x}_3 - (1 - \frac{\sigma_n^2}{\sigma_x^2} c_1)x_3 + \frac{\sigma_n^2}{\sigma_x^2} c_2 x_3 \right|^2 \\
- \frac{2\sigma_{n_2'n_2'}}{\sigma_{n_2'}^2} (\tilde{x}_2 - (1 - \frac{\sigma_n^2}{\sigma_x^2} c_1)x_2 + \frac{\sigma_n^2}{\sigma_x^2} c_2 x_3) (\tilde{x}_3 - (1 - \frac{\sigma_n^2}{\sigma_x^2} c_1)x_3 + \frac{\sigma_n^2}{\sigma_x^2} c_2 x_3) \right\} \tag{3.28}
\]

Finally, it should be noted that an advantage of quasi-orthogonal code of Table 3.2 over other nonorthogonal LD codes is its special structure which provides group orthogonality between the subspace spanned by first and forth columns of \( \mathbf{S} \) and the subspace spanned by second and third columns of \( \mathbf{S} \) [2]. This group orthogonality simplifies joint ML detection of the whole vector \( \hat{x} = [\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4]^T \) into separate detection of pairs \((\hat{x}_1, \hat{x}_4)\) and \((\hat{x}_2, \hat{x}_3)\). This simplification is regardless of whether the ML criterion is applied to \( \hat{x} \), as shown in (3.27) and (3.28), or the ML criterion is applied to the original observation \( y \), as shown in following detection rules, which are derived in [2],

\[
(\hat{x}_1, \hat{x}_4) = \arg \max_{(x_1,x_4)} \Re \{ \hat{x}_1 x_1^* + \hat{x}_4 x_4^* \} - \Re \{ x_1 x_4^* \} d - \frac{1}{2} (\sum_{i=1}^{4} |h_{i1}|^2)(|x_1|^2 + |x_4|^2) \tag{3.29}
\]

\[
(\hat{x}_2, \hat{x}_3) = \arg \max_{(x_2,x_3)} \Re \{ \hat{x}_2 x_2^* + \hat{x}_3 x_3^* \} + \Re \{ x_2 x_3^* \} d - \frac{1}{2} (\sum_{i=1}^{4} |h_{i1}|^2)(|x_2|^2 + |x_3|^2) \tag{3.30}
\]
where \( d = 2 \Re \{ h_{11}^* h_{41} - h_{21} h_{31}^* \} \), and

\[
\begin{align*}
\tilde{x}_1 &= h_{11}^* r_{11} + h_{21}^* r_{21} + h_{31}^* r_{31} + h_{41}^* r_{41}, \\
\tilde{x}_2 &= h_{21}^* r_{11} - h_{11}^* r_{21} + h_{41}^* r_{31} - h_{31}^* r_{41}, \\
\tilde{x}_3 &= h_{31}^* r_{11} + h_{41}^* r_{21} - h_{11}^* r_{31} - h_{21}^* r_{41}, \\
\tilde{x}_4 &= h_{41}^* r_{11} - h_{31}^* r_{21} - h_{21}^* r_{31} + h_{11}^* r_{41}.
\end{align*}
\] (3.31) (3.32) (3.33) (3.34)

Figure 3.2: The WL-MMSE receiver structure for quasi-orthogonal code of Table 3.2

### 3.2.4 Suboptimal WL-MMSE Receiver for Quasi-Orthogonal Codes

In Section 3.2.3, it was mentioned that due to the special structure of the quasi-orthogonal code in Table 3.2, both the proposed WL-MMSE receiver, given by (3.27) and (3.28), and the conventional ML receiver, given by (3.29) and (3.30), perform joint detection for pairs \((\hat{x}_1, \hat{x}_4)\) and \((\hat{x}_2, \hat{x}_3)\), rather than joint detection of the whole vector \(\hat{x} = [\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4]^T\). As a result, in both methods it is required to compare \(2L^2\), rather than \(L^4\), measurements in order to determine the most likely transmitted symbols, where \(L\) is the size of constellation used for mapping the symbols. However, these codes still suffer from the complexity of joint detection of two symbols. This joint processing becomes a troublesome issue when the constellation size \(L\) increases.

One possible solution to reduce this complexity is to ignore the correlation existing between the elements of each pair in (3.25) and (3.26); hence, ignoring the cross terms
Chapter 3. Widely-Linear MMSE Receivers for LD Codes

$x_1 x_4^*$ and $x_2 x_3^*$ in ML decision rules. This results in a decision rule which is no longer optimal in ML sense. This suboptimal solution is of particular interest in this thesis, in that it is compatible with our proposed two-stage structure of Fig. 1.2.

The resulting suboptimal receiver is illustrated in Fig. 3.3. It has the same WL-MMSE estimator in the first stage, but in the second stage the joint ML detectors of Fig. 3.3 are replaced with separate ML detectors which have the following decision rules:

$$\hat{x}_i = \arg \max_{x_i} \Re \{ \tilde{x}_i x_i^* \} - \frac{1}{2} \left( 1 - c_1 \frac{\sigma_n^2}{\sigma_x^2} \right) |x_i|^2 \quad \text{for } i = 1, \ldots, 4. \quad (3.35)$$

Similarly, we can ignore the the crossterms in conventional ML detectors given by (3.29) and (3.30). The resulting suboptimal detectors will have the following form

$$\hat{x}_i = \arg \max_{x_i} \Re \{ \tilde{x}_i x_i^* \} - \frac{1}{2} \left( \sum_{j=1}^{4} |h_{j1}|^2 \right) |x_i|^2 \quad \text{for } i = 1, \ldots, 4 \quad (3.36)$$

where the expressions for $\tilde{x}_i$ are given in (3.31)-(3.34).

In both suboptimal receivers of (3.35) and (3.36), substitution of joint detectors with separate detectors results in reducing the complexity of detection algorithm from $2L^2$ to $4L$, i.e., the decoding complexity will be reduced by a factor of $2/L$. However, in the next section it will be shown that the suboptimal receiver of (3.36) is unable to detect transmitted symbols reliably, while the suboptimal receiver of (3.35) detects these symbols with only a slight performance loss.

![Figure 3.3: Suboptimal WL-MMSE receiver for quasi-orthogonal code of Table 3.2](image-url)
3.3 Sufficiency of WL-MMSE Estimates for Improper Complex Gaussian Signals

Let $\mathbf{x} = [x_1, \ldots, x_N]$ be a Gaussian random vector and $\mathbf{y} = [y_1, \ldots, y_M]$ be the vector of observations with Gaussian distribution from which $\mathbf{x}$ is to be estimated, where $M$ and $N$ can be any arbitrary values. In [44], it has been shown that for scalar $x$ (i.e., $N = 1$), if $x$ and $y$ are jointly proper complex signals, linear MMSE estimator of $x$ is a minimal sufficient statistic for $x$. This section considers a more general case of $N \geq 1$ and proves that when the constraint on joint properness of $\mathbf{x}$ and $\mathbf{y}$ is relaxed, the widely-linear MMSE estimate of $\mathbf{x}$, denoted by $\tilde{\mathbf{x}}$, provides a minimal sufficient statistic for $\mathbf{x}$. Before giving the proof, we need to define the following Hilbert space which is a simple modification of the one discussed in [40].

Let $\mathbf{w} = [\mathbf{x}^T, \mathbf{y}^T]^T$ and $\mathbf{w}^H = [\mathbf{w}^T, \mathbf{w}^H]^T$. Then, $\mathcal{V}(\mathbf{w})$ denotes the set of all possible linear combinations of $x_i$, $x_i^*$, $y_j$, and $y_j^*$ for $i = 1, \ldots, N$ and $j = 1, \ldots, M$. This set together with the following inner-product operation forms a Hilbert space [40]: $\forall v_1, v_2 \in \mathcal{V}(\mathbf{w}): \langle v_1, v_2 \rangle = E\{v_1 v_2^*\}$. Similarly, we can define $\mathcal{V}(\mathbf{y})$ as a subspace of $\mathcal{V}(\mathbf{w})$.

Using this inner-product definition, two random variables $v_1$ and $v_2$ are called orthogonal to each other if $\langle v_1, v_2 \rangle = 0$. Note that unlike the real-valued Gaussian random variables, orthogonality of two complex Gaussian random variables does not necessitate independence of them. As a case in point, let $v_2 = v_1^*$. Obviously, $v_1$ and $v_2$ are not independent, but $\langle v_1, v_2 \rangle = 0$ when $v_1$ is a proper complex random variable. Up to the best of authors’ knowledge, this point is not clarified in the literature. Thus, we present the following lemma to explain the necessary and sufficient conditions for independence of two complex Gaussian random variables.

**Theorem 3.2** Let $v_1$ and $v_2$ be two jointly Gaussian complex random variables. These variables are independent if and only if

$$\langle v_1, v_2 \rangle = \langle v_1, v_2^* \rangle = 0$$

(3.37)
Proof: $v_1$ and $v_2$ are independent if and only if there exists a group independence between the sets $\{\Re\{v_1\},\Im\{v_1\}\}$ and $\{\Re\{v_2\},\Im\{v_2\}\}$ [51]. Since $v_1, v_2$ are complex Gaussian random variables, their real and imaginary parts are real-valued Gaussian random variables. As a result, the group independence of these two sets is equivalent to the orthogonality of them, i.e.,

$$\langle \Re\{v_1\}, \Re\{v_2\} \rangle = \langle \Re\{v_1\}, \Im\{v_2\} \rangle = \langle \Im\{v_1\}, \Re\{v_2\} \rangle = \langle \Im\{v_1\}, \Im\{v_2\} \rangle = 0.$$ (3.38)

These conditions are, indeed, expanded version of (3.37) into real and imaginary parts. Now, we can prove Theorem 3.1 as follows.

Proof of sufficiency: Let $\tilde{x} = [\tilde{x}_1, \ldots, \tilde{x}_N]$ be the WL-MMSE estimate of vector $x$ from $y$, where $\tilde{x}_i = f_i y + g_i y^*$, and $f_i$ and $g_i$ are derived by substituting $x$ in (3.2)-(3.3) with $x_i$. Since $x$ and $y$ are assumed to be jointly Gaussian, it can be concluded from [41] that $\tilde{x}_i = E\{x_i|y\}$. Indeed, $\tilde{x}_i$ is the projection of $x_i$ on $\mathcal{V}(y)$, and the estimation error $e_i = x_i - \tilde{x}_i$ is orthogonal to $\mathcal{V}(y)$ [40]. This implies that $\langle e_i, y \rangle = \langle e_i, y^* \rangle = \langle e_i, \tilde{x}_i \rangle = \langle e_i, \tilde{x}_i^* \rangle = 0$. Thus, according to Lemma 3.2, $e_i$ is independent from both $y$ and $\tilde{x}_i$. This, in turn, implies that the error vector $e = x - \tilde{x} = [e_1, \ldots, e_N]^T$ is independent of vectors $y$ and $\tilde{x}$. It follows that

$$f_{y|x,\tilde{x}}(y|x,\tilde{x}) = f_{y|e,\tilde{x}}(y|e,\tilde{x})$$
$$= \frac{f_{y,e|\tilde{x}}(y,e|\tilde{x})}{f_{e|\tilde{x}}(e|\tilde{x})}$$
$$= \frac{f_{y|\tilde{x}}(y|\tilde{x})f_e(e)}{f_e(e)}$$
$$= f_{y|\tilde{x}}(y|\tilde{x}).$$ (3.39)

The first equality in (3.39) is a direct result of the fact that $x = \tilde{x} + e$, which implies that $x$ and $\tilde{x}$ provide the same information as $e$ and $\tilde{x}$. The second equality is based on Bayes’ theorem, and the third equality is the consequence of independence of $e$ from $y$ and $\tilde{x}$. Since $x, y, \tilde{x},$ and $e$ are jointly Gaussian vectors, Equation (3.39) leads to $F_{y|x,\tilde{x}}(y|x,\tilde{x}) = F_{y|\tilde{x}}(y|\tilde{x})$, where $F_z(z)$ denotes the cumulative distribution function of $z$. Therefore, $\tilde{x}$ is a sufficient statistic for $x$ [52].
Chapter 3. Widely-Linear MMSE Receivers for LD Codes

Proof of minimality: Now, assume that $\tilde{x}'$ is another sufficient statistic for $x$. Then, $\tilde{x}'$ provides the same information about $x$ as $y$. As a result, $f_{x|\tilde{x}'}(x|y) = f_{x|x'}(x|\tilde{x}')$. Thus, given $\tilde{x}'$, the value of $\tilde{x}$ can be uniquely determined as a function of $\tilde{x}'$ as follows: $\tilde{x} = E\{x|y\} = E\{x|\tilde{x}'\}$. This is the necessary and sufficient condition for $\tilde{x}$ to be a minimal sufficient statistic.

3.4 Simulation Results

This section presents simulation results to compare the performance of the following receivers for orthogonal and quasi-orthogonal LD codes which were studied in Section 3.2:

- Proposed WL-MMSE receiver which applies the ML decision rule to $\tilde{x}$.
- Conventional ML receiver which applies the ML decision rule to $y$.
- Suboptimal WL-MMSE receiver which utilizes separate ML detectors on $\tilde{x}$.
- Suboptimal receiver which utilizes separate ML detectors on $y$.

As explained in Section 3.1, the rationale behind the proposed WL-MMSE receiver structure is the sufficiency of $\tilde{x}$ which was proved in Theorem 3.1. However, this theorem requires the vectors $x$ and $y$ to be jointly Gaussian, which is not the case in practice. Indeed, since the current realization of the channel is assumed to be known at the receiver, it can easily be shown that the assumption of jointly Gaussianity of vectors $x$ and $y$ in the MIMO system of Fig. 1.2 holds true if and only if the transmitted vector $x$ has a
Gaussian distribution. However, when a discrete constellation such as PSK or QAM is used for generating symbols $x_i$, the vector $x$ will have a multidimensional probability mass function, rather than a Gaussian distribution, and the condition of joint Gaussianity of $x$ and $y$ will not be satisfied. In order to study the effect of using realistic constellations for $x$ on the performance of the proposed receiver, this section examines the performance of aforementioned receivers for the following widely-used constellations: 4-PSK, 8-PSK, 16-PSK, 16-QAM, and 64-QAM. In all the simulations, the bit error rate of the decoded output is used for comparing the performance of these receivers.

3.4.1 Simulation Parameters

Table 3.3 summarizes the simulation parameters used for the transceiver of Fig. 1.2. Adopting the EDGE standard, we assume $T_s = 3.67\mu s$ and a carrier frequency of 1.8 GHz. This results in a normalized fading rate of $f_D T_s = 0.0061$ when the vehicular speed is approximately 100 km/h. Thus, the channel can be assumed to be a quasi-static fading channel. The path gains, which are modeled as Gaussian random variables with zero mean and variance of 0.5 on each of the real and imaginary parts, are generated using Jakes model for Rayleigh channel [33, 53]. The circularly symmetric Gaussian additive noise is assumed to be spatially and temporally white. It is assumed that the CSI is perfectly known at the receiver, and the receive filter is matched and synchronized with the transmit filter.

3.4.2 Performance Analysis for Orthogonal Codes

In Section 3.2.2, it was proved that for orthogonal code of Table 3.1, the WL-MMSE receiver of (3.18) is identical with the conventional ML receiver of (3.19)-(3.20), and both receivers have the same decision rules. Therefore, the performance and behavior of WL-MMSE receiver for this code is the same as conventional ML receiver, which is well studied in the literature. Also, it was shown that for orthogonal codes, the optimal
Table 3.3: Summary of Simulation Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>signal constellation</td>
<td>4-PSK, 8-PSK, and 16-QAM</td>
</tr>
<tr>
<td>space-time block-code</td>
<td>$4 \times 4$ quasi-orthogonal of Table 3.2</td>
</tr>
<tr>
<td># of transmit antennas</td>
<td>$M_t = 4$</td>
</tr>
<tr>
<td># of receive antennas</td>
<td>$M_r = 1$</td>
</tr>
<tr>
<td>symbol duration</td>
<td>$T_s = 3.67 \times 10^{-6}$ s</td>
</tr>
<tr>
<td>pulse shaping type</td>
<td>square root raised cosine</td>
</tr>
<tr>
<td>pulse shaping roll-off factor</td>
<td>$\alpha = 0.35$</td>
</tr>
<tr>
<td>pulse shaping filter order</td>
<td>40</td>
</tr>
<tr>
<td>center frequency</td>
<td>$f_c = 1.8 \times 10^9$ Hz</td>
</tr>
<tr>
<td>vehicular speed</td>
<td>$v = 100$ km/h</td>
</tr>
<tr>
<td>doppler frequency</td>
<td>$f_D = 166$ Hz</td>
</tr>
<tr>
<td>normalized fading rate</td>
<td>$f_D T_s = 0.00061$ (quasi-static)</td>
</tr>
</tbody>
</table>

receiver separately detects each $x_i$; hence, there is no need to consider a suboptimal solution for these codes. Consequently, we will not present any simulation for orthogonal codes.

### 3.4.3 Performance Analysis for Quasi-Orthogonal Codes

Consider the quasi-orthogonal code given in Table 3.2 with one receive antenna. Figures 3.5-3.9 examine different PSK and QAM constellations at the transmitter. In each figure, the performance of the following receivers are compared: the optimal WL-MMSE receiver given in (3.27) and (3.28), the conventional ML receiver in (3.29) and (3.30), the suboptimal WL-MMSE receiver in (3.35), and the suboptimal receiver in (3.36).

In these figures, it can be seen that the performance of optimal WL-MMSE receiver is the same as conventional ML receiver in all cases, despite the fact that the condition
Figure 3.5: Bit-error probability versus SNR for quasi-orthogonal code of Table 3.2, using 4-PSK modulation and one receive antenna.

Figure 3.6: Bit-error probability versus SNR for quasi-orthogonal code of Table 3.2, using 8-PSK modulation and one receive antenna.
Figure 3.7: Bit-error probability versus SNR for quasi-orthogonal code of Table 3.2, using 16-PSK modulation and one receive antenna.

of joint Gaussianity of $\mathbf{x}$ and $\mathbf{y}$ in Theorem 3.1 is not satisfied for PSK and QAM constellations. This justifies the utilization of $\tilde{x}$ in lieu of the sufficient statistic for $\mathbf{x}$ when PSK or QAM constellations are used together with LD codes.

It can also be seen that the suboptimal receiver of (3.36), derived directly from ML decision rule, has a poor performance for 4-PSK and has an error floor for all other cases. However, the proposed suboptimal WL-MMSE of (3.35) performs near optimal over a wide range of SNR, i.e., less than 1 dB away from the optimal receiver for 4-PSK, 16-QAM, and 64-QAM and 2 dB away for 8-PSK and 16-PSK. Note that these suboptimal receivers have considerably less complexity in comparison with optimal receivers, e.g., 1/32 of optimal receiver’s complexity for 64-QAM. By comparing decision rules of the suboptimal receiver of (3.36) and suboptimal WL-MMSE receiver of (3.35), this difference in performance can be justified as follows. The former receiver is a matched filter which uses $\hat{x}_i$, whereas the later receiver is a matched filter which uses $\tilde{x}_i$. Obviously, (3.35)
Figure 3.8: Bit-error probability versus SNR for quasi-orthogonal code of Table 3.2, using 16-QAM modulation and one receive antenna.

outperforms (3.36), owing to the fact that (3.35) deploys $\tilde{x}_i$ which has the minimum estimation error over all possible WL estimations of $x_i$, including $\hat{x}_i$ deployed by (3.36).

The performance loss resulted from using separate detectors instead of joint detector in proposed WL-MMSE receiver can be justified as follows. Consider the detection of $\hat{x}_1$ and $\hat{x}_4$. In joint detection of the pair $(\hat{x}_1, \hat{x}_4)$, the receiver takes advantage of the correlation existing between $\tilde{x}_1$ and $\tilde{x}_4$ in (3.25) as a side information for detecting both symbols; however, when $\hat{x}_1$ and $\hat{x}_4$ are detected separately, the receiver ignores this side information, which results in a loss in the performance of the receiver. It should be noted that the sensitivity of the detector’s performance to this side information varies for different constellation types. Figures 3.5-3.9 reveal that PSK detectors are much more sensitive in comparison with QAM detectors. Thus, the suboptimal WL-MMSE receiver of (3.35) is of particular interest for decoding the LD codes that are using QAM constellations.
Figure 3.9: Bit-error probability versus SNR for quasi-orthogonal code of Table 3.2, using 64-QAM modulation and one receive antenna.

3.5 Chapter Summary

In this chapter, we proposed a new receiver structure for LD codes, which applies ML detection criterion to the WL-MMSE estimates of the transmitted symbols. It was proved that for Gaussian signals, these estimates are minimal sufficient statistics for detection of transmitted symbols and the proposed receiver is an optimal (ML) receiver. Also, simulation results showed that the WL-MMSE estimates can be used in lieu of the sufficient statistics when the symbols are chosen from PSK or QAM constellations. Based on this new receiver structure, a suboptimal receiver was proposed which performs close to the optimal receiver while reducing the complexity of detection.

Among numerous existing LD codes, we studied two important examples of LD codes as representatives of orthogonal and quasi-orthogonal codes. The structures of the proposed optimal/suboptimal receivers were studied in detail for these two cases. For orthogonal codes, it was shown that the combining scheme proposed in [1] is a scaled
version of WL-MMSE estimation if certain conditions are satisfied. For quasi-orthogonal codes with QAM constellations, simulation results showed that the proposed suboptimal receiver performs less than 1 dB away from ML receiver while reducing the complexity of the detection.
Chapter 4

ML Detection in The Presence of Improper Noise

In the previous chapter, it was mentioned that after estimating the transmitted symbols, the resulting estimates should be passed through a maximum likelihood detector of the form \( \hat{x} = \arg \max_{x} f_{\tilde{x}|x}(\tilde{x}|x) \). In this chapter, it will be shown that \( f_{\tilde{x}|x}(\tilde{x}|x) \) is a complex Gaussian distribution but not necessarily a circularly-symmetric one. For non circularly-symmetric \( f_{\tilde{x}|x}(\tilde{x}|x) \), the conventional detectors are not optimal and do not yield ML performance. In order to take into account this case in the design of WL-MMSE receivers proposed in this thesis, we need to study the structure of optimal ML detector for this case.

This chapter proposes two alternative approaches for ML detection of \( x \) from improper observation \( \tilde{x} \). The first approach directly applies the ML criterion to the improper observation in order to find the decision rule for detection of \( x \). In the second approach, we propose a filter for converting improper signals into proper ones, called circularizing filter. This filter is then utilized together with conventional matched filter to construct the ML detector.
4.1 Need for ML Detection in The Presence of Improper Gaussian Noise

In Chapter 3, the following two detection strategies were studied:

- Conventional ML detection: This method applies the ML criterion directly to the vector $y$ according to the following decision rule: $\hat{x} = \arg \max_x f_{y|x}(y|x)$. Thus, the structure of a conventional detector depends on the conditional density $f_{y|x}(y|x)$. According to equation (1.3) the observation vector $y$ is composed of two components: a signal-of-interest, which is given by $y_s = [s_1 H, \ldots, s_N H]^T$, and an additive proper complex Gaussian noise, which is given by $y_n = [n_1, \ldots, n_N]^T$. Hence, the received signal can be modeled as $y = y_s + y_n$.

Since we are assuming that the channel coefficients are perfectly known at the receiver side, the value of $y_s$ can be uniquely determined if the transmitted vector is known. As a result, there are $M = \prod_{i=1}^{N} L_i$ possible values for the signal-of-interest $y_s$, where $L_i$ is the constellation size for $x_i$. In other words, $y_s \in \{y_s^0, \ldots, y_s^{M-1}\}$, where each $y_s^i$ represents a possible value for $y_s$. Therefore, $f_{y|x}(y|x) = f_{y|y_s}(y|y_s) = f_{y_n}(y - y_s)$, where $f_{y_n}(y_n)$ represents the pdf of the noise and is assumed to be circularly-symmetric Gaussian.

- WL-MMSE estimation followed by ML detection (Fig. 1.2): This detection method applies ML criterion to $\tilde{x}$ in order to detect the transmitted vector according to the following decision rule: $\hat{x} = \arg \max_x f_{\tilde{x}|x}(\tilde{x}|x)$. Therefore, the structure of the detector in this case depends on the conditional density $f_{\tilde{x}|x}(\tilde{x}|x)$. Similar to the observation $y$, the vector $\tilde{x}$ can be decomposed into two components as follows:

$$\tilde{x} = Fy + Gy^* = (Fy_s + Gy_s^*) + (Fy_n + Gy_n^*) = \tilde{x}_s + \tilde{x}_n,$$

where $F = [f_1; \ldots; f_N]$ and $G = [g_1; \ldots; g_N]$ are derived from equations (3.2) and (3.3). In Equation (4.1), the signal-of-interest is $\tilde{x}_s = (Fy_s + Gy_s^*) \in \{\tilde{x}_s^0, \ldots, \tilde{x}_s^{M-1}\}$,
and the additive noise component is \( \tilde{x}_n = (Fy_n + Gy^*_n) \).

Similar to the previous case, it can be easily shown that \( f_{\tilde{x}|x}(\tilde{x}|x) = f_{\tilde{x}_n}(\tilde{x} - \tilde{x}_s) \), where \( f_{\tilde{x}_n}(\tilde{x}_n) \) denotes the pdf of the noise \( \tilde{x}_n \). The vector \( \tilde{x}_n \) is a linear combination of two complex Gaussian vectors (i.e., \( y_n \) and \( y_n^* \)); thus, it is a complex Gaussian noise. However, it not necessarily a proper noise. In fact, the pseudo-covariance matrix of this noise is given by \( C_{\tilde{x}n\tilde{x}n} = \sigma_n^2(FG^T + G^TF) \), which is not necessarily zero. Consequently, depending on the values of \( F \) and \( G \), the estimated vector might include an improper additive noise component which results in a non circularly-symmetric distribution for \( f_{\tilde{x}|x}(\tilde{x}|x) \).

Since the additive noise component in the first strategy is a proper Gaussian noise, the corresponding ML detector in this case is a simple minimum distance detector. In the second strategy, however, the additive noise might be an improper Gaussian noise. This necessitates us to study the structure of maximum likelihood detectors in the presence of improper noise.

According to Equation (4.1), the vector \( \tilde{x} \) can be modeled using the following \( M \)-ary hypothesis:

\[
\begin{align*}
H_0 & : \quad \tilde{x} = \tilde{x}_s^0 + \tilde{x}_n \\
& \vdots \\
H_{M-1} & : \quad \tilde{x} = \tilde{x}_s^{M-1} + \tilde{x}_n
\end{align*}
\]

The ML detector of Fig. 1.2 tries to find \( \hat{x} \) based on the following ML criterion:

\[
\hat{x} = \arg \max_x f_{\tilde{x}|x}(\tilde{x}|x). \tag{4.3}
\]

Owing to the fact that the channel coefficients are known at the receiver, there exists a one-to-one relationship between the transmitted vector \( x \) and the vector \( \tilde{x}_s \); i.e., \( \tilde{x}_s = u(x) \) where \( u(\cdot) \) is an invertible function. Thus, the decision rule in (4.3) can be simplified to

\[
\hat{x} = u^{-1}(\hat{x}_s), \quad \text{where} \quad \hat{x}_s = \arg \max_{\tilde{x}_s} f_{\tilde{x}|\tilde{x}_s}(\tilde{x}|\tilde{x}_s). \tag{4.4}
\]
As a result, in this chapter, we study the problem of ML detection of $\hat{x}_s$ from $\tilde{x}$.

In order to provide simple formulations in the rest of this chapter, we will change the notation as follows. The signal-of-interest (formerly denoted by $\tilde{x}_s$) and the additive noise (formerly denoted by $\tilde{x}_n$) will be denoted by $s$ and $n$, respectively. Also, the observation vector, which was formerly denoted by $\tilde{x}$, will be denoted by $r$. As a result, the $M$-ary hypothesis of (4.2) can be expressed as follows:

$$H_0 : r = s^0 + n$$
$$\vdots$$
$$H_{M-1} : r = s^{M-1} + n$$

(4.5)

where $n$ is a zero-mean complex-valued Gaussian vector, which might be improper, and $s^i$ are fixed known vectors. Accordingly, the ML decision rule in (4.4) will be expressed as

$$\hat{s} = \arg \max_{s \in \{s^0, \ldots, s^{M-1}\}} f_{r|s}(r|s).$$

(4.6)

### 4.2 ML Decision Rule in The Presence of Improper Gaussian Noise

This section considers the $M$-ary hypothesis model given in (4.5), and studies the ML decision rule of Equation (4.6) in this case. Due to the simple additive model of noise in (4.5), we have $f_{r|s}(r|s) = f_n(r - s)$, where $f_n(n)$ is the pdf of the noise vector. Therefore, the ML decision rule in (4.6) reduces to

$$\hat{s} = \arg \max_{s \in \{s^0, \ldots, s^{M-1}\}} f_n(r - s)$$

(4.7)

Using the results of Section 2.2.2, we have

$$f_n(r - s) = \frac{1}{\sqrt{2\pi |C_{vv_nH}|}} \exp \left\{ -\frac{1}{2} (v_r - v_s)^H C_{vv_nH}^{-1} (v_r - v_s) \right\},$$

(4.8)
where \( \mathbf{v}_n = \frac{1}{\sqrt{2}} [\mathbf{s}^T, \mathbf{s}^H]^T \), \( \mathbf{v}_n = \frac{1}{\sqrt{2}} [\mathbf{n}^T, \mathbf{n}^H]^T \), \( \mathbf{v}_r = \frac{1}{\sqrt{2}} [\mathbf{r}^T, \mathbf{r}^H]^T \), and \( \mathbf{C}_{\mathbf{v}_n \mathbf{v}_n}^H \) is the covariance matrix of \( \mathbf{v}_n \) given by

\[
\mathbf{C}_{\mathbf{v}_n \mathbf{v}_n}^H = \begin{bmatrix} \mathbf{C}_{\mathbf{nn}}^H & \mathbf{C}_{\mathbf{nn}} \\ \mathbf{C}_{\mathbf{nn}}^* & \mathbf{C}_{\mathbf{nn}}^H \end{bmatrix} \Rightarrow \mathbf{C}_{\mathbf{v}_n \mathbf{v}_n}^{-1} = \begin{bmatrix} \mathbf{Q}_n^{-1} & -\mathbf{Q}_n^{-1} \mathbf{P}_n^* \\ -\mathbf{Q}_n^{-1} \mathbf{P}_n & \mathbf{Q}_n^{-1} \end{bmatrix}, \tag{4.9}
\]

where \( \mathbf{Q}_n = \mathbf{C}_{\mathbf{nn}}^* - \mathbf{C}_{\mathbf{nn}}^T \mathbf{C}_{\mathbf{nn}}^H \mathbf{C}_{\mathbf{nn}}^T \) and \( \mathbf{P}_n = \mathbf{C}_{\mathbf{nn}}^T \mathbf{C}_{\mathbf{nn}}^{-1} \). By substituting \( f_n(\mathbf{r} - \mathbf{s}) \) from (4.8) into (4.7) and after some manipulations, this decision rule will be reduced to

\[
\hat{s} = \arg \max_{\mathbf{s} \in \{\mathbf{s}^0, \ldots, \mathbf{s}^{M-1}\}} \mathbf{v}_s^H \mathbf{C}_{\mathbf{v}_n \mathbf{v}_n}^{-1} \mathbf{v}_r - \frac{1}{2} \mathbf{v}_s^H \mathbf{C}_{\mathbf{v}_n \mathbf{v}_n}^{-1} \mathbf{v}_s \tag{4.10}
\]

\[
= \arg \max_{\mathbf{s} \in \{\mathbf{s}^0, \ldots, \mathbf{s}^{M-1}\}} \Re \left\{ (\mathbf{s}^H \mathbf{Q}_n^{-*} \mathbf{z} - \frac{1}{2} \mathbf{s}^H \mathbf{Q}_n^{-*} \mathbf{s}) + (\mathbf{s}^T \mathbf{Q}_n^{-1} \mathbf{P}_n \mathbf{z} - \frac{1}{2} \mathbf{s}^T \mathbf{Q}_n^{-1} \mathbf{P}_n \mathbf{s}) \right\} \tag{4.11}
\]

The term \( \mathbf{s}^H \mathbf{Q}_n^{-*} \mathbf{z} \) in (4.11) finds the weighted correlation between the observation \( \mathbf{z} \) and each of the possible signals-of-interest \( \mathbf{s} \in \{\mathbf{s}^0, \ldots, \mathbf{s}^{M-1}\} \), while the term \( \mathbf{s}^T \mathbf{Q}_n^{-1} \mathbf{P}_n \mathbf{z} \) finds the weighted pseudo-correlation \([4]\) between \( \mathbf{z} \) and \( \mathbf{s} \). Then, the outputs of the correlator and pseudo-correlator are adjusted by \( \frac{1}{2} \mathbf{s}^H \mathbf{Q}_n^{-*} \mathbf{s} \) and \( \frac{1}{2} \mathbf{s}^T \mathbf{Q}_n^{-1} \mathbf{P}_n \mathbf{s} \), respectively, to mitigate the effect of the power of \( \mathbf{s} \) on the output of the correlator and pseudo-correlator. Finally, the ML decision will be based on the real part of this result.

Note that when the additive noise is proper, we have \( \mathbf{C}_{\mathbf{nn}} = \mathbf{0} \); hence, \( \mathbf{P}_n = \mathbf{0} \) and \( \mathbf{Q}_n = \mathbf{C}_{\mathbf{nn}}^* \). In this case, the output of pseudo-correlator is zero for all values of \( \mathbf{s} \); hence, the pseudo-correlator is ignored in the ML decision rule. Consequently, the ML decision rule in the presence of proper noise can be written as

\[
\hat{s} = \arg \max_{\mathbf{s} \in \{\mathbf{s}^0, \ldots, \mathbf{s}^{M-1}\}} \Re \left\{ \mathbf{s}^H \mathbf{C}_{\mathbf{nn}}^{-1} \mathbf{z} - \frac{1}{2} \mathbf{s}^H \mathbf{C}_{\mathbf{nn}}^{-1} \mathbf{s} \right\} \tag{4.12}
\]

which is the well-know ML detector for proper noise. This detector first decorrelates the noise component in the observation vector \( \mathbf{z} = \mathbf{r} + \mathbf{n} \) by multiplying it with \( \mathbf{C}_{\mathbf{nn}}^{-\frac{1}{2}} \). Then, it finds the correlation between this transformed observation and transformed version of signals-of-interest \( \mathbf{C}_{\mathbf{nn}}^{-\frac{1}{2}} \mathbf{s} \); hence, it calculates \( \mathbf{C}_{\mathbf{nn}}^{-\frac{1}{2}} \mathbf{s} \mathbf{C}_{\mathbf{nn}}^{-\frac{1}{2}} \mathbf{z} = \mathbf{s}^H \mathbf{C}_{\mathbf{nn}}^{-1} \mathbf{z} \). After
adjusting the output of this whitened correlator by \( \frac{1}{2}s^H C_{nn}^{-1}n \), the real part of the result is used for ML decision.

### 4.3 Circularizing Filter Followed by Conventional Detector

In Section 4.2, we studied the ML decision rule for the \( M \)-ary hypothesis model of Equation (4.5). It was shown that when the observation vector \( r = s + n \) has an improper noise component, conventional detectors are no longer optimal and should be substituted by the detector of (4.11). However, this replacement of the detectors is not desired in many practical cases. In this section, we propose to use a preprocessing step to convert \( r = s + n \) into a modified observation vector \( \tilde{r} = \tilde{s} + \tilde{n} \), where \( \tilde{n} \) is a proper Gaussian noise. Since \( C_{\tilde{n}\tilde{n}} = 0 \), the vector \( \tilde{r} \) can then be passed through a conventional detector of the following form

\[
\hat{\tilde{s}} = \arg \max_{\tilde{s} \in \{\tilde{s}_0, ..., \tilde{s}_{M-1}\}} \Re \left\{ \tilde{s}^H C_{\tilde{n}\tilde{n}}^{-1}\tilde{r} - \frac{1}{2}\tilde{s}^H C_{\tilde{n}\tilde{n}}^{-1} \tilde{n} \tilde{s} \right\}.
\] (4.13)

Provided that the preprocessing step is reversible, there will be a one-to-one relationship between \( s^m \) and \( \tilde{s}^m \) for \( m = 0, ..., M - 1 \), and the detector will be able to uniquely determine \( \hat{\tilde{s}} \) based on the value of \( \hat{\tilde{s}} \).

The following definition and theorem determine the structure and constraints of such a preprocessing filter.

**Definition 4.1** Given an improper complex-valued random vector \( z \in \mathbb{C}^K \), a widely-linear circularizing (or properizing) filter is defined as a filter which converts improper \( z \) into a proper random vector \( \tilde{z} = D_1 z + D_2 z^* \), where \( D_1, D_2 \in \mathbb{C}^{K \times K} \) satisfy the following condition:

\[
C_{\tilde{z}\tilde{z}}^T = D_1 C_{zz} D_1^T + D_1 C_{zz}^H D_2^T + D_2 C_{zz}^* D_1^T + D_2 C_{zz}^* D_2^T = 0
\] (4.14)
Theorem 4.1 A circularizing filter of the form $\mathbf{z} = D_1 \mathbf{z} + D_2 \mathbf{z}^*$ is reversible if and only if the following conditions are satisfied:

$$D_1 \neq 0, \quad D_2 \neq 0, \quad \det(D_1) \neq 0, \quad \text{and} \quad \det(D_1 - D_2 D_1^{-*} D_2^*) \neq 0. \quad (4.15)$$

Proof: see Appendix A.

In general, finding a set of matrices which satisfy the conditions of (4.14) and (4.15) is not easy. However, the following Lemma and Proposition, provide us with one solution for the circularizing filter.

Lemma 4.1 Let $\mathbf{z} \in \mathbb{C}^K$ be a complex-valued random vector. The eigen-decomposition of $\mathbf{C}_{\mathbf{z}\mathbf{z}^H}$ can be written as $\mathbf{C}_{\mathbf{z}\mathbf{z}^H} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H$, where

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \\ \mathbf{Q}_1^* & \mathbf{Q}_2^* \end{bmatrix} \quad \text{and} \quad \mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_{1,K} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_{K+1,2K} \end{bmatrix} \quad (4.16)$$

In this equation, $\mathbf{\Lambda}_{1,K} = \text{diag}(\lambda_1, \ldots, \lambda_K)$ and $\mathbf{\Lambda}_{K+1,2K} = \text{diag}(\lambda_{K+1}, \ldots, \lambda_{2K})$ are diagonal matrices including the eigen values of $\mathbf{C}_{\mathbf{z}\mathbf{z}^H}$. Also, $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are matrices of size $K \times K$.

Proof: see Appendix A.

Proposition 4.1 Consider $\mathbf{z} \in \mathbb{C}^K$ with $\mathbf{C}_{\mathbf{z}\mathbf{z}^H} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H$ as explained in Lemma 4.1. One solution for matrices $D_1$ and $D_2$ in Definition 4.1 is the following set of matrices:

$$D_1 = (\mathbf{\Lambda}_{1,K}^{-\frac{1}{2}} \mathbf{Q}_1^H + j\mathbf{\Lambda}_{K+1,2K}^{-\frac{1}{2}} \mathbf{Q}_2^H) \quad \text{and} \quad D_2 = (\mathbf{\Lambda}_{1,K}^{-\frac{1}{2}} \mathbf{Q}_1^T + j\mathbf{\Lambda}_{K+1,2K}^{-\frac{1}{2}} \mathbf{Q}_2^T) \quad (4.17)$$

Proof: see Appendix A.

Note that the solution for matrices $D_1$ and $D_2$ are not unique. Assume that $D_1$ and $D_2$ satisfy the conditions of (4.14) and (4.15). Then, for any arbitrary invertible matrix $\mathbf{E} \in \mathbb{C}^{K \times K}$, the following sets of matrices also satisfy the conditions of (4.14) and (4.15):

$$D'_1 = \mathbf{E}D_1 \quad \text{and} \quad D'_2 = \mathbf{E}D_2 \quad (4.18)$$
or

\[ D'_1 = ED_2^* \quad \text{and} \quad D'_2 = ED_1^* \] (4.19)

Moreover, it should be noted that the filter proposed above generates a whitened output (see the proof in Appendix A.5). However, if we need the filter to preserve the covariance matrix of the improper signal while properizing it, we can easily use the following modified matrices in the filter:

\[ D'_1 = C_{zz}^H (A_{1,K}^{-1} Q_1^H + jA_{K+1,2K}^{-1} Q_2^H) \quad \text{and} \quad D'_2 = C_{zz}^H (A_{1,K}^{-1} Q_1^T + jA_{K+1,2K}^{-1} Q_2^T) \] (4.20)

Now, let us make use of the filter proposed in Proposition 4.1 to convert the improper noise component of the observation vector \( r = s + n \) into a proper noise. Since this filter is used to circularize the vector \( n \), the matrices \( Q \) and \( \Lambda \) in Proposition (4.1) should be derived from \( C_{vnv_n^H} \). After determining the values of \( D_1 \) and \( D_2 \) from (4.17), we can pass the observation vector \( r \) through the resulting filter to get

\[ \tilde{r} = D_1 r + D_2 r^* = (D_1 s + D_2 s^*) + (D_1 n + D_2 n^*) = \tilde{s} + \tilde{n}, \] (4.21)

Since \( C_{nn^H} = 0 \), the output of the circularizing filter can be passed to the conventional detector given in (4.13). By substituting \( C_{nn^H} = 4I \) into (4.13), we get

\[ \hat{\tilde{s}} = \arg \max_{\tilde{s} \in \{\tilde{s}_0, \ldots, \tilde{s}_{M-1}\}} \Re \left\{ \tilde{s}^H \tilde{r} - \frac{1}{2} \tilde{s}^H \tilde{s} \right\} \] (4.22)

Using a similar procedure to Appendix A.5, it can be shown that

\[ v_{\tilde{s}} = 2T_K^H Q C_{vnv_n^H}^{-\frac{1}{2}} v_s \quad \text{and} \quad v_{\tilde{r}} = 2T_K^H Q C_{vnv_n^H}^{-\frac{1}{2}} v_r \] (4.23)

Therefore, the value of \( \tilde{s} \) can be easily determined from \( \hat{\tilde{s}} \), using the following equation

\[ v_{\tilde{s}} = \frac{1}{2} C_{vnv_n^H}^{\frac{1}{2}} Q^H T_K v_{\tilde{s}}. \] (4.24)

Finally, it should be noted that the decision rule given in (4.22) is equivalent to the decision rule given in (4.10). This can be simply proved by rewriting (4.22) as follows

\[ \hat{\tilde{s}} = \arg \max_{\tilde{s} \in \{\tilde{s}_0, \ldots, \tilde{s}_{M-1}\}} v_{\tilde{s}}^H v_{\tilde{r}} - \frac{1}{2} v_{\tilde{s}}^H v_{\tilde{s}} \] (4.25)
and replacing $v_\hat{s}$ and $v_r$ in (4.25) with their corresponding values in (4.23). In fact, circularization is a reversible preprocessing step and preserves all the data included in the observation vector $r$. Thus, applying the ML criterion to the preprocessed observation will yield the same result as applying the ML criterion to the original observation (see [54, Page 289]).

### 4.4 Binary ML Detection in the Presence of Scalar Improper Noise

In sections 4.2 and 4.3, two equivalent implementations of the ML detector in the presence of improper noise were studied. This section presents a more detailed analysis of the structure and performance of these proposed receivers for the following binary hypothesis:

$$
\begin{align*}
H_0 & : r = s^0 + n \\
H_1 & : r = s^1 + n
\end{align*}
$$

(4.26)

where $r, s^0, s^1, n \in \mathbb{C}$. In this model, $r$ is a scalar noisy observation of the unknown scalar signal-of-interest $s$ which might take either of the values $s^0$ or $s^1$ with equal probabilities.

#### 4.4.1 Detection Using Pseudo-Correlators and Correlators

In the scalar case, the second order characterization of the noise requires the knowledge of the following statistics

$$
\sigma_n^2 \triangleq \mathbb{E}\{|n - \bar{n}|^2\} \quad \text{and} \quad \gamma_n^2 \triangleq \mathbb{E}\{(n - \bar{n})^2\}.
$$

(4.27)

Substituting these values into (4.11), we get

$$
\Re \left\{ (r(s^1)^* - \frac{1}{2}|s^1|^2) - (rs^1 - \frac{1}{2}(s^1)^2) \alpha^*\right\}
\begin{cases}
\hat{s} = s^1 \\
\hat{s} = s^0
\end{cases}
\Re \left\{ (r(s^0)^* - \frac{1}{2}|s^0|^2) - (rs^0 - \frac{1}{2}(s^0)^2) \alpha^*\right\}.
$$

(4.28)
Figure 4.1: Binary ML detectors implementation as a combination of correlators and pseudo-correlators

The block diagram of this detector is shown in Fig. 4.1. In this detector, the term $rs$ correlates the observation $r$ with $s^j$, while the term $rs^j$ pseudo correlates $r$ with $s^j$.

After adjusting the outputs of the correlator and pseudo correlator by $\frac{1}{2}|s_j|^2$ and $\frac{1}{2}(s^j)^2$, respectively, they will be superposed according to the noncircularity coefficient $\alpha_n$, and the ML decision will be based on the real part of this result.

For proper noises, the outputs of pseudo correlators for both $s^1$ and $s^0$ are ignored (multiplied by $\alpha_n = 0$), and the decision only depends on the outputs of correlators. This leads to the following well-known decision rule:

$$\Re\left\{ (r(s^1)^* - \frac{1}{2}|s^1|^2) \right\} \begin{cases} \hat{s} = s^1 & M \\ \hat{s} = s^0 & \end{cases} \Re\left\{ (r(s^0)^* - \frac{1}{2}|s^0|^2) \right\}. \quad (4.29)$$

### 4.4.2 Detection Using a Circularizing Filter Followed by Conventional Detector

In this section, we apply the filter given in Proposition 4.1 to circularize the additive noise component of the observation. This modified observation will then be passed to a conventional detector which utilizes correlators.

Using Equation 4.17 together with the results of Lemma 2.1, we get the following
Figure 4.2: Binary ML detectors implementation as a combination of circularizing filter followed by conventional detector.

coefficients for the circularizing filter

\[
d_1 = \frac{e^{-j\varphi}}{\sqrt{1 + |\alpha_n|}} + j \frac{e^{-j\varphi} \alpha}{\sqrt{1 - |\alpha_n|}} = e^{-j\varphi} \left( \frac{1}{\sqrt{1 + |\alpha_n|}} + \frac{1}{\sqrt{1 - |\alpha_n|}} \right) \quad (4.30)
\]

\[
d_2 = \frac{e^{j\varphi}}{\sqrt{1 + |\alpha_n|}} + j \frac{e^{j\varphi} \alpha}{\sqrt{1 - |\alpha_n|}} = e^{j\varphi} \left( \frac{1}{\sqrt{1 + |\alpha_n|}} - \frac{1}{\sqrt{1 - |\alpha_n|}} \right) \quad (4.31)
\]

where \(\varphi\) is the phase of the noncircularity coefficient of the noise (\(\alpha_n\)). After passing the observation \(r\) through this filter, we get \(\tilde{r} = d_1 r + d_2 r^* = \tilde{s} + \tilde{n}\), where \(\tilde{n}\) is a proper noise. Thus, ML detection of \(\tilde{s}\) can be accomplished using a conventional detector with the following decision rule:

\[
\Re \left\{ \tilde{r} (\tilde{s}^\dagger)^* - \frac{1}{2} |\tilde{s}^\dagger|^2 \right\} \overset{\hat{s}=\tilde{s}^\dagger}{\Rightarrow} \Re \left\{ \tilde{r} (\tilde{s}^0)^* - \frac{1}{2} |\tilde{s}^0|^2 \right\} \quad (4.32)
\]

The structure of this detector is shown in Fig. 4.2.

In order to get a better insight into the circularization preprocessing step, let us rewrite Equation (4.23) in terms of \(w_{\tilde{r}}\) and \(w_{\tilde{s}}\)

\[
w_{\tilde{s}} = 2UC_{w_{\tilde{n}}w_{\tilde{n}}^T}^{-\frac{1}{2}} w_s = 2\Lambda^{-\frac{1}{2}} U^T w_s \quad (4.33)
\]

\[
v_{\tilde{r}} = 2UC_{w_{\tilde{n}}w_{\tilde{n}}^T}^{-\frac{1}{2}} w_r = 2\Lambda^{-\frac{1}{2}} U^T w_r \quad (4.34)
\]
where \( U \) and \( \Lambda \) are the matrices including eigenvectors and eigenvalues of \( C_{w_n w_n^T} \), respectively, (i.e. \( C_{w_n w_n^T} = U \Lambda U^T \)). In Lemma 2.1, it is shown that \( U \) is a rotation matrix with phase \( \frac{\varphi}{2} \), where \( \varphi \) is the phase of \( \alpha_n \). As a result, Equation (4.34) can be interpreted as follows. The circularizing filter first rotates the coordinate system and uses the eigenvectors of \( C_{w_n w_n^T} \) as the new coordinate system to remove the correlation between \( \Re \{ n \} \) and \( \Im \{ n \} \). Afterwards, the new coordinates is scaled separately by the inverse of the square root of corresponding eigenvalues to get a proper \( \tilde{n} \) with equal power on \( \Re \{ \tilde{n} \} \) and \( \Im \{ \tilde{n} \} \).

**4.4.3 Decision Regions of The Detector**

In this section, the decision regions of the binary ML detector will be determined. This result will be used in the next section to study the performance of this detector in terms of the probability of detection error (\( P_e \)).

According to the ML decision rule given in (4.7), the decision region for \( s^0 \) is a region in \( r \)-plane over which \( f_n(r - s^0) > f_n(r - s^1) \), and decision region for \( s^1 \) is the region where \( f_n(r - s^0) < f_n(r - s^1) \). Therefore, the boundary between these decision regions in \( r \)-plane is the solution of the following equation: \( f_n(r - s^0) = f_n(r - s^1) \). By substituting \( f_n(r - s^j) \) from (4.8) and after some manipulations, we will get the following equation for decision boundary in terms of \( x = \Re \{ r \} \) and \( y = \Im \{ r \} \):

\[
\left[ (1 - \Re \{ \alpha_n \}) \Re \{ \Delta s \} - \Im \{ \alpha_n \} \Im \{ \Delta s \} \right] \left( x - \Re \{ \bar{s} \} \right) \\
+ \left[ (1 + \Re \{ \alpha_n \}) \Im \{ \Delta s \} - \Im \{ \alpha_n \} \Re \{ \Delta s \} \right] \left( y - \Im \{ \bar{s} \} \right) = 0
\]

(4.35)

where \( \Delta s = s^1 - s^0 \) and \( \bar{s} = \frac{1}{2}(s^1 + s^0) \). As illustrated in Fig. 4.3.a, this boundary is a line in the \( r \)-plane which passes through the midpoint of the line segment \( s^0 s^1 \). This boundary becomes the perpendicular bisector of the line segment \( s^0 s^1 \) for proper noise (\( \alpha_n \neq 0 \)).

The following discussion gives a geometrical interpretation for this decision boundary.
In Section 2.3, it was shown that the pdf contours of an improper scalar signal have elliptical shape. Consequently, when the additive noise $n$ is improper, the contours of $f_i(r)$ are elliptical counters centered around the point $s^0$ or $s^1$, depending on the actual value of signal-of-interest. These contours are shown in Fig. 4.3.b for a simple example where $s^1 = -s^0 \in \mathbb{R}$. As it can be seen in this figure, the decision boundary is the locus of the points which have equal probability of occurrence in both hypotheses. In this example, this decision boundary can be formulated as $x = cy$, where

$$c = -\tan \psi = \frac{\Im \{\alpha_n\}}{1 - \Re \{\alpha_n\}} = \frac{\sigma_{xy}}{\sigma_y^2}.$$  \hspace{1cm} \text{(4.36)}$$

### 4.4.4 Probability of Detection Error

Using the results of the previous section, we can determine the probability of ML detection error for the binary hypothesis model of (4.26). Let us represent the observation $r = x + jy$ in a new coordinate system as shown in Fig. 4.3. The new coordinate system has the following axes: the line passing through the points $s^0$ and $s^1$, and the perpendicular bisector of the line segment $s^0s^1$. Assuming that the angle between the former line and x-axis is $\theta$, we have $[x', y']^T = R_w(-\theta) [x - \Re \{\bar{s}\}, y - \Im \{\bar{s}\}]^T$, where $R_w(\cdot)$ is the rotation matrix defined in Lemma 2.1.
Now, suppose that $H_1$ is active and our observation equals to $r = s^1 + n$, then the probability of having a detection error will be

$$P_e^1 = Pr \{ f_{r|s^1}(r|s^1) < f_{r|s^0}(r|s^0) \}$$

$$= \int \! \! \! \! \int_{f_{r|s^1}(r|s^1) < f_{r|s^0}(r|s^0)} f_{r|s^1}(r|s^1) dx' < c.y' = \int \! \! \! \! \int f_n(r' - s^1)$$

(4.37)

The value of $c$ in this equation can be determined according to (4.36) as

$$c = - \tan \psi' = \sigma_{n'_I} \sigma_{n'_Q}/\sigma_{n'_I}^2,$$

(4.38)

where $n'_I$ and $n'_Q$ are the inphase and quadrature parts of the noise in $x'$-$y'$ coordinates. Similarly, we can assume that $H_0$ is active and find the probability of detection error in this case ($P_e^0$). As it is proved in Appendix A.6, the average probability of detection error, denoted by $P_e = \frac{1}{2}(P_e^1 + P_e^0)$, can be simplified to

$$P_e = Q \left( \sqrt{\frac{|s^1 - s^0|^2}{2\sigma_n^2}} \times \frac{1 - |\alpha_n| \cos(\varphi - 2\theta)}{1 - |\alpha_n|^2} \right)$$

(4.39)

$$= Q \left( \sqrt{\frac{|s^1 - s^0|^2}{4\sigma_{n'_I}^2 (1 - \rho'^2)}} \right),$$

(4.40)

where $\rho' = \sigma_{n'_I} \sigma_{n'_Q}/\sqrt{\sigma_{n'_I}^2 \sigma_{n'_Q}^2}$ denotes the correlation factor of the noise in $x'$-$y'$ coordinates, and $\varphi$ denotes the phase of the noncircularity coefficient of noise in x-y coordinates.

In order to study the effect of $\alpha_n$ on the probability of detection error in (4.39), let us consider $\theta = 0$; i.e., $n'_I = n_I$, $n'_Q = n_Q$, and $s^0$ and $s^1$ are located on the x-axis in Fig. 4.3.\footnote{For $\theta \neq 0$, we can simply rotate the coordinate systems as shown in Fig. 4.3.} According to (4.39), the performance of detector not only does depend on value of SNR ($\frac{|s^1 - s^0|^2}{\sigma_n^2}$), but depends on the noncircularity coefficient of the noise ($\alpha_n$) as well. Obviously, the performance of the detector improves by increasing the SNR for a fixed $\alpha_n$. Thus, assume that the SNR is fixed. In this case, the error probability is dependent on both the magnitude and phase of $\alpha_n$. Fig. 4.4 illustrates the contours of constant $P_e$ in $\alpha_n$-plane. In Fig. 4.4, the value of $P_e$ is zero on the unit circle ($|\alpha_n| = 1$),
and $P_e$ increases as the radius of the contours decreases. In order to get a better insight into the changes of $P_e$ in this figure, consider the following two cases:

- $\Im \{\alpha_n\} = 0$: In this case, there exists no correlation between inphase and quadrature parts of the noise. Since the signals-of-interest are real valued ($\theta = 0$), the detector completely ignores the quadrature part of the observation in the detection process. As a result, the ML detector becomes a conventional detector with the following wellknown decision rule: $r - s^0 \overset{\hat{s} = s^1}{\gtrsim} r - s^1$. The probability of detection error for this case is

$$P_e = Q \left( \sqrt{\frac{|s^1 - s^0|^2}{4\sigma^2_n}} \right) = Q \left( \sqrt{\frac{|s^1 - s^0|^2}{2\sigma^2_n}} \times \frac{1}{1 + \Re \{\alpha_n\}} \right).$$  \hspace{1cm} (4.41)

Consequently, as $\alpha_n$ increases from $-1$ to $1$ on the real axis in Fig. 4.4, the probability of detection error increases.

- $\Re \{\alpha_n\} = 0$: In this case, the power of the noise is equally distributed between $n_I$ and $n_Q$, and the only source of impropriety is the correlation existing between $n_I$ and $n_Q$. The ML detector exploits this correlation as side information to reduce the effect of noise on detection process. Hence, the more correlation existing between
Figure 4.5: Error probability for the special example of \( s^1 = -s^0 = 1 \) with \( \Re\{\alpha_n\} = 0 \) and \( \Im\{\alpha_n\} = \rho/2 \).


\begin{align*}
n_I \text{ and } n_Q, \text{ the better the performance will be. In the extreme case that } \rho' = 1, \text{ the real part of the noise can be completely canceled using the imaginary part of the observation } r'; \text{ hence, we get } P_e = 0. \text{ In fact, the term } \sigma_{n_I}^2 (1 - \rho^2) \text{ in Equation (4.40) can be interpreted as the effective noise power in ML detection. The correlation factor of the noise for this case can be expressed as } \rho = 2\Im\{\alpha_n\}. \text{ Therefore, as } |\alpha_n| \text{ decreases from 1 to 0 on the imaginary axis in Fig. 4.4, the probability of detection error increases. Fig. 4.5 illustrates the performance of ML detector for a simple example where } s^1 = -s^0 = 1. \text{ Fig. 4.5 also illustrates the improvement which can be achieved by using ML detectors instead of conventional detectors as follows.}

A conventional MF, which is based on a proper noise model, assumes that there is no relevant data in the imaginary part of the observation \((y)\); hence, ignores \( \Im\{r\} \). The performance of such detector is well-known as \( Q(\sqrt{|s^1 - s^0|^2/4\sigma_{n_I}^2} ), \) which does not change for different values of \( \rho \) and is always equivalent to the case of \( \rho = 0 \) shown in Fig. 4.5. However, using the correct improper model for the
noise, the ML detector takes advantage of the correlation existing between the real and imaginary parts of observation \( r \) to give the optimal performance. Thus, its performance improves for improper noises when \( \rho \) increases.

It should be noted that the performance improvement in Fig. 4.5 for small values of \( |\alpha_n| \) is not significant. Therefore, if a slight loss in the performance can be tolerated at the receiver, we can still use conventional detectors for small values of \( |\alpha_n| \). As a rule of thumb, we suggest to use the optimal detector only if the noncircularity coefficient of the noise is more than 0.2. However, this threshold should be modified based on our application.

### 4.5 Chapter Summary

This chapter provided a more detailed study of the ML detection step of the WL-MMSE receiver proposed in the previous chapter. It was shown that under certain conditions, the output of the WL-MMSE estimator has an improper additive noise component. Thus, the receiver’s detector should take this fact into account in order to achieve optimal performance. We showed that, the optimal detector should make use of both correlators and pseudo-correlators for detecting the symbols.

As an alternative solution, we proposed a circularizing filter to convert the improper noise component of the observation, into a proper noise. Using this technique, we can equip conventional detectors, which are designed for proper noises, with this circularizing filter to yield the optimal performance.

The structure of both proposed methods was studied for a simple case of binary scalar detection. It was shown that the performance of the detector in this case depends on the noncircularity coefficient of the noise as well as the SNR. For some special cases, we analyzed the effect of the noncircularity coefficient of the noise on the performance of the receiver. Based on these results, we proposed to use optimal detectors instead of
conventional ones only if this coefficient passes a certain threshold.
Chapter 5

Conclusions

5.1 Research Summary

This thesis proposed a new receiver structure for decoding LD space-time block-codes. The proposed structure exploits the inherent impropriety of LD codes in order to extract the sufficient statistics for detection of transmitted symbols from the observation vector. We showed that when the transmitted and received signals are jointly Gaussian signals, these sufficient statistics can be generated through a widely-linear processing of the observation vector which provides WL-MMSE estimates of the transmitted symbols. Although the assumption of joint Gaussianity of transmitted and received signals does not hold true in practice, our simulation results showed that the sufficiency of WL-MMSE estimates still holds true for practical cases (e.g. PSK and QAM constellations).

In some special cases, the process of passing the observation vector through the widely-linear estimator might convert the additive noise component of the observation into an improper noise. Consequently, conventional detectors, which are designed based on the assumption of properness of the noise, cannot be used in these special cases. In response to this problem, we studied the structure of the ML detector in the presence of improper noise. It was shown that the ML detection requires us to change the structure of the
existing conventional detectors, which is not desired. In order to solve this problem, an alternative approach was proposed. In this approach, conventional detectors will be equipped with a preprocessing step that converts the improper noise component into a proper one.

After deriving the structure of the proposed receiver for general LD codes, we studied some special cases of LD codes in more detail. Due to the large extent of the LD codes, we only selected two important subsets of these codes, namely orthogonal and quasi-orthogonal codes. In the case of orthogonal codes, we showed that Alamouti’s combining scheme, which is commonly used for decoding the orthogonal codes, provides the WL-MMSE estimates of the symbols if certain conditions are satisfied. In the case of quasi-orthogonal codes, our studies led to a new suboptimal receiver which can considerably reduce the complexity of the receiver. Simulation results showed that this suboptimal receiver, performs reasonably close to the optimal receiver for QAM constellations.

Finally, it should be noted that this thesis has also introduced a new analytical tool for characterizing improper signals, namely the noncircularity matrix. This tool was used in the analysis of ML detectors in order to provide us with better insight into the structure and performance of ML detectors in the presence of improper noise.

5.2 Future Work

Due to the large number of the parameters that were involved in this work, we had made some specific assumptions and narrowed down our studies to particular cases. These assumptions can be relaxed in the future work as follows:

- One of the most important assumptions in this work, is the perfect knowledge of the channel state information at the receiver side. In practice, however, this assumption never holds true, and the channel coefficient are estimated at the receiver side. One possible extension of our work, is to study the sensitivity of the proposed
receiver to the channel estimation errors. Furthermore, the proposed structure can be combined with channel estimators to improve the overall performance of the system.

- In this thesis, we assumed that the channel is a flat quasi-static rayleigh channel. The sensitivity of the proposed receiver to the failure of these assumptions should be studied in the future.

- This work has proposed general optimal/suboptimal solutions for all possible LD codes; however, we had only studied its performance for two subsets of LD codes. In the future, the structure and performance of the proposed receiver should be studied for more cases.

- In this thesis, we only considered PSK and QAM constellations at the transmitter side since they are widely used in practice. However, in some applications other types of signaling might be desired, e.g., signaling methods with memory or constant modulus modulations. In those cases, the proposed structure should be changed to take into account the properties of the signaling scheme.

- The rationale behind this work was to exploit the inherent impropriety of LD codes in the receiver design. There exist some other space-time codes which have similar inherent impropriety. It could be an interesting area of research to study the possibility of developing similar receiver structure for those codes.
Appendix A

Proofs

A.1 Proof of Theorem 2.1

Let $\tilde{z} = Dz$, where $z \in \mathbb{C}^K$ and $D = \text{diag}(d_1, \ldots, d_K) \in \mathbb{R}^{K \times K}$. Also, let $C_{zz}^H = QAQ^H$, where $A$ is a diagonal matrix containing eigen values of $C_{zz}^H$ and $Q$ is a unitary matrix containing corresponding eigen vectors of $C_{zz}^H$. Then we get $C_{zz}^T = DC_{zz}D^T$ and $C_{zz}^H = DQAQ^HD^H$. Owing to the fact that $D$ is a real-valued diagonal matrix, it can be shown that $DQ = \tilde{Q}\tilde{D}$, where $\tilde{Q}$ is an orthogonal matrix and $\tilde{D}$ is a real-valued diagonal matrix. It follows that $C_{zz}^H = \tilde{Q}\tilde{D}A\tilde{D}Q^H$; hence $C_{zz}^{1/2} = \tilde{Q}\tilde{D}A^{1/2}$. Now, $A_{\tilde{z}}$ can be written as

$$
A_{\tilde{z}} = \left(C_{zz}^{1/2}C_{zz}C_{zz}^{-1/2}\right) - 1/2 \tilde{D}^{-1}Q^HDC_{zz}D^TQ^*\tilde{D}^{-1}A^{-1/2} - 1/2 \tilde{Q}^HC_{zz}^TQ^*A^{-1/2} - 1/2 \left(C_{zz}^{1/2}C_{zz}C_{zz}^{-1/2}\right)
$$

$$
= A_{z}.
$$

(A.1)
A.2 Proof of Theorem 2.2

According to (2.11) and (2.12), we have $\sigma_4^2 - |\gamma_4^2| = 4[\sigma_x^2\sigma_y^2 - \sigma_{xy}^2]$. From Schwartz inequality ($\sigma_x^2\sigma_y^2 - \sigma_{xy}^2 \geq 0$), it follows that $0 \leq |\gamma_4^2| \leq \sigma_4^2$, hence, $0 \leq |\gamma_2^2| \leq \sigma_2^2$. As a result, $0 \leq |\alpha_z| = \frac{|\gamma_2^2|}{\sigma_2^2} \leq 1$.

A.3 Proof of Theorem 4.1

Let $\tilde{z} = D_1z + D_2z^*$ be the output of the circularizing filter with $C_{\tilde{z}\tilde{z}} = 0$, where $z$ is the input of the filter with $C_{zz} \neq 0$. There are three possible cases for $D_1$ and $D_2$ as follows:

- If $D_2 = 0$: In this case, the output of the filter is $\tilde{z} = D_1z$, which should have a zero pseudo-covariance matrix $C_{\tilde{z}\tilde{z}} = D_1C_{zz}D_1^T = 0$. Now, assume that $D_1$ is invertible. This requires $D_1^{-1}C_{\tilde{z}\tilde{z}}D_1^{-T} = C_{zz} = 0$, which is in contradiction with the fact that $z$ is improper. Therefore, the assumption of invertibility of $D_1$ cannot be true.

- If $D_1 = 0$: In this case, the output of the filter is $\tilde{z} = D_2z^*$ with $C_{\tilde{z}\tilde{z}} = D_2C_{zz}D_2^T = 0$. Now, assume that $D_2$ is invertible. This requires $D_2^{-1}C_{\tilde{z}\tilde{z}}D_2^{-T} = C_{zz} = 0$, which is in contradiction with the fact that $z$ is improper. Therefore, the assumption of invertibility of $D_2$ cannot be true.

- $D_1 \neq 0 \neq D_2$: In this case, $\tilde{z} = [D_1 \ D_2]v_z$; therefore, the augmented vectors $v_{\tilde{z}}$ and $v_z$ are related to each other as follows:

$$v_{\tilde{z}} = \begin{bmatrix} D_1 & D_2 \\ D_2^* & D_1^* \end{bmatrix} v_z; \quad (A.2)$$

Obviously, the above transformation from $v_z$ to $v_{\tilde{z}}$ is reversible if and only if the determinant of the transformation matrix is nonzero, which requires $\det(D_1) \neq 0$ and $\det(D_1 - D_2D_1^{-*}D_2^*) \neq 0$. 

As a result, the circularizing filter is reversible if and only if

\[ D_1 \neq 0, \quad D_2 \neq 0, \quad \text{det}(D_1) \neq 0, \quad \text{and} \quad \text{det}(D_1 - D_2 D_1^{-1} D_2^*) \neq 0. \]  (A.3)

### A.4 Proof of Lemma 4.1

Suppose \( u \in \mathbb{C}^{2K \times 1} \) is an eigenvector of \( C_{v_z v_z}^H \) corresponding to the eigenvalue \( \lambda \in \mathbb{R} \), i.e., \( C_{v_z v_z}^H u = \lambda u \). Let \( u = [u_1^T, u_2^T]^T \), where \( u_1, u_2 \in \mathbb{C}^{K \times 1} \). It follows that

\[
\frac{1}{2} \begin{bmatrix}
C_{zz}^H & C_{zz}^T \\
C_{zz}^* & C_{zz}^{*H}
\end{bmatrix}
\begin{bmatrix}
u_1 \\
u_2
\end{bmatrix}
= \lambda
\begin{bmatrix}
u_1 \\
u_2
\end{bmatrix}
\]  (A.4)

By taking the complex conjugate of both sides in this equation and reordering the variables, we will get

\[
\frac{1}{2} \begin{bmatrix}
C_{zz}^H & C_{zz}^T \\
C_{zz}^* & C_{zz}^{*H}
\end{bmatrix}
\begin{bmatrix}
u_2^* \\
u_1^*
\end{bmatrix}
= \lambda
\begin{bmatrix}
u_2^* \\
u_1^*
\end{bmatrix}
\]  (A.5)

Thus, the vector \([u_2^H, u_1^H]^T\) is also an eigenvector of \( C_{v_z v_z}^H \) corresponding to eigenvalue \( \lambda \). Assuming that the eigenvalues of \( C_{zz}^H \) have no multiplicity, we get \([u_2^H, u_1^H]^T = c[u_1^T, u_2^T]^T\); hence, \( u_1 = u_2 \). As a result, matrix \( Q \) that contains all the eigenvectors of \( C_{v_z v_z}^H \) can be decomposed as follows:

\[
Q = \begin{bmatrix}
Q_1 & Q_2 \\
Q_1^* & Q_2^*
\end{bmatrix},
\]  (A.6)

where \( Q_1, Q_2 \in \mathbb{C}^{K \times K} \).
A.5 Proof of Proposition 4.1

Let $\tilde{z} = D_1 z + D_2 z^*$, where $D_1 = (\Lambda_{1,K}^{-\frac{1}{2}} Q_1^H + j\Lambda_{K+1,2K}^{-\frac{1}{2}} Q_2^H)$ and $D_2 = (\Lambda_{1,K}^{-\frac{1}{2}} Q_1^T + j\Lambda_{K+1,2K}^{-\frac{1}{2}} Q_2^T)$. Then, we have

$$v_{\tilde{z}} = \sqrt{2} \begin{bmatrix} I & jI \\ I & -jI \end{bmatrix} \begin{bmatrix} \Lambda_{1,K}^{-\frac{1}{2}} & 0 \\ 0 & \Lambda_{K+1,2K}^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} Q_1^H & Q_1^T \\ Q_2^H & Q_2^T \end{bmatrix} v_z$$

$$= 2T_K^H \Lambda^{-\frac{1}{2}} Q^H v_z$$

$$= 2T_K^H QC_{v_zv_z}^{-\frac{1}{2}} v_z$$

(A.7)

It follows that $C_{v_zv_z} = 4T_K^H C_{v_zv_z}^{-\frac{1}{2}} C_{v_zv_z}^{-\frac{1}{2}} C_{v_zv_z}^{-\frac{1}{2}} T_K = 4I$. Therefore, $C_{\tilde{z}\tilde{z}} = 0$, which implies that $\tilde{z}$ is a proper random vector. Furthermore, the covariance matrix of $\tilde{z}$ is an identity matrix ($C_{\tilde{z}\tilde{z}} = 4I$), implying that vector $\tilde{z}$ is whitened as well.

A.6 Derivation of Equation (4.40)

Substituting $[x', y']^T$ with

$$\begin{bmatrix} q \\ t \end{bmatrix} = R_u(-\psi') \begin{bmatrix} x' \\ y' \end{bmatrix},$$

the boundary line will lie on the t-axis, and the point $s'^1$ will get the following coordinates the point $s'^1$ in the new coordinate system will be

$$\begin{bmatrix} q^1 \\ t^1 \end{bmatrix} = R_u(-\psi') \begin{bmatrix} s'^1 \\ 0 \end{bmatrix} = s'^1 \begin{bmatrix} \cos \psi' \\ -\sin \psi' \end{bmatrix}$$

(A.8)

Hence, we get

$$f_n(r' - s'^1) = p_0 \exp\left\{ -\frac{1}{2} (w_{r'} - w_{s'^1})^T C_{w_{r'}w_{r'}}^{-1} (w_{r'} - w_{s'^1}) \right\}$$

$$= p_0 \exp\left\{ -\frac{1}{2} \begin{bmatrix} q - q^1 \\ t - t^1 \end{bmatrix} R_u(-\psi') C_{w_{r'}w_{r'}} R_u(\psi')^{-1} \begin{bmatrix} q - q^1 \\ t - t^1 \end{bmatrix} \right\}$$

(A.9)
where $p_0 = \frac{1}{2\pi \sqrt{|C_{w_n'w_n'}|}}$. Suppose $R_w(-\psi') C_{w_n'w_n'}^{-1} R_w(\psi') = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$, then

$$P_{e1} = \int_{-\infty}^{\infty} \int_{-\infty}^{c_y'} f_n(r' - s'^1) \, dx' \, dy'$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{0} p_0 \exp \left\{ -\frac{1}{2} \left[ a(q - q^1)^2 + 2b(q - q^1)(t - t^1) + c(t - t^1)^2 \right] \right\} \, dq \, dt$$

$$= p_0 \int_{-\infty}^{0} \exp \left\{ -\frac{1}{2} \left( \frac{a}{c} - \frac{b^2}{c} \right)(q - q^1)^2 \right\} \, dq$$

$$\times \left[ \int_{-\infty}^{\infty} \exp \left\{ -\frac{1}{2} \left( \sqrt{c}(t - t^1) + \frac{b}{\sqrt{c}}(q - q^1) \right)^2 \right\} \, dt \right] \, dq$$

$$= p_0 \sqrt{\frac{2\pi}{c}} \int_{-\infty}^{0} \exp \left\{ -\frac{1}{2} \left( \frac{a}{c} - \frac{b^2}{c} \right)(q - q^1)^2 \right\} \, dq$$

$$= \frac{2\pi p_0}{\sqrt{ac - b^2}} Q \left( \frac{q^1}{\sqrt{\frac{1}{ac - b^2}}} \right) \tag{A.10}$$

where

$$p_0 = \frac{1}{2\pi \sqrt{|C_{w_n'w_n'}|}} \tag{A.11}$$

$$ac - b^2 = \left| R_w(-\psi') C_{w_n'w_n'}^{-1} R_w(\psi') \right| = \left| C_{w_n'w_n'}^{-1} \right| \tag{A.12}$$

$$q^1 = s'^1 \cos \psi' \tag{A.13}$$

Also, it can be shown that

$$c = \frac{2 \cos^2 \psi'}{\sigma_{z'}^2 \left( 1 - \Re \{ \alpha' \} \right)} = \frac{\cos^2 \psi'}{\sigma_{y'}^2} \cdot \tag{A.14}$$
Thus, (A.10) will be reduced to

\[
P_e^1 = Q \left( \frac{s^1}{\sqrt{\frac{\sigma_x^2}{\sigma_y^2} - \frac{(\sigma_x^2 \rho')}{\sigma_y^2}}} \right) \quad (A.15)
\]

\[
= Q \left( \frac{|s^1 - s^0|^2}{4\sigma_n^2 (1 - \rho'^2)} \right) \quad (A.16)
\]

\[
= Q \left( \frac{|s^1 - s^0|^2}{2\sigma_n^2} \times \frac{1 - |\alpha_n| \cos(\varphi - 2\theta)}{1 - |\alpha_n|^2} \right) \quad (A.17)
\]
Appendix B

On widely-linearity of MMSE estimator for complex Gaussian vectors

Let $x$ and $y$ be two jointly Gaussian complex valued random vectors. Given an observation of $y$, MMSE estimator of $x$ generates an estimate $\tilde{x}$ which minimizes the mean squared estimation error $E\{\|x - \tilde{x}\|^2\}$. It can be shown that regardless of the joint distribution of $x$ and $y$, this estimator takes the form of $\tilde{x} = E\{x|y\}$ [52]. For real valued Gaussian random vectors, this posterior mean becomes a linear function of $y$. However, in this appendix it will be shown that for complex valued Gaussian random vectors this posterior mean is a widely-linear function of $y$.

Let $\mathbf{x} = [x^T, x^H]^T$, $\mathbf{y} = [y^T, y^H]^T$, $\mathbf{w} = [x^T, y^T]^T$, and $\mathbf{w} = [x^T, y^T]^T$. Using the results of [41], the conditional probability density function of $x$ given $y$ can be written as follows:

$$f_{x|y}(x|y) = \frac{f_{xy}(x,y)}{f_y(y)} = \frac{|2\pi C_{\mathbf{xw}^H}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \| \mathbf{w} - \mathbf{w} \|_{C_{\mathbf{xw}^H}}^{-1} \right\}}{|2\pi C_{\mathbf{yw}^H}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \| \mathbf{y} - \mathbf{y} \|_{C_{\mathbf{yw}^H}}^{-1} \right\}}$$

(B.1)
where $\|x\|_C \triangleq x^H C x$. The covariance matrix $C_{ww}^H$ has the following block structure

$$ C_{ww}^H = \begin{bmatrix} C_{yy}^H & C_{yx}^H \\ C_{xy}^H & C_{xx}^H \end{bmatrix}, \quad (B.2) $$

which results in $|C_{ww}^H| = |C_{yy}^H| |Q|$ and

$$ C_{ww}^{-1} = \begin{bmatrix} C_{yy}^{-1} + C_{yy}^{-1} C_{yx}^H Q^{-1} C_{yy}^{-1} C_{yx}^H & -C_{yy}^{-1} C_{yx}^H Q^{-1} \\ Q^{-1} C_{yx}^H C_{yy}^{-1} & Q^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} I -C_{yy}^{-1} C_{yx}^H & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} C_{yy}^{-1} & 0 \\ 0 & Q^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -C_{yy}^{-1} C_{yx}^H & I \end{bmatrix}, \quad (B.3) $$

where $Q = C_{yy} - C_{yy}^{-1} C_{yx}^H C_{yx}$. Thus, (B.1) can be rewritten as

$$ f_{x|y}(x|y) = |2\pi Q|^{-1/2} \exp \left\{ -\frac{1}{2} \left\| x - \bar{x} - C_{yy}^{-1} C_{yx}^H (y - \bar{y}) \right\|_Q^{-1} \right\}. \quad (B.4) $$

The following matrices can be substituted in (B.1)

$$ C_{xy}^H = \begin{bmatrix} C_{xy}^H \\ C_{yx}^H \end{bmatrix}, \quad (B.5) $$

and

$$ C_{yy}^{-1} = \begin{bmatrix} (Q')^{-1} & -C_{yy}^{-1} C_{yy}^T (Q')^{-*} \\ -C_{yy}^{-1} C_{yy}^T (Q')^{-1} & (Q')^{-*} \end{bmatrix} \quad (B.6) $$

where $Q' = C_{yy} - C_{yy}^{-1} C_{yy}^T C_{yy}$. After simple manipulations, we will get

$$ E\{x|y\} = \bar{x} + C_{xy}^H C_{yy}^{-1} (y - \bar{y}) = \bar{x} + F(y - \bar{y}) + G(y^* - \bar{y}^*) \quad (B.7) $$

where

$$ F = [C_{xy} - C_{xy}^T C_{yy}^{-1} C_{yy}^*] [C_{yy} - C_{yy}^T C_{yy}^{-1} C_{yy}^*]^{-1}, \quad (B.8) $$

$$ G = [C_{xy} - C_{xy}^T C_{yy}^{-1} C_{yy}^*] [C_{yy} - C_{yy}^T C_{yy}^{-1} C_{yy}^*]^{-*}. \quad (B.9) $$
Bibliography


   an architecture for realizing very high data rates over the rich-scattering wireless 
   channel. In *Int. Symp. Signals, Systems, and Electronics*, pages 295–300, Pisa, 

   processing for high spectral efficiency wireless communication employing multi- 

   time block coded transmission. In *IEEE Int. Conf. Communications, 2006.*, 

[40] B. Picinbono and P. Chevalier. Widely linear estimation with complex data. *IEEE 


   detection in improper complex gaussian noise. In *IEEE Int. Conf. Acoustics, Speech 


