Transceiver Designs and Analysis for LTI, LTV and Broadcast Channels - New Matrix Decompositions and Majorization Theory

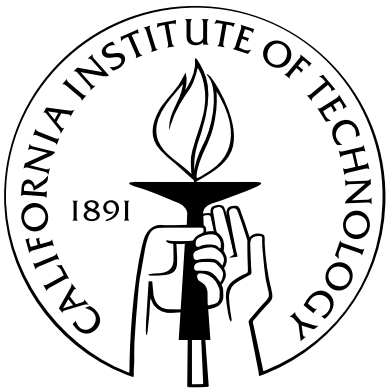
Thesis by

Chih-Hao Liu

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2013

(Defended May 21st, 2013)

Acknowledgments

Foremost, I would like to express my sincerest gratitude to my advisor Prof. P. P. Vaidyanathan for his guidance and support during my stay at Caltech. His office door has always been open, and I could always talk to him no matter how busy he was. He has taught me invaluable skills for doing research, writing papers, and making good presentations. His guidance and inspiration help me to develop confidence, persistence and many qualities of a good researcher. In every aspect, he is a fascinating professor and perfect role model whom I will continue to learn from.

I would also like to thank members of my defense and candidacy examining committee: Professor Yaser Abu-Mostafa, Professor Babak Hassibi, Professor Tracey Ho, Dr. Andre Tkacenko, and Dr. Kevin Quirk. Their knowledge and expertise have been instrumental to my study at Caltech. I studied machine learning from Yaser, and wireless communications from Babak; I learned error correction code from Tracey, communication theory with Kevin, and convex optimization with Andre. Generous support from Caltech and the Office of Naval Research is gratefully acknowledged.

Speaking of scholars, my appreciations go to my present and past labmates, Dr. Chun-Yang Chen, Dr. Ching-Chih Weng, Piya Pal and Srikanth Tenneti for their friendship and stimulating research discussions. In particular, Chun-Yang and Ching-Chih are like my brothers. I cherish their friendship, encouragement, and the joyful moments we share. Piya has been my office-mate for more than four years. I deeply value our academic and personal discussions.

Finally, I would like to thank my parents, Wen-Tin Liu and Mei-Hui Lai, for their unconditional love and support for my entire life. I was born in a farmers' village with scarce educational resources. Without their sacrifices and endeavors to provide me good education, I could not be in Caltech. I would also like to thank my brother, Chun-Cheng Liu, for taking care for our parents while I am abroad. My special thanks go to Louise Ho. I want to thank her for loving me and supporting me during this tumultuous time of my life. Also, Louise's parents, Chia-Tsang Ho and Mei-Ching Wang, have been more than supportive, treating me like their own son.

Abstract

Signal processing techniques play important roles in the design of digital communication systems. These include information manipulation, transmitter signal processing, channel estimation, channel equalization and receiver signal processing. By interacting with communication theory and system implementing technologies, signal processing specialists develop efficient schemes for various communication problems by wisely exploiting various mathematical tools such as analysis, probability theory, matrix theory, optimization theory, and many others. In recent years, researchers realized that multiple-input multiple-output (MIMO) channel models are applicable to a wide range of different physical communications channels. Using the elegant matrix-vector notations, many MIMO transceiver (including the precoder and equalizer) design problems can be solved by matrix and optimization theory. Furthermore, the researchers showed that the majorization theory and matrix decompositions, such as singular value decomposition (SVD), geometric mean decomposition (GMD) and generalized triangular decomposition (GTD), provide unified frameworks for solving many of the point-to-point MIMO transceiver design problems.

In this thesis, we consider the transceiver design problems for linear time invariant (LTI) flat MIMO channels, linear time-varying narrowband MIMO channels, flat MIMO broadcast channels, and doubly selective scalar channels. Additionally, the channel estimation problem is also considered. The main contributions of this dissertation are the development of new matrix decompositions, and the uses of the matrix decompositions and majorization theory toward the practical transmit-receive scheme designs for transceiver optimization problems. Elegant solutions are obtained, novel transceiver structures are developed, ingenious algorithms are proposed, and performance analyses are derived.

The first part of the thesis focuses on transceiver design with LTI flat MIMO channels. We propose a novel matrix decomposition which decomposes a complex matrix as a product of several sets of semi-unitary matrices and upper triangular matrices in an iterative manner. The complexity of

the new decomposition, generalized geometric mean decomposition (GGMD), is always less than or equal to that of geometric mean decomposition (GMD). The optimal GGMD parameters which yield the minimal complexity are derived. Based on the channel state information (CSI) at both the transmitter (CSIT) and receiver (CSIR), GGMD is used to design a butterfly structured decision feedback equalizer (DFE) MIMO transceiver which achieves the minimum average mean square error (MSE) under the total transmit power constraint. A novel iterative receiving detection algorithm for the specific receiver is also proposed. For the application to cyclic prefix (CP) systems in which the SVD of the equivalent channel matrix can be easily computed, the proposed GGMD transceiver has $K/\log_2(K)$ times complexity advantage over the GMD transceiver, where K is the number of data symbols per data block and is a power of 2. The performance analysis shows that the GGMD DFE transceiver can convert a MIMO channel into a set of parallel subchannels with the same bias and signal to interference plus noise ratios (SINRs). Hence, the average bit rate error (BER) is automatically minimized without the need for bit allocation. Moreover, the proposed transceiver can achieve the channel capacity simply by applying independent scalar Gaussian codes of the same rate at subchannels.

In the second part of the thesis, we focus on MIMO transceiver design for slowly time-varying MIMO channels with zero-forcing or MMSE criterion. Even though the GGMD/GMD DFE transceivers work for slowly time-varying MIMO channels by exploiting the instantaneous CSI at both ends, their performance is by no means optimal since the temporal diversity of the time-varying channels is not exploited. Based on the GTD, we develop space-time GTD (ST-GTD) for the decomposition of linear time-varying flat MIMO channels. Under the assumption that CSIT, CSIR and channel prediction are available, by using the proposed ST-GTD, we develop space-time geometric mean decomposition (ST-GMD) DFE transceivers under the zero-forcing or MMSE criterion. Under perfect channel prediction, the new system minimizes both the average MSE at the detector in each space-time (ST) block (which consists of several coherence blocks), and the average per ST-block BER in the moderate high SNR region. Moreover, the ST-GMD DFE transceiver designed under an MMSE criterion maximizes Gaussian mutual information over the equivalent channel seen by each ST-block. In general, the newly proposed transceivers perform better than the GGMD-based systems since the super-imposed temporal precoder is able to exploit the temporal diversity of time-varying channels. For practical applications, a novel ST-GTD based system which does not require channel prediction but shares the same asymptotic BER performance with the ST-GMD

DFE transceiver is also proposed.

The third part of the thesis considers two quality of service (QoS) transceiver design problems for flat MIMO broadcast channels. The first one is the power minimization problem (min-power) with a total bitrate constraint and per-stream BER constraints. The second problem is the rate maximization problem (max-rate) with a total transmit power constraint and per-stream BER constraints. Exploiting a particular class of joint triangularization (JT), we are able to jointly optimize the bit allocation and the broadcast DFE transceiver for the min-power and max-rate problems. The resulting optimal designs are called the minimum power JT broadcast DFE transceiver (MPJT) and maximum rate JT broadcast DFE transceiver (MRJT), respectively. In addition to the optimal designs, two suboptimal designs based on QR decomposition are proposed. They are realizable for arbitrary number of users.

Finally, we investigate the design of a discrete Fourier transform (DFT) modulated filterbank transceiver (DFT-FBT) with LTV scalar channels. For both cases with known LTV channels and unknown wide sense stationary uncorrelated scattering (WSSUS) statistical channels, we show how to optimize the transmitting and receiving prototypes of a DFT-FBT such that the SINR at the receiver is maximized. Also, a novel pilot-aided subspace channel estimation algorithm is proposed for the orthogonal frequency division multiplexing (OFDM) systems with quasi-stationary multipath Rayleigh fading channels. Using the concept of a difference co-array, the new technique can construct M^2 co-pilots from M physical pilot tones with *alternating pilot placement*. Subspace methods, such as MUSIC and ESPRIT, can be used to estimate the multipath delays and the number of identifiable paths is up to $O(M^2)$, theoretically. With the delay information, a MMSE estimator for frequency response is derived. It is shown through simulations that the proposed method outperforms the conventional subspace channel estimator when the number of multipaths is greater than or equal to the number of physical pilots minus one.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Multiple-input multiple-output (MIMO) channel models are unified models for a wide range of different physical communication channels such as multi-antenna wireless systems, wireline systems, or blocked transmission over linear time invariant (LTI) or linear time-varying (LTV) frequency selective scalar channels. The models can be handled by elegant matrix-vector notations. For point-to-point communications with LTI flat MIMO channels, signal processing researchers have proposed various linear transmit-receive processing schemes, and schemes with linear transmit precoder and decision-feedback receiver. The designs are obtained by solving various transceiver optimization problems in matrix-vector form with channel state information (CSI) at both transmitter and receiver. The point-to-point transceiver optimization is a collection of many different problems, depending upon the objective function to be optimized and the constraints to be used. The objective function can be mean square error (MSE), Gaussian mutual information, bit error rate (BER), transmitted power or data rate. Majorization theory [95], [97] and matrix decompositions, such as singular value decomposition (SVD), geometric mean decomposition (GMD) [17] or generalized triangular triangularization (GTD) [21], provide unified frameworks for solving many of the point-to-point transceiver design problems.

The main contributions of this dissertation are the developments of new matrix decompositions, and the uses of the matrix decompositions and the theory of majorization toward practical and efficient transmit-receive scheme designs for the transceiver optimization problems with point-to-point LTI and LTV MIMO channels. For minimum power and maximum rate DFE transceiver optimization problems with MIMO broadcast channels, we show the uses of the joint triangularization and majorization theories for elegant solutions. The transceiver design problem for LTV scalar channels and the channel estimation problem of OFDM systems with quasi-stationary sparse mul-

tipath Rayleigh fading channels will also be considered in this thesis. In the introductory chapter, we give an overview of transceiver optimization problems and review some important results of majorization theories. Every attempt is made to make the text as self-contained as possible. Due to the large volume of literature, the summary here is only directly related to the thesis topics and is by no means a complete treatment of all past work. The reader interested in more comprehensive treatments is referred to [91], [95], [97].

1.1 Channel Models

Transceiver designs depend on the types of channel models. We shall start with introducing the channel models considered in present text. The first channel model we consider is the point-to-point LTI narrowband MIMO channel depicted in Fig. 1.1(a). The input-output relation of the channel is given by

$$\mathbf{y}(n) = \mathbf{H}\mathbf{x}(n) + \mathbf{w}(n), \tag{1.1}$$

where n is the block time index, $\mathbf{H}$ is a $J \times N$ rank M channel matrix, and $\mathbf{x}(n)$ is an $N \times 1$ transmitted signal vector. The noise $\mathbf{w}(n)$ is a $J \times 1$ complex Gaussian random process vector with $E(\mathbf{w}(n)) = \mathbf{0}$ and $E(\mathbf{w}(n)\mathbf{w}^\dagger(n')) = \sigma_w^2\delta(n-n')\mathbf{I}_J$. The vector $\mathbf{y}(n)$ is the $J \times 1$ received signal vector. The LTI narrowband MIMO channel model is generic enough to model many communication scenarios including wireless multi-antenna channels, zero-padding or cylic-prefix systems on frequency selective scalar channels, wireline DSL systems, CDMA channels, and systems exploiting polarization diversity. Modeling a wireless multi-antenna system as a MIMO channel is physically natural. In (1.1), J is the number of receiving antennas, and N is the number of transmitting antennas. For a frequency selective finite impulse response (FIR) scalar channel, the *cyclic-prefix* or *zero-padding* mechanism can convert the channel into a MIMO channel [91]. The effect of the cyclic prefix is to convert a scalar FIR channel to a square circulant matrix $\mathbf{H}$ in (1.1). On the other hand, the zero-padding converts a scalar FIR channel into a full-banded Toeplitz matrix $\mathbf{H}$.

For the narrowband block fading MIMO channel model [98] in Fig. 1.1(b), the channel remains constant over the coherence period of N_c transmitted signal vectors and varies independently [98] or according to Jakes' model [29] across different coherence intervals. In this setting, we model the MIMO channel as a time dependent matrix $\mathbf{H}(n)$ instead of a constant matrix $\mathbf{H}$.

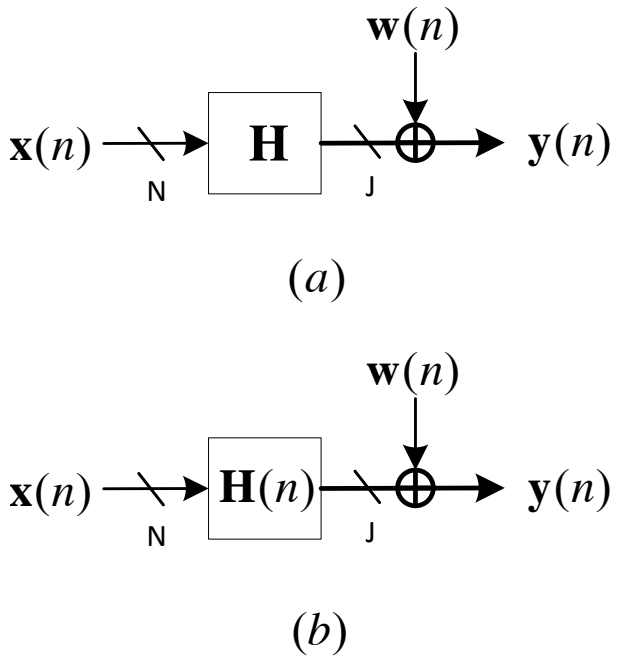


Figure 1.1: (a) LTI narrowband MIMO channel. (b) LTV narrowband MIMO channel.

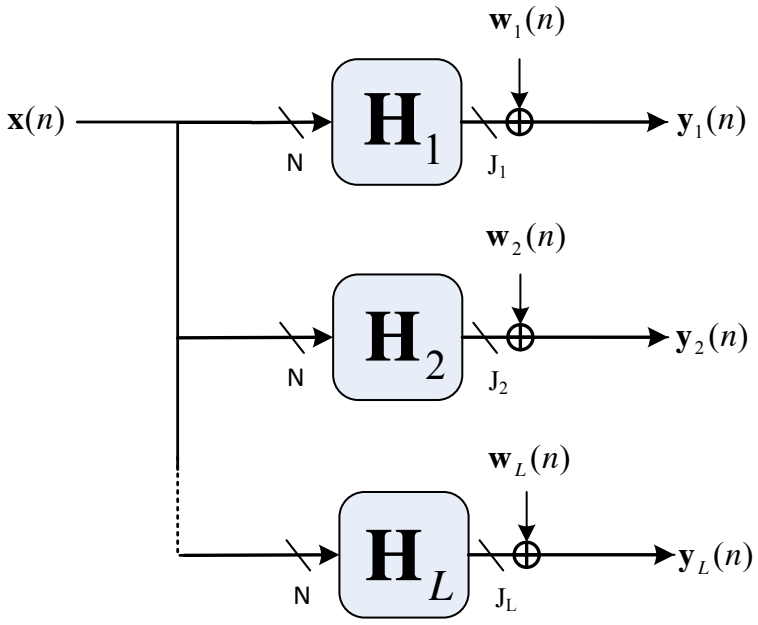


Figure 1.2: LTI narrowband MIMO broadcast channel.

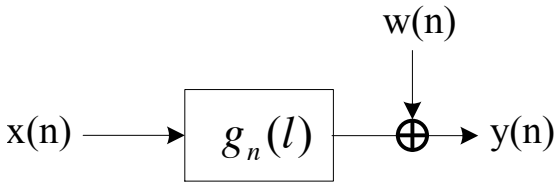


Figure 1.3: LTV scalar channel.

For the scenario that a base station broadcasts *common messages* to L subscribed users through MIMO narrowband channels as shown in Fig. 1.2, the broadcast channel is modeled as

$$\mathbf{y}_i(n) = \mathbf{H}_i\mathbf{x}(n) + \mathbf{w}_i(n),\ i = 1, 2 \ldots, L, \tag{1.2}$$

where $\mathbf{y}_i(n)$ is a $J_i \times 1$ received signal vector of the ith user, $\mathbf{x}(n)$ is an $N \times 1$ transmit signal vector, $\mathbf{H}_i$ is a $J_i \times N$ $(J_i \geq N)$ channel matrix for the ith user, and $\mathbf{w}_i(n)$ is a $J_i \times 1$ circular symmetric additive Gaussian noise vector for the ith user with its statistics given by $E(\mathbf{w}_i(n)) = \mathbf{0}$, $E(\mathbf{w}_i(n)\mathbf{w}_i^{\dagger}(n)) = \sigma_{w_i}^2\mathbf{I}_{J_i}$.

The linear time-varying (LTV) frequency selective scalar channel model is shown in Fig. 1.3. The channel is characterized by the impulse response $g_n(l)$ where $0 \leq l \leq L$. The input and output relation between the received signal $y(n)$ and the transmitted signal $x(n)$ is given by

$$y(n) = \sum_{l=0}^{L} g_n(l)x(n-l) + w(n). \tag{1.3}$$

The LTV channel $g_n(l)$ can be described equivalently by its discrete delay-Doppler function defined as

$$S(v, l) = \sum_{n=0}^{L} g_n(l)e^{-j2\pi vn}. \tag{1.4}$$

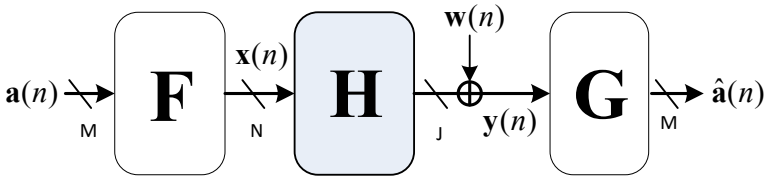


Figure 1.4: MIMO linear transceiver.

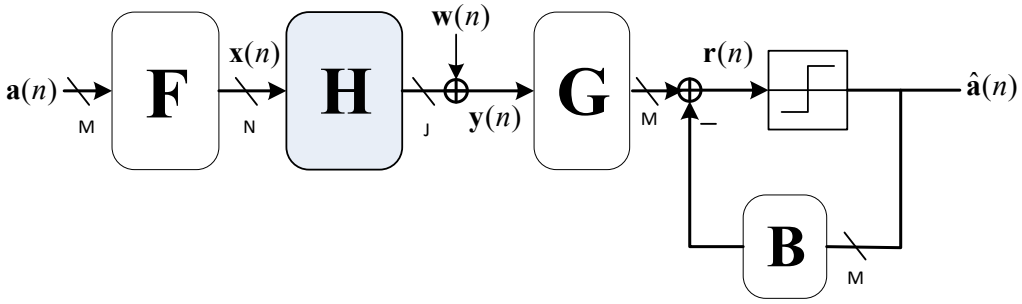


Figure 1.5: MIMO DFE transceiver.

1.2 Point-To-Point MIMO Transceiver Optimization

The problems considered in this thesis are closely related to transceiver optimization problems with point-to-point LTI MIMO channels, which are well studied in the literature. In this section, we shall review the optimization problems of the point-to-point MIMO linear transceiver shown in Fig. 1.4 and the point-to-point MIMO DFE transceiver shown in 1.5. Here $\mathbf{H}$ is a MIMO channel characterized by a $J \times N$ transfer matrix, and $\mathbf{w}(n)$ is the additive channel noise. In the MIMO linear transceiver, the transmitter has an $N \times M$ precoder $\mathbf{F}$ and the receiver has an $M \times J$ equalizer $\mathbf{G}$. The $M \times 1$ data symbol vector $\mathbf{a}(n)$ is precoded by the precoder $\mathbf{F}$ to form the $N \times 1$ transmitted signal vector $\mathbf{x}(n)$. The $J \times 1$ vector $\mathbf{y}(n)$ is the received signal from the MIMO channel, and $\mathbf{y}(n)$ is then equalized by $\mathbf{G}$ to get the $M \times 1$ estimated data symbol vector $\hat{\mathbf{a}}(n)$. The MIMO DFE transceiver has an additional feedback equalizer $\mathbf{B}$ of size $M \times M$. The successive detection and feedback is performed at the receiver, so the matrix $\mathbf{B}$ is restricted to be strictly upper triangular. The goal is to optimize $\{\mathbf{F}, \mathbf{G}\}$ for a linear transceiver or $\{\mathbf{F}, \mathbf{G}, \mathbf{B}\}$ for a DFE transceiver subject to appropriate constraints such that some measure of performance is optimized. This leads to a multitude of interesting problems depending upon what is being optimized and what the constraints are. For

example, one might want to minimize the total mean square reconstruction error under a power constraint. Or one might want to consider the quality of service problems which minimize transmit power or maximize bit rate under per-stream BER constraints. The optimization can be performed with or without the zero forcing constraint,

$$\mathbf{GHF} = \mathbf{I}, \tag{1.5}$$

for a linear transceiver, or

$$\mathbf{GHF} = \mathbf{I} + \mathbf{B}, \tag{1.6}$$

for a DFE transceiver under no error propagation assumption.

The MIMO linear transceiver design problem was considered as early as 1976 [1]. In 1985, Salz considered the joint optimization of the continuous time precoding MIMO filter $\mathbf{F}(jw)$ and MIMO equalizer $\mathbf{G}(jw)$ at the receiver under the average transmitting power constraint for the case where a discrete time sequence $\mathbf{a}(n)$ had to be transmitted over a continuous time channel $\mathbf{H}(jw)$. By Witsenhausen's theorem, which is based on the idea of Schur-convex functions, Saltz showed that the optimal solution can be obtained by diagonalizing the channel with singular value decomposition (SVD), and by solving for the optimal filters for the diagonal channel. The optimal equalizer is identified to be a Wiener filter and the precoding filter can be obtained by using KKT conditions [104]. Later on, diagonalization was also identified to be the key step in various MIMO linear transceiver optimization problems. In 1988, Malvar and Staelin considered the joint optimization of the MIMO linear transceiver in Fig. 1.4 for the case where $\{\mathbf{F}, \mathbf{H}, \mathbf{G}\}$ are only allowed to be rectangular constant matrices [3]. Furthermore, instead of imposing an average power constraint as in earlier work, they imposed per-antenna power constraints. In 1994, Yang and Roy addressed joint optimization of continuous time MIMO DFE transceiver by minimizing the geometric MSE [12].

The equalizer design problem of FIR scalar channels can be transformed into a MIMO transceiver optimization by converting FIR scalar channels into narrowband MIMO channels ($\mathbf{H}$) with the zero-padding or cyclic-prefix method. In a pioneering paper in 1997, Xia considered the equalization of FIR scalar channels with redundant precoder [4]. In 1999, Li and Ding considered the joint optimization of communication systems with redundant filter banks under power constraints [5]. The derivation is parallel to that of Salz in 1985.

One may realize that the MIMO transceiver design problem is a collection of many problems, depending on the objective function to be optimized and the constraints to be used. The objective function can be the average MSE, the Gaussian mutual information, the BER, or the signal to interference and noise ratio (SINR), etc. There may be a zero-forcing constraint, or not. The constraints can be an average power constraint or a peak power constraint, etc. In 2003, instead of considering each design criterion in a separate way, Palomar *et al.* introduced the majorization theory and used it to develop a unifying framework for MIMO linear transceiver designs [11]. The framework generalizes many of the existing results by considering two families of objective functions: Schur-concave and Schur-convex functions [95]. For Schur-concave objective functions, the channel diagonalization structure is always optimal, whereas for Schur-convex functions, an optimal solution diagonalizes the channel only after a very specific rotation of the transmitted symbols. Once the optimal stucture of the transmit-receive processing is known, the design problem simplifies and can be formulated within the powerful framework of convex optimization theory, in which a great number of interesting design criteria can be easily solved. Since this paper, the theory of majorization and convex optimization has became widely used in this field.

In 2005, Jiang *et al.* proposed MIMO DFE transceiver designs with and without a zero-forcing constraint based on the geometric mean decomposition (GMD) [17], [18]. The systems decompose a flat MIMO channel into multiple identical parallel subchannels, which can make it rather convenient to design modulation/demodulation and coding/decoding schemes. Similar MIMO DFE transceiver designs were also published by Zhang *et al.* [16] and Xu *et al.* [19] independently and at the same time. The designs are also known to minimize the arithmetic mean of the expected MSE at the input of the decision device and the average BER in high signal to noise ratio (SNR) [19]. Moreover, it was shown that the GMD-based system with a zero-forcing constraint achieves optimal channel throughput asymptotically in high SNR while the GMD-based system without a zero-forcing constraint maximizes Gaussian mutual information. A unifying framework for MIMO DFE transceiver design was proposed in [23], and independently by [22]. The framework can be regard as a parallel counterpart of [11] in 2003. The authors considered a broad range of design criteria that can be expressed as either Schur-convex or Schur-concave functions of the logarithm of the MSE of each data stream. For Schur-convex objectives, the GMD based MIMO MMSE DFE transceiver in [18], which results in data streams with equal MSEs, is optimal. Interestingly, for Schur-concave objectives, the MIMO linear transceiver is optimal. In 2009, Bergman *et al.* investi-

gated the joint design problem of MIMO DFE transceivers with bit allocation. It was shown that the optimal bit loading, which may take real bit numbers, will result in an orthogonalizing precoder design. Decision feedback (DF) becomes superfluous when the optimal bit allocation is realizable. DF may still be advantageous since it may allow us to redistribute the bit loading at very low cost such that the bit allocation becomes realizable.

It is sometimes desirable to optimize the MIMO transceiver with specific quality of service (QoS) constraints for each substream. In 2000, Lin and Phoong developed linear optimal zero-forcing (ZF) transceivers with bit allocation that minimize the total transmit power for a given target rate and meet per substream symbol error rate (SER) constraints [50], [51]. Optimal ZF solutions, with the aim of minimizing the total transmit power under unequal per substream SER or BER constraints, were considered by Pandaharipande and Dasgupta [53] in 2005 to support multiple data streams, such as voice, data and video, by using the majorization theory. Each of the data streams in general has different BER or SER requirements [52]-[53], [93]. In [54], linear MMSE transceivers were designed by minimizing the total transmit power with possibly different per substream *QoS* constraints, in terms of MSE. Linear transceiver designs with bit allocation which maximize total bit rate with power and SER constraints were considered in [55] and [56]. The proposed designs do not assume a given bit allocation. Rather, both the linear transceiver and bit allocation are jointly optimized. In 2010, Weng *et al.* considered both the min-power and max-rate design problems with possibly unequal SER constraints for zero-forcing MIMO DFE transceivers with bit allocation [57]. The min-power and max-rate transceiver design problems are commonly considered and seem to be closely related. In 2011, Li *el al.* showed the duality between these two problems for MIMO linear transceivers with bit allocation. If a MIMO linear transceiver is optimal for the min-power problem, it is also optimal for the max-rate problem, and the converse is true.

This thesis continues the line of research for point-to-point MIMO transceivers with decision feedback. We propose generalized geometric mean decompositions (GGMD) which generally requires fewer flops than GMD. Using the proposed GGMD, we are able to construct a novel MIMO DFE transceiver with a butterfly structure in both the precoder and receiver. The unique structure greatly reduces the complexity, especially for the application to CP systems. The performance of the proposed GGMD DFE transceiver, in terms of MSE, SER and Gaussian mutual information, are carefully investigated. In earlier works, the MIMO transceiver designs are based on LTI MIMO equivalent channels. We introduce the idea of space-time precoding which allows us to take ad-

vantage of the diversity offered by time-varying channels. Novel MIMO DFE transceiver designs with LTV MIMO channels are proposed for the cases with and without channel prediction. The QoS transceiver design problems are well studied for point-to-point MIMO channels. However, literature on DFE transceiver design with common message MIMO broadcast is very limited. This thesis also considers the broadcast DFE transceiver design with bit allocation for the min-power and max-rate QoS problem with MIMO broadcast channels. Moreover, the transceiver design and channel estimation problems for scalar LTV channels are presented in the present text.

1.3 Majorization Theory and Schur-Convexity

In this section, we shall briefly review the majorization theory and Schur-convexity on which many of the results in this thesis are based. To explain what Schur-convexity is, we first define the notion that the components of a vector $\mathbf{x}$ are less spread out than the components of a vector $\mathbf{y}$.

Definition 1.1: (*Additive Majorization.*) For any $\mathbf{x}, \mathbf{y} \in \mathcal{R}^n$, the vector $\mathbf{x}$ is said to be additively majorized by $\mathbf{y}$, and is denoted as $\mathbf{x} \prec_+ \mathbf{y}$ if and only if

$$\sum_{i=0}^{n-1} x_{[i]} = \sum_{i=0}^{n-1} y_{[i]}, \text{ and}$$

$$\sum_{i=0}^{k} x_{[i]} \leq \sum_{i=0}^{k} y_{[i]}, \text{ k=0,...,n-2,}$$

where $x_{[i]}$ and $y_{[i]}$ denote the $(i+1)$-th largest elements of $\mathbf{x}$ and $\mathbf{y}$, respectively. Moreover, $\mathbf{x}$ is said to be weakly additively majorized by $\mathbf{y}$, denoted $\mathbf{x} \prec_{+w} \mathbf{y}$, if and only if

$$\sum_{i=0}^{k} x_{[i]} \leq \sum_{i=0}^{k} y_{[i]}, \text{ k=0,...,n-1.}$$

□

The statement

$$\mathbf{x} \prec_+ \mathbf{y} \text{ on } \mathcal{S}$$

implies that $\mathbf{x}$ and $\mathbf{y}$ both belong to a subset $\mathcal{S}$ of real vectors, and $\mathbf{y}$ majorizes $\mathbf{x}$ in that subset.

An important tool to study majorization is defined in terms of a doubly stochastic matrix. Before introducing the tool, the formal definition of a doubly stochastic matrix is given as follows:

Definition 1.2: (*Doubly stochastic.*) An $n \times n$ matrix $\mathbf{P}$ is *doubly stochastic* if

$$P_{ij} \geq 0,$$

and

$$\sum_{i=0}^{n-1} P_{ij} = 1, \text{ for } 0 \leq j \leq n-1, \sum_{j=0}^{n-1} P_{ij} = 1 \text{ for } 0 \leq i \leq n-1.$$

□

An important theorem which characterizes additive majorization is stated as the following theorem.

Theorem 1.1: A necessary and sufficient condition that $\mathbf{x} \prec_+ \mathbf{y}$ is that there exists a doubly stochastic matrix $\mathbf{P} \in \mathcal{R}^{n \times n}$ such that $\mathbf{x} = \mathbf{P}\mathbf{y}$, where $\mathbf{x}, \mathbf{y} \in \mathcal{R}^n$. □

A direct but important result which follows from Theorem 1.1 is summarized in the following form:

Lemma 1.1: (*The two extreme vectors.*) Given an n-vector $\mathbf{y}$ whose components satisfy $y_i \geq 0$ and $\sum_i y_i = 1$, we have

$$\begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \end{bmatrix}^T \prec_+ \mathbf{y} \prec_+ \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}^T. \tag{1.7}$$

□

Functions that preserve the ordering of additive majorization are said to be Schur-convex.

Definition 1.3: (*Schur convexity.*) A real-valued function $f(\mathbf{x})$ defined on a set $\mathcal{A} \subset \mathcal{R}^n$ is said to be *Schur convex* on $\mathcal{A}$ if

$$\mathbf{x} \prec_+ \mathbf{y} \text{ on } \mathcal{A} \Rightarrow f(\mathbf{x}) \leq f(\mathbf{y}),$$

and *Schur concave* if

$$\mathbf{x} \prec_+ \mathbf{y} \text{ on } \mathcal{A} \Rightarrow f(\mathbf{x}) \geq f(\mathbf{y}).$$

□

Also note that $-f(\cdot)$ is *Schur concave* if and only if $f(\cdot)$ is *Schur-convex*. Although Schur-convex and Schur-concave functions do not form a partition of functional space, there are many examples of Schur-convex and Schur-concave functions that arise in transceiver optimization. Knowing that the objective is Schur-convex or Schur-concave allows researchers to apply powerful tools from majorization theory. The following theorems are very useful for identifying Schur-convex and Schur-concave functions. The first theorem shows the relationship between a convex function and a Schur-convex function [91].

Theorem 1.2: (*From convex to Schur-convex functions*) Let $g(x)$ be convex on some interval $\mathcal{I}$ of the real line. Then the function $f(\mathbf{x}) = \sum_{i=0}^{n-1} g(x_i)$ is Schur-convex on $\mathcal{I}^n$. □

Example 1.1: (*The average probability of error.*) For square QAM, the symbol error rate (SER) of the detector for AWGN channels is

$$P_e(y) = c_0 Q(A/\sqrt{y}),$$

where c_0 and A depend on the size of QAM constellation and the transmitting power, y is the error variance, and $Q(\cdot)$ is the Q-function. It was shown in [7] that $P_e(y)$ is convex when $y \leq A^2/3$ and concave when $y > A^2/3$. In MIMO systems, one often deals with the average SER of several data streams. For example, if we have M QAM symbol streams, the average SER is given by

$$P_e(\mathbf{y}) = \frac{c_0}{M} \sum_{i=0}^{M-1} Q\left(\frac{A}{\sqrt{y_i}}\right).$$

It follows from Theorem 1.2 that $P_e(\mathbf{y})$ is Schur-convex if $y_i \leq A^2/3$, and Schur-concave if $y_i > A^2/3$ for all i. □

The following theorem is related to compositions that involve Schur-convex or Schur-concave functions [95].

Theorem 1.3: (*Composition of Schur-convex functions.*) Consider compositions of the form

$$\psi(\mathbf{x}) = h(\phi_1(\mathbf{x}), \ldots, \phi_k(\mathbf{x})),$$

where $h(\cdot)$ is a real-valued function defined on $\mathcal{R}^k$, and the real functions, $\phi_i(\cdot)$, have common

domain $\mathcal{A} \subset \mathcal{R}^n$. If each ϕ_i is Schur-convex on $\mathcal{A}$ and h is increasing on $\mathcal{R}^k$ then $\psi(\mathbf{x})$ is Schur-convex. More exhaustively,

$$\begin{aligned}
\text{Each } \phi_i(\mathbf{x}) \text{ is Schur-convex and } h \text{ is increasing} &\implies \psi(\mathbf{x}) \text{ is Schur-convex;} \\
\text{Each } \phi_i(\mathbf{x}) \text{ is Schur-convex and } h \text{ is decreasing} &\implies \psi(\mathbf{x}) \text{ is Schur-concave;} \\
\text{Each } \phi_i(\mathbf{x}) \text{ is Schur-concave and } h \text{ is increasing} &\implies \psi(\mathbf{x}) \text{ is Schur-concave;} \\
\text{Each } \phi_i(\mathbf{x}) \text{ is Schur-concave and } h \text{ is decreasing} &\implies \psi(\mathbf{x}) \text{ is Schur-convex.}
\end{aligned}$$

□

Parallel to the additive majorization is the notion of multiplicative majorization [97].

Definition 1.4: (*Multiplicative majorization.*) The vector $\mathbf{x} \in \mathcal{R}_+^n$ is multiplicatively majorized by $\mathbf{y} \in \mathcal{R}_+^n$, denoted by $\mathbf{x} \prec_\times \mathbf{y}$, if and only if

$$\begin{aligned}
&\prod_{i=0}^{n-1} x_{[i]} = \prod_{i=0}^{n-1} y_{[i]}, \\
&\prod_{i=0}^{k} x_{[i]} \le \prod_{i=0}^{k} y_{[i]}, \quad \text{for } 0 \le k < n-1.
\end{aligned}$$

□

Similar to the definition of the Schur-convex/Schur-concave function, it is natural to define a multiplicatively Schur-convex/Schur-concave function.

Definition 1.5: (*Multiplicative Schur convexity.*) A real-valued function $f(\mathbf{x})$ defined on a set $\mathcal{A} \subset \mathcal{R}^n$ is said to be *multiplicatively Schur convex* on $\mathcal{A}$ if

$$\mathbf{x} \prec_\times \mathbf{y} \text{ on } \mathcal{A} \Rightarrow f(\mathbf{x}) \le f(\mathbf{y}),$$

and *multiplicatively Schur concave* if

$$\mathbf{x} \prec_\times \mathbf{y} \text{ on } \mathcal{A} \Rightarrow f(\mathbf{x}) \ge f(\mathbf{y}).$$

□

However, it is not necessary to use the notion of multiplicatively Schur-convex/Schur-concave functions since $\mathbf{x} \prec_\times \mathbf{y}$ if and only if $\log(\mathbf{x}) \prec_+ \log(\mathbf{y})$.

1.4 Outline of the Thesis

This thesis covers transceiver designs with linear time invariant MIMO channels, linear time-varying MIMO channels, MIMO broadcast channels and time-varying frequency selective scalar channels. In Chapter 2, a novel matrix decomposition is proposed and applied to construct a new MIMO DFE transceiver for LTI flat MIMO channels. The new transceiver has less design and implementation complexity compared to the conventional GMD DFE transceiver. The performance of the new MIMO DFE transceiver is studied and comparisons are made with well-known transceivers in Chapter 3. The new DFE transceiver is shown to achieve the same optimality as the conventional GMD DFE transceiver. Chapter 4 and Chapter 5 deals with the transceiver designs with linear time-varying flat MIMO channels. By exploiting the proposed space time decomposition, we develop novel DFE transceivers with temporal and spatial precoders which can harvest both the temporal and spatial diversity of channels. In Chapter 6, the max-rate and min-power quality of service design problems over MIMO broadcast channels are investigated. By using a particular kind of joint triangularization, we are able to solve both QoS problems analytically. Chapter 7 considers transceiver design for scalar LTV channels. Iterative design algorithms which maximize the receiver SINR are proposed. Chapter 8 focuses on the development of a subspace channel estimation algorithm for OFDM systems with quasi-stationary sparse multipath Rayleigh fading channels. The proposed subspace algorithm can identify up to $O(M^2)$ multipath delays by using M physical pilots. In this section, we will briefly introduce the scope and capture the major results of each chapter.

1.4.1 Generalized Geometric Mean Decomposition and DFE Transceiver Design – Chapter 2

As described in Sec. 1.2, geometric mean decomposition (GMD) is used to design MIMO DFE transceivers which decouple the point-to-point LTI flat MIMO channels into independent subchannels with equal signal to noise plus interference ratios (SINRs) [17], [18]. In Chapter 2, we propose a new matrix decomposition, namely, the generalized geometric mean decomposition (GGMD), which decomposes a complex matrix as a product of several sets of semi-unitary matrices and upper triangular matrices in an iterative manner. The inner most triangular matrix has its diagonal elements equal to the geometric mean of the singular values of the complex matrix. The complexity of GGMD is less than or equal to that of GMD. Based on the proposed GGMD, we are able to

construct a novel MIMO DFE transceiver with butterfly structure. Along with the new butterfly structure, a new iterative interference cancellation algorithm is proposed. For application to cyclic systems, in which the equivalent MIMO channel matrices are circulant and the SVD can be easily computed, the proposed GGMD DFE transceiver has $K/\log_2(K)$ times complexity advantage over the GMD DFE transceiver, where K is the number of data symbols per data block and is a power of 2.

1.4.2 Performance Analysis of GGMD DFE Transceiver – Chapter 3

In Chapter 2, we propose GGMD and demonstrate the uses of GGMD in MIMO transceiver design with decision feedback. Chapter 3 analyzes the performance of the GGMD DFE transceiver for LTI flat MIMO channels without zero-forcing constraint. It is shown that the GGMD DFE transceiver converts MIMO channels into parallel equivalent subchannels and each subchannel has the same bias and equal SINRs. Hence, no bit allocation is required to minimize the average bit error rate (BER). In fact, the transceiver is optimal for symbol error rate (SER) in the moderate high SINR region. Moreover, the GGMD DFE transceiver can achieve MIMO channel capacity simply by independently applying the scalar Gaussian code of the same rate to each subchannel. The analysis in mean square error (MSE) shows that a GGMD DFE transceiver achieves the same optimal average MSE that a GMD DFE MMSE transceiver does. For application to CP systems, the analyses suggest that the MSE and SER performance of the GGMD DFE transceiver is better than an orthogonal frequency division multiplexing (OFDM) or a single carrier cyclic prefix (SC-CP) system. In terms of complexity and performance, the proposed GGMD DFE transceiver has the best of both worlds.

1.4.3 Zero-forcing DFE Transceiver Design over Slowly Time-Varying MIMO Channels Using ST-GTD – Chapter 4

In Chapter 2, Chapter 3, and most of the literature, the research on MIMO transceiver design focuses on linear time invariant MIMO channels. This chapter considers transceiver designs for the block fading MIMO channel, in which the channel is constant over the coherence (block) interval of N_c symbol vectors. The channel varies across different coherence intervals [96]. It is assumed that both instantaneous channel state information (CSI) at the transmitter (CSIT) and the receiver (CSIR) are available. Exploiting the CSIT and CSIR, the GGMD-based transceiver in Chapter 2 or the GMD-based transceiver in [17] can be applied directly. However, the performance measures

are not optimized since different data blocks pass through different MIMO channels and there is no temporal processing among different data blocks. Based on the generalized triangular decomposition (GTD) [21], we develop space-time GTD for the decomposition of linear time-varying flat MIMO channels. Under the assumption that CSIT, CSIR and channel prediction are available, we develop a space-time geometric mean decomposition (ST-GMD) DFE transceiver with the zero-forcing constraint based on ST-GTD. Under perfect channel prediction, the new system minimizes both the average MSE at the detector in each space-time (ST) block (which consists of several coherence blocks), and the average per ST-block BER in the moderate high SNR region. The new transceiver has better performance than the GMD-based system mainly because the super-imposed temporal precoder can exploit the temporal diversity of time-varying channels. For practical applications, a novel ST-GTD based system which does not require channel prediction but shares the same asymptotic BER performance with the ST-GMD DFE transceiver is also proposed.

1.4.4 MMSE DFE Transceiver Design over Slowly Time-Varying MIMO Channels Using ST-GTD – Chapter 5

In Chapter 4, the ST-GTD DFE transceiver with zero-forcing constraint is designed for time-varying MIMO channels. Due to the zero-forcing constraint, the transceiver may suffer from channel nulls and is Gaussian mutual information lossy. In this chapter, we consider the design of DFE transceivers without zero-forcing constraint for the same channel model. In the first part, we propose a GMD based MMSE DFE transceiver with a channel independent temporal precoder (GMDM-TP). A constant unitary temporal precoder is superimposed on the block-wise GMD-based MMSE transceiver. Therefore, it only requires instantaneous CSIT and CSIR. In the second part, ST-GTD is applied to the design of a ST-GMD MMSE DFE transceiver which can maximize Gaussian mutual information over the equivalent channel seen by each ST-block. This is not possible for the ST-GMD zero-forcing (ZF) DFE transceiver introduced in Chapter 4. Furthermore, analyses show that the ST-GMD MMSE DFE transceiver outperforms the ST-GMD ZF DFE transceiver in terms of average MSE and BER.

1.4.5 MIMO Broadcast DFE Transceivers with QoS Constraints: Min-Power and Max-Rate Solutions – Chapter 6

In Chapter 2, we study DFE transceiver design for point-to-point MIMO channels. MIMO broadcasting channels have aroused much interest in recent years. In the common message broadcast scenario, the base station has only one precoder to encode a common message while the subscribed users can decode the message independently. The nature of the problem prevents one from adapting SVD as in a point-to-point MIMO linear transceiver or GMD as in a point-to-point MIMO DFE transceiver. The QR decomposition based transceiver fails to achieve channel capacity with scalar Gaussian code. In Chapter 6, the QoS transceiver design problems for common message MIMO broadcast channels are investigated. The first problem is a power minimization problem with a total bit rate constraint and per-stream BER constraints. The second problem is a data rate maximization problem with a total transmitting power constraint and per-stream BER constraints. In both QoS problems, we jointly optimize the common precoder and the bit allocation at the base station, and feed-forward and feedback equalizers at the receivers. By exploiting a particular class of joint triangularization (JT), we propose the minimum power JT broadcast DFE transceiver (MPJT) for the min-power QoS problem, and the maximum rate JT broadcast DFE transceiver (MRJT) for the max-rate QoS problem. This chapter also shows the duality of the proposed MPJT and MRJT transceivers. Moreover, two suboptimal QR-based MIMO broadcast DFE transceivers are also proposed for the min-power and the max-rate QoS problems.

1.4.6 Optimized DFT-Filterbank Transceivers over LTV Channels – Chapter 7

In Chapter 7, we consider blocked transceiver optimization with linear time-varying (LTV) multipath scalar channels in which channels vary continuously within one block time. For practical applications, it had been reported by Kozek *et al.* in 1998 [75] and Liu *et al.* in 2004 [74] that windowed Fourier functions in the form of $f(n - lN)\exp(j2\pi kn/M)$ serve as good approximate eigenfunctions of LTV multipath channels. The blocked transceiver based on these eigenfunctions can be represented as a discrete Fourier transform modulated filterbank transceiver (DFT-FBT). The chirped OFDM proposed by Barbarossa *et al.* in 2001 [68] and the Affine OFDM systems proposed by Erseghe *et al.* in 2005 are special cases of DFT-FBT in which they use the chirped modulated Fourier basis. Both the chirped and Affine OFDM systems can diagonalize an LTV multipath scalar channel into a set of parallel memoryless subchannels if the support of the spreading function

$S(\nu, \tau)$ of the LTV channel is mainly concentrated on a straight line. Our goal is to design the transmitting and receiving filters of a general DFT-FBT such that the receiving SINR is maximized. We develop an iterative SINR maximization algorithm for the design based on the perfect knowledge of channel state information. A corresponding iterative algorithm is also developed for the case where only the channel statistics is available.

1.4.7 Channel Estimation for OFDM Systems with Co-Pilots – Chapter 8

In Chapter 8, we consider the channel estimation problem for OFDM systems with sparse multipath Rayleigh fading channels. Using the concept of a difference co-array, we propose a new technique which can construct M^2 co-pilots from M physical pilot tones with *alternating pilot placement*. Subspace methods, such as MUSIC and ESPRIT, can theoretically be used to estimate the multipath delays and the number of identifiable paths is up to $O(M^2)$. With the delay information, an MMSE estimator for frequency response is derived.

1.5 Notations

In this section, we define the notations used in this thesis. Upper case bold letters are reserved for matrices and lower case bold letters for vectors. Superscript $(\cdot)^T$ and $(\cdot)^\dagger$ denote the transpose and the conjugate transpose, respectively. The expression x_i or $[\mathbf{x}]_i$ denotes the ith element of a vector $\mathbf{x}$; $A_{i,j}$ or $[\mathbf{A}]_{i,j}$ denotes the (i,j)th element of a matrix $\mathbf{A}$. The notation $\mathbf{I}_M$ denotes the $M \times M$ identity matrix and $\mathbf{1}$ stands for a vector with each entry equal to one. Notation $E(\cdot)$ stands for expectation. The notation $diag(\mathbf{x})$ is a diagonal matrix with the entries of $\mathbf{x}$ on the diagonal. The notation Tr() denotes the trace of the matrix $\mathbf{A}$. The notation $(n \mod m)$ represents the remainder of n divided by m. The notation vec($\mathbf{A}$) denotes a vector formed by stacking the columns of the matrix $\mathbf{A}$ sequentially. The notations $\odot$ and $\otimes$ denote the Khatri-Rao product and Kronecker product, respectively. For two real vectors $\mathbf{x}$ and $\mathbf{y}$, $\mathbf{x} \prec_+ \mathbf{y}$ denotes that $\mathbf{x}$ is additively majorized by $\mathbf{y}$. For two complex vectors $\mathbf{u}$ and $\mathbf{v}$, $\mathbf{u} \prec_\times \mathbf{v}$ and $\mathbf{u} \prec_{\times w} \mathbf{v}$ denote multiplicative and weakly multiplicative majorization, respectively [97], [95]. A *semi-unitary* matrix $\mathbf{A}$ is a matrix such that $\mathbf{A}^\dagger\mathbf{A} = \mathbf{I}$, i.e., all its columns are orthonormal. In figures, "$\uparrow N$" and "$\downarrow N$" denote the signal upsampler and downsampler, respectively [92].

Chapter 2

Generalized Geometric Mean Decomposition and DFE Transceiver Design

Geometric mean decomposition (GMD) was used by Jiang *et al.* to construct MIMO DFE transceivers [17], [18] which decompose a flat MIMO channel into multiple identical parallel subchannels and support simple modulation/demodulation and coding/decoding schemes. This chapter considers a new matrix decomposition which decomposes a complex matrix as a product of several sets of semi-unitary matrices and upper triangular matrices in an iterative manner. The inner most triangular matrix has its diagonal elements equal to the geometric mean of the singular values of the target complex matrix. The complexity[1] of the new decomposition, generalized geometric mean decomposition (GGMD), depends on its parameters, but is always less than or equal to that of GMD. The optimal parameters which yield the minimal complexity are derived. The chapter also shows how to use GGMD to design an optimal decision feedback equalizer (DFE) transceiver for multi-input-multi-output (MIMO) channels without zero-forcing constraint. A novel iterative receiving detection algorithm for the specific receiver is also proposed. For the application to cyclic prefix systems in which the SVD of the equivalent channel matrix can be easily computed, the proposed GGMD transceiver has $K/\log_2(K)$ times complexity advantage over the GMD transceiver, where K is the number of data symbols per data block and is a power of 2. The performance analyses of the proposed GGMD transceiver is deferred to the next chapter. Most of the results in this chapter have been reported in our recent journal paper [27].

[1]The complexity is defined in terms of the number of floating point operations.

2.1 Introduction

In multi-input-multi-output (MIMO) communications, matrix decompositions, such as singular value decompostion (SVD), QR decompostion, gemetric mean decompostion (GMD), generalized triangular decompostion (GTD), etc., are often used to design transceivers which decouple the MIMO channels into independent subchannels [98], [15]-[21]. For a multiuser MIMO broadcast scenario, block diagonal geometric mean decomposition (BD-GMD) [24] can be used for the design instead. The decoupling allows communication systems to achieve the channel capacity[2] by simply applying independent scalar Gaussian codes on individual subchannels.

For the transmission over frequency selective channels, cyclic prefix (CP) systems are found to be useful in mitigating the inter-symbol interference [91]. The CP system converts frequency selective channels into MIMO circulant channels provided the cyclic prefix length is greater than or equal to the order of the dispersive channels. Orthogonal frequency division multiplexing (OFDM) [8], and discrete multitone (DMT) [9] are popular SVD based linear transceivers for CP systems. In order to maximize the bit rate [35] or minimize the average bit error rate (BER) [25], bit allocation is required. However, bit allocation not only makes modulation rather complicated but also may result in bit rate loss or BER performance loss due to the constellation granularity [35], [17], [25]. The GMD decision feedback equalizer (DFE) minimum mean square error (MMSE) transceiver [18], [19] has an advantage over SVD based transceivers in that all its effective subchannels have equal signal to noise plus interference ratios (SINR), and thus may support codebooks of the same rate, avoiding the need for a bit allocation mechanism.

In an OFDM system, the complexity of the transmitter and receiver is mostly that of a Discrete Fourier Transform (DFT) matrix or an Inverse DFT (IDFT) matrix. For both matrices, Fast Fourier Transform (FFT) can be used to reduce the complexity from $O(K^2)$ to $O(K \log_2(K))$ [106], where K is the size of the data block in the underlying CP system and is often chosen to be a power of 2 in practice. We will also assume K to being a power of 2 throughout this chapter. Since DFT and IDFT matrices can convert MIMO circulant channels into parallel subchannels, the main *design complexity* of a GMD DFE MMSE transceiver for a CP system is not from the singular value decompositions (SVD) of a channel matrix, but comes from the GMD of a diagonal matrix with positive ordered diagonal elements, which is $O(K^2)$ [18], [20], [21]. Moreover, the *implementation complexity* of a GMD DFE MMSE transceiver for a CP system is $O(K^2)$. From the perspective of implementation,

[2] In the QR case, the channel capacity is MIMO channel capacity without channel state information at the transmitter.

introducing GMD DFE MMSE transceivers in CP systems does not seem to be practical because K is usually large, and hence it might lead to formidable *design* and *implementation complexity*.

In this chapter, we propose the generalized geometric mean decomposition (GGMD) which decomposes a complex matrix as the product of several sets of *semi-unitary* matrices and upper triangular matrices in an iterative manner. The inner most triangular matrix has its diagonal elements equal to the geometric mean of the singular values of the complex matrix. The class of GGMD is characterized by a set of GGMD parameters and its complexity is completely determined by the parameters. We derive the optimal GGMD parameters which minimize the number of floating point operations (flops) required by the decomposition. In general, GGMD requires fewer flops than GMD. Based on the proposed GGMD, we also propose GGMD DFE transceivers for the general MIMO channels. Using the GGMD decomposition, we can design a precoder of butterfly structure which can be implemented efficiently. Along with the newly proposed iterative interference canceler at the receiver, the transceiver decouples the MIMO channel into parallel subchannels and all its subchannels have equal SINRs akin to a GMD transceiver.

For application to CP systems, the *design* and *implementation complexity* advantages of a GGMD transceiver as opposed to a GMD transceiver are especially prominent. The *design complexity* of a GGMD transceiver comes from K-point fast Fourier transform (FFT), which takes $O(K \log_2(K))$ flops, and applying GGMD to a diagonal matrix with positive ordered diagonal elements, which takes $O(K)$ flops. Moreover, a GGMD DFE transceiver enjoys lower *implementation complexity*, $O(K \log_2(K))$ than a GMD DFE MMSE transceiver, $O(K^2)$.

2.2 Outline

The sections in this chapter are structured as follows. In Section 2.3, we review the time dispersive channel model, CP systems, and the GMD-based MMSE transceiver without zero-forcing constraint [18], [19]. In Section 2.4, we propose generalized geometric decomposition and discuss its computational complexity. Section 2.5 is devoted to the development of the optimal GGMD MMSE transceiver. Concluding remarks are given in Section 2.6.

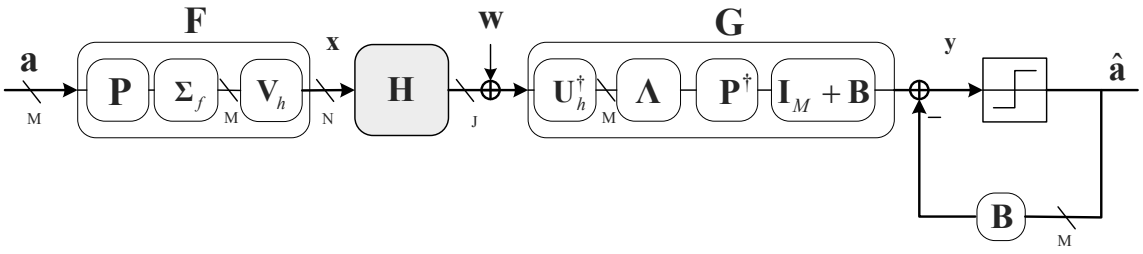


Figure 2.1: The conventional GMD DFE MMSE Transceiver.

2.3 Preliminaries

In CP systems, the purpose of a cyclic prefix is to convert a finite impulse response channel

$$H(z) = \sum_{n=0}^{N_h} h(n) z^{-n}, \tag{2.1}$$

into an $M \times M$ circulant MIMO channel $\mathbf{H}_{cir}$, where M is the data block size. Recall the fact that circulants can be diagonalized by a DFT matrix [91]. That is,

$$\mathbf{\Phi}_\theta \mathbf{\Phi} = \mathbf{W} \mathbf{H}_{cir} \mathbf{W}^\dagger,$$

where $\mathbf{W}$ is a $M \times M$ DFT matrix, $\mathbf{\Phi}$ and $\mathbf{\Phi}_\theta$ are diagonal matrices with the DFT coefficients $\Phi_{m,m} = |H(e^{\frac{j2m\pi}{M}})|$ and $[\mathbf{\Phi}_\theta]_{m,m} = e^{j\angle H(e^{j2m\pi/M})}$. The DFT coefficients can be computed using FFT algorithms, which takes $M \log_2(M)$ flops when M is a power of 2. It is possible to determine the SVD of $\mathbf{H}_{cir}$ without actually running the SVD algorithm of complexity $O(M^3)$. The SVD of $\mathbf{H}_{cir}$ is given by

$$\mathbf{H}_{cir} = \mathbf{U}_h \mathbf{\Sigma}_h \mathbf{V}_h^\dagger, \tag{2.2}$$

where $\mathbf{U}_h = \mathbf{W}^\dagger \mathbf{\Phi}_\theta \mathbf{\Pi}$, $\mathbf{V}_h^\dagger = \mathbf{\Pi}^\dagger \mathbf{W}$,

$$\mathbf{\Sigma}_h = diag(\sigma_{h,0}, \sigma_{h,1}, \ldots, \sigma_{h,M-1}),$$

$\sigma_{h,i} = \Phi_{[i,i]}$ and $\Phi_{[i,i]}$ denotes the ith largest diagonal element of $\mathbf{\Phi}$. $\mathbf{\Pi}$ is a permutation matrix chosen such that $\mathbf{\Pi}^\dagger \mathbf{\Phi} \mathbf{\Pi} = \mathbf{\Sigma}_h$. Well studied sorting algorithms, such as merge sort or heap sort [107] which take $O(M \log_2(M))$ comparisons, can be exploited to sort $\Phi_{m,m}$ for $\mathbf{\Sigma}_h$, and determine

the permutation matrix $\mathbf{\Pi}$. Next, a crucial theorem is reviewed for the transceiver design.

Theorem 2.1: (*The geometric mean decomposition* [20]) Let $\mathbf{H} \in \mathbb{C}^{J\times N}$ have rank M with non-zero singular values $\mathbf{d}_h = [\sigma_{h,0}, \sigma_{h,1}, \cdots, \sigma_{h,M-1}]^T$. Then there exists an upper triangular matrix $\mathbf{R} \in \mathbb{C}^{M\times M}$, and *semi-unitary* matrices $\mathbf{P} \in \mathbb{C}^{N\times M}$ and $\mathbf{Q} \in \mathbb{C}^{J\times M}$, such that

$$\mathbf{H} = \mathbf{Q}\mathbf{R}\mathbf{P}^{\dagger}, \tag{2.3}$$

where $R_{ii} = (\prod_{i=0}^{M-1} \sigma_{h,i})^{1/M}$. A *semi-unitary* matrix is a matrix with orthonormal columns. □

2.3.1 GMD MMSE Transceiver for MIMO channels

A point to point MIMO channel can be modeled as:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{2.4}$$

where $\mathbf{y}$ is a $J \times 1$ channel output vector, $\mathbf{x}$ is an $N \times 1$ channel input vector subject to the total power constraint $E(Tr(\mathbf{x}\mathbf{x}^{\dagger})) = P_0$, $\mathbf{H}$ is a $J \times N$ channel matrix of rank M, and $\mathbf{w}$ is a $J \times 1$ circular symmetric additive Gaussian noise vector with mean and variance given by

$$E(\mathbf{w}) = \mathbf{0},\ E(\mathbf{w}\mathbf{w}^{\dagger}) = \sigma_w^2\mathbf{I}_J. \tag{2.5}$$

The GMD can be applied to design the optimal GMD DFE MMSE transceiver [18], [19] in Fig. 2.1 for the channel $\mathbf{H}$. The details can be found in [39], [18] and [19]. For the completeness of this chapter, the design procedures shall be summarized briefly in the following. For the application to CP systems, one just needs to replace $\mathbf{H}$ with $\mathbf{H}_{cir}$, given by, (2.1) in the design procedures.

It is assumed that the channel state information $\mathbf{H}$ ($\mathbf{H}_{cir}$) is available both at the transmitter and the receiver. The data symbol vector $\mathbf{a}$, which represents a data block, is an $M \times 1$ vector with each element a_i chosen from the alphabet χ of finite size. We assume that

$$E(\mathbf{a}) = \mathbf{0},\ E(\mathbf{a}\mathbf{a}^{\dagger}) = \sigma_a^2\mathbf{I} \text{ and } E(\mathbf{a}\mathbf{w}^{\dagger}) = \mathbf{0}_{M\times M}, \tag{2.6}$$

where channel noise $\mathbf{w}$ is given by (2.5). The diagonal channel matrix $\mathbf{\Sigma}_h \in \mathcal{C}^{M\times M}$, the *semi-unitary* matrix $\mathbf{V}_h \in \mathcal{C}^{J\times M}$ in the precoder $\mathbf{F}$ and *semi-unitary* $\mathbf{U}_h \in \mathcal{C}^{N\times M}$ in the feedforward matrix $\mathbf{G}$

are given by the SVD of $\mathbf{H}$

$$\mathbf{H} = \mathbf{U}_h \mathbf{\Sigma}_h \mathbf{V}_h^\dagger. \tag{2.7}$$

For a CP system, they can be easily computed by (2.2). The power loading matrix $\mathbf{\Sigma}_f$ is a diagonal matrix determined by the water-filling formula in (5.4). To determine the $M \times M$ strictly upper triangular feedback matrix $\mathbf{B}$ and unitary precoder $\mathbf{P} \in \mathcal{C}^{M \times M}$, we first construct the $M \times M$ diagonal design matrix

$$\mathbf{\Theta} = \big(\mathbf{I}_M + \frac{\sigma_a^2}{\sigma_w^2} \mathbf{\Sigma}_f^2 \mathbf{\Sigma}_h^2\big)^{1/2}, \tag{2.8}$$

and apply GMD on $\mathbf{\Theta}$, which is

$$\mathbf{\Theta} = \sigma \mathbf{Q} (\mathbf{I}_M + \mathbf{B}) \mathbf{P}^\dagger, \tag{2.9}$$

where

$$\sigma = \big(\prod_{i=0}^{M-1} \sigma_i \big)^{1/M}, \ \sigma_i = \Theta_{i,i}. \tag{2.10}$$

The $M \times M$ diagonal matrix $\mathbf{\Lambda}$ in $\mathbf{G}$ can be obtained by

$$\mathbf{\Lambda} = \mathbf{\Sigma}_f \mathbf{\Sigma}_h \big(\mathbf{\Sigma}_f^2 \mathbf{\Sigma}_h^2 + \frac{\sigma_w^2}{\sigma_a^2} \mathbf{I}\big)^{-1}. \tag{2.11}$$

2.4 Generalized Geometric Mean Decomposition

In this section, we propose the generalized geometric mean decomposition (GGMD) which is a powerful tool, like GMD in transceiver design [20], [17], [16]. It is defined for any $J \times N$ complex matrix $\mathbf{H}$ of rank M. It decomposes the matrix $\mathbf{H}$ as the product of several sets of *semi-unitary* matrices and upper triangular matrices and hence is a generalization of GMD. GGMD requires significantly less computational complexity than GMD especially for application to diagonal matrices with highly factorable rank number M.

Both GMD and GGMD consist of two stages, SVD and triangularization. The SVD stage is common to both of the decompositions. Let us first take a simple example, where $\mathbf{H}$ has rank $M = 6$, to illustrate the idea of GGMD. The number M can be factorized as the product of positive

integer numbers greater than or equal to 2. Suppose that we factorize M as $M = l_0 \times l_1 = 2 \times 3$; the SVD of $\mathbf{H}$ is done and the singular values are $\{1, 2, 3, 4, 5, 6\}$. The GGMD, with parameters $(l_0, l_1) = (2, 3)$, is as follows. Note that the parameters (l_0, l_1) are determined by the factorization of M and one is free to choose any one of the factorizations. The $(2, 3)$-GGMD consists of two levels of triangularization. For the first level, we construct a diagonal matrix consisting of the singular values of $\mathbf{H}$ on its diagonal as

$$\mathbf{\Sigma}^{(0)} = diag(6, 5, 4, 3, 2, 1). \tag{2.12}$$

Then, we take $l_0 \times l_0$ diagonal submatrices of $\mathbf{\Sigma}^{(0)}$ and decompose them with GMD [20] respectively which can be expressed as

$$\underbrace{\begin{bmatrix} 6 & 0 \\ 0 & 5 \end{bmatrix}}_{\tilde{\mathbf{\Sigma}}_0^{(0)}} = \underbrace{\sqrt{30}}_{\sigma_0^{(1)}} \cdot \tilde{\mathbf{Q}}_0^{(0)} \underbrace{\begin{bmatrix} 1 & \times \\ 0 & 1 \end{bmatrix}}_{\mathbf{I}+\tilde{\mathbf{B}}_0^{(0)}} \tilde{\mathbf{P}}_0^{(0)\dagger}, \tag{2.13}$$

$$\underbrace{\begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}}_{\tilde{\mathbf{\Sigma}}_1^{(0)}} = \underbrace{\sqrt{12}}_{\sigma_1^{(1)}} \cdot \tilde{\mathbf{Q}}_1^{(0)} \underbrace{\begin{bmatrix} 1 & \times \\ 0 & 1 \end{bmatrix}}_{\mathbf{I}+\tilde{\mathbf{B}}_1^{(0)}} \tilde{\mathbf{P}}_1^{(0)\dagger}, \tag{2.14}$$

$$\underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}}_{\tilde{\mathbf{\Sigma}}_2^{(0)}} = \underbrace{\sqrt{2}}_{\sigma_2^{(1)}} \cdot \tilde{\mathbf{Q}}_2^{(0)} \underbrace{\begin{bmatrix} 1 & \times \\ 0 & 1 \end{bmatrix}}_{\mathbf{I}+\tilde{\mathbf{B}}_2^{(0)}} \tilde{\mathbf{P}}_2^{(0)\dagger}. \tag{2.15}$$

The number $\sigma_n^{(1)}$ is the geometric mean of the diagonal elements of $\tilde{\mathbf{\Sigma}}_n^{(0)}$, and $\tilde{\mathbf{Q}}_n^{(0)}$ and $\tilde{\mathbf{P}}_n^{(0)}$ are $l_0 \times l_0$ unitary matrices. Putting (2.13)-(2.15) in a compact matrix form, we have

$$\mathbf{\Sigma}^{(0)} = \mathbf{Q}^{(0)}(\mathbf{\Sigma}^{(1)} \otimes \mathbf{I}_2)(\mathbf{I} + \mathbf{B}^{(0)})\mathbf{P}^{(0)\dagger}, \tag{2.16}$$

where $\mathbf{Q}^{(0)} = diag(\tilde{\mathbf{Q}}_0^{(0)}, \tilde{\mathbf{Q}}_1^{(0)}, \tilde{\mathbf{Q}}_2^{(0)})$, $\mathbf{P}^{(0)} = diag(\tilde{\mathbf{P}}_0^{(0)}, \tilde{\mathbf{P}}_1^{(0)}, \tilde{\mathbf{P}}_2^{(0)})$, $\mathbf{B}^{(0)} = diag(\tilde{\mathbf{B}}_0^{(0)}, \tilde{\mathbf{B}}_1^{(0)}, \tilde{\mathbf{B}}_2^{(0)})$, and $\mathbf{\Sigma}^{(1)} = diag(\sigma_0^{(1)}, \sigma_1^{(1)}, \sigma_2^{(1)})$. This ends the first level of decomposition. For the second level of

decomposition, we take the diagonal matrix $\mathbf{\Sigma}^{(1)}$ in (2.16) and perform GMD as

$$\underbrace{\begin{bmatrix} \sqrt{30} & 0 & 0 \\ 0 & \sqrt{12} & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix}}_{\mathbf{\Sigma}^{(1)}} = \underbrace{\sqrt[6]{720}}_{\Sigma^{(2)}} \cdot \mathbf{Q}^{(1)} \underbrace{\begin{bmatrix} 1 & \times & \times \\ 0 & 1 & \times \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{I}+\mathbf{B}^{(1)}} \mathbf{P}^{(1)\dagger}.$$

The number $\Sigma^{(2)}$ is the geometric mean of the diagonal elements of $\mathbf{\Sigma}^{(1)}$, and $\mathbf{Q}^{(1)}$ and $\mathbf{P}^{(1)}$ are $l_1 \times l_1$ unitary matrices. These operations complete the GGMD for this simple example. From this example, one can see the differences of the triangularization stages of GMD and GGMD while both of them share the same SVD stage. We also show the detailed triangularization results of GMD and GGMD for this example in (2.17) and (2.18), respectively, to demonstrate the differences.

2.4.1 The GGMD Algorithm

For a general $J \times N$ matrix $\mathbf{H}$ of rank M, the following summarizes the procedure for the GGMD with parameters $(l_0, l_1, \ldots, l_{L-1})$ where L is the number of levels in the GGMD and l_i denotes the parameter of GMD in the ith level. The numbers l_i and L could be any positive integers satisfying

$$M = \prod_{i=0}^{L-1} l_i, \tag{2.19}$$

where $l_i \geq 2$. The choice of l_i and L determines the computational complexity of GGMD which will be detailed in the next subsection.

The $(l_0, l_1, \ldots, l_{L-1})$-GGMD consists of two major stages, the SVD and the iterative triangularization.

1. SVD: the SVD decomposes $\mathbf{H}$ as

$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\dagger}, \tag{2.20}$$

where $\mathbf{U}$ is a $J \times M$ *semi-unitary* matrix, $\mathbf{V}$ is an $N \times M$ *semi-unitary* matrix, and $\mathbf{\Sigma}$ is an $M \times M$ diagonal matrix.

2. Triangularization: the triangularization consist of L levels. For $i = 0, \ldots, L-2$, we define

$$N_{i+1} = \frac{N_i}{l_i} = \prod_{n=i+1}^{L-1} l_n, \tag{2.21}$$

with $N_0 = M$ and let $\mathbf{\Sigma}^{(i)}$ be some diagonal matrix

$$\mathbf{\Sigma}^{(i)} = diag(\sigma_0^{(i)}, \sigma_1^{(i)}, \ldots, \sigma_{N_i-1}^{(i)}). \tag{2.22}$$

The matrix $\mathbf{\Sigma}^{(0)}$ is initialized as $\mathbf{\Sigma}^{(0)} = \mathbf{\Sigma}$ where $\mathbf{\Sigma}$ is from (2.20) and the other $\mathbf{\Sigma}^{(i)}$ shall be determined in the subsequent steps.

$$\mathbf{\Sigma}^{(0)} = \begin{bmatrix} 0 & 0 & 0 & 0 & .89 & -.45 \\ 0 & 0 & .53 & -.82 & .10 & .19 \\ 0 & .86 & -.44 & -.26 & .03 & .06 \\ 1 & -.03 & 0 & 3 & 0 & .01 \\ .06 & .51 & .73 & .44 & -.05 & -.10 \\ 0 & 0 & 0 & .25 & .43 & .86 \end{bmatrix} \begin{bmatrix} 3 & -.11 & -.09 & -.02 & .02 & .01 \\ 0 & 3 & -1.86 & -.40 & .42 & .21 \\ 0 & 0 & 3 & -1.16 & 1.22 & .61 \\ 0 & 0 & 0 & 3 & -2.8 & -1.40 \\ 0 & 0 & 0 & 0 & 3 & -4.5 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$
$$\cdot \begin{bmatrix} 0 & 0 & 0 & 1 & .09 & 0 \\ 0 & 0 & .64 & -.07 & .76 & 0 \\ 0 & .31 & -.72 & -.05 & .61 & 0 \\ 0 & -.61 & -.16 & -.01 & .13 & .76 \\ .44 & .64 & .17 & .01 & -.14 & -.07 \\ -.89 & .32 & .08 & .01 & -.07 & .29 \end{bmatrix} \tag{2.17}$$

$$\mathbf{\Sigma}^{(0)} = \begin{bmatrix} .74 & -.67 & 0 & 0 & 0 & 0 \\ .67 & .73 & 0 & 0 & 0 & 0 \\ 0 & 0 & .76 & -.65 & 0 & 0 \\ 0 & 0 & .65 & .75 & 0 & 0 \\ 0 & 0 & 0 & 0 & .82 & -.58 \\ 0 & 0 & 0 & 0 & .58 & .82 \end{bmatrix} \Big(\underbrace{\begin{bmatrix} 5.48 & 0 & 0 \\ 0 & 3.46 & 0 \\ 0 & 0 & 1.41 \end{bmatrix}}_{\mathbf{\Sigma}^{(1)}} \otimes \mathbf{I}_2 \Big)$$
$$\begin{bmatrix} 1 & -.18 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -.29 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -.71 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} .67 & .73 & 0 & 0 & 0 & 0 \\ -.74 & .67 & 0 & 0 & 0 & 0 \\ 0 & 0 & .65 & .76 & 0 & 0 \\ 0 & 0 & -.75 & .65 & 0 & 0 \\ 0 & 0 & 0 & 0 & .57 & .81 \\ 0 & 0 & 0 & 0 & -.81 & .58 \end{bmatrix},$$

$$\underbrace{\begin{bmatrix} 5.48 & 0 & 0 \\ 0 & 3.46 & 0 \\ 0 & 0 & 1.41 \end{bmatrix}}_{\mathbf{\Sigma}^{(1)}} = 2.99 \begin{bmatrix} 0 & .88 & -.48 \\ .97 & -.12 & -.23 \\ .26 & .46 & .84 \end{bmatrix} \begin{bmatrix} 1 & -.45 & -.25 \\ 0 & 1 & -1.28 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & .83 & .55 \\ .48 & -.48 & .73 \\ -.88 & -.26 & .40 \end{bmatrix} \tag{2.18}$$

(a) Let $i = 0$.

(b) Apply GMD to the N_{i+1} $l_i \times l_i$ diagonal submatrices of $\mathbf{\Sigma}^{(i)}$, i.e., for $n = 0, 1, \ldots, N_{i+1}-1$,

$$\tilde{\mathbf{\Sigma}}_n^{(i)} = \sigma_n^{(i+1)} \tilde{\mathbf{Q}}_n^{(i)} (\mathbf{I} + \tilde{\mathbf{B}}_n^{(i)}) \tilde{\mathbf{P}}_n^{(i)\dagger}, \tag{2.23}$$

where $\tilde{\mathbf{\Sigma}}_n^{(i)} = diag(\sigma_{nl_i}^{(i)}, \ldots \sigma_{nl_i+l_i-1}^{(i)})$, $\tilde{\mathbf{Q}}_n^{(i)}$ and $\tilde{\mathbf{P}}_n^{(i)}$ are $l_i \times l_i$ unitary matrices, and $\tilde{\mathbf{B}}_n^{(i)}$ is $l_i \times l_i$ strictly upper triangular matrix. By Theorem 2.1,

$$\sigma_n^{(i+1)} = \Big(\prod_{k=0}^{l_i-1} \sigma_{l_i n+k}^{(i)} \Big)^{1/l_i}. \tag{2.24}$$

(c) If $i < L-1$ then set $i = i+1$ and go to (b), else stop.

From (2.23), we can identify the diagonal matrix $\mathbf{\Sigma}^{(i+1)}$ for the next stage since $\sigma_n^{(i+1)}$ are now available. Stacking $\tilde{\mathbf{\Sigma}}_n^{(i)}$ to form $\mathbf{\Sigma}^{(i)}$ in (2.22), equation (2.23) can be re-formulated as the *GGMD iterative equation*,

$$\mathbf{\Sigma}^{(i)} = \mathbf{Q}^{(i)} (\mathbf{\Sigma}^{(i+1)} \otimes \mathbf{I}_{l_i}) (\mathbf{I} + \mathbf{B}^{(i)}) \mathbf{P}^{(i)\dagger}, \tag{2.25}$$

where $\mathbf{Q}^{(i)} = diag(\tilde{\mathbf{Q}}_0^{(i)}, \ldots, \tilde{\mathbf{Q}}_{N_{i+1}-1}^{(i)})$, $\mathbf{P}^{(i)} = diag(\tilde{\mathbf{P}}_0^{(i)}, \ldots, \tilde{\mathbf{P}}_{N_{i+1}-1}^{(i)})$, and $\mathbf{B}^{(i)} = diag(\tilde{\mathbf{B}}_0^{(i)}, \ldots, \tilde{\mathbf{B}}_{N_{i+1}-1}^{(i)})$. Using (2.25), we have

$$\Sigma^{(L)} = \Big(\prod_{i=0}^{M-1} \sigma_i^{(0)} \Big)^{1/M}. \tag{2.26}$$

The scalar $\Sigma^{(L)}$ is the geometric mean of the singular values of $\mathbf{H}$.

In summary, GMD decomposes $\mathbf{H}$ as the product of $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{P}$ whereas GGMD decomposes $\mathbf{H}$ as the product of $\Sigma^{(L)}$, and L sets of $\mathbf{Q}^{(i)}$, $\mathbf{I} + \mathbf{B}^{(i)}$ and $\mathbf{P}^{(i)}$ are given by (2.25), i.e.,

$$\mathbf{H} = \mathbf{U}\mathbf{Q}'\Sigma^{(L)}\mathbf{A}'\mathbf{V}^{\dagger}, \tag{2.27}$$

where

$$\mathbf{Q}' = \prod_{i=0}^{L-1} \mathbf{Q}^{(i)} \otimes \mathbf{I}_{l'_{i-1}},$$
$$\mathbf{A}' = \prod_{i=0}^{L-1} (\mathbf{I} + \mathbf{B}^{(L-1-i)}) \mathbf{P}^{(L-1-i)\dagger} \otimes \mathbf{I}_{l'_{L-2-i}},$$
$$l'_n = \prod_{i=0}^{n} l_i \text{ for } n \geq 0,\ l'_{-1} = 1.$$

Also notice that the GMD is a special case of GGMD when we set $L = 1$ and $l_0 = M$.

2.4.2 Complexity

The SVD stage requires $O(MJN)$ flops which are independent of the GGMD parameters $(l_0, l_1, \ldots, l_{L-1})$. Given the SVD of $\mathbf{H}$, the GMD on each $l_i \times l_i$ diagonal matrix requires $7l_i^2 + 18l_i - 35$ flops [21] where $l_i \geq 2$. For $\mathbf{\Sigma}^{(i)}$, the algorithm needs to apply GMDs on N_{i+1} $l_i \times l_i$ diagonal matrices. Hence, the triangularization step of GGMD takes

$$\sum_{i=0}^{L-1} T(l_i)(\prod_{n=i+1}^{L-1} l_n) \tag{2.28}$$

flops where

$$T(x) = 7x^2 + 18x - 35. \tag{2.29}$$

Define

$$f(\mathbf{l}, L) = \sum_{i=0}^{L-1} \Big(T(l_i)(\prod_{n=i+1}^{L-1} l_n) \Big), \tag{2.30}$$

where $\mathbf{l} = [l_0, l_1, \ldots, l_{L-1}]^T$. We can minimize the complexity of the triangularization by solving

$$\min_{\mathbf{l}, L} f(\mathbf{l}, L) \tag{2.31}$$
$$\text{subject to } \prod_{n=0}^{L-1} l_n = M,\ l_n \geq 2,\ 1 \leq L \leq \log_2(M).$$

This is a combinatorial optimization problem, solving it directly is difficult. The optimal solution is given in Theorem 2.2 which uses the results in the following lemmas.

Lemma 2.1: The GGMD with parameters $(l_0, \ldots, l_j, \ldots, l_{L-1})$ requires more floating point operations than the GGMD with parameters $(l_0, \ldots, m_0, m_1, \ldots, l_{L-1})$ if there exists a l_j which can be further be factorized as $l_j = m_0 m_1$, where $m_0, m_1 \geq 2$ and $l_j \geq 4$. □

Proof: See appendix. □

Lemma 2.1 shows that increasing the levels, L, of GGMD reduces the complexity. In other words, one shall factorize M as much as possible.

Lemma 2.2: The GGMD with parameters $(l_0, \ldots, l_j, l_{j+1} \ldots, l_{L-1})$ requires fewer floating point operations than the GGMD with parameters $(l_0, \ldots, l_{j+1}, l_j, \ldots, l_{L-1})$ where $l_j < l_{j+1}$, $l_j, l_{j+1} \geq 2$ and $l_j l_{j+1} = \beta$. □

Proof: See appendix. □

Lemma 2.2 suggests that given L, arranging the parameters in increasing order, $(l_{(0)}, l_{(1)}, \ldots, l_{(L-1)})$, yields the minimum complexity where $l_{(n)}$ denotes the nth smallest element in $\mathbf{l} \in \mathcal{R}^L$. Using Lemma 2.1 and Lemma 2.2, we have the following theorem.

Theorem 2.2: The solution to the optimization problem given in (2.31) is given by

$$\mathbf{l}^* = (l_0^*, l_1^*, \ldots, l_{L^*-1}^*), \tag{2.32}$$

where l_n^* are prime numbers, $\prod_{n=0}^{L^*-1} l_n^* = M$, $2 \leq l_n^*$, and $l_i^* \leq l_j^*$ for $0 \leq i \leq j \leq L^* - 1$. □

Proof: If l_n is not a prime number, by Lemma 2.1, we can always reduce the complexity by factorizing l_n. Moreover, if $l_j > l_{j+1}$, by Lemma 2.2, swapping l_j and l_{j+1} shall reduce the complexity. In the same manner as bubble sort [107], one shall get the optimal ordering as given in (2.32) after $L^* - 1$ runs of swapping scans. □

We call the GGMD with the optimal GGMD parameters $(l_0^*, l_1^*, \ldots, l_{L^*-1}^*)$, in particular, the *optimal* GGMD. In the applications where $M = K = 2^\gamma$, Theorem 2.2 says that the complexity of the triangularization step is minimized by choosing $l_i = 2$ and $L = \gamma$. The optimal value of problem (2.31) is

$$f^*(\mathbf{l}^*, L^*) = 29(K - 1).$$

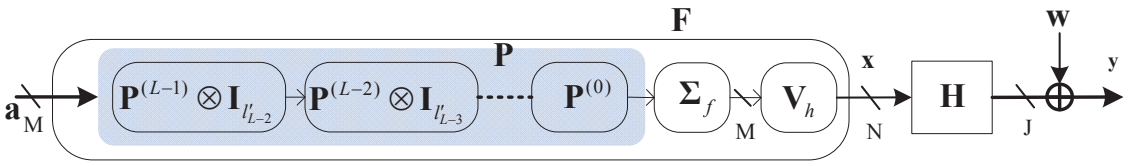


Figure 2.2: The transmitter of the GGMD DFE transceiver and the channel.

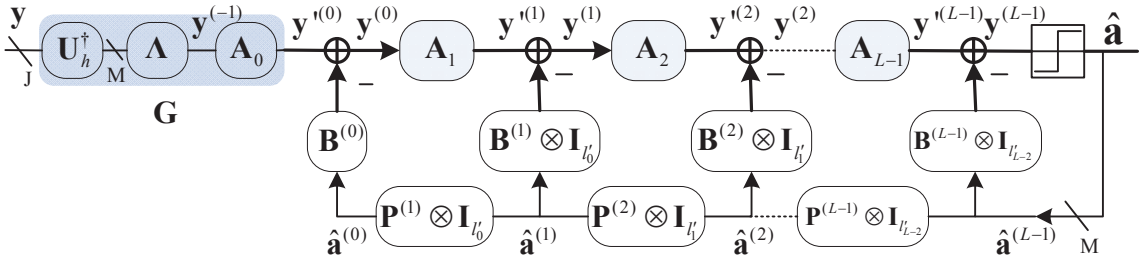


Figure 2.3: The receiver of the GGMD DFE transceiver.

So, the total complexity of the *optimal* GGMD is $O(KJN) + O(29K)$. For the GMD proposed in [20], SVD is required as well. In the second stage, the algorithm of GMD decomposes the diagonal matrix $\mathbf{\Sigma}$ as

$$\mathbf{\Sigma} = \mathbf{Q}'\mathbf{R}\mathbf{P}'^{\dagger}, \tag{2.33}$$

where $\mathbf{R}$ is a $K \times K$ upper triangular matrix, $\mathbf{P}'$ and $\mathbf{Q}'$ are $K \times K$ unitary matrices. This requires $T(K)$ flops. To obtain the form $\mathbf{H} = \mathbf{Q}\mathbf{R}\mathbf{P}^{\dagger}$, we take $\mathbf{R}$, $\mathbf{Q}'$ and $\mathbf{P}'$ from (2.33), and compute $\mathbf{Q} = \mathbf{U}\mathbf{Q}'$ and $\mathbf{P} = \mathbf{V}\mathbf{P}'$. So, GMD has the total complexity of $O(KJN) + O(7K^2 + 2(J+N)K)$. For the applications in which the SVD of $\mathbf{H}$ has negligible complexity, the *optimal* GGMD requires much lower complexity, $O(K)$, than GMD, which is $O(K^2)$.

For the value of M where M not a prime number, the *optimal* GGMD is a different kind of decomposition to GMD, since the decomposition in (2.27), in general, cannot be re-grouped as the product of $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{P}$ as in (2.3) where $R_{i,i}$ equals the geometric mean of the singular values of $\mathbf{H}$.

2.5 GGMD DFE Transceiver

In this section, we propose the GGMD DFE transceiver for general MIMO channels by using the *optimal* GGMD introduced in Sec. 2.4. It can be directly applied to CP systems in which we will see

obvious *design* and *implementation complexity* advantages over the GMD DFE MMSE transceiver. The proposed precoder and receiver of a GGMD transceiver are shown in Fig. 2.2 and Fig. 2.3, respectively. The equivalent MIMO channel is given by a $J \times N$ matrix $\mathbf{H}$ as in (2.4). For a CP system, the channel matrix $\mathbf{H} = \mathbf{H}_{cir}$ where $\mathbf{H}_{cir}$ is defined in Sec. 2.3. The channel state information $\mathbf{H}$ is assumed to be available both at the transmitter and the receiver. The data symbol vector $\mathbf{a}$ and the channel noise $\mathbf{w}$ are given by (2.6) and (2.5), respectively. The diagonal channel matrix $\mathbf{\Sigma}_h$, the matrix $\mathbf{V}_h$ in the precoder $\mathbf{F}$ and $\mathbf{U}_h^\dagger$ in the feedforward matrix $\mathbf{G}$ are from the SVD of channel matrix $\mathbf{H}$ as in (2.7). The power loading matrix $\mathbf{\Sigma}_f$ is a diagonal matrix described in Sec. 2.3.1 and the diagonal matrix $\mathbf{\Lambda}$ in feedforward decoder $\mathbf{G}$ is given by (2.11). To design the rest of the transceiver, we take the $M \times M$ diagonal matrix $\mathbf{\Theta}$ in (2.8) and apply the proposed GGMD in Sec. 2.4 to obtain $\mathbf{B}^{(i)}$ and $\mathbf{P}^{(i)}$. The GMMD parameters are given by the factorization of $M = \prod_{i=0}^{L-1} l_i$ as in (2.19) and the factors l_i are chosen to be prime numbers in increasing order so that the number of flops is minimized. The partial product of l_i is defined as

$$l'_n = \prod_{i=0}^{n} l_i. \tag{2.34}$$

By setting

$$\mathbf{\Sigma}^{(0)} = \mathbf{\Theta}, \tag{2.35}$$

the feedback matrices $\mathbf{B}^{(i)}$ and precoding matrices $\mathbf{P}^{(i)}$ can be computed via the GGMD algorithm in Sec. 2.4.1, i.e.,

$$\mathbf{\Sigma}^{(i)} = \mathbf{Q}^{(i)}(\mathbf{\Sigma}^{(i+1)} \otimes \mathbf{I}_{l_i})(\mathbf{I} + \mathbf{B}^{(i)})\mathbf{P}^{(i)\dagger}. \tag{2.36}$$

Note that the SVD stage is not required since $\mathbf{\Theta}$ is already a diagonal matrix with positive diagonal elements in decreasing order. The matrices $\mathbf{A}_i$ are given by

$$\mathbf{A}_i = (\mathbf{I} + \mathbf{B}^{(i)})\mathbf{P}^{(i)\dagger} \otimes \mathbf{I}_{l'_{i-1}}. \tag{2.37}$$

We also define $\mathbf{P}^{(L)} = 1$.

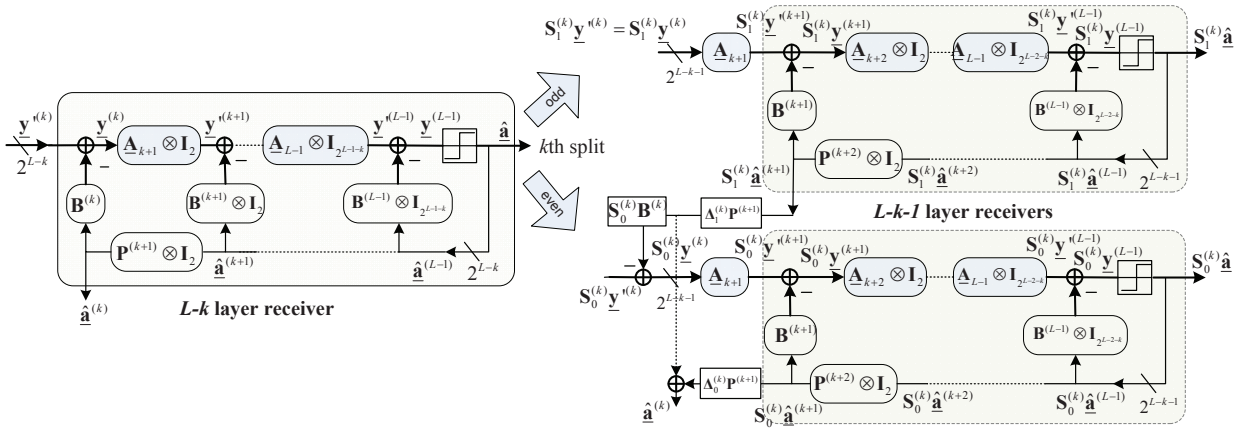


Figure 2.4: The decomposition of an equivalent $(L-k)$-layer receiver into two $(L-k-1)$-layer receivers

2.5.1 Receiver Detection Algorithm

Fig. 2.3 is a conceptual figure of the receiver which specifies the input and output relationships of the received signals and the detected symbols, just like the receiver of a GMD MMSE transceiver [19]. The actual symbol detection must follow a particular ordering such that the feedback parts of the system do not violate causality. The estimated symbol vector $\hat{\mathbf{a}}$ is obtained by decoding the output of $\mathbf{G}$, $\mathbf{y}'^{(0)}$, with the iterative cancellation detection algorithm to be described later. Before giving the algorithm, we need some notations, then, the algorithm is presented. Following the algorithm, a simple example is given. The detailed explanations of the algorithm are deferred to Sec. 2.5.2.

Define $f(z)$ as a function which sets $z \in \mathcal{C}$ to the element in the alphabet χ with the minimum Euclidean distance. The downsampling operator $\mathbf{S}_i^{(m)}$ is a downsampler with the downsampling factor l_m and shift i. It is defined on a $N_m \times 1$ vector $\mathbf{x}$ such that the output is a $N_{m+1} \times 1$ vector given by

$$\mathbf{y} = \mathbf{S}_i^{(m)}\mathbf{x},$$

where $y_n = x_{l_m n+i}$, $i = 0, 1, \ldots, l_m - 1$, $0 \le n \le N_{m+1} - 1$ and N_m is given by (2.21). In matrix form, it can be represented as

$$\mathbf{S}_i^{(m)} = \mathbf{I}_{N_{m+1}} \otimes \mathbf{e}_i^{(m)T}, \tag{2.38}$$

where $\mathbf{e}_i^{(m)}$ is an $l_m \times 1$ vector containing all zeros except for an 1 in the ith entry. We also define the upsampling operator $\mathbf{\Delta}_i^{(m)}$ with the upsampling factor l_m and shift i on a $N_{m+1} \times 1$ vector $\mathbf{x}$ such that the $N_m \times 1$ output vector $\mathbf{y}$ is

$$\mathbf{y} = \mathbf{\Delta}_i^{(m)} \mathbf{x},$$

where $y_{l_m n+i} = x_n$ and the other entries of $\mathbf{y}$ are zeros. In matrix form, $\mathbf{\Delta}_i^{(m)}$ can be expressed as

$$\mathbf{\Delta}_i^{(m)} = \mathbf{I}_{N_{m+1}} \otimes \mathbf{e}_i^{(m)}.$$

The bijective function $j = g(i)$ maps a positive number $0 \leq i \leq M-1$ to a positive number $0 \leq j \leq M-1$. Given i, j is computed via the iterative equations

$$q_{-1} = i, \tag{2.39}$$

$$q_{k-1} = q_k l_k + r_k, \tag{2.40}$$

where $0 \leq k \leq L-1$, $0 \leq r_k \leq l_k - 1$, L and l_k are given by (2.19). We first set $q_{-1} = i$, and compute

$$q_k = \Big\lfloor \frac{q_{k-1}}{l_k} \Big\rfloor,$$
$$r_k = q_{k-1} \mod l_k,$$

iteratively for $k = 0, 1, \ldots, L-1$. After obtaining all r_k, the number j is given by

$$j = \sum_{k=0}^{L-1} \Big((l_k - 1 - r_k) \cdot \prod_{m=k+1}^{L-1} l_m \Big). \tag{2.41}$$

For example, if $M = K = 2^L$ and we choose $l_k = 2$, then the function $j = g(i)$ takes the positive number $i = (b_{L-1} b_{L-2} \ldots b_1 b_0)_2$ and maps it to

$$j = g(i) = (\bar{b}_0 \bar{b}_1 \ldots \bar{b}_{L-2} \bar{b}_{L-1})_2. \tag{2.42}$$

The detection algorithm is described by the pseudo code as follows:

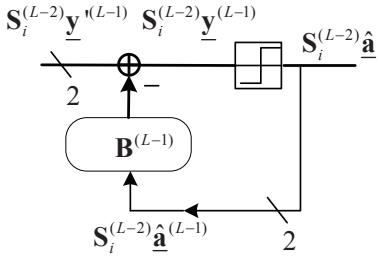


Figure 2.5: Equivalent 1-layer receiver

Algorithm 2.5.1: ITERATIVEDETECTION($\mathbf{y}'^{(0)}, L$)

procedure DETECT($k, \mathbf{y}, v, \hat{\mathbf{a}}, L, \mathbf{l}$)
for $i \leftarrow 0$ **to** $l_k - 1$ $do\{$
$n \leftarrow l_k - 1 - i$
if $i = 0$
then $\begin{cases} \mathbf{x} \leftarrow \mathbf{0}_{N_m \times 1} \\ \mathbf{z}' \leftarrow \mathbf{S}_n^{(k)} \mathbf{y} \end{cases}$
else $\left\{ \mathbf{z}' \leftarrow \mathbf{S}_n^{(k)} \left(\mathbf{y} - \mathbf{B}^{(k)} \mathbf{x} \right) \right.$ (ii)
if $k < L - 1$
then $\begin{cases} \mathbf{z} \leftarrow (\mathbf{I} + \mathbf{B}^{(k+1)}) \mathbf{P}^{(k+1)\dagger} \mathbf{z}' & \text{(iii)} \\ (\hat{\mathbf{a}}, \tilde{\mathbf{a}}, v) \leftarrow \text{DETECT}(k+1, \mathbf{z}, v, \hat{\mathbf{a}}, L, \mathbf{l}) & \text{(iv)} \end{cases}$
else $\begin{cases} \hat{a}_{g^{-1}(v)} = \tilde{\mathbf{a}} \leftarrow f(\mathbf{z}') & \text{(v)} \\ v \leftarrow v + 1 \end{cases}$
$\mathbf{x} \leftarrow \mathbf{x} + \mathbf{\Delta}_n^{(k)} \mathbf{P}^{(k+1)} \tilde{\mathbf{a}}$ (vi)
$\}$
$\tilde{\mathbf{a}} \leftarrow \mathbf{x}$ (vii)
return $(\hat{\mathbf{a}}, \tilde{\mathbf{a}}, v)$ (viii)

main
$k \leftarrow 0$
$v \leftarrow 0$
$\mathbf{l} \leftarrow (l_0, l_1, \ldots, l_{L-1})$
$\hat{\mathbf{a}} \leftarrow \mathbf{0}_{M \times 1}$
$(\hat{\mathbf{a}}, \tilde{\mathbf{a}}, v) \leftarrow \text{DETECT}(k, \mathbf{y}'^{(0)}, v, \hat{\mathbf{a}}, L, \mathbf{l})$ (ix)
output $(\hat{\mathbf{a}})$

The left arrow "$\leftarrow$" indicates an assignment of a variable. "**for** $i \leftarrow 0$ **to** $l_k - 1$" denotes the loop with index variable i taking the value 0, then 1 and so on up to $l_k - 1$. "**main**" and "**output**" indicate the main procedure and the output of a program, respectively. DETECT(k, $\mathbf{y}$, v, $\hat{\mathbf{a}}$, L, $\mathbf{l}$) denotes the

procedure DETECT with input variable k, $\mathbf{y}$, v, $\hat{\mathbf{a}}$, L and $\mathbf{l}$. For ease of justification, we will explain the algorithm for the case where $M = K = 2^L$ and $l_k = 2$. The procedure DETECT(k,$\mathbf{y}$, v, $\hat{\mathbf{a}}$, L, $\mathbf{l}$) can be used to decode $\underline{\hat{\mathbf{a}}}$ and $\underline{\hat{\mathbf{a}}}^{(k)}$ from input signal $\underline{\mathbf{y}}'^{(k)}$ for the equivalent $(L-k)$-layer receiver in Fig. 2.4. The matrices $\underline{\mathbf{A}}_i$ in Fig. 2.4 are defined as

$$\underline{\mathbf{A}}_i = (\mathbf{I} + \mathbf{B}^{(i)})\mathbf{P}^{(i)\dagger}.$$

In this case, the input signal of the receiver, $\underline{\mathbf{y}}'^{(k)}$, corresponds to the input variable $\mathbf{y}$ of the procedure DETECT, and the output signals $\underline{\hat{\mathbf{a}}}$ and $\underline{\hat{\mathbf{a}}}^{(k)}$ correspond to the output variables $\hat{\mathbf{a}}$ and $\tilde{\mathbf{a}}$ in line (viii), respectively. The input variables k and L of DETECT indicate that the procedure is doing the decoding for an $(L-k)$-layer receiver.

The algorithm below starts by calling DETECT in line (ix) to decode $\hat{\mathbf{a}}$ with input signal $\mathbf{y}'^{(0)}$ for the (L)-layer GGMD receiver in Fig. 2.3. When $i = 0$, the procedure extracts the odd indexed elements of $\mathbf{y}^{(0)}$ (line i) and feeds them into DETECT with parameter $k = 1$ (line iv), which does the decoding for the $(L-1)$-layer receiver. It repeats the same actions with even indexed elements of $\mathbf{y}'^{(0)}$ when $i = 1$. This is because from the perspectives of the odd and even indexed components of $\mathbf{y}'^{(0)}$, they both see equivalent $(L-1)$-layer receivers as on the right side of Fig. 2.4. Note that the decoding of even indexed elements has to wait until the odd indexed elements are fully decoded. The procedure continues to break the signals into even and odd until the signals see one layer ($m = 1$) receivers (Fig. 2.5) in which the decoding is as simple as a V-BLAST [15] with two subchannels. Then, the decoding of $\hat{\mathbf{a}}$ follows the ordering defined by (2.42) in line (v). More elaborations on the algorithm can be found later in Sec. 2.5.2.

To understand the proposed detection algorithm better, we shall go through the detection algorithm with a simple example where $M = K = 4$ and $l_k = 2$ in the following, i.e., the GGMD DFE receiver in Fig. 2.3 has $L = 2$ and $M = 4$. After the reception of a block $\mathbf{y}$, the receiver computes $\mathbf{y}'^{(0)} = \mathbf{G}\mathbf{y}$ and calls the procedure DETECT in "**main**" to decode $\hat{\mathbf{a}}$. The detailed operations are summarized as follows:

1) Initialize $k = 0$, the counter $v = 0$, $\mathbf{l} = [2, 2]$, $\hat{\mathbf{a}} = \mathbf{0}_{K\times 1}$.

2) Call procedure DETECT with input k, $\mathbf{y}'^{(0)}$, v and $\hat{\mathbf{a}}$ to decode the 2 layer receiver on the left side of Fig. 2.4.

1. $i = 0$, $[y_1^{(0)} y_3^{(0)}]^T = [y_1'^{(0)} y_3'^{(0)}]^T$. (extract odd indexed signal)

2. $\mathbf{z}^{(0)} = [y_1'^{(1)} y_3'^{(1)}]^T = (\mathbf{I} + \mathbf{B}^{(1)})\mathbf{P}^{(1)\dagger}[y_1^{(0)} y_3^{(0)}]^T$. (line (iii))

3. Call precedure DETECT with input $k = 1$, $\mathbf{z}^{(0)}$, v, and $\hat{\mathbf{a}}$ to decode the equivalent 1 layer receiver as on the upper right side of Fig. 2.4, and in Fig. 2.5.

a. $i = 0$, $y_3^{(1)} = y_3'^{(1)}$. (extract odd indexed signal in line (i))

b. $\hat{a}_3 = f(y_3^{(1)})$ and $v = v + 1$. (decode $\hat{a}_3$ in line (v))

c. $\mathbf{x}^{(1)} = [0, \hat{\mathbf{a}}_3]^T$. (line (vi))

d. $i = 1$, $y_1^{(1)} = \mathbf{S}_0^{(1)}([y_1'^{(1)} y_3'^{(1)}]^T - \mathbf{B}^{(1)}\mathbf{x}^{(1)})$. (line (ii))

e. $\hat{a}_1 = f(y_1^{(1)})$ and $v = v + 1$. (decode $\hat{a}_1$, line (v))

f. $\tilde{\mathbf{a}}^{(1)} = [\hat{a}_1, \hat{a}_3]^T$. (return to previous layer of DETECT)

4. $\mathbf{x}^{(0)} = \mathbf{\Delta}_1^{(0)}\mathbf{P}^{(1)}\tilde{\mathbf{a}}^{(1)}$. (line (vi))

5. $i = 1$, $[y_0^{(0)} y_2^{(0)}]^T = \mathbf{S}_0^{(0)}(\mathbf{y}'^{(0)} - \mathbf{B}^{(0)}\mathbf{x}^{(0)})$. (even indexed signal, line (ii))

6. $\mathbf{z}'^{(0)} = [y_0'^{(1)} y_2'^{(1)}]^T = (\mathbf{I} + \mathbf{B}^{(1)})\mathbf{P}^{(1)\dagger}[y_0^{(0)} y_2^{(0)}]^T$. (line (iii))

7. Call procedure DETECT with input $k = 1$, $\mathbf{z}'^{(0)}$, v, and $\hat{\mathbf{a}}$ to decode the equivalent 1 layer receiver as in the lower right side of Fig. 2.4.

a. $i = 0$, $y_2^{(1)} = y_2'^{(0)}$.

b. $\hat{a}_2 = f(y_2^{(1)})$ and $v = v + 1$. (decode $\hat{a}_2$, line (v))

c. $\mathbf{x}^{(1)} = [0, \hat{\mathbf{a}}_2]^T$

d. $i = 1$, $y_0^{(1)} = \mathbf{S}_0^{(1)}([y_0'^{(1)} y_2'^{(1)}]^T - \mathbf{B}^{(1)}\mathbf{x}^{(1)})$.

e. $\hat{a}_0 = f(y_0^{(1)})$ and $v = v + 1$. (decode $\hat{a}_0$, line (v))

f. $\tilde{\mathbf{a}}^{(1)} = [\hat{a}_0, \hat{a}_2]^T$. (done and return)

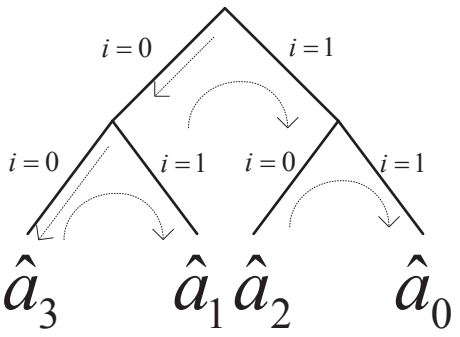


Figure 2.6: The flow of detection algorithm for $L = 2$

Fig. 2.6 illustrates the flow of the detection algorithm. The depth of the tree corresponds to the depth of the function call. The detection sequence of $\hat{a}_n$ is in bit-reversed order of the L bits one's complement of n. In our example, the symbol $\hat{a}_n$, where $n = (b_1 b_0)_2$ in binary representation, is the $(\bar{b}_0 \bar{b}_1)_2$th symbol to be detected.

2.5.2 The Validity of the Detection Algorithm

In the following paragraphs, we will first show that an $(L-k)$-layer receiver can be split into l_k $(L-k-1)$-layer receivers, where $k = 0, 1, \ldots, L-2$, as illustrated in Fig. 2.4, and then, we shall explain how the algorithm successfully exploits this to decode $\hat{\mathbf{a}}$ in Fig. 2.3. For ease of justification, we consider the GGMD transceiver with $M = K = 2^L$ and $l_i = 2$. Similar justifications can be extended to GGMD transceivers with M not being a power of 2. The following lemma is useful in justifying the algorithm.

Lemma 2.3: The $N_{m+1} \times N_m$ matrix $\mathbf{S}_i^{(m)}$ defined in (2.38) satisfies

$$\mathbf{S}_i^{(m)}(\mathbf{Q}^{(m)} \otimes \mathbf{I}_{l_m}) = \mathbf{Q}^{(m)}\mathbf{S}_i^{(m)} \tag{2.43}$$

for any $\mathbf{Q}^{(m)} \in \mathcal{C}^{N_{m+1} \times N_{m+1}}$. □

Proof:

$$\begin{aligned}\mathbf{S}_i^{(m)}(\mathbf{Q}^{(m)} \otimes \mathbf{I}_{l_m}) &= (\mathbf{I}_{N_{m+1}} \otimes \mathbf{e}_i^{(m)T})(\mathbf{Q}^{(m)} \otimes \mathbf{I}_{l_m}) \\ &= \mathbf{Q}_i^{(m)} \otimes \mathbf{e}_i^{(m)T} = \mathbf{Q}^{(m)}\mathbf{S}_i^{(m)}.\end{aligned}$$

□

For decoding, the algorithm perform a total of L splits and each split breaks the receiver into two sub-receivers at a time. The idea of receiver splits is demonstrated in Fig. 2.4. The splits start at $k = 0$, in which the receiver on the left side of Fig. 2.4 corresponds to the L-layer receiver in Fig. 2.3. We have $\underline{\mathbf{y}}^{(n)} = \mathbf{y}^{(n)}$, $\underline{\mathbf{y}}'^{(n)} = \mathbf{y}'^{(n)}$, and $\underline{\hat{\mathbf{a}}}^{(n)} = \hat{\mathbf{a}}^{(n)}$ for $0 \le n \le L-1$. After the split, we have two $(L-1)$-layer receivers as shown on the right side of Fig. 2.4. The splits continue for k going from 0 to $L-1$. From the kth to $(k+1)$th split, for notational convenience, we set $\underline{\mathbf{y}}^{(n)}$, $\underline{\mathbf{y}}'^{(n)}$ and $\underline{\hat{\mathbf{a}}}^{(n)}$ on the left hand side of Fig. 2.4 in the $(k+1)$th split as

$$\underline{\mathbf{y}}^{(n)} = \mathbf{S}_i^{(k)}\underline{\mathbf{y}}^{(n)}, \underline{\mathbf{y}}'^{(n)} = \mathbf{S}_i^{(k)}\underline{\mathbf{y}}'^{(n)}, \underline{\hat{\mathbf{a}}}^{(n)} = \mathbf{S}_i^{(k)}\underline{\hat{\mathbf{a}}}^{(n)},$$

where $\mathbf{S}_i^{(k)}\underline{\mathbf{y}}^{(n)}$, $\mathbf{S}_i^{(k)}\underline{\mathbf{y}}'^{(n)}$, and $\mathbf{S}_i^{(k)}\underline{\hat{\mathbf{a}}}^{(n)}$ are from the right side of Fig. 2.4 in kth split, and $k+1 \le n \le L-1$. Eventually, the splits break the original receiver into 2^{L-1} one layer receivers as shown in Fig. 2.5. The decoding of $\mathbf{S}_i^{(L-2)}\underline{\hat{\mathbf{a}}}$ from $\mathbf{S}_i^{(L-2)}\underline{\mathbf{y}}'^{(L-1)}$ is essentially the same as the GMD MMSE

receiver in Fig. 2.1 with $M = 2$.

The split is possible due to the special structure of $\mathbf{B}^{(k)}$,

$$\mathbf{B}^{(k)} = diag(\tilde{\mathbf{B}}_0^{(k)}, \tilde{\mathbf{B}}_1^{(k)}, \ldots, \tilde{\mathbf{B}}_{2^{L-k-1}-1}^{(k)}),$$

where $\tilde{\mathbf{B}}_n^{(k)}$ are 2×2 strictly upper triangular matrices. To exploit the special structure of $\mathbf{B}^{(k)}$, we break $\underline{\mathbf{y}}^{(k)}$ in Fig. 2.4 into odd and even indexed signal components, $\mathbf{S}_1^{(k)}\underline{\mathbf{y}}^{(k)}$ and $\mathbf{S}_0^{(k)}\underline{\mathbf{y}}^{(0)}$, respectively. They can be computed via

$$\mathbf{S}_1^{(k)}\underline{\mathbf{y}}^{(k)} = \mathbf{S}_1^{(k)}\underline{\mathbf{y}}'^{(k)}, \tag{2.44}$$

$$\mathbf{S}_0^{(k)}\underline{\mathbf{y}}^{(k)} = \mathbf{S}_0^{(k)}\underline{\mathbf{y}}'^{(k)} - \mathbf{S}_0^{(k)}\mathbf{B}^{(k)}\underline{\hat{\mathbf{a}}}^{(k)}. \tag{2.45}$$

The component $\mathbf{S}_0^{(k)}\underline{\mathbf{y}}^{(k)}$ takes the feedback component $\mathbf{S}_0^{(k)}\mathbf{B}^{(k)}\underline{\hat{\mathbf{a}}}^{(k)}$ which can further be factored as

$$\mathbf{S}_0^{(k)}\mathbf{B}^{(k)}\underline{\hat{\mathbf{a}}}^{(k)} = \mathbf{S}_0^{(k)}\mathbf{B}^{(k)}\mathbf{\Delta}_1^{(k)}\mathbf{P}^{(k+1)}\mathbf{S}_1^{(k)}\underline{\hat{\mathbf{a}}}^{(k+1)}. \tag{2.46}$$

Here, we use the fact $\mathbf{S}_0^{(k)}\mathbf{B}^{(k)} = \mathbf{S}_0^{(k)}\mathbf{B}^{(k)}\Delta_1^{(k)}\mathbf{S}_1^{(k)}$ and Lemma 2.3.

From the left side of Fig. 2.4, the $\underline{\mathbf{y}}^{(n)}$ and $\underline{\mathbf{y}}'^{(n)}$ have the relationships such that

$$\begin{aligned}
\underline{\mathbf{y}}'^{(n+1)} &= \Big((\mathbf{I} + \mathbf{B}^{(n+1)})\mathbf{P}^{(n+1)\dagger} \otimes \mathbf{I}_{2^{n+1-k}}\Big)\underline{\mathbf{y}}^{(n)}, \\
\underline{\mathbf{y}}^{(n+1)} &= \underline{\mathbf{y}}'^{(n+1)} - (\mathbf{B}^{(n+1)} \otimes \mathbf{I}_{2^{n+1-k}})\underline{\hat{\mathbf{a}}}^{(n+1)},
\end{aligned} \tag{2.47}$$

for $n = k, k+1, \ldots, L-2$. Multiplying $\mathbf{S}_i^{(k)}$ on both sides and applying Lemma 2.3, we have

$$\begin{aligned}
\mathbf{S}_i^{(k)}\underline{\mathbf{y}}'^{(n+1)} &= \Big((\mathbf{I} + \mathbf{B}^{(n+1)})\mathbf{P}^{(n+1)\dagger} \otimes \mathbf{I}_{2^{n-k}}\Big)\mathbf{S}_i^{(k)}\underline{\mathbf{y}}^{(n)}, \\
\mathbf{S}_i^{(k)}\underline{\mathbf{y}}^{(n+1)} &= \mathbf{S}_i^{(k)}\underline{\mathbf{y}}'^{(n+1)} - (\mathbf{B}^{(n+1)} \otimes \mathbf{I}_{2^{n-k}})\mathbf{S}_i^{(k)}\underline{\hat{\mathbf{a}}}^{(n+1)}.
\end{aligned} \tag{2.48}$$

Using Lemma 2.3, one can show that

$$\mathbf{S}_i^{(k)}\underline{\hat{\mathbf{a}}}^{(n)} = \Big(\mathbf{P}^{(n+1)} \otimes \mathbf{I}_{2^{n-k}}\Big)\mathbf{S}_i^{(k)}\underline{\hat{\mathbf{a}}}^{(n+1)}. \tag{2.49}$$

From (2.48) and (2.49), we can conclude that both $\mathbf{S}_1^{(k)}\underline{\mathbf{y}}^{(k)}$ (odd) and $\mathbf{S}_0^{(k)}\underline{\mathbf{y}}^{(k)}$ (even) respectively see the equivalent $(L-k-1)$-layer receivers as shown on the right side of Fig. 2.4. Moreover, the

Table 2.1: Design complexity (K is a power of 2)

	GMD Transceiver	GGMD Transceiver
$\mathbf{\Sigma}_h$, $\mathbf{V}_h$, $\mathbf{U}_h^\dagger$	$O(K\log_2(K))$	$O(K\log_2(K))$
$\mathbf{\Sigma}_f$	$O(K)$	$O(K)$
$\mathbf{\Theta}$	$O(K)$	$O(K)$
Decompose $\mathbf{\Theta}$	$O(K^2)$	$O(K)$
Total	$O(K^2)$	$O(K\log_2(K))$

output $\underline{\hat{\mathbf{a}}}^{(k)}$ of the $(L-k)$-layer can be written as

$$\begin{aligned}\underline{\hat{\mathbf{a}}}^{(k)} &= \left(\sum_{i=0}^{1}\mathbf{\Delta}_i^{(k)}\mathbf{S}_i^{(k)}\right)\left(\mathbf{P}^{(k+1)}\otimes\mathbf{I}_2\right)\hat{\mathbf{a}}^{(k+1)} \\ &= \sum_{i=0}^{1}\mathbf{\Delta}_i^{(k)}\mathbf{P}^{(k+1)}\mathbf{S}_i^{(k)}\hat{\mathbf{a}}^{(k+1)},\end{aligned} \tag{2.50}$$

where $\sum_{i=0}^{1}\mathbf{\Delta}_i^{(k)}\mathbf{S}_i^{(k)} = \mathbf{I}_{N_k}$ and the second equality follows from Lemma 2.3.

Now, we are ready to justify how the algorithm works. Suppose that the procedure DETECT works for an $(L-k-1)$-layer receiver and we want to show that the procedure DETECT can decode $\underline{\hat{\mathbf{a}}}$ and $\underline{\hat{\mathbf{a}}}^{(k)}$ from $\underline{\mathbf{y}}'^{(k)}$ for the $(L-k)$-layer receiver in Fig. 2.4. At iteration $i=0$, line (i) computes $\mathbf{S}_1^{(k)}\underline{\mathbf{y}}^{(k)}$ and the signal $\mathbf{S}_1^{(k)}\underline{\mathbf{y}}^{(k)}$ sees the equivalent $(L-k-1)$-layer receiver shown on the upper right side of Fig. 2.4. Line (iii) computes $\mathbf{S}_1^{(k)}\underline{\mathbf{y}}'^{(k+1)}$. Since the procedure DETECT works for a $(L-k-1)$-layer receiver, line (iv) can successfully decode $\mathbf{S}_1^{(k)}\underline{\hat{\mathbf{a}}}$ and $\mathbf{S}_1^{(k)}\underline{\hat{\mathbf{a}}}^{(k+1)}$ from $\mathbf{S}_1^{(k)}\underline{\mathbf{y}}'^{(k+1)}$. After that, at iteration $i=1$, $\mathbf{S}_1^{(k)}\underline{\hat{\mathbf{a}}}^{(k+1)}$ can be removed from $\mathbf{S}_0^{(k)}\underline{\mathbf{y}}'^{(k)}$ by (2.45) and (2.46) (line (vi) and (ii)) and the lower receiver can again use DETECT (line (iv)) to decode $\mathbf{S}_0^{(k)}\underline{\hat{\mathbf{a}}}$ and $\mathbf{S}_0^{(k)}\underline{\hat{\mathbf{a}}}^{(k+1)}$ from $\mathbf{S}_0^{(k)}\underline{\mathbf{y}}'^{(k+1)}$. Using (2.50), line (vi) and (vii) compute $\hat{\mathbf{a}}^{(k+1)}$ from $\mathbf{S}_1^{(k)}\underline{\hat{\mathbf{a}}}^{(k+1)}$ and $\mathbf{S}_0^{(k)}\underline{\hat{\mathbf{a}}}^{(k+1)}$. So, we have shown that DETECT works for an $(L-k)$-layer receiver based on the assumption that it works for an $(L-k-1)$-layer receiver. For equivalent 1-layer receivers in Fig. 2.5 which the splits shall eventually break down to, line (i), (ii), (v), (vi), and (vii) decode $\mathbf{S}_i^{(L-2)}\underline{\hat{\mathbf{a}}}$ and $\mathbf{S}_i^{(L-2)}\underline{\hat{\mathbf{a}}}^{(L-1)}$ sequentially from $\mathbf{S}_i^{(L-2)}\underline{\mathbf{y}}'^{(L-1)}$ just like a GMD MMSE receiver [18], [19]. By induction, we can conclude that Algorithm 2.5.1 can successfully decode $\hat{\mathbf{a}}$ in Fig. 2.3.

2.5.3 Complexity

The proposed GGMD DFE transceiver is applicable to general MIMO channels and has less or equal *design* and *implementation complexity* as compared to the GMD MMSE transceiver. However,

Table 2.2: Implementation complexity (K is a power of 2)

	GMD Transceiver	GGMD Transceiver
$\mathbf{P}$	$O(K^2)$	$O(K\log_2(K))$
$\mathbf{\Sigma}_f$	$O(K)$	$O(K)$
$\mathbf{V}_h$	$O(K\log_2(K))$	$O(K\log_2(K))$
$\mathbf{G}$	$O(K^2)$	$O(K\log_2(K))$
$\mathbf{B}$	$O(K^2)$	
$\mathbf{A}_i$, $\mathbf{P}^{(i)}\otimes\mathbf{I}_{2^i}$, and $\mathbf{B}^{(i)}\otimes\mathbf{I}_{2^i}$		$O(K\log_2(K))$
Total	$O(K^2)$	$O(K\log_2(K))$

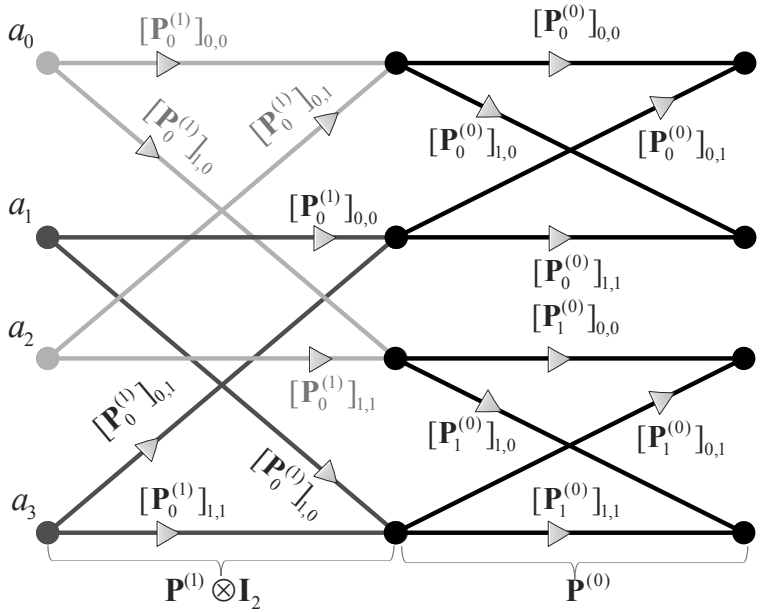


Figure 2.7: Structure of $\mathbf{P}$ for $K = 4$

the *design complexity* is dominated by the SVD step, which takes $O(JNM)$ [102]; the *implementation complexity* is dominated by the full matrices $\mathbf{V}_h$ and $\mathbf{U}_h^\dagger$ in Fig. 2.2 and 2.3. So, the complexity gain of a GGMD transceiver over a GMD transceiver is not obvious. It becomes prominent when we apply it to CP systems. So, in the following paragraphs, we will only discuss the complexity of the GGMD transceiver for CP systems and comparisons will be made with the GMD transceiver.

In CP systems, the number of subchannels, M, is often chosen to be a power of 2 which allows efficient implementation, i.e., $M = K = 2^L$. Here, we also use this convention. The *design complexity* of both transceivers is shown in Table 2.1. From Sec. 2.3, we see that obtaining $\mathbf{\Sigma}_h$, $\mathbf{V}_h$ and $\mathbf{U}_h^\dagger$

takes $O(K\log_2(K))$ flops. Since $\mathbf{\Theta}$ is already a diagonal matrix with positive diagonal elements in decreasing order, the complexity of decomposing $\mathbf{\Theta}$ only comes from the triangularization parts of GGMD and GMD, respectively. According to Sec. 2.4.2, the GGMD has lower complexity, $O(K)$, than the complexity of GMD, $O(K^2)$. Hence, the total *design complexity* of a GGMD transceiver is $O(K\log_2(K))$ while that of a GMD transceiver is $O(K^2)$.

To compare the *implementation complexity*, we let these systems process one $K\times 1$ data vector $\mathbf{a}$ and compare the number of floating point operations. The results are summarized in Table 2.2. For the transmitter part, the complexity of the GMD transceiver is $O(K^2)$ which comes from the $K\times K$ full-size precoder $\mathbf{P}$ in Fig. 2.1. The precoder $\mathbf{P}$ of the GGMD transmitter can be implemented efficiently by the butterfly structure illustrated in Fig. 2.7, which has complexity $O(K\log_2(K))$. The structure has $\log_2(K)$ levels of $\mathbf{P}^{(i)}\otimes\mathbf{I}_{2^i}$ and each level requires $2K$ multiplications and K additions. $\mathbf{V}_h$ can be implemented by fast Fourier transform with complexity $O(K\log_2(K))$.

Next, we compare the receivers. For the GMD transceiver, the feedforward matrix $\mathbf{G}$ has complexity $O(K^2)$ and the successive cancelation algorithm for implementing [17] the feedback matrix $\mathbf{B}$ has complexity $O(K^2)$. So, the total *implementation complexity* of the GMD transceiver is $O(K^2)$. For the GGMD receiver, the feedforward matrix $\mathbf{G}$ has the complexity $O(K\log_2(K))$ since $\mathbf{U}_h$ is virtually a DFT matrix whereas $\mathbf{A}_0$ is a block diagonal matrix consisting of $(K/2)$ 2×2 submatrices. The iterative detection algorithm, Algorithm 2.5.1, is a *divide-n-conquer* algorithm [103] and the number of flops required for the detection of an $L-k$-layer in Fig. 2.4 can be described by the recurrence

$$\Phi(L-k)=\begin{cases} O(1), & k=L-1,\\ 2\Phi(L-k-1)+O(2^{L-k}), & 0\le k<L-1,\end{cases} \tag{2.51}$$

where $O(1)$ is *base case* complexity for the detection of 1-layer receiver in Fig. 2.5, $2\Phi(L-k-1)$ is the complexity of conquering subproblems, and $O(2^{L-k})$ is from the *dividing* and *combining* steps. Lines (i)-(iii) are the dividing steps and Line (vi) is the combining step of the *divide-n-conquer* algorithm and the number of flops is determined by $\mathbf{B}^{(k)}$ and $\mathbf{P}^{(k+1)}$. Since $\mathbf{P}^{(k+1)}$ and $\mathbf{B}^{(k)}$ are block diagonal matrices with 2×2 submatrices on their diagonals, the *dividing* and *combining* steps take $O(2^{L-k})$ flops. Solving the recurrence (2.51) [103], we have the complexity of iterative detection algorithm

$$\Phi(L)=O(K\log_2 K).$$

One can also verify that the cost of multiplications and additions to implement the iterative detection is approximately equal to implementing the matrices in Fig. 2.3 besides $\mathbf{G}$. Since $\mathbf{P}^{(i)}$ and $\mathbf{B}^{(i)}$ are block diagonal matrices with 2×2 submatrices on their diagonals, implementing $\mathbf{A}_i$, $\mathbf{P}^{(i)} \otimes \mathbf{I}_{2^i}$ and $\mathbf{B}^{(i)} \otimes \mathbf{I}_{2^i}$ requires complexity $O(K)$. And $\log_2(K)$ sets of them add up the complexity to $O(K \log_2(K))$. Hence, we can conclude that the *implementation complexity* of a GGMD transceiver is $O(K \log_2(K))$, which is within the same order as the DFT/IDFT matrix, and $K/\log_2(K)$ times smaller than the *implementation complexity* of GMD transceiver.

2.6 Conclusions

In this chapter, we have proposed a new matrix decomposition GGMD which decomposes a complex matrix as a product of several sets of semi-unitary matrices and upper triangular matrices in an iterative manner. The optimal GGMD parameters which minimize the number of flops are derived. Along with the snovel iterative detection algorithm, we have shown that the GGMD can be used to design the GGMD DFE transceiver for MIMO channels. We found that the GGMD DFE transceiver is the most useful for CP systems since its *design* and *implementation complexity* are much less than that of a GMD MMSE transceiver. In the next chapter, we will analyze the performance of a GGMD transceiver and show that it reaches the same optimality that a GMD MMSE transceiver can possibly achieve.

2.7 Appendix

2.7.1 Proof of Lemma 2.1

Since the SVD steps of both decompositions have the same complexity, we only consider the complexity of the triangularization steps. By (2.28), the difference of complexity between $(l_0, \ldots, l_j, \ldots, l_{L-1})$-GGMD and $(l_0, \ldots, m_0, m_1, \ldots, l_{L-1})$-GGMD is given by

$$\begin{aligned} d &= \big(T(m_0)m_1 + T(m_1) - T(l_j)\big)\Big(\prod_{n=j+1}^{L-1} l_n\Big) \\ &= \frac{\prod_{n=j+1}^{L-1} l_n}{m_1}(7m_1^3 - 17m_1^2 - 7l_j^2 m_1 + 7l_j^2). \end{aligned} \tag{2.52}$$

The second equality follows from the direct substitution of $T(\cdot)$ from (2.29). By the assumptions $m_0, m_1 \geq 2$ and $m_0 m_1 = l_j$, we have

$$2 \leq m_1 \leq l_j/2. \tag{2.53}$$

Consider the function

$$g(x) = 7x^3 - 17x^2 - 7l_j^2 x + 7l_j^2, \tag{2.54}$$

for $x \in \{\mathcal{R} : 2 \leq x \leq l_j/2\}$. Its second derivative is given by

$$\frac{d^2 g}{dx^2} = 42x - 34 > 0.$$

The last inequality follows from $x \geq 2$. Hence, $g(x)$ is a convex function for $x \in \{\mathcal{R} : 2 \leq x \leq l_j/2\}$. Now, the values of $g(x)$ at boundary points are given by

$$\begin{aligned} g(2) &= -12 - 7l_j^2 < 0, \\ g(l_j/2) &= l_j^2(-\frac{21}{8} l_j + \frac{22}{8}) < 0. \end{aligned}$$

One can write $m_1 = 2\alpha + (1-\alpha) l_j/2$ for some $\alpha \in \{\mathcal{R}_+ : 0 \leq \alpha \leq 1\}$. Because $g(x)$ is a convex function, by Jensen's inequality [104], we have

$$g(m_1) \leq \alpha g(2) + (1-\alpha) g(l_j/2) < 0.$$

By substituting $g(m_1)$ back in (2.52), we have $d < 0$ and this concludes the proof.

2.7.2 Proof of Lemma 2.2

By (2.28), the difference of complexity between $(\ldots, l_j, l_{j+1}, \ldots)$-GGMD and $(\ldots, l_{j+1}, l_j, \ldots)$-GGMD is given by

$$\begin{aligned} d &= \Big(T(l_j) l_{j+1} + T(l_{j+1}) - T(l_{j+1}) l_j - T(l_j)\Big) N_{j+2} \\ &= \frac{N_{j+2}(l_{j+1} - l_j)}{l_j} \big(7l_j^2 - (7\beta + 17) l_j + 7\beta\big), \end{aligned} \tag{2.55}$$

where N_{j+2} is given by (2.21), $\beta = l_j l_{j+1}$ and the last equality is from the direct substitution of $T(\cdot)$ in (2.29). Consider the function

$$h(x) = 7x^2 - (7\beta + 17)x + 7\beta,$$

which is a convex function in x. Now,

$$\begin{aligned} h(2) &= -7\beta - 6 < 0, \\ h(\frac{\beta}{2}) &= -\frac{7}{4}\beta^2 - \frac{3}{2}\beta < 0. \end{aligned} \tag{2.56}$$

Since $2 \leq l_j \leq \beta/2$, one can express

$$l_j = 2\alpha + (1-\alpha)\frac{\beta}{2},$$

for some $0 \leq \alpha \leq 1$. Because $h(x)$ is a convex function, by Jensen's inequality [104] and (2.56), we have

$$h(l_j) = \alpha h(2) + (1-\alpha)h(\frac{\beta}{2}) < 0.$$

Therefore, $d < 0$ for $l_j < l_{j+1}$ and this proves the lemma.

Chapter 3

Performance Analysis of Generalized Geometric Mean Decomposition DFE Transceiver

In Chapter 2, the generalized geometric mean decomposition (GGMD) was proposed and used to design GGMD decision feedback equalizer (DFE) transceivers for arbitrary multi-input-multi-output (MIMO) channels without zero-forcing constraint. For the application to cyclic prefix (CP) systems, the GGMD DFE transceiver has the greatest advantage over the GMD DFE MMSE transceiver in terms of design and implementation complexity. This chapter presents the performance analysis for the GGMD DFE transceiver implementation proposed in Chapter 2. The arithmetic mean square error (MSE), symbol error rate (SER) and Gaussian mutual information of the proposed system are investigated. The performance advantages of the GGMD DFE transceiver over popular orthogonal frequency division multiplexing (OFDM) and single-carrier CP MMSE systems are shown analytically and verified by numerical simulations.

3.1 Introduction

In Chapter 2, we introduced a new matrix decomposition, generalized geometric mean decomposition (GGMD). We demonstrated the potential uses of GGMD in DFE transceiver design without zero-forcing constraint. Along with the proposed iterative detection algorithm, a GGMD DFE transceiver with lower *design* and *implementation complexity* than a GMD DFE MMSE transceiver [18], [19] can be constructed. The GGMD DFE transceiver is not just an implementation efficient version of a GMD DFE MMSE transceiver. It is a new DFE transceiver constructed from the newly

proposed GGMD and, as we shall explain in Sec. 3.4, it has a different input and output relationship from the conventional GMD transceiver. So, an in-depth investigation of its performance is needed.

The analysis in this chapter shows that a GGMD DFE transceiver achieves the same optimal arithmetic MSE as a GMD DFE MMSE transceiver [19] does. Then it shows that the proposed transceiver converts the MIMO channels into parallel equivalent subchannels and each subchannel has the same bias and equal signal to noise plus interference ratios (SINRs). Hence, no bit allocation is required to minimize the average bit error rate (BER) [35] or maximize the bitrate [25]. Moreover, we show that the proposed GGMD DFE transceiver can achieve MIMO channel capacity by independently applying the scalar Gaussian code of the same rate to each subchannel. The investigation also shows that the proposed transceiver minimizes the average symbol error rate (SER) in the moderate high signal to noise plus interference ratio (SINR) region.

In Chapter 2, it is shown that the *design* and *implementation* of a GGMD DFE transceiver for CP systems takes $O(K \log_2(K))$ flops, which is of the same order as the *implementation complexity* of a single carrier cyclic prefix (SC-CP) or an orthogonal frequency division multiplexing (OFDM) MMSE system. The analysis in this chapter explicitly shows the advantages of the proposed transceiver over OFDM and SC-CP MMSE systems in terms of arithmetic MSE and average SER. We show that a GGMD transceiver has lower SER than a SC-CP MMSE system [10] for all SINRs, and lower SER than an OFDM MMSE system in the moderate high SINR region. Numerical simulations are also conducted to verify the analytical results. Furthermore, the simulations demonstrate the robustness of GGMD DFE transceivers to channel spectral nulls while the average bit error rates (BERs) of SC-CP and OFDM MMSE systems degrade significantly. Most of the results in this chapter have been reported in our recent journal paper [28].

3.2 Outline

The sections in this chapter are structured as follows. In Section 3.3, we briefly review the performance of a GMD DFE MMSE transceiver and the design of a GGMD DFE transceiver [27]. In Section 3.4, we analyze the performance of the proposed transceiver in terms of arithmetic MSE, average SER and Gaussian mutual information. Analytical comparisons of the GGMD DFE transceiver, the GMD DFE MMSE transceiver, OFDM and SC-CP MMSE systems on their average MSEs and SERs are made in Section 3.5. Numerical simulations of BER performances are given in

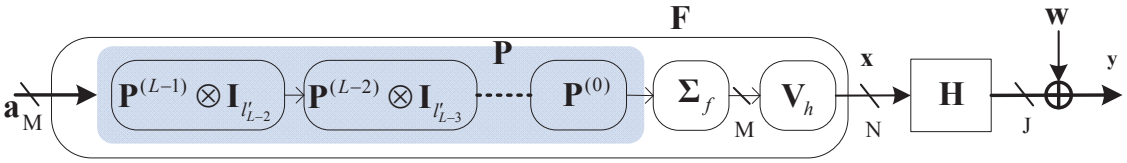


Figure 3.1: The transmitter of the GGMD DFE transceiver and the channel.

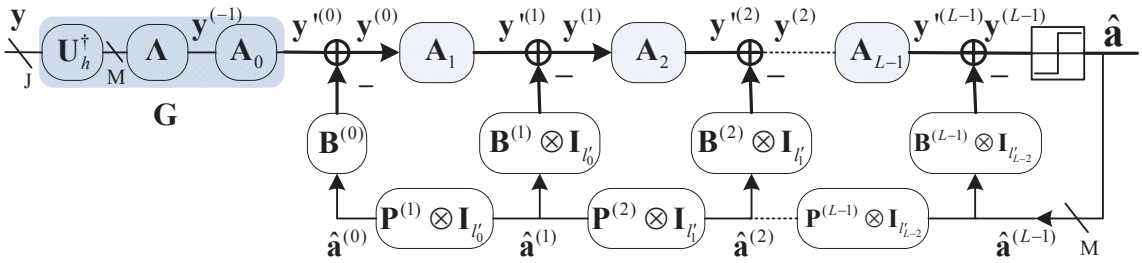


Figure 3.2: The receiver of the GGMD DFE transceiver.

Section 3.6. Concluding remarks are given in Section 3.7.

3.3 Preliminaries

In this section, we shall first review the performance of the conventional GMD DFE MMSE transceiver [18], [19] shown in Fig. 2.1 of Chapter 2. The error component vector in front of the detector is given by $\mathbf{e} = \mathbf{y} - \mathbf{a}$ where $\mathbf{y}$ and $\mathbf{a}$ are defined in Fig. 2.1 of Chapter 2. Under the no error propagation assumption, the error covariance matrix [19] can be written as

$$\mathbf{R}_{ee}^{gmd} = E(\mathbf{e}\mathbf{e}^\dagger) = \sigma_a^2(\mathbf{I} + \mathbf{B})\mathbf{P}^\dagger\mathbf{\Theta}^{-2}\mathbf{P}(\mathbf{I} + \mathbf{B})^\dagger, \tag{3.1}$$

where the design matrix

$$\mathbf{\Theta} = \big(\mathbf{I}_M + \frac{\sigma_a^2}{\sigma_w^2}\mathbf{\Sigma}_f^2\mathbf{\Sigma}_h^2\big)^{1/2}, \tag{3.2}$$

σ_a^2, σ_w^2, $\mathbf{\Sigma}_f$ and $\mathbf{\Sigma}_h$ are defined in Sec. 2.3 of Chapter 2. The $M \times M$ strictly upper triangular feedback matrix $\mathbf{B}$ and unitary precoding matrix $\mathbf{P}$ are obtained from the GMD of $\mathbf{\Theta}$, which is

$$\mathbf{\Theta} = \mathbf{Q}\Big(\prod_{k=0}^{M-1}\sigma_i\Big)^{1/M}(\mathbf{I}_M + \mathbf{B})\mathbf{P}^\dagger, \tag{3.3}$$

where $\sigma_k = [\mathbf{\Theta}]_{k,k}$. Using (3.3), (3.1) can be reduced to

$$\mathbf{R}_{ee}^{gmd} = E(\mathbf{e}\mathbf{e}^{\dagger}) = \sigma_a^2 \Big(\prod_{k=0}^{M-1} \sigma_k^{-2} \Big)^{1/M} \mathbf{I}_M. \tag{3.4}$$

One can see that $\mathbf{R}_{ee}^{gmd}$ is determined by the diagonal elements of the upper triangular matrix from the GMD of $\mathbf{\Theta}$ in (3.3). For square QAM [94], the average uncoded SER of a GMD MMSE transceiver [19] is

$$P_{gmd} = c_0 Q(c_1 \sqrt{\gamma_{gmd} - 1}), \tag{3.5}$$

where

$$\gamma_{gmd} = \frac{\sigma_a^2}{[\mathbf{R}_{ee}^{gmd}]_{k,k}}, \tag{3.6}$$

c_0 and c_1 depend on the number of bits b associated with each QAM symbol as

$$c_0 = 4(1 - 2^{-b/2}) \text{ and } c_1 = \sqrt{3/(2^b - 1)}. \tag{3.7}$$

Fig. 3.1 and Fig. 3.2 show the transmitter and receiver of the proposed GGMD DFE transceiver, respectively. The details of the design are given in Chapter 2. The feedback matrices $\mathbf{B}^{(i)}$ and precoding matrices $\mathbf{P}^{(i)}$ are obtained from the *optimal* GGMD of the design matrix $\mathbf{\Theta}$ in (3.2),

$$\mathbf{\Sigma}^{(i)} = \mathbf{Q}^{(i)} (\mathbf{\Sigma}^{(i+1)} \otimes \mathbf{I}_{l_i}) (\mathbf{I} + \mathbf{B}^{(i)}) \mathbf{P}^{(i)\dagger}, \tag{3.8}$$

where $\mathbf{\Sigma}^{(0)} = \mathbf{\Theta}$, l_i are GGMD parameters defined in Chapter 2 and $0 \leq i \leq L-1$. The matrices $\mathbf{A}_i$ are defined as

$$\mathbf{A}_i = (\mathbf{I} + \mathbf{B}^{(i)}) \mathbf{P}^{(i)\dagger} \otimes \mathbf{I}_{l'_{i-1}}, \tag{3.9}$$

where $l'_i = \prod_{n=0}^{i} l_n$.

As mentioned in Chapter 2, the *optimal* GGMD and GMD, in general, are two different kinds of decomposition for non-prime M since GGMD cannot be written as the product of $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{P}^{\dagger}$, where the upper triangular matrix $\mathbf{R}$ has diagonal elements equal to the geometric mean of the singular values of the matrix it is decomposing. Moreover, the GGMD DFE transceiver has a

completely different receiver structure than that of a GMD DFE MMSE transceiver. One cannot compute the error covariance matrix of the GGMD DFE transceiver using results from (3.1), (3.3) and (3.4). It is not clear how the error covariance matrix depends on the matrices obtained from GGMD in (3.8). Hence, we shall carefully analyze the performance of the proposed GGMD DFE transceiver based on the new receiver structure in Fig. 3.2 and the *optimal* GGMD in (3.8).

3.4 Performance Analysis of GGMD DFE Transceiver

In this section, we derive the input-output relationship of the GGMD DFE transceiver in Fig. 3.1 and 3.2. And then, we will analyze the performance of the proposed GGMD transceiver in terms of total MSE, uncoded average SER and Gaussian mutual information. All the analysis results are based on the assumption that there is no error propagation in the decision feedback loop, i.e. $\hat{\mathbf{a}} = \mathbf{a}$, which is a standard assumption for performance analysis of DFE equalizers [18], [19], [26] and [17].

For notational convenience, we define

$$\begin{aligned}
\mathbf{A}'(m) &= \prod_{i=0}^{L-1-m} \mathbf{A}_{L-1-i}, \\
\mathbf{P}'(m) &= \prod_{i=m}^{L-1} \mathbf{P}^{(i)} \otimes \mathbf{I}_{l'_{i-1}}, \\
\mathbf{Q}'(m) &= \prod_{i=m}^{L-1} \mathbf{Q}^{(i)} \otimes \mathbf{I}_{l'_{i-1}},
\end{aligned} \tag{3.10}$$

$\mathbf{A}'(L) = \mathbf{P}'(L) = \mathbf{Q}'(L) = \mathbf{I}_M$, where $0 \leq m \leq L-1$ and l'_{i-1} is defined in (3.9). So, the GGMD in (3.8) can be written as

$$\mathbf{\Sigma}^{(0)} = \mathbf{Q}'(0)\Sigma^{(L)}\mathbf{A}'(0). \tag{3.11}$$

In Fig. 3.1 and 3.2, the input-output relationship of the GGMD transceiver (from input vector $\mathbf{a}$ to the signal vector $\mathbf{y}^{(L-1)}$), has the form

$$\mathbf{y}^{(L-1)} = \mathbf{r} = \mathbf{T}_{ggmd}\mathbf{a} + \mathbf{w}', \tag{3.12}$$

where the equivalent noise $\mathbf{w}' = \mathbf{A}'(0)\mathbf{\Lambda}\mathbf{U}_h^\dagger\mathbf{w}$ and the equivalent channel matrix $\mathbf{T}_{ggmd}$ is given by

$$\mathbf{T}_{ggmd} = \Big(\mathbf{A}'(0)\mathbf{\Lambda}\mathbf{\Sigma}_h\mathbf{\Sigma}_f\mathbf{P}'(0) - \sum_{n=1}^{L}\mathbf{A}'(n)(\mathbf{B}^{(n-1)} \otimes \mathbf{I}_{l'_{n-2}})\mathbf{P}'(n)\Big). \tag{3.13}$$

For the GMD transceiver in Fig. 2.1 of Chapter 2, the input-output relationship can be expressed as

$$\mathbf{y} = \mathbf{T}_{gmd}\mathbf{a} + (\mathbf{I}_M + \mathbf{B})\mathbf{P}^\dagger\mathbf{\Lambda}\mathbf{U}_h^\dagger\mathbf{w}, \tag{3.14}$$

where the equivalent channel matrix

$$\mathbf{T}_{gmd} = (\mathbf{I}_M + \mathbf{B})\mathbf{P}^\dagger\mathbf{\Lambda}\mathbf{\Sigma}_h\mathbf{\Sigma}_f\mathbf{P} - \mathbf{B}. \tag{3.15}$$

Since $\mathbf{A}'(m)$, $\mathbf{P}'^{(m)}$ and $\mathbf{B}^{(m)} \otimes \mathbf{I}_{l'_{m-1}}$ in (3.12) are obtained from the GGMD as in (3.11), they all have special parallel butterfly structures as shown in Fig. 2.7 of Chapter 2, while $\mathbf{P}$ and $\mathbf{B}$ obtained from GMD do not have such structures. So, in general, $\mathbf{A}'(0)$ in (3.12) is not equal to $(\mathbf{I} + \mathbf{B})\mathbf{P}^\dagger$ in (3.14), which means that the GGMD and GMD receivers have different noise components in front of the detector even for the same channel noise $\mathbf{w}$. Moreover, $\mathbf{T}_{ggmd}$ is generally not equal to $\mathbf{T}_{gmd}$. GGMD and GMD transceivers have different equivalent channels and noise components in front of the detectors, unless M is prime number so that GGMD reduces to GMD. Therefore, we cannot expect both transceivers to output the same $\hat{\mathbf{a}}$ even for the same input data vector $\mathbf{a}$ and channel noise $\mathbf{w}$. The GGMD DFE transceiver is essentially a different transceiver from the GMD DFE MMSE transceiver. Hence, a thorough investigation of its performance is needed to verify the usefulness of a GGMD transceiver.

3.4.1 Mean Square Error

To understand the mean square error of the proposed transceiver in Fig. 3.1 and Fig. 3.2, we shall first analyze the error component in front of the detector,

$$\mathbf{e} = \mathbf{y}^{(L-1)} - \mathbf{a}, \tag{3.16}$$

and then derive the error covariance matrix. From Fig. 3.2, we have

$$\mathbf{y}^{(m)} = \mathbf{A}_m \mathbf{y}^{(m-1)} - (\mathbf{B}^{(m)} \otimes \mathbf{I}_{l'_{m-1}}) \mathbf{P}'(m+1)\mathbf{a}, \tag{3.17}$$

for $0 \leq m \leq L-1$ where $\mathbf{P}'(m+1)$ is given by (3.10). Using this relation of $\mathbf{y}^{(m)}$ for consecutive m, we have the following lemma for the error component.

Lemma 3.1: The error component $\mathbf{e}$ defined as in (3.16) can be expressed as

$$\mathbf{e} = \mathbf{A}'(m)\Big(\mathbf{y}^{(m-1)} - \mathbf{P}'(m)\mathbf{a}\Big), \tag{3.18}$$

for $1 \leq m \leq L-1$. □

Proof: See appendix. □

From Fig. 3.2, it can be shown that

$$\mathbf{y}^{(0)} = \mathbf{GHFa} - \mathbf{B}^{(0)}\mathbf{P}'(1)\mathbf{a} + \mathbf{Gw}. \tag{3.19}$$

The channel matrix $\mathbf{H}$ is decomposed by SVD in Chapter 2 as

$$\mathbf{H} = \mathbf{U}_h \mathbf{\Sigma}_h \mathbf{V}_h^{\dagger}, \tag{3.20}$$

where $\mathbf{\Sigma}_h$ is an $M \times M$ diagonal matrix, $\mathbf{U}_h$ is a $J \times M$ semi-unitary matrix and $\mathbf{V}_h$ is an $N \times M$ unitary matrix. Substituting (3.19) and (3.20) into (3.18) for $m = 1$ yields

$$\mathbf{e} = \mathbf{A}'(1)\Big(\big(\mathbf{G}\mathbf{U}_h\mathbf{\Sigma}_h\mathbf{\Sigma}_f\mathbf{P}^{(0)} - \mathbf{I} - \mathbf{B}^{(0)}\big)\mathbf{P}'(1)\mathbf{a} + \mathbf{Gw}\Big). \tag{3.21}$$

Following similar procedures as in Sec. 19.C [91], the error covariance matrix can be shown to be

$$\begin{aligned} \mathbf{R}_{ee}^{ggmd} &= E[\mathbf{e}\mathbf{e}^{\dagger}] \\ &= \sigma_a^2 \mathbf{A}'(0)\mathbf{\Sigma}^{(0)^{-2}}\mathbf{A}'(0)^{\dagger}, \end{aligned} \tag{3.22}$$

where $\mathbf{A}'(0)$ is given by (3.10) and $\mathbf{\Sigma}^{(0)}$ is given by (3.11). Applying (3.11) to (3.22) , one can finally

have

$$\mathbf{R}_{ee}^{ggmd} = \sigma_a^2 \Big(\prod_{k=0}^{M-1} \sigma_k^{-2} \Big)^{1/M} \mathbf{I}_M, \tag{3.23}$$

where σ_k is given in (3.3). From the expression above, the error covariance matrix $\mathbf{R}_{ee}^{ggmd}$ here is identified to be exactly the same as that of the GMD MMSE transceiver given in (3.4). The total MSEs of both the GGMD and GMD DFE transceivers over a data block are given by

$$\xi_{ggmd} = \xi_{gmd} = Tr(\mathbf{R}_{ee}^{ggmd}) = M\sigma_a^2 \Big(\prod_{k=0}^{M-1} \sigma_k^{-2} \Big)^{1/M}. \tag{3.24}$$

This implies that the GGMD transceiver is optimal in terms of arithmetic MSE, just as the GMD MMSE transceiver.

3.4.2 Bias of the Detector

In practice, whenever MMSE receivers are used, the bias is estimated and removed before the symbol detection. A careful treatment of bias is required before we can analyze the uncoded average SER performance of the proposed GGMD DFE transceiver. The GGMD iterative equations in (3.8) can be rewritten as

$$\mathbf{\Sigma}^{(i)} = \mathbf{Q}^{(i)} \mathbf{D}^{(i)} \mathbf{B}_1^{(i)} \mathbf{P}^{(i)\dagger}, \tag{3.25}$$

where $\mathbf{B}_1^{(i)} = \mathbf{I} + \mathbf{B}^{(i)}$ and

$$\mathbf{D}^{(i)} = \mathbf{\Sigma}^{(i+1)} \otimes \mathbf{I}_{l_i}. \tag{3.26}$$

The following two lemmas shall be useful in deriving the bias of the GGMD DFE receiver.

Lemma 3.2: The product $\mathbf{\Sigma}_f \mathbf{\Sigma}_h \mathbf{\Lambda}$ can be reduced to

$$\mathbf{\Sigma}_f \mathbf{\Sigma}_h \mathbf{\Lambda} = (\mathbf{I} - \mathbf{P}^{(0)} {\mathbf{B}_1^{(0)}}^{-1} {\mathbf{D}^{(0)}}^{-2} {\mathbf{B}_1^{(0)}}^{-\dagger} \mathbf{P}^{(0)\dagger}). \tag{3.27}$$

□

Proof: See appendix. □

Lemma 3.3: Suppose

$$\mathbf{E} = (\mathbf{A} \otimes \mathbf{I}_l)\Big(\sum_{m=1}^{l-1} \mathbf{D}_m \mathbf{S}^m\Big)(\mathbf{B} \otimes \mathbf{I}_l),$$

where $\mathbf{D}_m$ are any $nl \times nl$ diagonal matrices, $\mathbf{A}$ and $\mathbf{B}$ are any $n \times n$ complex matrices, and $\mathbf{S}$ is a $ln \times ln$ circular shift operator of the form

$$\mathbf{S} = \begin{pmatrix} 0 & 0 & \cdots & 1 \\ 1 & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}. \tag{3.28}$$

Then, the diagonal elements of $\mathbf{E}$ are all zeros, i.e., $E_{ii} = 0$. □

Proof: See appendix. □

To determine the bias of the detector, we shall re-express each component of $\mathbf{r}$ in (3.12) as

$$r_i = \alpha_i a_i + \tau_i, \tag{3.29}$$

where $0 \leq i \leq M-1$, α_i is the bias, τ_i is the noise plus interference component such that $E(a_i \tau_i^*) = 0$. The following theorem will be helpful in reducing (3.12) to (3.29).

Theorem 3.1: The received signal $\mathbf{r}$ in (3.12) can be expressed as

$$\begin{aligned} \mathbf{r} = {} & \mathbf{A}'(m)\Big(\mathbf{I} - \mathbf{D}^{(m-1)^{-2}} \mathbf{B}_1^{(m-1)^{-\dagger}} \otimes \mathbf{I}_{l'_{m-2}}\Big)\mathbf{P}'(m)\mathbf{a} \\ & - \sum_{n=1+m}^{L} \mathbf{A}'(n)\Big(\mathbf{B}^{(n-1)} \otimes \mathbf{I}_{l'_{n-2}}\Big)\mathbf{P}'(n)\mathbf{a} - \sum_{n=2}^{m} \tilde{\mathbf{E}}_n \mathbf{a} + \mathbf{w}', \end{aligned} \tag{3.30}$$

where

$$\begin{aligned} \tilde{\mathbf{E}}_n = {} & \mathbf{A}'(n)\Big\{\Big(\mathbf{D}^{(n-1)^{-2}} \mathbf{B}_1^{(n-1)^{-\dagger}} \mathbf{P}^{(n-1)\dagger} \otimes \mathbf{I}_{l_{n-2}}\Big) \\ & \cdot \mathbf{L}^{(n-2)} \otimes \mathbf{I}_{l'_{n-3}}\Big\}\mathbf{P}'(n-1), \end{aligned} \tag{3.31}$$

$$\mathbf{L}^{(n)} = \mathbf{B}_1^{(n)^{-\dagger}} - \mathbf{I}, \tag{3.32}$$

and $1 \leq m \leq L$. □

Proof: Using the result from Lemma 3.2 and (3.12), $\mathbf{r}$ can be written as

$$\begin{aligned}\mathbf{r} = \mathbf{A}'(1)\Big(\mathbf{I} - {\mathbf{D}^{(0)}}^{-2}{\mathbf{B}_1^{(0)}}^{-\dagger}\Big)\mathbf{P}'(1)\mathbf{a} \\ - \sum_{n=2}^{L}\mathbf{A}'(n)\Big(\mathbf{B}^{(n-1)} \otimes \mathbf{I}_{l'_{n-2}}\Big)\mathbf{P}'(n)\mathbf{a} + \mathbf{w}'.\end{aligned} \tag{3.33}$$

So, (3.30) holds for $m = 1$. Suppose (3.30) holds for $m = k$, i.e.,

$$\begin{aligned}\mathbf{r} &= \mathbf{A}'(k)\Big(\mathbf{I} - {\mathbf{D}^{(k-1)}}^{-2}{\mathbf{B}_1^{(k-1)}}^{-\dagger} \otimes \mathbf{I}_{l'_{k-2}}\Big)\mathbf{P}'(k)\mathbf{a} \\ &- \sum_{n=1+k}^{L}\mathbf{A}'(n)\Big(\mathbf{B}^{(n-1)} \otimes \mathbf{I}_{l'_{n-2}}\Big)\mathbf{P}'(n)\mathbf{a} - \sum_{n=2}^{k}\tilde{\mathbf{E}}_n\mathbf{a} + \mathbf{w}' \\ &= \mathbf{A}'(k)\Big[\mathbf{I} - (\mathbf{P}^{(k)}{\mathbf{B}_1^{(k)}}^{-1}{\mathbf{D}^{(k)}}^{-2}{\mathbf{B}_1^{(k)}}^{-\dagger}\mathbf{P}^{(k)\dagger} \otimes \mathbf{I}_{l_{k-1}})(\mathbf{I} \\ &+ \mathbf{L}^{(k-1)}) \otimes \mathbf{I}_{l'_{k-2}}\Big] \cdot \mathbf{P}'(k)\mathbf{a} - \sum_{n=2}^{k}\tilde{\mathbf{E}}_n\mathbf{a} + \mathbf{w}' \\ &- \sum_{n=1+k}^{L}\mathbf{A}'(n)\Big(\mathbf{B}^{(n-1)} \otimes \mathbf{I}_{l'_{n-2}}\Big)\mathbf{P}'(n)\mathbf{a} \\ &= \mathbf{A}'(k+1)\Big(\mathbf{I} - {\mathbf{D}^{(k)}}^{-2}{\mathbf{B}_1^{(k)}}^{-\dagger} \otimes \mathbf{I}_{l'_{k-1}}\Big)\mathbf{P}'(k+1)\mathbf{a} \\ &- \sum_{n=2+k}^{L}\mathbf{A}'(n)\Big(\mathbf{B}^{(n-1)} \otimes \mathbf{I}_{l'_{n-2}}\Big)\mathbf{P}'(n)\mathbf{a} - \sum_{n=2}^{k+1}\tilde{\mathbf{E}}_n\mathbf{a} + \mathbf{w}'.\end{aligned}$$

The second equality follows from (3.25) and (3.26) and the third equality is obtained by factoring $\mathbf{A}'(k)$ as $\mathbf{A}'(k) = \mathbf{A}'(k+1)\mathbf{A}_k$. So, (3.30) holds for $m = k+1$. Therefore, we have proven the theorem by induction. □

In particular for $m = L$, Theorem 3.1 reduces to

$$\begin{aligned}\mathbf{r} =& (\mathbf{I} - {\mathbf{D}^{(L-1)}}^{-2} \otimes \mathbf{I}_{l'_{L-2}})\mathbf{a} - \sum_{n=2}^{L}\tilde{\mathbf{E}}_n\mathbf{a} \\ &- ({\mathbf{D}^{(L-1)}}^{-2}\mathbf{L}^{(L-1)} \otimes \mathbf{I}_{l'_{L-2}})\mathbf{a} + \mathbf{w}'.\end{aligned} \tag{3.34}$$

Note that ${\mathbf{D}^{(L-1)}}^{-2}\mathbf{L}^{(L-1)} \otimes \mathbf{I}_{l'_{L-2}}$ is a strictly lower triangular matrix. The following lemma shows that $\tilde{\mathbf{E}}_n$ have zero diagonal elements. Along with the elements a_i in $\mathbf{a}$ being uncorrelated, (3.34) can be written as (3.29). The component $\alpha_i a_i$ in r_i only comes from the first term in (3.34). By (3.8),

(3.26), we have

$$\mathbf{D}^{(L-1)} = \Big(\prod_{k=0}^{M-1} \sigma_k\Big)^{1/M} \mathbf{I}_{l_{L-1}}. \tag{3.35}$$

Hence, the bias α_i is given by

$$\alpha = \alpha_i = 1 - \Big(\prod_{k=0}^{M-1} \frac{1}{\sigma_k^2}\Big)^{1/M}. \tag{3.36}$$

All the equivalent subchannels of a GGMD transceiver have the same bias α.

Lemma 3.4: The matrices $\tilde{\mathbf{E}}_n$ in (3.31) have zero diagonal elements for $n = 2, 3, \ldots, L$. □

Proof: See appendix. □

3.4.3 Symbol Error Rate Performance

For square QAM modulations [94], the uncoded SER of the ith subchannel of GGMD DFE transceiver is

$$P_{ggmd,i} = c_0 Q(c_1 \sqrt{\gamma_{ggmd,br,i}}), \tag{3.37}$$

where c_0 and c_1 are given by (3.7) and $\gamma_{ggmd,br,i}$ is the signal to noise plus interference ratio (SINR) in the ith subchannel of the detector after bias removal. Given the type of modulation, the SER is completely determined by the SINR $\gamma_{ggmd,br,i}$. To compute the $\gamma_{ggmd,br,i}$, we need to determine the covariance of τ_i in (3.29). The ith component of $\mathbf{e}$ in (3.16) is given by

$$e_i = r_i - a_i = (\alpha - 1)a_i + \tau_i, \tag{3.38}$$

where the second equality follows by direct substitution of (3.29). Then we can compute the error covariance of e_i as

$$\begin{aligned}[\mathbf{R}_{ee}^{ggmd}]_{i,i} = [E(\mathbf{e}\mathbf{e}^\dagger)]_{i,i} &= (1-\alpha)^2\sigma_a^2 + [\mathbf{R}_{\tau\tau}]_{i,i} \\ &= (1-\alpha)\sigma_a^2,\end{aligned} \tag{3.39}$$

where last equality is from the result in (3.23). By algebraic manipulation, one can rewrite such that

$$[\mathbf{R}_{\tau\tau}]_{i,i} = \alpha(1-\alpha)\sigma_a^2. \tag{3.40}$$

The SINR for the ith subchannel is given by

$$\gamma_{ggmd,br,i} = \gamma_{ggmd,br} = \frac{\sigma_a^2}{[\mathbf{R}_{\tau\tau}]_{i,i}/\alpha^2} = \frac{\alpha}{1-\alpha}, \tag{3.41}$$

which is constant for all subchannels. Averaging over all M subchannels, we have the uncoded average SER of the GGMD DFE transceiver,

$$P_{ggmd} = \frac{c_0}{M}\sum_{i=0}^{M-1} Q(c_1\sqrt{\gamma_{ggmd,br,i}}) = c_0 Q(c_1\sqrt{\gamma_{ggmd,br}}). \tag{3.42}$$

The following is an important theorem connecting the average SER performances of GGMD and GMD DFE transceivers.

Theorem 3.2: Both GGMD and GMD DFE MMSE transceivers share the same average symbol error rate. □

Proof: The average uncoded SER of the GMD MMSE transceiver in [19] is given by (3.5). As shown earlier in this section, both the GGMD and GMD MMSE transceivers have the same error covariance matrices given by (3.23). So, by (3.39) and (3.6), we have

$$\gamma_{gmd} = \frac{\sigma_a^2}{(1-\alpha)\sigma_a^2} = \frac{1}{1-\alpha}. \tag{3.43}$$

Then, it follows that

$$\gamma_{ggmd,br} = \gamma_{gmd} - 1,$$

which also implies

$$P_{ggmd} = P_{gmd}. \tag{3.44}$$

□

By equalizing the SINRs of all subchannels, the GGMD DFE transceiver also minimizes the average SER in the moderate high SINR region [19].

Using the Gray encoding scheme [94], we can assume that there is only one bit error for each occurrence of the symbol error in the moderate high SNR region. The uncoded bit error rate (BER) of the ith subchannel of a GGMD DFE transceiver can be approximated by

$$P_{b,i}^{ggmd} \approx P_{ggmd,i}/b, \tag{3.45}$$

where b is given by (3.7). It follows that the average BER is given by

$$P_b^{ggmd} = \frac{1}{M}\sum_{i=0}^{M-1} P_{b,i}^{ggmd} = P_{ggmd}/b. \tag{3.46}$$

We use the fact $P_{ggmd} = P_{ggmd,i}$, from (3.37) and (3.42), for the second equality. By Theorem 3.2, we can conclude that both GMD and GGMD transceivers have the same average BER performance, i.e.,

$$P_b^{ggmd} = P_b^{gmd}. \tag{3.47}$$

3.4.4 Gaussian Mutual Information

The GGMD DFE MMSE transceiver converts the MIMO channel $\mathbf{H}$ into M equivalent parallel subchannels and each subchannel has equal SINR given by (3.41). The Gaussian mutual information of the GGMD DFE MMSE transceiver in Fig. 3.1 and 3.2 is

$$\begin{aligned} I(\mathbf{y}^{(L-1)};\mathbf{a}) &= M\log_2(1+\gamma_{ggmd,br}) \\ &= M\log_2(\frac{1}{1-\alpha}) \\ &= M\log_2\Big(\prod_{k=0}^{M-1}\sigma_k^2\Big)^{1/M}. \end{aligned} \tag{3.48}$$

The channel capacity of the MIMO channel $\mathbf{H}$ in Fig. 3.1 is as follows [98]:

$$C = \max_{Tr(\mathbf{R_{xx}})\le P_0} \log_2\det\Big(\mathbf{I}_N + \frac{\mathbf{H}\mathbf{R_{xx}}\mathbf{H}^\dagger}{\sigma_w^2}\Big), \tag{3.49}$$

where $\mathbf{R}_{\mathbf{xx}} = E(\mathbf{x}\mathbf{x}^{\dagger})$. Using properties of unitary matrices and waterfilling [91], the optimization problem in (3.49) can be solved and one can obtain the channel capacity

$$C = \log_2 \det\Big(\mathbf{I}_N + \frac{\sigma_a^2}{\sigma_w^2}\mathbf{\Sigma}_f^2\mathbf{\Sigma}_h^2\Big) = M\log_2\Big(\prod_{k=0}^{M-1}\sigma_k^2\Big)^{1/M}, \tag{3.50}$$

where $\mathbf{\Sigma}_h$ and $\mathbf{\Sigma}_f$ are given in Sec. 2.3 of Chapter 2. The last equality follows from (3.2) and (3.3). The analysis shows that

$$C = I(\mathbf{y}^{(L-1)}; \mathbf{a}),$$

which implies the GGMD DFE transceiver is *Gaussian mutual information lossless*. Along with the fact that all the equivalent subchannels of the GGMD transceiver have the same SINR, identical and independent Gaussian channel codes can be applied so that the proposed system achieves channel capacity. In other words, the GGMD DFE transceiver is an efficent capacity achieving structure for MIMO communications.

In summary, we have shown that the proposed GGMD DFE transceiver has the same performance as the GMD DFE MMSE transceiver in terms of arithmetic MSE, average SER and Gaussian mutual information for generic MIMO channels. To compare the performance with linear transceivers [91], one can just refer to the comparison results of the GMD DFE MMSE transceiver as in Ch. 19 [91].

3.5 Performance Comparison

In Chapter 2, we show that the *design* and *implementation* of a GGMD DFE MMSE transceiver for a CP system takes $O(K\log_2(K))$ flops, where K is the size of a data block and is a power of 2. Also, it is well known [106] that the *implementation complexity* of an OFDM or SC-CP MMSE system is $O(K\log_2(K))$. Without increasing the complexity too much, the proposed GGMD DFE transceiver will be a good candidate for CP systems if it demonstrates superior performance over an OFDM or SC-CP MMSE system. In this section, we compare the performance of the proposed GGMD DFE transceiver with the popular OFDM and SC-CP MMSE transceiver over CP systems. Here, the MIMO equivalent channel is a circulant matrix $\mathbf{H}_{cir}$. Comparisons in terms of average MSE and SER are made in the following two subsections.

3.5.1 Mean Square Error

To facilitate the performance comparisons with OFDM and SC-CP MMSE systems in which uniform power allocation is applied, we consider the versions of GGMD and GMD DFE MMSE transceiver with uniform power allocation (GGMDU and GMDU). The power allocation matrix takes the form

$$\mathbf{\Sigma}_f = \sqrt{P_0/M}\mathbf{I}_M. \tag{3.51}$$

The corresponding error covariance matrices are given by

$$\mathbf{R}_{ee}^{ggmdu} = \mathbf{R}_{ee}^{gmdu} = \sigma_a^2\Big(\prod_{k=0}^{M-1}\overline{\sigma}_k^{-2}\Big)^{1/M}\mathbf{I}_M, \tag{3.52}$$

where

$$\overline{\sigma}_k = \Big[\big(\mathbf{I}_M + \frac{\sigma_a^2 P_0}{\sigma_w^2 M}\mathbf{\Sigma}_h^2\big)^{1/2}\Big]_{k,k}, \tag{3.53}$$

and the total MSEs are

$$\xi_{ggmdu} = \xi_{gmdu} = M\sigma_a^2\Big(\prod_{k=0}^{M-1}\overline{\sigma}_k^{-2}\Big)^{1/M}. \tag{3.54}$$

It is clear from [19] that

$$\xi_{ggmdu} = \xi_{gmdu} \geq \xi_{ggmd} = \xi_{gmd}, \tag{3.55}$$

since the power loading matrix $\mathbf{\Sigma}_f$ in the GMD DFE MMSE transceiver is the optimal solution that minimizes the total MSE.

Both OFDM and SC-CP MMSE systems are special cases of the DFE MMSE transceiver in Fig. 3.1 and Fig. 3.2. A GGMD DFE transceiver reduces to an OFDM MMSE system if we make $\mathbf{P}^{(i)}$ and $\mathbf{A}_i$ identity matrices, $\mathbf{B}^{(i)}$ zero matrices, and $\mathbf{\Sigma}_f$ as in (3.51). So, its error covariance follows directly from (3.22) and is of the form

$$\mathbf{R}_{ee}^{ofdm} = \sigma_a^2\Big(\mathbf{I}_M + \frac{\sigma_a^2 P_0}{\sigma_w^2 M}\mathbf{\Sigma}_h^2\Big)^{-1}. \tag{3.56}$$

The total MSE is

$$\xi_{ofdm} = \sigma_a^2 \sum_{k=0}^{M-1} \frac{1}{\overline{\sigma}_k^2}, \tag{3.57}$$

where $\overline{\sigma}_k$ is given in (3.53). To make the GGMD DFE MMSE transceiver a SC-CP MMSE system, we let $\mathbf{P}^{(0)} = \mathbf{W}$ and $\mathbf{A}_0 = \mathbf{W}^\dagger$, where $\mathbf{W}$ is an $M \times M$ DFT matrix. All other $\mathbf{P}^{(i)}$ and $\mathbf{A}_i$ should be identity matrices, and $\mathbf{B}^{(i)}$ are zero matrices. Thus, the error covariance of a SC-CP MMSE system can be shown to be

$$\mathbf{R}_{ee}^{sccp} = \sigma_a^2 \mathbf{W}^\dagger \Big(\mathbf{I}_M + \frac{\sigma_a^2 P_0}{\sigma_w^2 M} \mathbf{\Sigma}_h^2 \Big)^{-1} \mathbf{W}. \tag{3.58}$$

One can verify that $\mathbf{R}_{ee}^{sccp}$ has equal diagonal elements [6], [7] where

$$[\mathbf{R}_{ee}^{sccp}]_{i,i} = \frac{\sigma_a^2}{M} \sum_{k=0}^{M-1} \frac{1}{\overline{\sigma}_k^2}. \tag{3.59}$$

The total MSE of the SC-CP MMSE system is given by

$$\xi_{sccp} = \sigma_a^2 \sum_{k=0}^{M-1} \frac{1}{\overline{\sigma}_k^2}, \tag{3.60}$$

which is equal to ξ_{ofdm}.

By AM-GM inequality and (3.55), we can conclude that

$$\xi_{ofdm} = \xi_{sccp} \geq \xi_{ggmdu} = \xi_{gmdu} \geq \xi_{ggmd} = \xi_{gmd}. \tag{3.61}$$

3.5.2 Symbol Error Rate

For the GGMD and GMD DFE MMSE transceiver with uniform power allocation (GGMDU and GMDU), the average SERs are given by

$$P_{ggmdu} = P_{gmdu} = c_0 Q(c_1 \sqrt{\gamma_{ggmdu} - 1}), \tag{3.62}$$

where

$$\gamma_{ggmdu} = \frac{\sigma_a^2}{[\mathbf{R}_{ee}^{ggmdu}]_{i,i}} = \frac{M\sigma_a^2}{\xi_{ggmdu}},$$

and $\mathbf{R}_{ee}^{ggmdu}$ is given by (3.52). Since P_{ggmd} and P_{ggmdu} are increasing functions of MSE [11], [39], it follows from (3.55) that

$$P_{gmd} = P_{ggmd} \leq P_{gmdu} = P_{ggmdu}, \tag{3.63}$$

when $\gamma_{ggmdu} \geq 1$.

The comparisons of GGMD, OFDM and SC-CP MMSE transceivers in terms of SERs are summarized in Theorem 3.3. Before that, we define SINRs of the ith subchannel for GGMD, GGMDU, OFDM, SC-CP MMSE transceivers as

$$\begin{aligned} \Gamma_i^{ggmd} &= \frac{\sigma_a^2}{[\mathbf{R}_{ee}^{ggmd}]_{i,i}},\ \Gamma_i^{ggmdu} = \frac{\sigma_a^2}{[\mathbf{R}_{ee}^{ggmdu}]_{i,i}}, \\ \Gamma_i^{ofdm} &= \frac{\sigma_a^2}{[\mathbf{R}_{ee}^{ofdm}]_{i,i}},\ \Gamma_i^{sccp} = \frac{\sigma_a^2}{[\mathbf{R}_{ee}^{sccp}]_{i,i}}, \end{aligned} \tag{3.64}$$

respectively. In what follows "convex region" means the SINRs are such that

$$\Gamma_i^{ggmdu}, \Gamma_i^{ofdm}, \Gamma_i^{sccp} \in \{\mathcal{R}_{cvx} \bigcap \hat{\mathcal{R}}_{cvx}\}, \tag{3.65}$$

and "concave region" means

$$\Gamma_i^{ggmdu}, \Gamma_i^{ofdm}, \Gamma_i^{sccp} \in \{\mathcal{R}_{ccv} \bigcap \hat{\mathcal{R}}_{ccv}\}. \tag{3.66}$$

The sets, $\mathcal{R}_{cvx}$, $\hat{\mathcal{R}}_{cvx}, \mathcal{R}_{ccv}$, and $\hat{\mathcal{R}}_{ccv}$, depend on the choice of modulation schemes and are defined in Sec. VI of [39]. To make this chapter self-contained, we summarize the case for square 2^b-QAM as follows: The number b only takes even integers. Let $\mathcal{R}_{++} = \{x \in \mathcal{R} : x \geq 1\}$. For $b = 2$, we define $\mathcal{R}_{cvx} = \{\Gamma \in \mathcal{R}_{++}\}$. For $b \geq 4$,

$$\mathcal{R}_{cvx} = \{\Gamma \in \mathcal{R}_{++} : \Gamma \leq t_1 \text{ or } \Gamma \geq t_2\}, \tag{3.67}$$

in which $t_1 \leq t_2$ and

$$t_1, t_2 = \frac{8}{K_1 + 3 \pm \sqrt{K_1^2 - 10K_1 + 9}},$$

$K_1 = 3/(2^b - 1)$, and $\mathcal{R}_{ccv} = \mathcal{R}_{++} \setminus \mathcal{R}_{cvx}$. Moreover, we define $\hat{\mathcal{R}}_{cvx} = \{\Gamma \in \mathcal{R}_{++}\}$ for $b = 2, 4$. For

$b \geq 6$,

$$\hat{\mathcal{R}}_{cvx} = \{\Gamma \in \mathcal{R}_{++} : \Gamma \leq t_1' \text{ or } \Gamma \geq t_2'\}, \tag{3.68}$$

in which $t_1' \leq t_2'$

$$t_1', t_2' = \frac{4}{K_1 + 1 \pm \sqrt{K_1^2 - 6K_1 + 1}},$$

and $\hat{\mathcal{R}}_{ccv} = \mathcal{R}_{++} \setminus \hat{\mathcal{R}}_{cvx}$. The convex region usually means the very low and moderate high SINR region and the concave region means medium SINR region. For QPSK modulation, the convex region contains all SINR values.

Theorem 3.3: Let P_{ofdm} be the average SER of the OFDM MMSE system, P_{sccp} be the average SER of the SC-CP MMSE system. Then,

$$P_{ggmdu} \leq P_{sccp} \leq P_{ofdm}, \text{ in the convex region,}$$
$$P_{ofdm} \leq P_{ggmdu} = P_{gmdu} \leq P_{sccp}, \text{ in the concave region,}$$
$$\text{and } P_{ggmd} \leq P_{ggmdu}.$$

□

Proof: The proof consists of four parts. Firstly, we show that $P_{ggmdu} \leq P_{sccp}$ for all SINRs. The second part is the proof that $P_{ggmdu} \leq P_{ofdm}$ in the convex region and the other way around in the concave region. The third part is the proof that $P_{sccp} \leq P_{ofdm}$ for $\Gamma_i^{sccp}, \Gamma_i^{ofdm} \in \mathcal{R}_{cvx}$ and the other way around for $\Gamma_i^{sccp}, \Gamma_i^{ofdm} \in \mathcal{R}_{ccv}$. This part had been proven in [7] and a similar proof can be found in Theorem 2 of [39]. The last part $P_{ggmd} \leq P_{ggmdu}$ follows from (3.63).

To prove the first part, we define SER function in terms of MSE x as

$$\psi(x) = c_0 Q(c_1 \sqrt{\frac{\sigma_a^2}{x} - 1}),$$

so that

$$P_{ggmdu} = \psi(\xi_{ggmdu}/M) \text{ and } P_{sccp} = \psi(\xi_{sccp}/M),$$

where ξ_{ggmdu} and ξ_{sccp} are given by (3.54) and (3.60), respectively. Since $\psi(x)$ is a monotone increasing function [11], [39], the fact that $\xi_{ggmdu} \leq \xi_{sccp}$ in (3.61) leads to $P_{ggmdu} \leq P_{sccp}$.

For the proof of second part, we define

$$g(\mathbf{y}) = \frac{1}{M} \sum_{k=0}^{M-1} c_0 Q(c_1 \sqrt{e^{2y_k} - 1}), \tag{3.69}$$

where $\mathbf{y}$ is a $M \times 1$ vector. So that by (3.52) and (3.56)

$$P_{ggmdu} = g(\mathbf{y}^{ggmdu}),$$
$$P_{ofdm} = g(\mathbf{y}^{ofdm}).$$

The elements of $\mathbf{y}^{ggmdu}$ and $\mathbf{y}^{ofdm}$ are given by

$$y_i^{ggmdu} = \log(\sqrt{\Gamma_i^{ggmdu}}) = \frac{1}{M} \sum_{k=0}^{M-1} \log \overline{\sigma}_k,$$
$$y_i^{ofdm} = \log(\sqrt{\Gamma_i^{ofdm}}) = \log \overline{\sigma}_i,$$

where $\overline{\sigma}_k$ are defined in (3.53). By the definition of majorization [95], one can verify that $\mathbf{y}^{ggmdu}$ is additively majorized by $\mathbf{y}^{ofdm}$, denoted as

$$\mathbf{y}^{ggmdu} \prec_+ \mathbf{y}^{ofdm}. \tag{3.70}$$

By Lemma 2 in [39] and Proposition C.1 in [95], it follows that $g(\mathbf{y})$ is Schur-convex when $e^{2y_i} \in \hat{\mathcal{R}}_{cvx}$, and Schur-concave when $e^{2y_i} \in \hat{\mathcal{R}}_{ccv}$. For $\Gamma_i^{ggmdu}, \Gamma_i^{ofdm} \in \hat{\mathcal{R}}_{cvx}$; $\mathbf{y}^{ggmdu}$ and $\mathbf{y}^{ofdm}$ are in the domain where $g(\mathbf{y})$ is Schur-convex. Hence, by (3.70), we can conclude that $g(\mathbf{y}^{ggmdu}) \leq g(\mathbf{y}^{ofdm})$. Similarly, it follows that $g(\mathbf{y}^{ggmdu}) \geq g(\mathbf{y}^{ofdm})$ for $\Gamma_i^{ggmdu}, \Gamma_i^{ofdm} \in \hat{\mathcal{R}}_{ccv}$. □

3.6 Numerical Results

In this section, we present the simulation results on the average BERs of the proposed GGMD DFE transceiver, the GMD DFE MMSE transceiver [18], [19], the SC-CP MMSE system, and OFDM MMSE systems [7], [6], [106]. For OFDM MMSE systems, both cases, with and without bit allocation, are considered. The bit allocation is done according to Ch. 8 in [106] with slightly modifications for the OFDM MMSE system. We assume that the channel noise is AWGN with variance

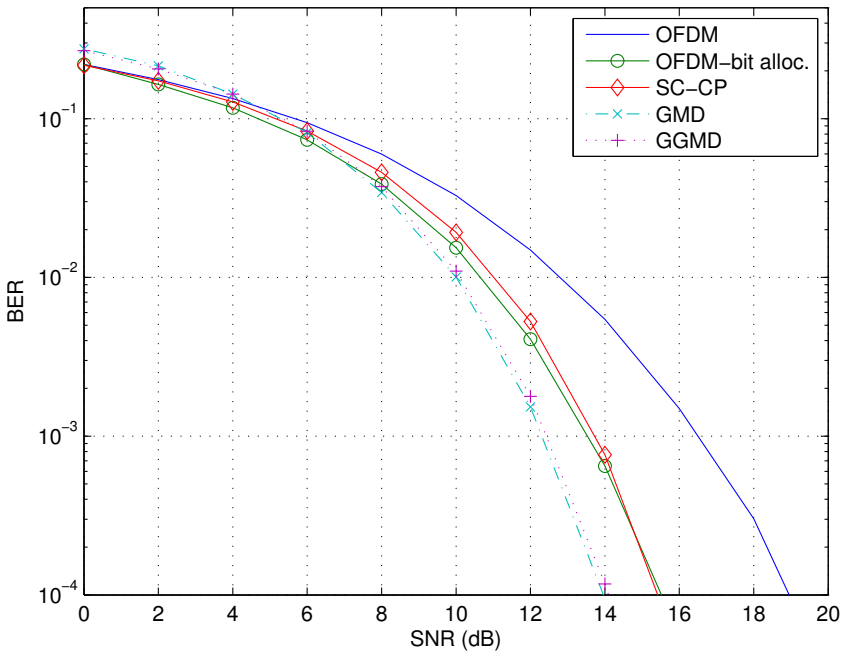


Figure 3.3: The BER performance comparison for the channel $\mathbf{h}_1$.

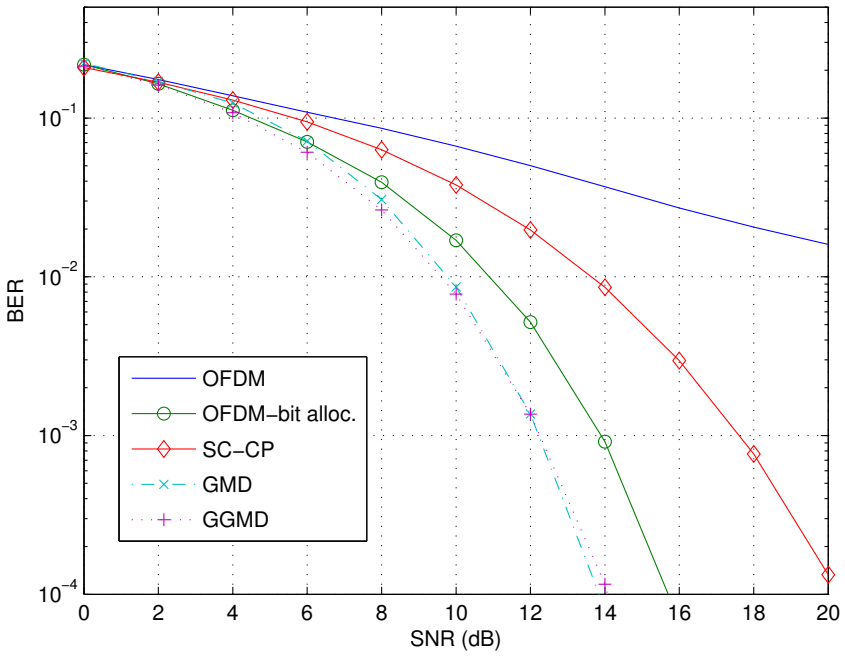


Figure 3.4: The BER performance comparison for the channel $\mathbf{h}_2$.

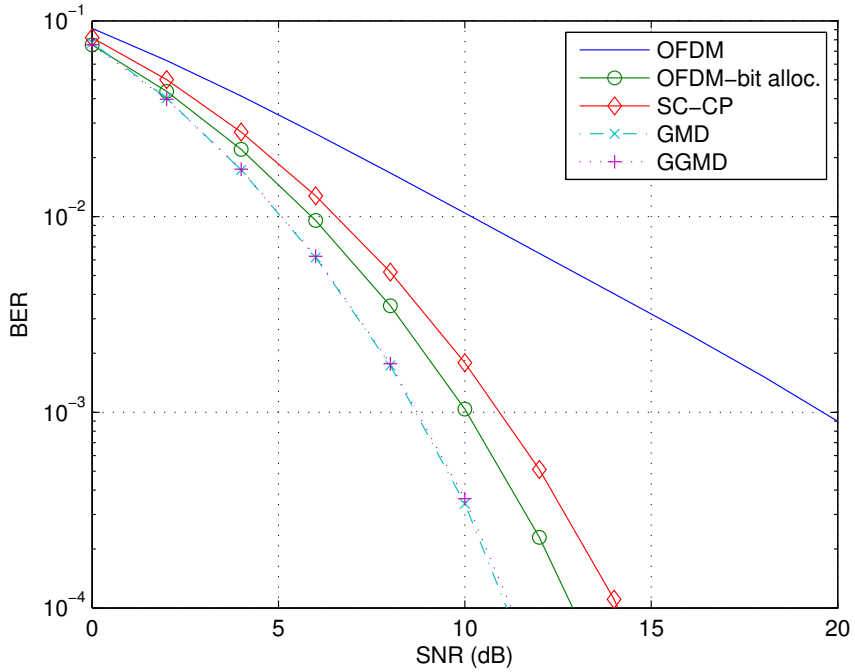


Figure 3.5: The comparison of BERs

σ_w^2 as in Sec. 3.3. The number of subchannels $M = K = 64$ and the length of the cyclic prefix is equal to the order of the channels. The SNR is defined as $SNR = E_{av}/\sigma_w^2$, where E_{av} is the average transmitting symbol power. For moderate high SNR and Gray encoding schemes, one can assume one bit error per symbol error. Then, $BER = SER/b$ where b is the number of bits per symbol [94]. So, the relationships of average SERs for different transceivers roughly reflects those of the BERs. The average BERs are evaluated over fixed channels and multipath slowly fading channels in Example 1 and Example 2, respectively.

Example 1: Two channels with four coefficients ($N_h = 3$) as in [7] are used in this example:

$$H_1(z) = \mathbf{z}\mathbf{h}_1,\ H_2(z) = \mathbf{z}\mathbf{h}_2, \tag{3.71}$$

where

$$\mathbf{h}_1 = \begin{bmatrix} 0.3903 + j0.1049 \\ 0.6050 + j0.1422 \\ 0.4402 + j0.0368 \\ 0.0714 + j0.5002 \end{bmatrix}, \quad \mathbf{h}_2 = \begin{bmatrix} -0.3699 - j0.5782 \\ -0.4053 - j0.5750 \\ -0.0834 - j0.0406 \\ 0.1587 - j0.0156 \end{bmatrix},$$

and $\mathbf{z} = [1, z^{-1}, z^{-2}, z^{-3}]$. The second channel has a spectral null at 0.9π. For the proposed GGMD DFE transceiver, the GMD MMSE transceiver , the SC-CP MMSE system, and OFDM MMSE systems, the modulated symbols are standard QPSK symbols with Gray encoding [94] and two bits are allocated to each subchannel. So, the convex region in Theorem 3.3 contains all SNRs in $\mathcal{R}_{++}$ whenever the theorem applies. For the OFDM MMSE system with bit allocation, an average of two bits is allocated to each subcarrier and each substream only takes square QAM [94] symbols with Gray encoding. Hence, the bit allocation algorithm [106] assigns even number of bits to each substream. The BERs are evaluated by sending 10000 64×1 data blocks for each system. Fig. 3.3 and Fig. 3.4 show that both the GGMD and GMD DFE MMSE transceivers have almost the same BER performance, which conforms to (3.47). Both figures also show that $P_{ggmd} \leq P_{sccp} \leq P_{ofdm}$, except for the low SNR region where error propagation is significant. Otherwise, the results follow the analysis result in Theorem 3.3 nicely. Moreover, both GGMD and GMD DFE MMSE transceivers are robust to spectral nulls while the BERs of OFDM and SC-CP MMSE systems degrade significantly if there is a spectral null as illustrated in Fig. 3.4. Although the OFDM MMSE system with bit allocation performs considerably well when compared with SC-CP and OFDM systems without bit allocation, the simulations show that it suffers from BER loss due to the granularity issue of bit allocation.

Example 2: The multipath slowly fading channel is given by

$$H(z) = \sum_{n=0}^{N_h} h(n) z^{-n}. \tag{3.72}$$

The channel order $N_h = 3$, the channel taps $h(n)$ are assumed to be complex Gaussian variables satisfying $E[h(n)] = 0$ and $E[h(n)h^*(n-k)] = \delta(k)$. The cyclic prefix length is chosen to be 3. The modulation and bit allocation schemes for all the systems are the same as those in Example 1. To evaluate the BERs, we have 10000 64×1 data blocks processed by each system for one channel realization. The average BERs are obtained by averaging over 100 channel realizations. The average BERs are shown in Fig. 3.5. The figure shows that both the GGMD and GMD DFE MMSE transceivers have almost the same BER performance. Moreover, the GGMD and GMD transceivers greatly reduce the average BERs in the CP system when compared with linear MMSE receivers, SC-CP and OFDM. The result, $P_{ggmd} \leq P_{sccp} \leq P_{ofdm}$, is consistent with Theroem 3.3. Moreover, it is observed that both GGMD and GMD DFE MMSE transceivers outperform OFDM MMSE systems with bit allocation.

3.7 Conclusions

The GGMD transceiver achieves the same optimality as the optimal GMD MMSE transceiver, in that it maximizes *Gaussian mutual information*, and minimizes the arithmetic mean square error and average SER. For CP systems, the SER analysis shows that the GGMD DFE transceiver outperforms the SC-CP MMSE system for all SINRs, and the OFDM MMSE system in moderate high SINR region. Moreover, our simulations demonstrate the robustness of the GGMD DFE transceiver when there are spectral nulls, and verify its performance advantage over OFDM and SC-CP MMSE systems. The proposed GGMD DFE transceiver is a good transceiver candidate for CP systems with competitive *complexity* and excellent SER performance in moderate high SINR.

3.8 Appendix

3.8.1 Proof of Lemma 3.1

By (3.17), the error component defined in (3.16) can be re-expressed as

$$\begin{aligned} \mathbf{e} &= \mathbf{A}_{L-1}\mathbf{y}^{(L-2)} - (\mathbf{B}^{(L-1)} \otimes \mathbf{I}_{l'_{L-2}})\mathbf{a} - \mathbf{a} \\ &= \mathbf{A}_{L-1}\Big(\mathbf{y}^{(L-2)} - (\mathbf{P}^{(L-1)} \otimes \mathbf{I}_{l'_{L-2}})\mathbf{a}\Big). \end{aligned}$$

The last equality follows from (3.9). So, (3.18) holds when $m = L-1$. Suppose (3.18) holds for $m = k$ as well, i.e.,

$$\mathbf{e} = \mathbf{A}'(k)\Big(\mathbf{y}^{(k-1)} - \mathbf{P}'(k)\mathbf{a}\Big). \tag{3.73}$$

Substituting $\mathbf{y}^{(k-1)}$ in (3.73) with (3.17), we have

$$\begin{aligned} \mathbf{e} &= \mathbf{A}'(k)\Big(\mathbf{A}_{k-1}\mathbf{y}^{(k-2)} - \big((\mathbf{B}^{(k-1)} + \mathbf{I}) \otimes \mathbf{I}_{l'_{k-2}}\big)\mathbf{P}'(k)\mathbf{a}\Big) \\ &= \mathbf{A}'(k-1)\Big(\mathbf{y}^{(k-2)} - \mathbf{P}'(k-1)\mathbf{a}\Big). \end{aligned}$$

The last equality follows from (3.9). Hence, (3.18) holds when $m = k-1$. By induction, we can conclude the proof.

3.8.2 Proof of Lemma 3.2

By the definition of $\mathbf{\Lambda}$ in Chapter 2, we have

$$\mathbf{\Sigma}_f\mathbf{\Sigma}_h\mathbf{\Lambda} = \frac{\sigma_a^2}{\sigma_w^2}\mathbf{\Sigma}_f^2\mathbf{\Sigma}_h^2\Big(\frac{\sigma_a^2}{\sigma_w^2}\mathbf{\Sigma}_f^2\mathbf{\Sigma}_h^2 + \mathbf{I}\Big)^{-1}.$$

Applying the matrix inversion lemma [105], one can have

$$\mathbf{\Sigma}_f\mathbf{\Sigma}_h\mathbf{\Lambda} = \mathbf{I} - \Big(\mathbf{I} + \frac{\sigma_a^2}{\sigma_w^2}\mathbf{\Sigma}_f^2\mathbf{\Sigma}_h^2\Big)^{-1}. \tag{3.74}$$

By (3.2), (3.26) and the GGMD iterative equations in (3.25), $(\mathbf{I} + \sigma_a^2/\sigma_w^2\mathbf{\Sigma}_f^2\mathbf{\Sigma}_h^2)^{1/2}$ can be factored as

$$\Big(\mathbf{I} + \frac{\sigma_a^2}{\sigma_w^2}\mathbf{\Sigma}_f^2\mathbf{\Sigma}_h^2\Big)^{1/2} = \mathbf{Q}^{(0)}\mathbf{D}^{(0)}\mathbf{B}_1^{(0)}\mathbf{P}^{(0)\dagger}. \tag{3.75}$$

Substituting (3.75) into (3.74), we obtain (3.27).

3.8.3 Proof of Lemma 3.3

Denote

$$\mathbf{E}_m = (\mathbf{A}\otimes\mathbf{I}_l)\mathbf{D}_m\mathbf{S}^m(\mathbf{B}\otimes\mathbf{I}_l),$$

so that $\mathbf{E} = \sum_{m=1}^{l-1}\mathbf{E}_m$. Let $\mathbf{G} = (\mathbf{A}\otimes\mathbf{I}_l)\mathbf{D}_m$ and $\mathbf{F} = \mathbf{S}^m(\mathbf{B}\otimes\mathbf{I}_l)$. For $0 \le i_1, i_2 \le n-1$ and $0 \le j_1, j_2 \le l-1$, one can observe that

$$\begin{aligned} &G_{i_1l+j_1,i_2l+j_2} = 0, \text{ for } |j_1 - j_2| \neq 0, \\ &F_{i_1l+j_1,i_2l+j_2} = 0, \text{ for } |j_1 - j_2| \neq m. \end{aligned} \tag{3.76}$$

The diagonal elements of $\mathbf{E}_m$ can be expressed as

$$[\mathbf{E}_m]_{i_1l+j_1,i_1l+j_1} = \sum_{j=0}^{nl-1} G_{i_1l+j_1,j}F_{j,i_1l+j_1}. \tag{3.77}$$

Renumbering the indices and using (3.76), we have

$$[\mathbf{E}_m]_{i_1 l+j_1, i_1 l+j_1} = \sum_{i_2=0}^{n-1} \Big(\sum_{j_2=0, j_2 \neq j_1}^{l-1} G_{i_1 l+j_1, i_2 l+j_2} F_{i_2 l+j_2, i_1 l+j_1} + G_{i_1 l+j_1, i_2 l+j_1} F_{i_2 l+j_1, i_1 l+j_1} \Big) = 0.$$

Hence, the diagonal elements of $\mathbf{E}$ are given by

$$E_{i,i} = \sum_{m=1}^{l-1} [\mathbf{E}_m]_{i,i} = 0. \tag{3.78}$$

3.8.4 Proof of Lemma 3.4

Taking out the common factor $\mathbf{I}_{l'_{n-3}}$ from $\mathbf{A}'(n)$, $\mathbf{P}'(n-1)$ and the bracketed term in (3.31), $\tilde{\mathbf{E}}_n$ can be re-expressed as

$$\tilde{\mathbf{E}}_n = \Big\{ (\mathbf{A} \otimes \mathbf{I}_{l_{n-2}}) \mathbf{L}^{(n-2)} (\mathbf{B} \otimes \mathbf{I}_{l_{n-2}}) \Big\} \otimes \mathbf{I}_{l'_{n-3}}, \tag{3.79}$$

where

$$\begin{aligned} \mathbf{A} &= \Big(\prod_{i=0}^{L-n-1} (\mathbf{B}_1^{(L-1-i)} \mathbf{P}^{(L-1-i)\dagger} \otimes \mathbf{I}_{\prod_{m=n-1}^{L-2-i} l_m}) \Big) \\ &\quad \cdot \mathbf{D}^{(n-1)^{-2}} \mathbf{B}_1^{(n-1)^{-\dagger}} \mathbf{P}^{(n-1)\dagger}, \\ \mathbf{B} &= \prod_{i=n-1}^{L-1} \mathbf{P}^{(i)} \otimes \mathbf{I}_{\prod_{m=n-1}^{i-1} l_m}. \end{aligned}$$

Since $\mathbf{B}^{(n-2)}$ is a strictly upper triangular matrix with $l_{n-2} \times l_{n-2}$ submatrices on its diagonal, using (3.32), we can show that $\mathbf{L}^{(n-2)}$ is a strictly lower triangular matrix with $l_{n-2} \times l_{n-2}$ strictly lower submatrices on its diagonal. So, $\mathbf{L}^{(n-2)}$ can be written in the form:

$$\mathbf{L}^{(n-2)} = \sum_{m=1}^{l_{n-2}-1} \mathbf{D}_m \mathbf{S}^m, \tag{3.80}$$

where $\mathbf{D}_m$ are diagonal matrices of size $\prod_{j=n-2}^{L-2} l_j$ and $\mathbf{S}$ is of the same size having the same form as (3.28). Using Lemma 3.3, we finally have $[\tilde{\mathbf{E}}_n]_{i,i} = 0$.

Chapter 4

Zero-Forcing DFE Transceiver Design over Slowly Time-Varying MIMO Channels Using ST-GTD

This chapter considers the optimization of transceivers with decision feedback equalizers (DFE) for slowly time-varying memoryless multi-input multi-output (MIMO) channels. The data vectors are grouped into space-time blocks (ST-blocks) for the spatial and temporal precoding to take advantage of the diversity offered by time-varying channels. Space-time generalized triangular decomposition (ST-GTD) is proposed for application in time-varying channels. Under the assumption that the instantaneous channel state information at the transmitter (CSIT) and receiver (CSIR), and the channel prediction are available, we also propose the space-time geometric mean decomposition (ST-GMD) system based on ST-GTD. Under perfect channel prediction, the system minimizes both the arithmetic MSE at the feedback detector, and the average un-coded bit error rate (BER) in the moderate high signal to noise ratio (SNR) region. For practical applications, a novel ST-GTD based system which does not require channel prediction but shares the same asymptotic BER performance with the ST-GMD system is also proposed. At the moderate high SNR region, our analysis and numerical results show that all the proposed systems have better BER performance than the conventional GMD-based systems over time-varying channels; the average BERs of the proposed systems are non-increasing functions of the ST-block size. Most of the results in this chapter have been reported in our journal paper [38].

4.1 Introduction

In recent years, multi-input multi-output (MIMO) transceiver design has received a great deal of attention [12]-[19]. Most of the research on MIMO transceiver design focuses on time-invariant channels. In practice, the wireless channels are time-varying due to users' mobility. In this chapter, we consider transceiver design based on the block fading model in which the MIMO channel is constant over the coherence (block) interval of N_c symbol vectors. The channel varies across different coherence intervals independently or according to Jakes' model [29], [98]. Zero-forcing constraint is assumed throughout the chapter.

To exploit the array gain for the full channel capacity, both channel state information (CSI) at the transmitter (CSIT) and the receiver (CSIR) are required [14], [98]. When the channel varies at a much lower rate compared to the data rate of the systems, CSIT can be obtained from the receiver via feedback mechanism. However, the overhead becomes too large if the channel is varying at a faster rate. In time division duplex (TDD) systems, the uplink and downlink are multiplexed on the same channel, so channel reciprocity holds. Hence, the transmitter can estimate its own CSI at the current time slot using the received signal from the reverse link, and use the estimated CSI to transmit data at the next time slot provided that the channel does not change significantly [30]-[34]. Wiener filter prediction can further be exploited to improve the accuracy of CSIT [29], [30]. Both feedback and TDD schemes can offer instantaneous CSI at transmitter and receiver.

For time invariant channels, the geometric mean decomposition (GMD) based systems with "zero-forcing" and "minimum mean square error (MSE)" decision feedback structures [16], [17], [18] and [19], and the GGMD-based transceiver proposed in Chapter 2, are known to minimize the arithmetic mean (over the spatial domain) of the expected MSE at the input of the decision device and the average bit error rate (BER) in high signal to noise ratio (SNR). Moreover, [17] shows that the GMD-based system with zero-forcing constraint achieves optimal channel throughput asymptotically in high SNR. Unlike the singular value decomposition based systems which require bit allocation to achieve the optimal average BER [35], the GMD-based systems do not require bit allocation since all the effective subchannels have the same SINR [17].

In the case of time-varying channels, different data blocks pass through MIMO channels with different channel coefficients. If instantaneous CSIT and CSIR are available, the GMD-based system can be applied directly to time-varying channels. However, its average BER is not minimized since different coherence blocks have different arithmetic MSEs at the feedback loop detector. In [36],

we proposed the GMD transceiver with a superimposed channel-independent temporal precoder (GMD-TP) which also only requires instantaneous CSIT and CSIR. We took the space-GMD and introduced the channel-independent temporal precoder to construct the GMD-TP. The temporal precoder equalizes the MSEs and hence SNRs across different coherence blocks (intervals) so that the average BER per space-time block (ST-block) is minimized.

In this chapter, based on the generalized triangular decomposition (GTD) [21], we develop space-time GTD (ST-GTD) for the decomposition of time-varying MIMO channels which does GMD on the spatial domain and GTD on the temporal domain. Using the special case of ST-GTD, namely ST-GMD which does GMD on both spatial and temporal domains, we develop the ST-GMD transceivers with zero-forcing constraint. The design of ST-GTD transceivers requires instantaneous CSIR, CSIT and channel prediction. The required prediction length depends on the size of an ST-block. Similar issues of channel prediction have been studied in several papers, e.g., [29], [37]. The Wiener filter theory is usually adopted for the prediction of future channel coefficients based on the previous channel estimations. The accuracy of prediction depends highly on the channel model. Under the perfect channel prediction assumption, the ST-GMD transceiver is shown to jointly minimize the arithmetic MSE in each ST-block (which consists of several coherence blocks), and minimize the average per ST-block BER in high SNR. Next, in consideration of the feasibility of channel prediction, a causal ST-GTD based transceiver (CST-GTD) with stationary temporal processing is also proposed here. It does not require channel prediction because its temporal precoder is stationary. It is shown that the CST-GTD has smaller arithmetic MSE and average BER than the conventional GMD-based system in the high SNR region. The simulation also shows that the BER performance of the CST-GTD approximates that of the ST-GMD transceiver asymptotically. In any case, the ST-GMD transceiver serves as a performance benchmark for the general class of ST-GTD transceiver, including CST-GTD.

The novelty of the ST-GTD transceiver is the incorporation of the temporal precoder and the newly proposed "nested-loop" receiver structure. For each ST-block, these two components not only redistribute the MSEs among blocks, but also reduce the arithmetic MSE. This is in contrast to the linear block precoder in [6] and [7], and the temporal precoder in [36], which keep the same arithmetic MSE while equalizing the MSEs. At the moderate high SNR region, our analysis and numerical results show that all the proposed systems have better BER performance than the conventional GMD-based systems over time-varying channels; the average BERs of the proposed sys-

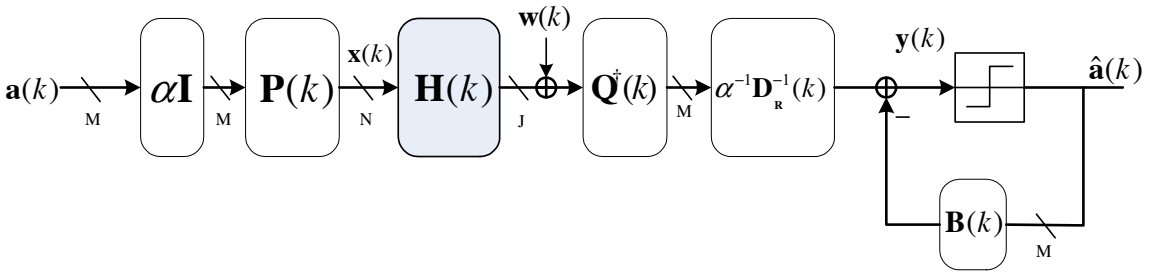


Figure 4.1: The GMD-based system

tems are non-increasing functions of the ST-block size. Moreover, our analysis shows that if the block size is a power of two, i.e., 2^n, then the average BERs of the proposed systems are non-increasing functions of n at the high SNR region and non-decreasing functions of n at the low SNR region. Our numerical studies in Sec. 4.7 also demonstrate the use of channel prediction for ST-GMD transceivers in the Jakes' channel model. In the cases in which channel prediction is accurate enough, the performance of ST-GMD transceivers with imperfect channel prediction is still very close to that with perfect channel prediction.

4.2 Outline

The sections are structured as follows. In Section 4.3, we introduce the time-varying channel model and review the GTD theorem [21]. The GMD-based DFE transceiver with the zero-forcing constraint [17], [19] is reviewed. In Section 4.4.1, we develop space-time GTD based on the spatial GMD. Section 4.4.2 is devoted to the derivation the optimal ST-GTD transceiver which minimizes the arithmetic MSE. A practical suboptimal ST-GTD transceiver which does not require channel prediction is proposed in Section 4.5. In Section 4.6, we analyze the performance of the proposed transceivers. Numerical examples of BER performances are given in Section 4.7. Concluding remarks are given in Section 4.8.

4.3 Preliminaries and Reviews

4.3.1 System Model

In this chapter, we consider the narrowband block fading MIMO channel model [98]. The channel remains constant over the coherence period of N_c transmitted signal vectors and varies indepen-

dently [98] or according to Jakes' model [29] across different coherence intervals. For simplicity of analysis, we just pick one transmitted signal vector from each coherence block since the transmitted signal vectors in the same block go through the same channel. The channel model is given by

$$\mathbf{y}(k) = \mathbf{H}(k)\mathbf{x}(k) + \mathbf{w}(k), \tag{4.1}$$

where k is the coherence block index, $\mathbf{H}(k)$ is a $J \times N$ rank M channel matrix, and $\mathbf{x}(k)$ is an $N \times 1$ transmitted signal vector. The elements of $\mathbf{H}(k)$ are i.i.d. Gaussian random variables. In Jakes' model, the (i,j)-th coefficients of the channel matrices from different k are related by

$$E[H_{i,j}(k)H_{i,j}^*(k')] = J_0(2\pi f_d|k-k'|N_cT_s), \tag{4.2}$$

where $J_0(\cdot)$ is the zeroth order Bessel function of first kind, f_d the Doppler spread and T_s the symbol period. The noise $\mathbf{w}(k)$ is a $J \times 1$ Gaussian random process vector with $E(\mathbf{w}(k)) = \mathbf{0}$ and $E(\mathbf{w}(k)\mathbf{w}^\dagger(k')) = \sigma_w^2\delta(k-k')\mathbf{I}_J$, and $\mathbf{y}(k)$ is the $J \times 1$ received signal vector. At each coherence interval, $\mathbf{H}(k)$ is assumed to be known to the transmitter and receiver.

4.3.2 GTD Decomposition

In the following, we give a brief review of GTD theorem [21] and its application for the design of the GMD-based zero-forcing DFE transceiver [17].

Theorem 4.1: (*The generalized triangular decomposition.* [21]) Let $\mathbf{H} \in \mathbb{C}^{J\times N}$ have rank M with non-zero singular values $\mathbf{d}_H = [\sigma_{H,0}, \sigma_{H,1}, \cdots, \sigma_{H,M-1}]^T$. Then there exists an upper triangular matrix $\mathbf{R} \in \mathbb{C}^{M\times M}$, and semi-unitary matrices $\mathbf{P} \in \mathbb{C}^{N\times M}$ and $\mathbf{Q} \in \mathbb{C}^{J\times M}$ in which all columns are orthonormal, such that $\mathbf{H} = \mathbf{QRP}^\dagger$ if and only if $\mathbf{r} \prec_\times \mathbf{d}_H$ where $r_i = |R_{ii}|$. □

Suppose $R_{ii} = |R_{ii}|e^{j\theta_i}$. Without loss of generality, we can make the diagonal entries of $\mathbf{R}$ real and positive by extracting $e^{j\theta_i}$ from the ith row of $\mathbf{R}$ and multiplying the ith column of $\mathbf{Q}$ by $e^{j\theta_i}$. If one chooses $R_{ii} = (\prod_{i=0}^{M-1}\sigma_{H,i})^{1/M}$, then GTD is reduced to GMD.

4.3.3 GMD-based Transceivers with Zero-forcing Constraint

Fig. 4.1 shows the GMD transceiver, which has been shown to be optimal in average BER at high SNR for linear time invariant channels [19]. Since both the transmitter and receiver have perfect

CSI at current block time k, the $N \times M$ precoding matrix $\mathbf{P}(k)$, and the $J \times M$ feedforward matrix $\mathbf{Q}(k)$ can be obtained from the GMD of $\mathbf{H}(k)$ which is

$$\mathbf{H}(k) = \mathbf{Q}(k)\mathbf{R}(k)\mathbf{P}^{\dagger}(k), \tag{4.3}$$

where $\mathbf{P}(k) \in \mathbb{C}^{N\times M}$ and $\mathbf{Q}(k) \in \mathbb{C}^{J\times M}$ with orthonormal columns. $\mathbf{R}(k) \in \mathbb{C}^{M\times M}$ is an upper triangular matrix with $\mathbf{r}(k)$ on the diagonal. The $M \times 1$ vector $\mathbf{r}(k)$ has equal elements

$$r_i(k) = \sigma_k = \big(\prod_{i=0}^{M-1} \sigma_{H,i}(k) \big)^{1/M}, \tag{4.4}$$

where $\sigma_{H,i}(k)$ is the ith singular value of $\mathbf{H}(k)$. The $M \times M$ feedback matrix $\mathbf{B}(k)$ is given by

$$\mathbf{B}(k) = \mathbf{D}_{\mathbf{R}}^{-\mathbf{1}}(k)\mathbf{R}(k) - \mathbf{I}_M, \tag{4.5}$$

where $\mathbf{D}_{\mathbf{R}}(k) = diag(\mathbf{r}(k))$. $\mathbf{a}(k)$ is an $M \times 1$ symbol vector from the kth block with each element $[\mathbf{a}(k)]_i$ chosen from the alphabet χ of finite size. We assume $E(\mathbf{a}(k)\mathbf{a}^{\dagger}(k')) = \sigma_a^2\delta(k-k')\mathbf{I}_M$. The gain α is chosen to satisfy the total transmitting power constraint

$$P_0 = \mathrm{Tr}\Big(E\big(\mathbf{x}(k)\mathbf{x}^{\dagger}(k)\big)\Big), \tag{4.6}$$

and hence satisfies $\alpha = \sqrt{P_0/M\sigma_a^2}$.

If there is no error propagation in the DFE loop, the received signal vector in front of the detector is given by

$$\mathbf{y}(k) = \mathbf{a}(k) + \mathbf{e}(k), \tag{4.7}$$

where $\mathbf{e}(k) = \alpha^{-1}{\mathbf{D}_{\mathbf{R}}}^{-1}(k)\mathbf{Q}^{\dagger}(k)\mathbf{w}(k)$. The error covariance of $\mathbf{e}(k)$ is

$$\mathbf{R}_{\mathbf{ee}} = \frac{\sigma_w^2}{\alpha^2} diag(|\mathbf{r}(k)|)^{-2}. \tag{4.8}$$

The total MSE of the kth block at the detector is

$$\xi_{gmd}(k) = tr(\mathbf{R}_{\mathbf{ee}}) = \frac{\sigma_w^2}{\alpha^2} M \big(\prod_{i=0}^{M-1} \frac{1}{\sigma_{H,i}^2(k)} \big)^{1/M}. \tag{4.9}$$

4.4 Space-Time GTD Transceivers

4.4.1 Space-Time GTD

To facilitate space-time processing for the later sections, K blocks of symbol vectors are grouped into one space-time block as

$$\mathbf{a}_m = \left[\begin{array}{ccc} \mathbf{a}^T(mK) & \ldots & \mathbf{a}^T(mK+K-1) \end{array} \right]^T, \tag{4.10}$$

where m is the ST-block index. The symbols, m and mK, will be omitted for convenience. The equivalent MIMO channel matrix for the mth ST-block is a $KJ \times KN$ block diagonal matrix given by

$$\mathbf{H} = diag\big(\mathbf{H}(0), \mathbf{H}(1), \ldots, \mathbf{H}(K-1)\big). \tag{4.11}$$

Let $0 \leq k \leq K-1$. If GMD is applied to each $\mathbf{H}(k)$ separately (in spatial domain), we have $\mathbf{H}(k) = \mathbf{Q}(k)\mathbf{R}(k)\mathbf{P}^\dagger(k)$ as (4.3). $\mathbf{H}$ can be decomposed as

$$\mathbf{H} = \mathbf{Q}\mathbf{R}\mathbf{P}^\dagger, \tag{4.12}$$

where $\mathbf{Q}$, $\mathbf{P}$ and $\mathbf{R}$ are block diagonal matrices with $\mathbf{Q}(k)$, $\mathbf{P}(k)$ and $\mathbf{R}(k)$ on the diagonals, respectively. Let

$$\mathbf{d} = [\sigma_0, \ldots, \sigma_{K-1}]^T, \tag{4.13}$$

where σ_i is defined in (4.4). $\mathbf{R}$ can be expressed as

$$\mathbf{R} = \mathbf{D}_\mathbf{R}(\mathbf{I}_{KM} + \mathbf{B}), \tag{4.14}$$

where $\mathbf{D}_\mathbf{R} = \mathbf{\Sigma} \otimes \mathbf{I}_M$, $\mathbf{\Sigma} = diag(\mathbf{d})$, and $\mathbf{B}$ is a block diagonal matrix with $\mathbf{B}(k)$ on the diagonal. $\mathbf{B}(k)$ are strictly upper triangular $M \times M$ matrices given by (4.5).

Since $\mathbf{\Sigma}$ is a diagonal matrix consisting of positive entries σ_k, these are also the singular values. Therefore, by Theorem 4.1, we can decompose $\mathbf{\Sigma}$ by GTD as

$$\mathbf{\Sigma} = \mathbf{Q}_1\mathbf{R}_1\mathbf{P}_1^\dagger, \tag{4.15}$$

where $\mathbf{P}_1$ and $\mathbf{Q}_1$ are $K \times K$ unitary matrices, and $\mathbf{R}_1$ is a $K \times K$ upper triangular matrix. The necessary and sufficient condition for the GTD in (4.15) to be possible is

$$\mathbf{r}_1 \prec_{\times} \mathbf{d}, \tag{4.16}$$

in which $\mathbf{r}_1$ is a $K \times 1$ vector consisting of diagonal elements of $\mathbf{R}_1$ and $\mathbf{d}$ is given by (4.13). We refer to the GTD of $\mathbf{\Sigma}$ as the temporal domain GTD because the σ_k depend on $\mathbf{H}(k)$ and the decomposition needs all σ_k all at once. By (4.12), (4.14) and (4.15), the rank MK block diagonal matrix $\mathbf{H}$ of the from (4.11) can be decomposed as

$$\mathbf{H} = \mathbf{Q}\big((\mathbf{Q}_1\mathbf{R}_1\mathbf{P}_1^{\dagger} \otimes \mathbf{I}_M)(\mathbf{I}_{MK} + \mathbf{B})\big)\mathbf{P}^{\dagger}, \tag{4.17}$$

if and only if $\mathbf{r}_1 \prec_{\times} \mathbf{d}$. The decomposition taking this form is referred to as the space-time GTD (ST-GTD). We denote

$$\mathbf{D}_{\mathbf{R_1}} = diag(\mathbf{r}_1), \tag{4.18}$$

$$\mathbf{B}_1 = {\mathbf{D}_{\mathbf{R_1}}}^{-1}\mathbf{R}_1 - \mathbf{I}_K. \tag{4.19}$$

When the entries of $\mathbf{r}_1$ equal $(\prod_{k=0}^{K-1} \sigma_k)^{1/K}$, the time domain GTD in (4.15) reduces to the time domain GMD. We name this kind of ST-GTD, in particular, as ST-GMD.

The ST-GTD has some advantages over directly applying GTD on big matrix $\mathbf{H}$. Both algorithms first compute the SVD of $\mathbf{H}$, and do the decompositions on the diagonal matrix consisting of all singular values. The block diagonal structure of $\mathbf{H}$ helps to reduce the complexity in the SVD stage, from $O(K^3MNJ)$ to $O(KMNJ)$. Assuming that SVD of $\mathbf{H}$ is given, ST-GTD requires lower computational complexity, $O(KM(N+J))+O(K^2)$, than the complexity of directly applying GTD on $\mathbf{H}$, which is $O(2K^2M^2)$ [21]. Moreover, ST-GTD decouples precoding into spatial and temporal domains. So, ST-GTD can be chosen in such a way that channel prediction is not necessary, as we show later in Section 4.5.

4.4.2 Space-Time GTD Transceivers

In this subsection, we propose the ST-GTD ZF-DFE transceiver based on the ST-GTD introduced in the preceding subsection. The proposed precoder of the ST-GTD transceiver is shown in Fig. 4.2.

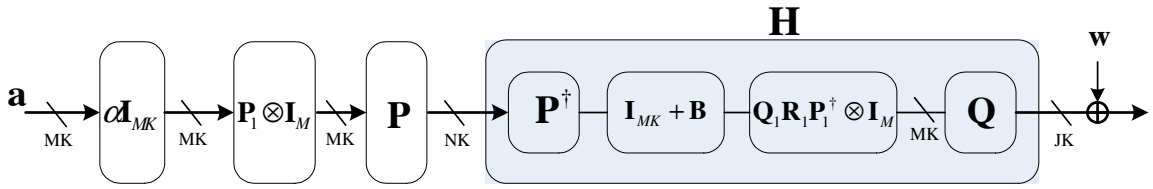


Figure 4.2: The transmitter of the ST-GTD transceiver and the channel

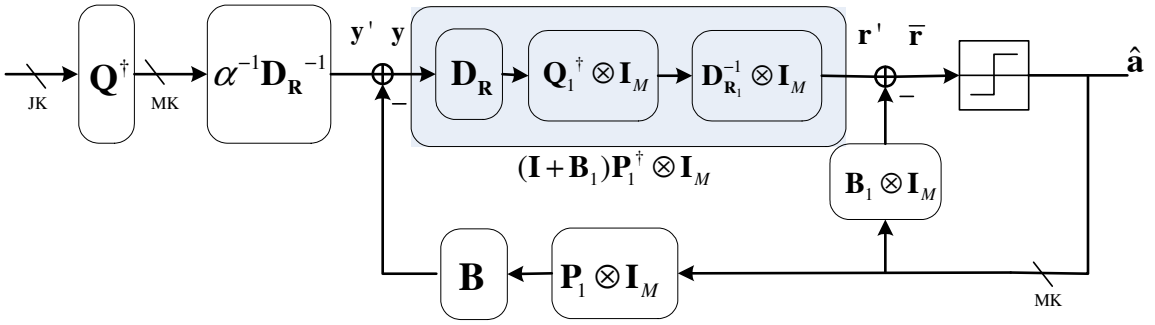


Figure 4.3: The receiver of the ST-GTD transceiver

The proposed receiver is in Fig. 4.3 and its operation will be explained later. Here, it is assumed that the transmitter could predict the channels $\mathbf{H}(k)$ for $0 \leq k \leq K-1$ before sending an ST-block $\mathbf{a}$. We also assume that the receiver can track the channels perfectly and the decoding follows after the reception of a whole ST-block. There are well-studied methods [29], [37] which we can exploit here for the channel prediction. In any case, this system performance serves as a benchmark for performance comparisons and the theoretical foundation for the development of the transceiver which does not require channel prediction in next section.

Channel prediction is applicable when $\mathbf{H}(k)$ for different k are correlated [29], [37]. Before precoding an ST-block, one can apply a Wiener filter to predict $\mathbf{H}(1), \ldots, \mathbf{H}(K-1)$ based upon previous P channel matrices as

$$\hat{H}_{i,j}(k) = \mathbf{w}_{i,j,k}^{\dagger} \mathbf{h}_{i,j}, \tag{4.20}$$

where $\hat{H}_{i,j}(k)$ is the (i,j)-th element of the predicted channel matrix $\hat{\mathbf{H}}(k)$ of $\mathbf{H}(k)$ for $k = 1, \ldots, K-1$, $\mathbf{w}_{i,j,k} = [w_{i,j,k}(0), w_{i,j,k}(1), \ldots, w_{i,j,k}(P-1)]^T$ and $\mathbf{h}_{i,j} = [H_{i,j}(0), H_{i,j}(-1), \ldots, H_{i,j}(-P+1)]$. Suppose Jakes' model in (4.2) is used, then $\mathbf{w}_{i,j,k}$ is given by

$$\mathbf{w}_{i,j,k} = {\mathbf{R}_{\mathbf{h}}}^{-1} \mathbf{r}_{\mathbf{h}}(k), \tag{4.21}$$

where

$$[\mathbf{R_h}]_{m,n} = J_0(2\pi f_d N_c T_s |n-m|),$$
$$[\mathbf{r_h}(k)]_m = J_0(2\pi f_d N_c T_s |m+k|),$$

for $0 \leq m, n \leq P-1$. Interested readers can refer to [29] and [37] for more details.

If channel prediction is perfect, then the transmitter and receiver have perfect CSI of $\mathbf{H}$. ST-GTD can be applied to decompose $\mathbf{H}$ as (4.17) to get $\mathbf{P}$, $\mathbf{P}_1$, $\mathbf{Q}$, $\mathbf{D_R}$, $\mathbf{B}$ and $\mathbf{B}_1$ for the ST-GTD transceiver. Note that $\mathbf{D_R}$ should be chosen as $diag(\mathbf{d}) \otimes \mathbf{I}_M$ and $\mathbf{D}_{\mathbf{R}1} = diag(\mathbf{r}_1)$ where $\mathbf{r}_1 \prec_\times \mathbf{d}$. If ST-GMD is applied instead of ST-GTD, we name the transceiver, in particular, the ST-GMD transceiver.

However, in practice, perfect channel prediction is not possible. So, the design of ST-GTD transceivers based on the predicted channel may not match the actual channel due to channel prediction error. To alleviate the mismatch, we modify the design procedure. Note that channel prediction is not required by the receiver to get $\mathbf{H}$ since the receiver can store the signal until the whole ST-block is received before it starts to decode. The only part that depends on channel prediction is the temporal precoder $\mathbf{P}_1$ at the transmitter. The computation of precoding matrix $\mathbf{P}_1$ requires knowledge of the singular values of $\mathbf{H}(0), \ldots, \mathbf{H}(K-1)$. At time $k=0$, the precoder already needs $\mathbf{P}_1$ to precode one block for transmission. The implementation of the spatial precoder $\mathbf{P}$ without channel prediction is not a problem, since $\mathbf{P}$ is a block diagonal matrix consisting of $\mathbf{P}(k)$ and the computation of $\mathbf{P}(k)$ requires only the current CSI $\mathbf{H}(k)$. Letting $\mathbf{a}' = (\mathbf{P}_1 \otimes \mathbf{I}_M)\mathbf{a}$, the precoded block at time k is given by

$$[\mathbf{t}]_{kM:kM+M-1} = \mathbf{P}(k)[\mathbf{a}']_{kM:kM+M-1}. \tag{4.22}$$

To design $\mathbf{P}_1$, we firstly apply the Wiener prediction filter in (4.20) for channel prediction and construct the predicted channel matrix

$$\hat{\mathbf{H}} = diag(\mathbf{H}(0), \hat{\mathbf{H}}(1), \ldots, \hat{\mathbf{H}}(K-1)), \tag{4.23}$$

where $\hat{\mathbf{H}}(k)$ are given by (4.20). Then, $\mathbf{P}_1$ is obtained from the ST-GTD or ST-GMD of $\hat{\mathbf{H}}$. In the case of perfect channel prediction, $\mathbf{Q}_1$, $\mathbf{P}_1$ and $\mathbf{R}_1$ are obtained from the temporal domain GTD of $\mathbf{\Sigma}$ which is obtained from $\mathbf{H}$, i.e., $\mathbf{\Sigma} = \mathbf{Q}_1\mathbf{R}_1\mathbf{P}_1^\dagger$. Here, $\mathbf{Q}_1$ and $\mathbf{R}_1$ are obtained from the QR

decomposition of $\mathbf{\Sigma P}_1$ as

$$\mathbf{\Sigma P}_1 = \mathbf{Q}_1 \mathbf{R}_1, \tag{4.24}$$

where $\mathbf{P}_1$ is from the ST-GTD of $\hat{\mathbf{H}}$, $\mathbf{Q}_1$ is a $K \times K$ unitary matrix and $\mathbf{R}_1$ is a $K \times K$ upper triangular matrix. Without loss of generality, we can make the vector $\mathbf{r}_1$, which consists of the diagonal entries of $\mathbf{R}_1$, positive and real. One can also verify that if the channel prediction is perfect both design procedures lead to the same ST-GTD or ST-GMD transceiver.

For each ST-block time, $\mathbf{a}$ is precoded by the linear precoders, transmitted through the channel $\mathbf{H}$ and pre-processed at the receiver by $\mathbf{Q}^\dagger$ and $\alpha^{-1}{\mathbf{D}_\mathbf{R}}^{-1}$. The estimation of $\mathbf{a}$ is obtained by the successive cancelation algorithm described next.

4.4.3 Successive Cancellation Detection Algorithm for ST-GTD Transceivers

Before summarizing the detection algorithm at the receiver in Fig. 4.3, we need some notations. For $0 \leq i \leq M-1$, define the $M \times M$ diagonal matrix $\mathbf{E}_i$ as

$$\mathbf{E}_i = diag([\underbrace{0, \ldots, 0}_{M-i}, \underbrace{1, \ldots, 1}_{i}]), \tag{4.25}$$

and the $M \times 1$ vector $\mathbf{s}_i$ in which all the entries are zero except the ith entry, which is 1. Based on $\mathbf{E}_i$ and $\mathbf{s}_i$, we define two operators, $\mathbf{\Theta}_i = \mathbf{I}_K \otimes \mathbf{E}_i$ and $\mathbf{S}_i = \mathbf{I}_K \otimes \mathbf{s}_i^T$, on the ST-block vector $\mathbf{a}$, which has the form (4.10). $\mathbf{\Theta}_i$ retains the last i symbols of each $\mathbf{a}(k)$ and makes the other symbols of $\mathbf{a}(k)$ zero. And $\mathbf{S}_i$ is such that

$$\mathbf{S}_i \mathbf{a} = [a_i(0), \ldots, a_i(K-1)]^T.$$

The detection algorithm for the receiver in Fig. 4.3 is as follows:

1. Initialize: $i = 0$.

2. *Outer loop feedback (space domain):* Calculate

   $$\mathbf{S}_{M-i-1}\mathbf{y} = \mathbf{S}_{M-i-1}\big(\mathbf{y}' - \mathbf{B}(\mathbf{P}_1 \otimes \mathbf{I}_M)\mathbf{\Theta}_i \hat{\mathbf{a}}\big).$$

3. *Inner loop feedback and detection (time domain):* Compute

$\mathbf{z}^{(i)} = \mathbf{S}_{M-i-1}\mathbf{r}' = (\mathbf{I}_K + \mathbf{B}_1)\mathbf{P}_1^{\dagger}\mathbf{S}_{M-i-1}\mathbf{y}$; $\mathbf{S}_{M-i-1}\hat{\mathbf{a}}$ can be decoded sequentially with the following procedures:

(a) $\hat{a}_{(K-1)M+M-1-i} = Qt(z_{K-1}^{(i)})$ where the function $q = Qt(t)$ sets q to the element in χ such that it is closest to t in Euclidean norm.

(b) For $c = 2, \ldots, K$, $\hat{a}_{(K-c)M+M-1-i} = Qt\big(z_{K-c}^{(i)} - \sum_{m=K-c+1}^{K-1}[\mathbf{B}_1]_{K-c,m}\hat{a}_{mM+M-i-1}\big)$.

4. If $i = M-1$, then stop, else set $i = i+1$ and go to 2.

Step 1-2 are clear by direct substitution. To justify step 3, we assume that there is no error propagation, i.e., $\mathbf{\Theta}_i\hat{\mathbf{a}} = \mathbf{\Theta}_i\mathbf{a}$. By substitution, we have

$$\mathbf{y}' = (\mathbf{I}_{MK} + \mathbf{B})(\mathbf{P}_1 \otimes \mathbf{I}_M)\mathbf{a} + \mathbf{w}', \tag{4.26}$$

where

$$\mathbf{w}' = (1/\alpha){\mathbf{D}_{\mathbf{R}}}^{-1}\mathbf{Q}^{\dagger}\mathbf{w},$$
$$\mathbf{w} = [\mathbf{w}(0)^T, \ldots, \mathbf{w}(K-1)^T]^T.$$

Then

$$\begin{aligned}\mathbf{S}_{M-i-1}\mathbf{y} &= \mathbf{S}_{M-i-1}(\mathbf{y}' - \mathbf{B}(\mathbf{P}_1 \otimes \mathbf{I}_M)\mathbf{\Theta}_i\hat{\mathbf{a}}) \\ &= \mathbf{P}_1\mathbf{S}_{M-i-1}\mathbf{a} + \mathbf{S}_{M-i-1}\mathbf{w}'.\end{aligned}$$

Substituting $\mathbf{S}_{M-i-1}\mathbf{y}$ into $\mathbf{S}_{M-i-1}\mathbf{r}'$, we have

$$\mathbf{S}_{M-i-1}\mathbf{r}' = (\mathbf{I}_K + \mathbf{B}_1)\mathbf{S}_{M-i-1}\mathbf{a} + \mathbf{w}_2, \tag{4.27}$$

where $\mathbf{w}_2 = \mathbf{S}_{M-i-1}((\mathbf{I}_K + \mathbf{B}_1)\mathbf{P}_1^{\dagger} \otimes \mathbf{I}_M)\mathbf{w}'$. Observe that the equivalent channel between $\mathbf{S}_{M-i-1}\mathbf{r}'$ and $\mathbf{S}_{M-i-1}\mathbf{a}$ is an upper triangular matrix. Hence, $\mathbf{S}_{M-i-1}\mathbf{a}$ can be detected sequentially by the VBLAST-like algorithm in the steps 3a and 3b above.

4.4.4 Mean Square Error at the Detector

To analyze the performance of the ST-GTD transceiver in Fig. 4.3, we assume perfect channel prediction. The performance of the ST-GTD transceiver mainly depends on the noise component in

$\bar{\mathbf{r}}$. To characterize the performance of the detector, we calculate the error covariance matrix of the noise component. We assume that there is no error propagation so that $\hat{\mathbf{a}} = \mathbf{a}$, which is a legitimate assumption at high SNR. Under this assumption, the signal vector $\bar{\mathbf{r}}$ can be expressed as

$$\bar{\mathbf{r}} = \mathbf{a} + \mathbf{e}, \tag{4.28}$$

where

$$\mathbf{e} = \frac{1}{\alpha}(\mathbf{D}_{\mathbf{R}_1}{}^{-1}\mathbf{Q}_1^{\dagger} \otimes \mathbf{I}_M)\mathbf{q}.$$

The entries of the $MK \times 1$ vector $\mathbf{q} = \mathbf{Q}^{\dagger}\mathbf{w}$ are i.i.d. complex Gaussian with zero mean and variance σ_w^2. The error covariance matrix $\mathbf{R_{ee}}$ is given by

$$\mathbf{R_{ee}} = E(\mathbf{e}\mathbf{e}^{\dagger}) = \frac{\sigma_w^2}{\alpha^2}(\mathbf{D}_{\mathbf{R}_1}{}^{-1}\mathbf{D}_{\mathbf{R}_1}{}^{-\dagger} \otimes \mathbf{I}_M), \tag{4.29}$$

Denote the vector $\mathbf{r}_1$ which consists of the diagonal elements of $\mathbf{R}_1$ as $\mathbf{r}_1 = [\eta_0, \eta_1, \ldots, \eta_{K-1}]^T$. The total MSE of the ST-GTD transceiver over an ST-block is

$$\xi_{st-gtd} = \frac{\sigma_w^2 M}{\alpha^2}\sum_{k=0}^{K-1}\frac{1}{|\eta_k|^2} \geq \frac{\sigma_w^2 MK}{\alpha^2}\Big(\prod_{k=0}^{K-1}\frac{1}{\sigma_k^2}\Big)^{1/K}, \tag{4.30}$$

where $\eta_k = [\mathbf{r}_1]_k$. The last inequality comes from AM-GM inequality and $\mathbf{r_1} \prec_{\times} \mathbf{d}$. The equality holds when $|\eta_0| = |\eta_1| = \ldots = |\eta_{K-1}|$ and $\mathbf{r_1} \prec_{\times} \mathbf{d}$. In particular, if we choose

$$\eta_k = \Big(\prod_{i=0}^{K-1}\sigma_i\Big)^{1/K}, \tag{4.31}$$

then $\mathbf{r}_1 \prec_{\times} \mathbf{d}$ is also satisfied, making the ST-GTD possible. This is the case when the ST-GMD is applied. We call this class of ST-GTD transceiver the ST-GMD transceiver. The total mean square error of the ST-GMD transceiver is given by

$$\begin{aligned}\xi_{st-gmd} &= \frac{\sigma_w^2 MK}{\alpha^2}\Big(\prod_{k=0}^{K-1}\frac{1}{\sigma_k^2}\Big)^{1/K} \\ &= \frac{\sigma_w^2 MK}{\alpha^2}\Big(\prod_{k=0}^{K-1}\prod_{i=0}^{M-1}\frac{1}{\sigma_{H,i}^2(k)}\Big)^{1/MK}.\end{aligned} \tag{4.32}$$

The class of ST-GMD transceivers is the optimal subclass of ST-GTD transceivers in terms of total mean square error. Notice that the ST-GMD allows the ST-GTD transceiver to reach the optimal MSE in (4.32), which is the smallest achievable MSE possible from directly applying the GMD to the big matrix $\mathbf{H}$. Furthermore, the error covariance matrix of ST-GMD transceiver has equal diagonal elements. Hence, for every ST-block, the ST-GMD transceiver minimizes both the arithmetic and geometric MSE, and the average un-coded BER at the high SNR region according to [19].

4.5 Space-Time GTD Transceivers with Fixed Temporal Precoder

In the previous section, the design of ST-GTD transceivers relies on channel prediction. However, the channel prediction might not always be that accurate when the MIMO channels $\mathbf{H}(k)$ from block to block become more independent. The performance of the transceiver degrades when the predicted CSI at the transmitter is unreliable. In this section, we develop an ST-GTD transceiver which does not use channel prediction. We say that the transmitter is "causal".

As mentioned in section 4.4.2, the computation of a precoding matrix $\mathbf{P}_1$ requires knowledge of the singular values of $\mathbf{H}(0), \ldots, \mathbf{H}(K-1)$. Without channel prediction, the precoder only has the CSI at the current and previous times, and it is impossible to compute $\mathbf{P}_1$. To make the precoder causal, one can let $\mathbf{P}_1$ be a constant unitary matrix $\mathbf{W}$. In [7] and [6], the DFT or Hadamard matrix is chosen as the channel independent precoder for the OFDM system to equalize the MSEs over subchannels and hence minimize average BER. This motivates us to choose $\mathbf{W}$ to be a DFT or Hadamard matrix. $\mathbf{Q}_1$ and $\mathbf{R}_1$ are obtained from the QR decomposition of $\mathbf{\Sigma W}$ as

$$\mathbf{\Sigma W} = \mathbf{Q}_1 \mathbf{R}_1, \tag{4.33}$$

where $\mathbf{Q}_1$ is also a $K \times K$ unitary matrix and $\mathbf{R}_1$ is a $K \times K$ upper triangular matrix. We call this kind of transceiver the causal ST-GTD transceiver (CST-GTD). It is in fact a subclass of ST-GTDs with perfect channel prediction. The error covariance matrix of the noise signal in front of the detector is given by (4.29), and the total mean square error $\xi_{cst-gtd}$ is the same as (4.30).

4.5.1 Comparison of Mean Square Error

Now, we compare the performance of the conventional GMD-based system [17], the ST-GMD transceiver with perfect channel prediction, and the CST-GTD transceiver. The total MSE of the

GMD-based system in one ST-block is

$$\xi_{gmd} = \frac{\sigma_w^2}{\alpha^2} M \sum_{k=0}^{K-1} \frac{1}{\sigma_k^2}. \tag{4.34}$$

The comparison of the three transceivers is given in Theorem 4.2.

Theorem 4.2: The total mean square errors over one ST-block for the three transceivers are such that

$$\xi_{st-gmd} \leq \xi_{st-gtd} \leq \xi_{gmd}. \tag{4.35}$$

□

Proof: The first inequality follows from (4.30). To prove the second inequality, we firstly define a function

$$f(\mathbf{x}) = \frac{\sigma_w^2 M}{\alpha^2} \sum_{k=0}^{K-1} e^{-2x_k}, \tag{4.36}$$

where $f(\cdot) : \mathbb{R}^K \mapsto \mathbb{R}$ and $\mathbf{x} = [x_0, x_1, \ldots, x_{K-1}]^T$. Since e^{-2x}, $x \in \mathbb{R}$ is a convex function, $f(\mathbf{x})$ is a Schur-Convex function by proposition 3.C.1 in [95]. Let

$$\begin{aligned} \mathbf{\Theta} &= [\log|\eta_0|, \log|\eta_1|, \ldots, \log|\eta_{K-1}|]^T, \\ \mathbf{\Gamma} &= [\log|\sigma_0|, \log|\sigma_1|, \ldots, \log|\sigma_{K-1}|]^T. \end{aligned} \tag{4.37}$$

Since $\mathbf{r}_1 \prec_\times \mathbf{d}$, then $\mathbf{\Theta} \prec_+ \mathbf{\Gamma}$. ξ_{st-gtd} in (4.30) and ξ_{gmd} in (4.34) can be expressed in term of $\mathbf{\Gamma}$ and $\mathbf{\Theta}$ respectively as

$$\xi_{st-gtd} = f(\mathbf{\Theta}), \tag{4.38}$$

$$\xi_{gmd} = f(\mathbf{\Gamma}). \tag{4.39}$$

Since $\mathbf{\Theta} \prec_+ \mathbf{\Gamma}$, by the definition of Schur-convex function in [95], we have $\xi_{st-gtd} \leq \xi_{gmd}$. □

Therefore, we have proven that the class of ST-GTD transceivers has performance superior to the conventional GMD-based system over time-varying channels in terms of total MSE within one

ST-block or arithmetic MSE. In particular, for CST-GTD, we have

$$\xi_{st-gmd} \leq \xi_{cst-gtd} \leq \xi_{gmd}.$$

Also, note that an ST-GMD transceiver with imperfect channel prediction can be treated as an ST-GTD transceiver with perfect channel prediction due to the channel mismatch caused by prediction error. Hence, the total mean square error of ST-GMD transceiver with imperfect channel prediction $\xi_{st-gmdic}$ is such that

$$\xi_{st-gmd} \leq \xi_{st-gmdic} \leq \xi_{gmd}.$$

The temporal precoder $\mathbf{P}_1$ or $\mathbf{W}$, and the "nested-feedback-loop" receiver in ST-GTD or CST-GTD transceiver not only redistribute the MSEs of the blocks in each ST-block but also reduce the arithmetic MSE per ST-block. This is in contrast to the linear block precoder in [6] and [7], which keeps the same arithmetic MSE while equalizing the MSEs. Also, note that the conventional GMD-based system is actually a subclass of CST-GTD with the constant temporal precoding matrix $\mathbf{W}_K = \mathbf{I}_K$.

4.5.2 Comparison of Complexity

In this section, we compare the complexity of the conventional GMD-based system and the ST-GTD transceiver. We let these systems process one ST-block and compare the number of multiplications and additions. For the transmitter part, the complexity of the GMD-based system is $O(MNK)$ which comes from the spatial precoder $\mathbf{P}(k)$ in Fig. 4.1. Since the transmitters of ST-GTD incorporate an additional temporal precoder $\mathbf{P}_1$ as in Fig. 4.2, it has complexity $O(MNK + K^2M)$.

Next, we compare the receivers. For the GMD-based system in Fig. 4.1, the feedforward matrix $\mathbf{Q}^\dagger(k)$ has complexity $O(JM)$ and the feedback matrix $\mathbf{B}(k)$ has $O(M^2)$. The total complexity of the GMD-based receiver is $O(JMK + M^2K)$. The ST-GTD transceiver contains two additional temporal precoders in the feedback loop and an additional temporal feedback loop with feedback matrix $\mathbf{B}_1 \otimes \mathbf{I}_M$ which has complexity $O(MK^2)$. Hence, its total complexity is $O(JMK + MK^2 + M^2K)$.

4.6 Performance Analysis

In this section, we will compare the average BER and the ergodic channel capacity of the conventional GMD-based system, ST-GTD and ST-GMD transceivers. For both ST-GTD and ST-GMD transceivers, perfect channel prediction is assumed. We assume N' uses of the time-varying channels:

$$\mathbf{H}(n), \text{ for } 0 \leq n \leq N' - 1. \tag{4.40}$$

Every K successive uses constitute one ST-block. So the mth ST-block uses the channels:

$$\mathbf{H}(mK + k), \text{ for } 0 \leq k \leq K - 1, \tag{4.41}$$

where $0 \leq m \leq \lceil N'/K \rceil - 1$. The number of blocks, N', is assumed to be a large number and a multiple of the ST-block size K. An even number of bits, b, are allocated for every symbol $a_i(k)$ of each ST-block. For square QAM [94], the BER for each symbol in the kth block of the mth ST-block, assuming that there is just one bit error per symbol error, is approximately

$$P_e \approx cQ\Big(\frac{A}{\sqrt{\nu(mK + k)}}\Big), \tag{4.42}$$

where $Q(\cdot)$ is the Q-function defined in [94], $A = \sqrt{(3E_{av})/(2^b - 1)}$, E_{av} is the average symbol power, $\nu(mK+k)$ is the per symbol MSE of the kth block in the mth ST-block and $c = (4/b)(1-2^{\frac{-b}{2}})$. Notice that the symbol error rate (SER) equals bP_e. The average BER over the entire transmission is hence given by

$$\mathcal{P} = \frac{1}{N'} \sum_{m=0}^{(N'/K)-1} \sum_{k=0}^{K-1} cQ(\frac{A}{\sqrt{\nu(mK + k)}}). \tag{4.43}$$

The function $Q(A/\sqrt{y})$ for $y \in \mathbb{R}$ plays a crucial role in BER analysis. An important property of it is restated as the following lemma.

Lemma 4.1: The function $f(y) = Q(A/\sqrt{y})$ is monotone increasing. It is convex when $y \leq A^2/3$ and concave when $y > A^2/3$. □

Proof: See [7]. □

We define the SNR of the nth block as $\Gamma(n)$. The SNR expressions for the ST-GTD and ST-GMD transceivers are given respectively by

$$\Gamma_{st-gtd}(n) = \frac{\alpha^2 E_{av}|\eta_n^2|}{\sigma_w^2},$$
$$\Gamma_{st-gmd}(n) = \frac{\alpha^2 E_{av}}{\sigma_w^2(\prod_{i=0}^{K-1}\frac{1}{\sigma^2(mK+i)})^{1/K}}, \tag{4.44}$$

which follow from (4.29). We also define two SNR regions:

$$\mathcal{R}_{high} = \{\Gamma : \Gamma \geq \frac{2^b-1}{3}\},\ \mathcal{R}_{low} = \{\Gamma : \Gamma < \frac{2^b-1}{3}\}. \tag{4.45}$$

Before starting the analysis, we prove another useful lemma:

Lemma 4.2: If $\mathbf{A} \in \mathbb{R}_+^{n\times n}$, $\mathbf{B} \in \mathbb{R}_+^{m\times m}$ are doubly stochastic matrices, the $\mathbf{C} = \mathbf{A} \otimes \mathbf{B}$ is an $mn \times mn$ doubly stochastic matrix. □

Proof: See Appendix 4.9.1. □

4.6.1 BER Performance Comparison of the Transceivers

Now, we will compare the BER of the entire class of ST-GTD transceivers including ST-GMD, ST-GMD with imperfect channel prediction and CST-GTD transceivers with the conventional GMD-based system. The following lemma is helpful for further analysis:

Lemma 4.3: The function $\Delta(y) = Q(c_2 \exp(y))$ is monotone decreasing where $c_2 = \alpha A/\sigma_w > 0$. It is convex when $c_2^2 e^{2y} \geq 1$ and concave when $c_2^2 e^{2y} < 1$. □

Proof: The proof is similar to Lemma 4.1. □

In the following theorems, "high SNR" means the SNRs $\Gamma(n)$ of the transceivers are such that $\Gamma(n) \in \mathcal{R}_{high}$ and "low SNR" means $\Gamma(n) \in \mathcal{R}_{low}$.

Theorem 4.3: Let $\mathcal{P}_{st-gmd}$, $\mathcal{P}_{st-gtd}$ and $\mathcal{P}_{gmd}$ be the average BER of ST-GMD, ST-GTD and the

conventional GMD-based transceivers, respectively. Then,

$$\begin{aligned}
&\mathcal{P}_{st-gmd} \leq \mathcal{P}_{st-gtd} \leq \mathcal{P}_{gmd}, \text{ at high SNR,} \\
&\mathcal{P}_{st-gmd} \geq \mathcal{P}_{st-gtd} \geq \mathcal{P}_{gmd}, \text{ at low SNR.}
\end{aligned} \tag{4.46}$$

□

Proof: We first prove the second inequality for both high and low SNR. Let

$$g(\mathbf{z}) = \frac{1}{K} \sum_{k=0}^{K-1} Q(c_2 e^{z_k}), \tag{4.47}$$

where $g(\cdot) : \mathbb{R}^K \longmapsto \mathbb{R}$, $\mathbf{z} = [z_0, \ldots, z_{K-1}]^T$ and $c_2 = \alpha A/\sigma_w$. By Lemma 4.3 and Theorem 1.2, $g(\mathbf{z})$ is Schur-convex when $c_2^2 e^{2z_k} \geq 1$ and Schur-concave when $c_2^2 e^{2z_k} \leq 1$ for all $0 \leq k \leq K-1$. For $0 \leq m \leq (N'/K) - 1$, define $K \times 1$ vectors as

$$\mathbf{x}(m) = [\log(\sigma_{mK}), \ldots, \log(\sigma_{mK+K-1})]^T, \tag{4.48}$$

$$\mathbf{y}(m) = [\log(|\eta_{mK}|), \ldots, \log(|\eta_{mK+K-1}|)]^T. \tag{4.49}$$

So, the BERs of the GMD and the ST-GTD transceivers are given respectively by

$$\mathcal{P}_{gmd} = \frac{cK}{N'} \sum_{m=0}^{(N'/K)-1} g(\mathbf{x}(m)), \tag{4.50}$$

$$\mathcal{P}_{st-gtd} = \frac{cK}{N'} \sum_{m=0}^{(N'/K)-1} g(\mathbf{y}(m)). \tag{4.51}$$

At the high SNR region, where $\Gamma_{gmd}(n), \Gamma_{st-gtd}(n) \in \mathcal{R}_{high}$, we have $c_2^2 e^{2x_k(m)} \geq 1$ and $c_2^2 e^{2y_k(m)} \geq 1$ for all k, m. In this domain, the function $g(\cdot)$ is Schur-convex. It is known that $\mathbf{r}_1 \prec_\times \mathbf{d}$, so $\mathbf{y}(m) \prec_+ \mathbf{x}(m)$; hence, we have $g(\mathbf{y}(m)) \leq g(\mathbf{x}(m))$ for $\forall m$. Therefore, $\mathcal{P}_{st-gtd} \leq \mathcal{P}_{gmd}$. At the low SNR region, where $\Gamma_{gmd}(n)$, $\Gamma_{st-gmd}(n) \in \mathcal{R}_{low}$, we can prove $\mathcal{P}_{st-gtd} \geq \mathcal{P}_{gmd}$ similarly. The first inequality can also been proven by following similar steps. □

Let $\mathcal{P}_{st-gmdic}$ and $\mathcal{P}_{cst-gtd}$ denote the average BER of the ST-GMD transceiver with imperfect channel prediction, and CST-GTD, respectively. At the high SNR region, from Theorem 4.3, we can

conclude, in particular, that

$$\mathcal{P}_{st-gmd} \leq \mathcal{P}_{cst-gtd} \leq \mathcal{P}_{gmd}, \tag{4.52}$$

$$\mathcal{P}_{st-gmd} \leq \mathcal{P}_{st-gmdic} \leq \mathcal{P}_{gmd}. \tag{4.53}$$

4.6.2 Block Size and the BER Performance

In this subsection, the relationship between the size of ST-block and the BER performance is explored.

Theorem 4.4: Let $\mathcal{P}_{st-gmd}^{(K)}$ denote the BER of the ST-GMD transceiver with ST-block size K and $\mathcal{P}_{st-gmd}^{(qK)}$ denote that with ST-block size qK, for $q, K \in \mathbb{N}$. The number of blocks transmitted, N', is assumed to be the multiple of qK. Then,

$$\mathcal{P}_{st-gmd}^{(qK)} \begin{cases} \leq \mathcal{P}_{st-gmd}^{(K)} & \text{, at high SNR,} \\ \geq \mathcal{P}_{st-gmd}^{(K)} & \text{, at low SNR.} \end{cases} \tag{4.54}$$

□

Proof: Let

$$\psi(\mathbf{w}) = \frac{c}{N'} \sum_{n=0}^{N'-1} Q(c_2 \exp(w_n)), \tag{4.55}$$

where $\mathbf{w} = [w_0, w_1, \ldots, w_{N'-1}]^T$ and $c_2 = \alpha A/\sigma_w$. $\psi(\mathbf{w})$ is Schur-convex if all $c_2^2 e^{2w_n} \geq 1$, Schur-concave if all $c_2^2 e^{2w_n} < 1$. Let $\mathbf{x} = [x_0, x_1, \ldots, x_{N'-1}]^T$ where

$$x_n = \log \sigma_n. \tag{4.56}$$

For the ST-block size K,

$$\mathcal{P}_{st-gmd}^{(K)} = \psi(\mathbf{y}), \tag{4.57}$$

where $\mathbf{y} = [y_0, y_1, \ldots, y_{N'-1}]^T$ and

$$y_{mK+k} = \frac{1}{K} \sum_{i=0}^{K-1} \log(\sigma_{mK+i}), \tag{4.58}$$

for $0 \leq k \leq K-1$ and $0 \leq m \leq (N'/K)-1$. From (4.56) and (4.58),

$$\mathbf{y} = (\mathbf{I}_{N'/K} \otimes \mathbf{L}_K)\mathbf{x}, \tag{4.59}$$

where $\mathbf{L}_K$ is a $K \times K$ matrix with equal elements, $1/K$.

For the ST-block size qK,

$$\mathcal{P}_{st-gmd}^{(qK)} = \psi(\mathbf{z}), \tag{4.60}$$

where $\mathbf{z} = [z_0, z_1, \ldots, z_{N'-1}]^T$ and

$$z_{m'qK+k'} = \frac{1}{qK} \sum_{i=0}^{qK-1} log(\sigma_{m'qK+i}), \tag{4.61}$$

for $0 \leq k' \leq qK-1$ and $0 \leq m' \leq N'/(qK)-1$. From (4.59), (4.61),

$$\begin{aligned} \mathbf{z} &= (\mathbf{I}_{N'/(qK)} \otimes \mathbf{L}_q \otimes \mathbf{L}_K)\mathbf{x} \\ &= [(\mathbf{I}_{N'/(qK)} \otimes \mathbf{L}_q) \otimes \mathbf{I}_K][\mathbf{I}_{N'/K} \otimes \mathbf{L}_K]\mathbf{x} \\ &= (\mathbf{I}_{N'/(qK)} \otimes \mathbf{L}_q \otimes \mathbf{I}_K)\mathbf{y}. \end{aligned} \tag{4.62}$$

By Lemma 4.2, $\mathbf{I}_{N'/(qK)} \otimes \mathbf{L}_q \otimes \mathbf{I}_K$ is a doubly stochastic matrix; So, $\mathbf{z} \prec_+ \mathbf{y}$ [95].

At the high SNR region, where $\Gamma_{st-gmd}^{(K)}(n), \Gamma_{st-gmd}^{(qK)}(n) \in \mathcal{R}_{high}$, we have $c_2^2 e^{2y_n} \geq 1$ and $c_2^2 e^{2z_n} \geq 1$ for all n. In this domain, the function $\psi(\cdot)$ is Schur-convex. Since $\mathbf{z} \prec_+ \mathbf{y}$, one can conclude that $\psi(\mathbf{z}) \leq \psi(\mathbf{y})$. At the low SNR region, where $\Gamma_{st-gmd}^{(K)}(n), \Gamma_{st-gmd}^{(qK)}(n) \in \mathcal{R}_{low}$, we can prove $\psi(\mathbf{y}) \leq \psi(\mathbf{z})$ similarly. □

At the high SNR region, from Theorem 4.4, we can conclude that $\mathcal{P}_{st-gmd}^{(qK)}$ is a non-increasing function of q. As the ST-block size gets larger, the BER performance of ST-GMD improves monotonically. Larger ST-block size is more favorable because it gains more diversity from the time-varying channels. However, it implies a longer decoding delay at the receiver. At the low SNR region, the relationship is the other way around, so it is better to have small ST-block size.

4.6.3 Performance Comparison in Capacity

In the conventional GMD-based system, the average BER per ST-block is dominated by the block with the largest MSE. To achieve the optimal per ST-block average BER and hence minimize the average BER, bit allocation is required. The proposed ST-GMD transceiver does not require bit allocation among blocks since all SNRs of different blocks in each ST-block are the same. In this subsection, from the perspective of capacity, we will show the asymptotic optimality of ST-GMD transceiver.

With uniform power loading, the ergodic channel capacity [96] for the equivalent channel $\mathbf{H}$ in (4.11) of an ST-block is given by

$$\begin{aligned} C_{upl} &= E_{\mathbf{H}}\{\log\big(\det(\mathbf{I}+\rho\mathbf{H}^H\mathbf{H})\big)\} \\ &= E_{\mathbf{H}}\{\sum_{k=0}^{K-1}\sum_{i=0}^{M-1}\log\big(1+\rho\sigma_{H,i}^2(k)\big)\}, \end{aligned} \tag{4.63}$$

where $\rho = (\sigma_a\alpha)^2/\sigma_w^2$ and $\sigma_{H,i}(k)$ is given in (4.4). In the conventional GMD-based system and the ST-GMD transceiver, the channel $\mathbf{H}$ is converted into equivalent parallel subchannels. Hence, the ergodic channel capacities of the equivalent subchannels obtained by using the conventional GMD-based system and the ST-GMD transceiver are respectively given by

$$\begin{aligned} C_{gmd} &= ME_{\mathbf{H}}\{\sum_{k=0}^{K-1}\log\big(1+\rho\sigma_k^2\big)\}, \\ C_{st-gmd} &= MKE_{\mathbf{H}}\{\log\big(1+\rho(\prod_{k=0}^{K-1}\sigma_k^2)^{1/K}\big)\}, \end{aligned} \tag{4.64}$$

where σ_k is given by (4.4). For high SNR,

$$\lim_{\rho\to\infty} C_{upl} - C_{gmd} = \sum_{k=0}^{K-1} E_{\mathbf{H}}\{\log(\frac{\prod_i \rho\sigma_{H,i}^2(k)}{\prod_i \rho\sigma_{H,i}^2(k)})\} = 0. \tag{4.65}$$

So the GMD-based system does not have capacity loss. For ST-GMD transceivers, $\lim_{\rho\to\infty} C_{gmd} - C_{st-gmd} = ME_{\mathbf{H}}\{\log(\prod_k \rho\sigma_k^2/\prod_k \rho\sigma_k^2)\} = 0$. Together with (4.65), we have

$$\lim_{\rho\to\infty} C_{upl} - C_{st-gmd} = 0. \tag{4.66}$$

Therefore, for high SNR, the ST-GMD transceiver is asymptotically optimal in capacity and per

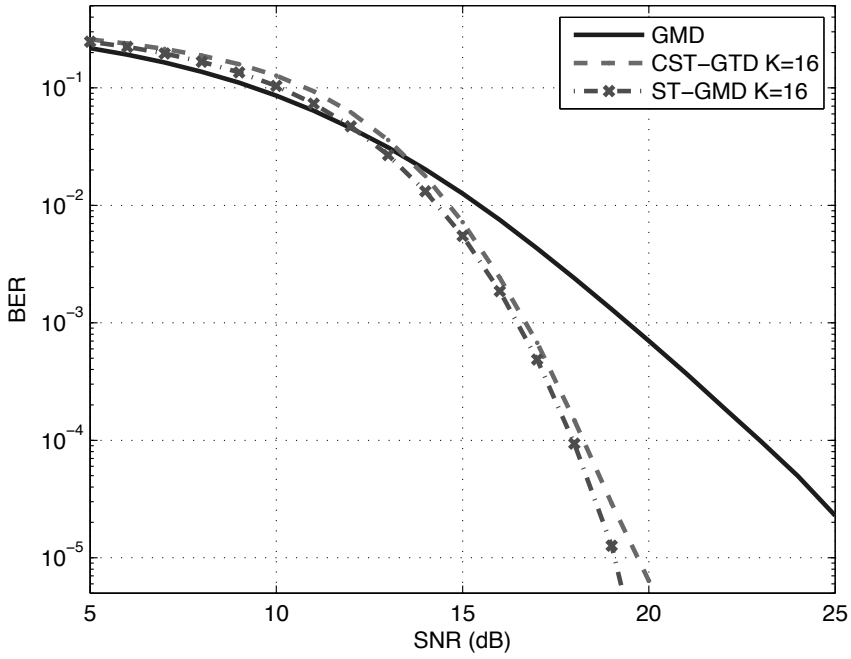


Figure 4.4: BER performance of GMD, ST-GMD and CST-GTD.

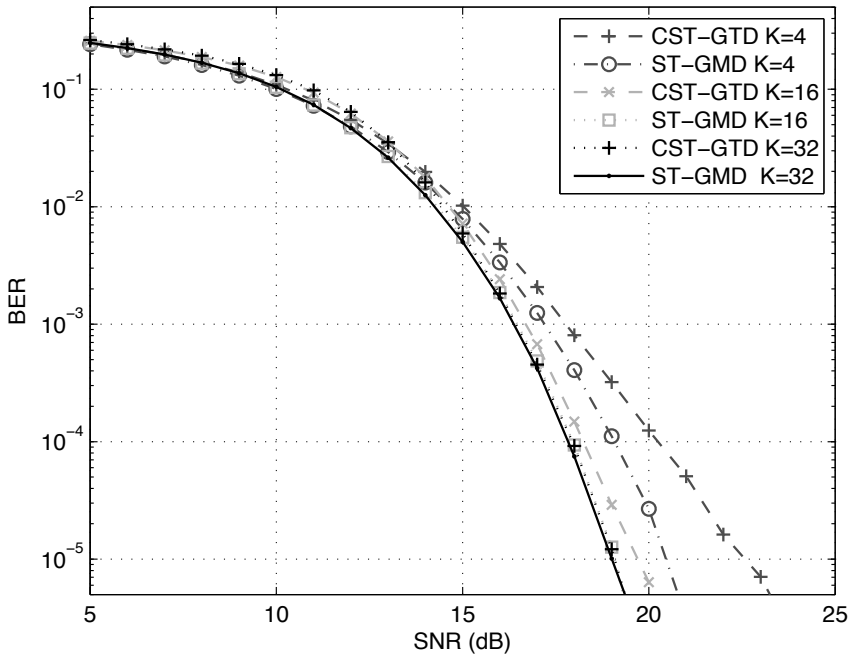


Figure 4.5: BER performance v.s. block size for ST-GMD and CST-GTD.

ST-block average BER simultaneously. Note that the design of the ST-GMD transceiver is possible only when the channel prediction is good.

4.7 Numerical Results

In this section, we present the numerical results on the average BERs of the GMD-based system, ST-GMD and CST-GTD transceivers. We also demonstrate how the ST-block size affects the BER performance. The channel model described in section 4.3.1 is adopted. The noise is AWGN. The MIMO channel matrices $\mathbf{H}(k)$ are 3×3 complex Gaussian random matrices. The elements of $\mathbf{H}(k)$ are i.i.d. complex Gaussian random variables with zero mean and unit variance. Uniform bit allocation is adopted with $b = 4$ for each subchannel. The modulation scheme is 16-QAM. Both transmitter and receiver have perfect CSI at current time k. We assume that perfect channel prediction is available only for the ST-GMD transceiver. The temporal precoding matrix of the CST-GTD is a $K \times K$ DFT matrix. $N' = 2^{20}$ data blocks are sent through the channels for BER performance evaluation.

Example 1: The ST-block size is $K = 16$. $\mathbf{H}(k)$ are independent for different k. Fig. 4.4 shows the BER performance of the conventional GMD-based system, CST-GTD and ST-GMD transceivers for different SNRs. For the high SNR region, Fig. 4.4 satisfies $\mathcal{P}_{st-gmd} < \mathcal{P}_{cst-gtd} < \mathcal{P}_{gmd}$, which verifies Theorem 4.3. At BER 10^{-4}, the SNR gains of the ST-GMD and the CST-GTD over the GMD-based system are 5 dB and 4.7 dB, respectively. The performance of CST-GTD is close to ST-GMD. At BER 10^{-5}, the SNR gain of the ST-GMD transceiver over the CST-GTD transceiver is about 0.6 dB. At the low SNR region, $\mathcal{P}_{st-gmd}$ and $\mathcal{P}_{cst-gtd}$ are greater than $\mathcal{P}_{gmd}$. This is because of the error propagation. For the space-time processing at these receivers, the errors might propagate through the entire ST-block, i.e., K blocks.

Example 2: In this example, various choices of ST-block size are compared. $\mathbf{H}(k)$ are independent for different k. Fig. 4.5 shows $\mathcal{P}_{st-gmd}$ and $\mathcal{P}_{cst-gtd}$. At the low SNR region, $\mathcal{P}_{st-gmd}$ and $\mathcal{P}_{cst-gtd}$ increase with respect to K as in shown Fig. 4.7. But at the high SNR region, $\mathcal{P}_{st-gmd}$ and $\mathcal{P}_{cst-gtd}$ decrease with respect to K which is best illustrated by Fig. 4.6. These results verify Theorem 4.4. Notice that the CST-GTD transceiver almost has the same performance as the ST-GMD transceiver when $K = 32$. The SNR gap is only 0.08 dB at BER 10^{-5}.

Example 3: Here, the BER performances of the three transceivers are evaluated using Jakes' channel model. $\mathbf{H}(k)$ for different k are correlated and the cross-correlation is given by (4.2). Fig. 4.8

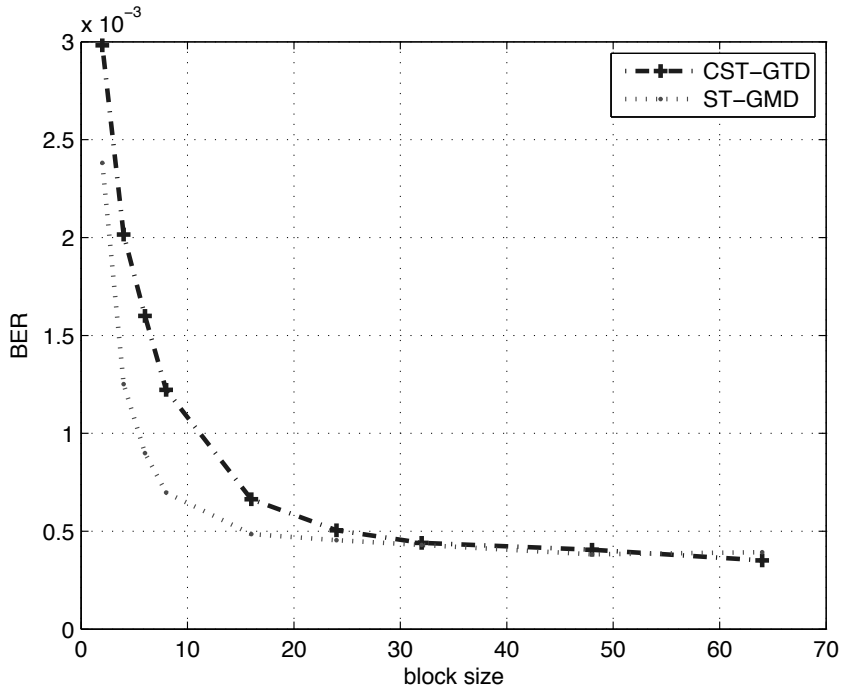


Figure 4.6: BER performance v.s. block size K at SNR=17 dB.

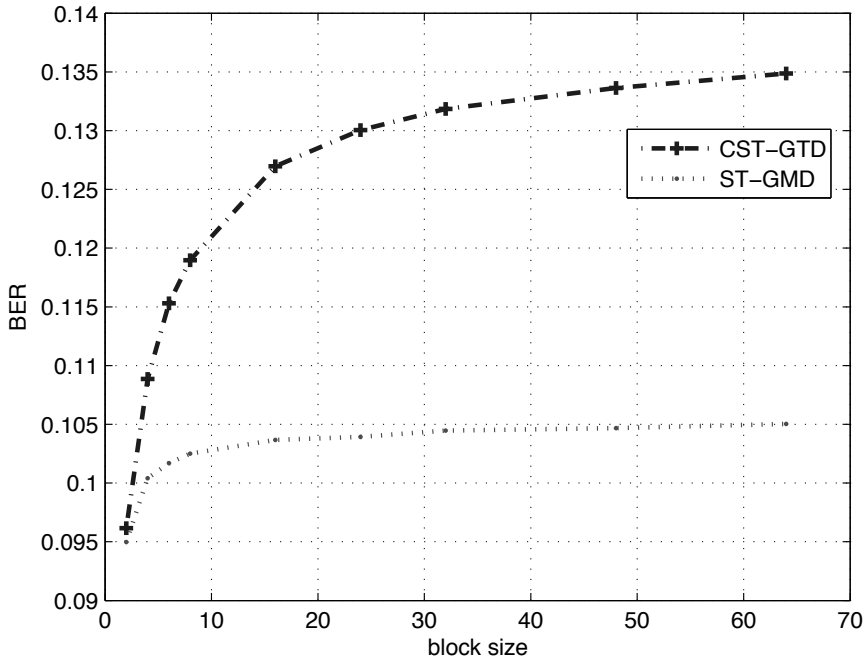


Figure 4.7: BER performance v.s. block size K at SNR=10 dB.

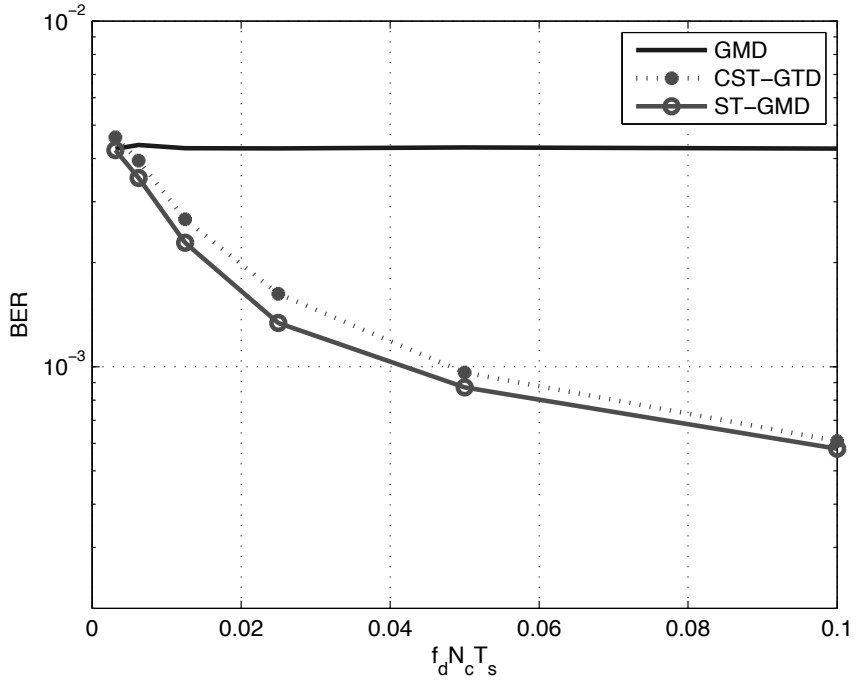


Figure 4.8: BER performance v.s. $f_dN_cT_s$ at $SNR = 17$ dB.

shows the average BER performances of the conventional GMD-based system, CST-GTD and ST-GMD transceivers for different values of the product $f_dN_cT_s$, which appears in (4.2). As $f_dN_cT_s$ gets larger, the channels are changing at faster rates, and $\mathbf{H}(k)$ for different k become more uncorrelated. The ST-block size is $K = 16$ and the BERs are evaluated at $SNR = 17$ dB. For small $f_dN_cT_s$, the channels are almost like time invariant channels; the BER improvements of CST-GTD and ST-GMD transceivers over the conventional GMD-based system are small since there is not much temporal diversity for the temporal precoders to exploit. As $f_dN_cT_s$ increases, the average BERs of CST-GTD and ST-GMD transceivers drop quickly due to the rich temporal diversity offered by the time-varying channels.

Example 4: This example demonstrates the BER performance of ST-GMD transceiver with imperfect channel prediction. $\mathbf{H}(k)$ follows the Jakes' model in Sec. 4.3.1. The ST-block size $K = 4$, and $f_dN_cT_s = 0.1$. The ST-GMD transceiver is designed according to the procedure for the case of imperfect channel prediction in Sec. 4.4.1. Fig. 4.9 illustrates the BER performance of the ST-GMD transceiver based on channel prediction. Its BER curve follows the curve of the ST-GMD transceiver with perfect channel prediction closely for most of the SNR values, and deviates at the very high

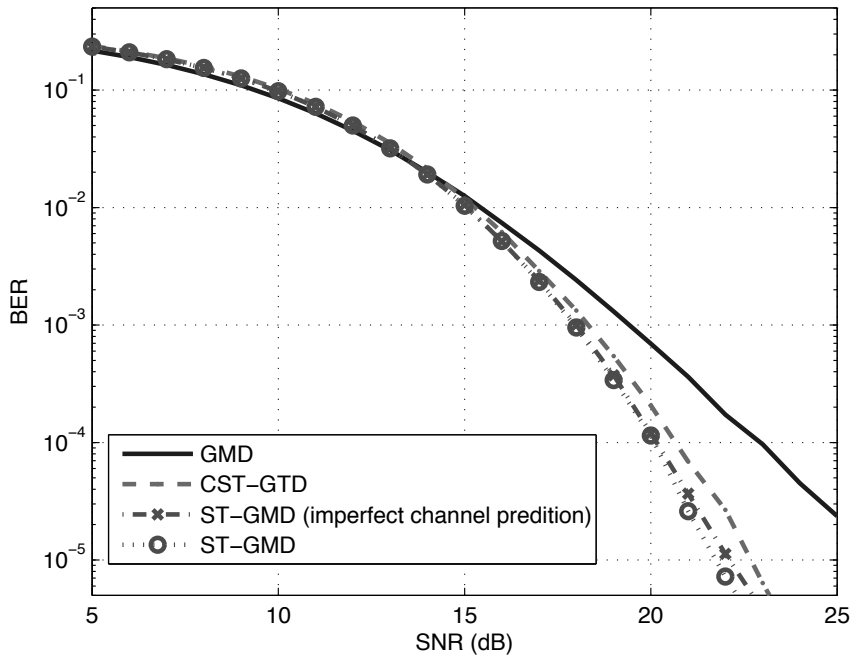


Figure 4.9: BER performance of ST-GMD with channel prediction.

SNR region. The BER degradation results from the channel prediction error. At BER 10^{-5}, the SNR loss from imperfect channel prediction is 0.35 dB.

4.8 Conclusions

We have proposed two MIMO transceivers with zero-forcing decision feedback structure for MIMO slowly time-varying channels. They harvest the rich temporal diversity due to the time-varying nature of the channels to minimize the average BER. The issue of available CSIT for slowly time-varying channels can be addressed using feedback mechanisms or TDD schemes. Under the assumption of perfect channel prediction, the ST-GMD transceiver is shown to be the best in terms of arithmetic MSE and average BER in high SNR. The ST-GMD transceiver serves as a benchmark for performance. The CST-GTD transceivers only require the instantaneous CSIT and CSIR, as the GMD-based systems does. It has the same asymptotic BER performance as the ST-GMD transceiver and has smaller arithmetic MSE than the conventional GMD-based systems. The dependency of BER on the ST-block size has also been analyzed. Simulations show that only moderate ST-block size is required for good average BER performance.

4.9 Appendix

4.9.1 Proof of Lemma 4.2

Let $[\mathbf{A}]_{ij} = a_{ij}$, $[\mathbf{B}]_{ij} = b_{ij}$ and $[\mathbf{C}]_{ij} = c_{ij}$. By the definition of Kronecker product,

$$c_{i_1 m+j_1, i_2 m+j_2} = a_{i_1,i_2} b_{j_1,j_2} \geq 0,$$

where $0 \leq i_1, i_2 \leq n-1$ and $0 \leq j_1, j_2 \leq m-1$. Since $\mathbf{A}$ and $\mathbf{B}$ are doubly stochastic,

$$\sum_{i=0}^{n-1} a_{ij} = \sum_{j=0}^{n-1} a_{ij} = \sum_{i=0}^{m-1} b_{ij} = \sum_{j=0}^{m-1} b_{ij} = 1. \tag{4.67}$$

Consider

$$\sum_{i=0}^{mn-1} c_{i,j} = \sum_{i_1=0}^{n-1} \sum_{j_1=0}^{m-1} c_{i_1 m+j_1, i_2 m+j_2} = \sum_{i_1} \sum_{j_1} a_{i_1,i_2} b_{j_1,j_2},$$

where $j = i_2 m + j_2$. By (4.67), we have

$$\sum_{i=0}^{mn-1} c_{i,j} = \sum_{i_1} a_{i_1,i_2} \sum_{j_1} b_{j_1,j_2} = 1. \tag{4.68}$$

Similarly, we can prove $\sum_{j=0}^{mn} c_{i,j} = 1$. Therefore, $\mathbf{C}$ is also a doubly stochastic matrix.

Chapter 5

MMSE DFE Transceiver Design over Slowly Time-varying MIMO Channels Using ST-GTD

In the previous chapter, we studied the zero-forcing (ZF) transceiver with decision feedback equalizer (DFE) over slowly time-varying narrowband multi-input multi-output (MIMO) channels. This chapter addresses the design problem of a DFE transceiver without zero-forcing constraint. In the first part, a channel independent temporal precoder is superimposed on the conventional block-wise GMD-based minimum mean square error (MMSE) DFE transceiver to take advantage of the temporal diversity. In the second part, ST-GTD is applied for the design of MMSE DFE transceivers. With accurate channel prediction and space-time powerloading, the proposed ST-GMD MMSE transceiver minimizes the arithmetic MSE at the feedack detector, and maximizes Gaussian mutual information. For practical applications, the ST-GTD MMSE transceiver, which does not require channel prediction but shares the same asymptotic BER performance with the ST-GMD MMSE system, is also developed. In the convex region, our analysis shows that the proposed MMSE transceiver has better BER performance than the conventional GMD-based MMSE transceiver; the average BERs of the proposed systems are non-increasing functions of the ST-block size. The superior performance of thes ST-GMD MMSE transceiver over the ST-GMD ZF transceiver is also verified analytically. Most of the results in this chapter have been reported in our journal paper [39].

5.1 Introduction

In Chapter 4 and [36], we considered the design of zero-forcing decision feedback equalizer (ZF-DFE) transceivers over slowly time-varying multi-input multi-output (MIMO) memoryless channels. We used the block fading MIMO channel model in which the channel is constant over the coherence interval of N_c blocks and varies across different coherence intervals according to Jakes' model [29], [98]. It is assumed that both instantaneous channel state information (CSI) at the transmitter (CSIT) and the receiver (CSIR) are available.

In [36], we proposed the geometric mean decomposition (GMD) ZF transceiver with a channel independent temporal precoder (GMD-TP). The GMD-TP only requires instantaneous CSIR and CSIT. The temporal precoder equalizes the mean square errors (MSEs) in each space-time block (ST-block) (which consists of several coherence blocks) such that the average BER per ST-block is minimized. However, the temporal precoder does not alter the per ST-block total MSE. In Chapter 4, based on the proposed space-time generalized triangular decomposition (ST-GTD), we were able to develop an ST-GMD ZF transceiver which minimizes both the arithmetic MSE, and the average BER in high signal to noise ratio (SNR) per ST-block. The design of the ST-GTD transceiver requires instantaneous CSIT, CSIR and channel prediction. For practical implementation, a causal ST-GTD ZF transceiver with constant temporal precoder which does not require channel prediction was proposed. For all ST-GTD transceivers, the temporal precoder and the newly proposed "nested loop" receiver structure not only redistribute the MSEs in each ST-block, but also reduce the per ST-block arithmetic MSE. However, the proposed transceivers are mutual information lossy because of the zero-forcing constraint.

In this chapter, we consider the design of DFE transceivers without zero-forcing constraint. For linear time invariant (LTI) MIMO channels, the GMD-based MMSE system (GMDM) [19], [18] is shown in [19] to minimize the arithmetic MSE, the geometric MSE at the feedback detector, and BER at high SINR region. Moreover, the design also maximizes the Gaussian mutual information. Unlike the singular value decomposition based systems which require bit allocation to achieve the the optimal average BER [35], GMD-based systems do not require bit allocation since all the effective subchannels have the same SINR [17]. However, in the case of time-varying channels, different data blocks pass through different MIMO channels. If the GMD-based MMSE system is directly applied to the time-varying channels, the performance measures are not optimized.

In the first part, we propose the GMD MMSE transceiver with a channel independent temporal

precoder (GMDM-TP). A constant unitary temporal precoder is superimposed on the block-wise GMD-based MMSE transceiver. Only instantaneous CSIT and CSIR are required. The temporal precoder equalizes the MSEs for each ST-block so that the average per block BER is minimized in the convex region as GMD-TP [36]. The convex region here corresponds to moderate high SINR, or to all SINR values depending on the modulation constellation. However, the arithmetic MSE of GMDM-TP is smaller than that of GMD-TP [36] since the underlying GMD MMSE system has smaller total MSE than the zero-forcing case in each block [91]. In the second part, ST-GTD is applied for the design of ST-GTD MMSE transceivers. Inherited from the ST-GTD ZF transceiver, the ST-GMD MMSE transceiver requires instantaneous CSIT, CSIR, and channel prediction. With perfect channel prediction, the proposed space-time GMD MMSE transceiver (ST-GMDM), with space-time power-loading, maximizes Gaussian mutual information over the equivalent channel seen by each ST-block which is not possible for the ST-GMD ZF transceiver in Chapter 4. Moreover, it jointly minimizes the arithmetic MSEs for each ST-block and the per ST-block average BER in the convex region (see below). Hence, the average BER over the entire transmission is reduced. Furthermore, we show that the ST-GMDM transceiver outperforms ST-GMD ZF transceiver in terms of arithmetic MSE and average BER. For the applications where channel prediction is not practical, we propose the "causal" ST-GTD MMSE transceiver (CST-GTDM) which does not require channel prediction. The simulation shows the CST-GTDM has asymptotic BER performance as the ST-GMD MMSE transceiver when the ST-block size increases. In any case, ST-GMD MMSE transceiver serves as a performance benchmark for the general class of ST-GTD MMSE transceiver.

In the convex region, which corresponds to very low and moderate high SINR, or all SINR values, depending on the modulation constellation, the analysis shows that the proposed MMSE transceiver has better BER performance than the conventional GMD-based MMSE transceiver. Moreover, if the ST-block size is 2^n, then the average BERs of the proposed systems are non-increasing functions of n in the convex region and non-decreasing functions of n in the concave region. Simulations also demonstrate the superior performance of the proposed MMSE transceivers over the corresponding ZF transceivers.

5.2 Outline

The sections are structured as follows. In Section 5.3, we introduce the time-varying channel model and review the GMD-based MMSE transceiver without zero-forcing constraint [19], [18]. In Section

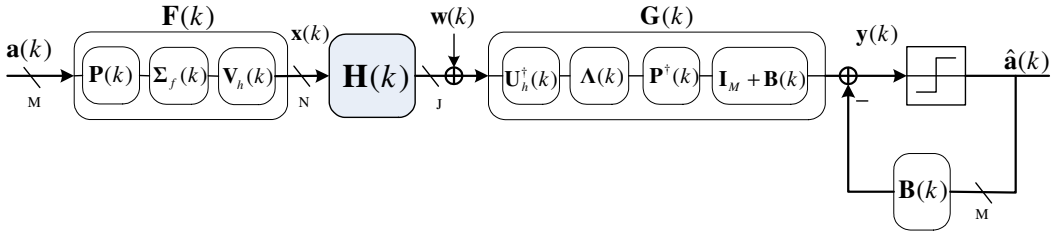


Figure 5.1: The GMD-based system

5.4, we derive the optimal channel independent temporal precoder for the GMD-based MMSE transceiver. Section 5.5 is devoted to the derivation of the optimal ST-GTD MMSE transceiver which minimizes the arithmetic MSE and average BER. A practical suboptimal ST-GTD MMSE transceiver which does not require channel prediction is proposed in Section 5.6. In Section 5.7, we analyze the performance of the proposed transceivers. Numerical examples of BER performances are given in Section 5.8. Concluding remarks are given in Section 5.9.

5.3 Preliminaries

In this chapter, we use the same narrowband block fading MIMO channel model as described in Sec. 4.3.1 of Chapter 4.

5.3.1 GMD-Based MMSE Transceivers

Fig. 1 shows the optimal GMD MMSE transceiver, which has been shown to be optimal in average BER at high SNR and maximizes the Gaussian mutual information for LTI channels [19] and [18]. Now, we apply the structure directly to the time-varying channel $\mathbf{H}(k)$ and refer to it as the "conventional" GMD-based MMSE transceiver. It is assumed that $E(\mathbf{a}(k)\mathbf{a}^\dagger(k')) = \sigma_a^2\delta(k-k')\mathbf{I}_M$. Since both the transmitter and receiver have perfect CSI at current block time k, the $N \times M$ precoding matrix $\mathbf{F}(k)$, the $M \times J$ feedforward matrix $\mathbf{G}(k)$ and the $M \times M$ feedback matrix $\mathbf{B}(k)$ are determined by $\mathbf{H}(k)$. The $J \times M$ matrix $\mathbf{U}_h(k)$ and the $N \times M$ matrix $\mathbf{V}_h(k)$ can be obtained from the singular value decomposition (SVD) of $\mathbf{H}(k)$,

$$\mathbf{H}(k) = \mathbf{U}_h(k)\mathbf{\Sigma}_h(k)\mathbf{V}_h^\dagger(k), \tag{5.1}$$

where $\mathbf{U}_h^{\dagger}(k)\mathbf{U}_h(k) = \mathbf{I}_M$, $\mathbf{V}_h^{\dagger}(k)\mathbf{V}_h(k) = \mathbf{I}_M$ and

$$\mathbf{\Sigma}_h(k) = diag\big(\sigma_{h,0}(k), \sigma_{h,1}(k), \ldots, \sigma_{h,M-1}(k)\big). \tag{5.2}$$

Here, $\sigma_{h,i}(k)$ is the ith singular value of $\mathbf{H}(k)$. The power loading matrix $\mathbf{\Sigma}_f(k)$ is a diagonal matrix given by

$$\mathbf{\Sigma}_f(k) = diag\big(\sigma_{f,0}(k), \sigma_{f,1}(k), \ldots, \sigma_{f,M-1}(k)\big), \tag{5.3}$$

where $\sigma_{f,i}^2(k)$ are determined by the water-filling formula

$$\sigma_{f,i}^2(k) = \begin{cases} \frac{1}{\lambda(k)} - \frac{\sigma_w^2}{\sigma_a^2 \sigma_{h,i}^2(k)}, & 0 \le i \le L(k)-1 \\ 0, & L(k) \le i \le M-1. \end{cases} \tag{5.4}$$

Here, $L(k)-1$ is the largest number such that $1/\lambda(k) - \sigma_w^2/(\sigma_a^2 \sigma_{h,L(k)-1}^2(k))$ is positive, and

$$\frac{1}{\lambda(k)} = \frac{1}{L(k)\sigma_a^2}\Big(P_0 + \sigma_w^2 \sum_{i=0}^{L(k)-1} \frac{1}{\sigma_{h,i}^2(k)}\Big), \tag{5.5}$$

where $P_0 = Tr(E(\mathbf{x}(k)\mathbf{x}^{\dagger}(k)))$. To determine the $M \times M$ matrices $\mathbf{R}(k)$ and $\mathbf{P}(k)$, we first construct the $M \times M$ diagonal matrix

$$\mathbf{\Theta}(k) = \big(\mathbf{I}_M + \frac{\sigma_a^2}{\sigma_w^2}\mathbf{\Sigma}_f^2(k)\mathbf{\Sigma}_h^2(k)\big)^{1/2}. \tag{5.6}$$

$\mathbf{R}(k)$ and $\mathbf{P}(k)$ can be obtained by the GMD [17] of $\mathbf{\Theta}(k)$, which is

$$\mathbf{\Theta}(k) = \mathbf{Q}(k)\mathbf{R}(k)\mathbf{P}^{\dagger}(k), \tag{5.7}$$

where $\mathbf{P}(k)$ and $\mathbf{Q}(k)$ are unitary matrices, and $\mathbf{R}(k) \in \mathbb{C}^{M \times M}$ is an upper triangular matrix with $\mathbf{r}(k)$ on the diagonal. The $M \times 1$ vector $\mathbf{r}(k)$ has equal elements

$$r_l(k) = \sigma_k = \big(\prod_{i=0}^{M-1} [\mathbf{\Theta}(k)]_{i,i}\big)^{1/M}. \tag{5.8}$$

Then, the $M \times M$ feedback matrix $\mathbf{B}(k)$ is given by

$$\mathbf{B}(k) = \mathbf{D}_{\mathbf{R}}^{-\mathbf{1}}(k)\mathbf{R}(k) - \mathbf{I}_M, \tag{5.9}$$

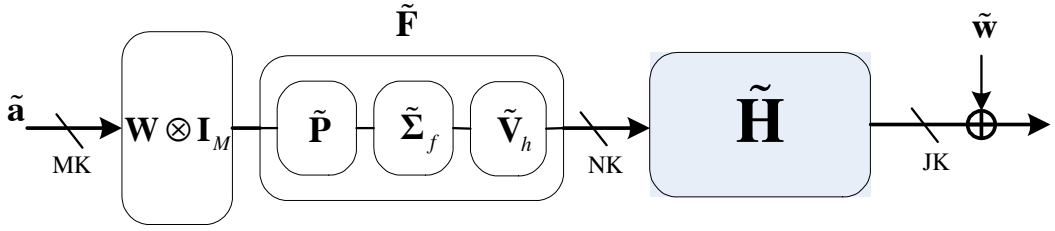


Figure 5.2: The transmitter of GMDM-TP and the channel

where $\mathbf{D_R}(k) = diag(\mathbf{r}(k))$. The $M \times M$ diagonal matrix $\mathbf{\Lambda}(k)$ in $\mathbf{G}(k)$ can be obtained by

$$\mathbf{\Lambda}(k) = \mathbf{\Sigma}_f(k)\mathbf{\Sigma}_h(k)\big(\mathbf{\Sigma}_f^2(k)\mathbf{\Sigma}_h^2(k) + \frac{\sigma_w^2}{\sigma_a^2}\mathbf{I}\big)^{-1}. \tag{5.10}$$

If there is no error propagation in the decision feedback loop, the received signal vector in front of the detector is given by

$$\mathbf{y}(k) = \mathbf{a}(k) + \mathbf{e}(k), \tag{5.11}$$

where

$$\mathbf{e}(k) = (\mathbf{G}(k)\mathbf{H}(k)\mathbf{F}(k) - \mathbf{B}(k) - \mathbf{I}_M)\mathbf{a}(k) + \mathbf{G}(k)\mathbf{w}(k).$$

Following the same derivation as in [19], the error covariance matrix of $\mathbf{e}(k)$ can be written as

$$\mathbf{R_{ee}}(k) = E(\mathbf{e}(k)\mathbf{e}^{\dagger}(k)) = \sigma_a^2 diag(|\mathbf{r}(k)|)^{-2}. \tag{5.12}$$

The total MSE of the kth block at the detector is

$$\begin{aligned} \xi_{gmd}(k) &= Tr\big(\mathbf{R_{ee}}(k)\big) \\ &= M\sigma_a^2\big(\frac{\sigma_w^2\lambda(k)}{\sigma_a^2}\big)^{\frac{L(k)}{M}}\big(\prod_{i=0}^{L(k)-1}\frac{1}{\sigma_{h,i}^2(k)}\big)^{1/M}. \end{aligned} \tag{5.13}$$

5.4 GMD MMSE Transceivers with Temporal Precoder

In this section, we develop the GMD MMSE transceiver with channel independent temporal precoder (GMDM-TP) based on the conventional GMD-based MMSE system, [19] and [18], by incor-

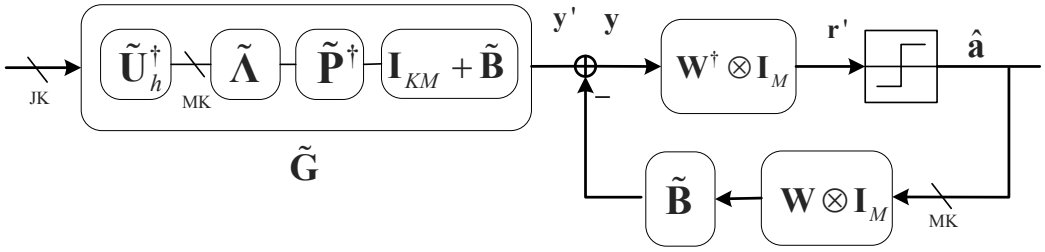


Figure 5.3: The receiver of GMDM-TP

porating a constant $K \times K$ unitary matrix $\mathbf{W}$ as the temporal precoder. The purpose of $\mathbf{W}$ is to equalize the MSEs (at the input of decision device) of different blocks.

The GMDM-TP scheme is shown in Fig. 5.2 and Fig. 5.3. Except for the temporal precoding matrix $\mathbf{W}$, the transceiver has similar structure as the GMD-based system as described in the last section. To facilitate space-time processing, K blocks of symbol vectors are grouped into one space-time block as

$$\tilde{\mathbf{a}}_m = \left[\begin{array}{ccc} \mathbf{a}^T(mK) & \ldots & \mathbf{a}^T(mK+K-1) \end{array} \right]^T, \tag{5.14}$$

where m is the ST-block index. The symbols, m and mK, will be omitted for convenience. The equivalent MIMO channel matrix for the mth ST-block is a $KJ \times KN$ block diagonal matrix given by

$$\tilde{\mathbf{H}} = diag\big(\mathbf{H}(0), \mathbf{H}(1), \ldots, \mathbf{H}(K-1)\big). \tag{5.15}$$

In Fig. 5.2, the noise vector $\tilde{\mathbf{w}} = [\mathbf{w}^T(0), \ldots, \mathbf{w}^T(K-1)]^T$. Let $0 \le k \le K-1$. If SVD is applied to $\mathbf{H}(k)$ separately (in the spatial domain), we have $\mathbf{H}(k) = \mathbf{U}_h(k)\mathbf{\Sigma}_h(k)\mathbf{V}_h^{\dagger}(k)$ as in (5.1). $\tilde{\mathbf{H}}$ can be decomposed as

$$\tilde{\mathbf{H}} = \tilde{\mathbf{U}}_h \tilde{\mathbf{\Sigma}}_h \tilde{\mathbf{V}}_h^{\dagger}, \tag{5.16}$$

where $\tilde{\mathbf{U}}_h$, $\tilde{\mathbf{V}}_h$ and $\tilde{\mathbf{\Sigma}}_h$ are block diagonal matrices with $\mathbf{U}_h(k)$, $\mathbf{V}_h(k)$ and $\mathbf{\Sigma}_h(k)$ on the diagonals, respectively. The power loading matrix $\tilde{\mathbf{\Sigma}}_f$ in the transmitter is a $KM \times KM$ diagonal matrix

$$\tilde{\mathbf{\Sigma}}_f = diag(\mathbf{\Sigma}_f(0), \mathbf{\Sigma}_f(1), \ldots, \mathbf{\Sigma}_f(K-1)), \tag{5.17}$$

where $\mathbf{\Sigma}_f(k)$ is given by (5.3). The $KM \times KM$ block diagonal matrix $\tilde{\mathbf{P}}$ is given by

$$\tilde{\mathbf{P}} = diag\big(\mathbf{P}(0), \mathbf{P}(1), \ldots, \mathbf{P}(K-1)\big), \tag{5.18}$$

where $\mathbf{P}(k)$ is determined by the GMD of $\mathbf{\Theta}(k)$ as in (5.7). The $KM \times KM$ feedback block diagonal matrix $\tilde{\mathbf{B}}$ consists of $\mathbf{B}(k)$ on the diagonal where $\mathbf{B}(k)$ is as (5.9). The $MK \times MK$ diagonal matrix $\tilde{\mathbf{\Lambda}}$ in Fig. 5.3 can be written as

$$\tilde{\mathbf{\Lambda}} = diag\big(\mathbf{\Lambda}(0), \mathbf{\Lambda}(1), \ldots, \mathbf{\Lambda}(K-1)\big), \tag{5.19}$$

where $\mathbf{\Lambda}(k)$ is given by (5.10).

At the transmitter, every ST-block $\tilde{\mathbf{a}}$ is first processed by the temporal precoder $\mathbf{W} \otimes \mathbf{I}_M$, precoded by the spatial precoder $\tilde{\mathbf{F}}$ and then sent through the channel $\tilde{\mathbf{H}}$. At the receiver, the received blocks are first processed with feedforward matrix $\tilde{\mathbf{G}}$ and decoded by the temporal precoder $\mathbf{W}^\dagger \otimes \mathbf{I}_M$. The detector detects the symbols in the ST-block $\tilde{\mathbf{a}}$ sequentially according to the successive cancellation algorithm described in the following. Before summarizing the detection algorithm of the receiver, we need some notations. For $0 \le i \le M-1$, define the $M \times M$ diagonal matrix $\mathbf{E}_i$ as

$$\mathbf{E}_i = diag([\underbrace{0, \ldots, 0}_{M-i}, \underbrace{1, \ldots, 1}_{i}]), \tag{5.20}$$

and the $M \times 1$ vector $\mathbf{s}_i$ in which all the entries are zero except the ith entry which is 1. Based on $\mathbf{E}_i$ and $\mathbf{s}_i$, we define two operators, $\mathbf{\Theta}_i = \mathbf{I}_K \otimes \mathbf{E}_i$ and $\mathbf{S}_i = \mathbf{I}_K \otimes \mathbf{s}_i^T$, on the ST-block vector $\mathbf{a}$ which has the form (5.14). $\mathbf{\Theta}_i$ retains the last i symbols of each $\mathbf{a}(k)$ and makes the other symbols of $\mathbf{a}(k)$ zero. And note that $\mathbf{S}_i\mathbf{a} = [a_i(0), \ldots, a_i(K-1)]^T$.

The successive cancellation algorithm:

1. Initialize: $i = 0$.

2. *Feedback:* Calculate

   $\mathbf{S}_{M-i-1}\mathbf{y} = \mathbf{S}_{M-i-1}(\mathbf{y}' - \mathbf{B}(\mathbf{W} \otimes \mathbf{I}_M)\mathbf{\Theta}_i\hat{\mathbf{a}})$.

3. *Detection:* Calculate $\mathbf{z}^{(i)} = \mathbf{S}_{M-i-1}\mathbf{r}' = \mathbf{W}^\dagger\mathbf{S}_{M-i-1}\mathbf{y}$.

   Detect $\mathbf{S}_{M-i-1}\hat{\mathbf{a}} = Qt(\mathbf{z}^{(i)})$, where the function $\mathbf{q} = Qt(\mathbf{t})$ sets $[\mathbf{q}]_k$ to the element in χ such

that it is closest to $[\mathbf{t}]_k$ in Euclidean norm.

4. If $i = M - 1$, then stop, else set $i = i + 1$ and go to 2.

The algorithm can be justified by direct substitution under the no error propagation assumption, i.e., $\mathbf{\Theta}_i\hat{\mathbf{a}} = \mathbf{\Theta}_i\mathbf{a}$. To characterize the performance of the detector, we consider the error covariance matrix of the error signal component, $\mathbf{e} = \mathbf{r}' - \mathbf{a}$, in front of the detector. If there is no error propagation, i.e., $\hat{\mathbf{a}} = \mathbf{a}$, $\mathbf{e}$ is given by

$$\mathbf{e} = (\mathbf{W}^\dagger \otimes \mathbf{I}_M)\big((\tilde{\mathbf{G}}\tilde{\mathbf{H}}\tilde{\mathbf{F}} - \tilde{\mathbf{B}} - \mathbf{I})(\mathbf{W} \otimes \mathbf{I}_M)\mathbf{a} + \tilde{\mathbf{G}}\tilde{\mathbf{w}}\big), \tag{5.21}$$

in which the elements of $\tilde{\mathbf{w}}$ are i.i.d complex circular Gaussian, with zero mean and variance σ_w^2. So, we have the error covariance matrix

$$\mathbf{R}_{\mathbf{ee}}(\mathbf{W}) = E(\mathbf{e}\mathbf{e}^\dagger) = \sigma_a^2(\mathbf{W}^\dagger\mathbf{\Sigma}^{-2}\mathbf{W} \otimes \mathbf{I}_M), \tag{5.22}$$

where

$$\mathbf{\Sigma} = diag(\sigma_0, \sigma_1, \ldots, \sigma_{K-1}), \tag{5.23}$$

and σ_k is given by (5.8).

Let $\nu_W^{(i)}(k)$ denote the MSE when the detection device detects the ith symbol of the kth block in an ST-block. It is given by

$$\nu_W^{(i)}(k) = [\mathbf{R}_{\mathbf{ee}}(\mathbf{W})]_{kM+i,kM+i}, \tag{5.24}$$

where $0 \leq i \leq M - 1$ and $0 \leq k \leq K - 1$. From (5.22), one can observe that all the equivalent subchannels of the MIMO channel $\mathbf{H}(k)$ have the same MSE, $\nu_W(k)$, i.e.,

$$\nu_W(k) = \nu_W^{(i)}(k) = \sigma_a^2[\mathbf{W}^\dagger\mathbf{\Sigma}^{-2}\mathbf{W}]_{k,k}, \tag{5.25}$$

for all i. This is because GMD is applied for each $\mathbf{H}(k)$. We denote the MSE vector with the temporal precoder $\mathbf{W}$ as

$$\mathbf{v}_W = [\nu_W(0), \nu_W(1), \ldots, \nu_W(K-1)]^T. \tag{5.26}$$

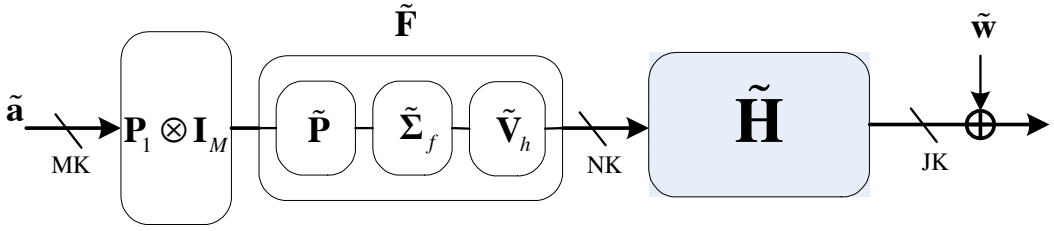


Figure 5.4: The transmitter of the ST-GTD transceiver and the channel

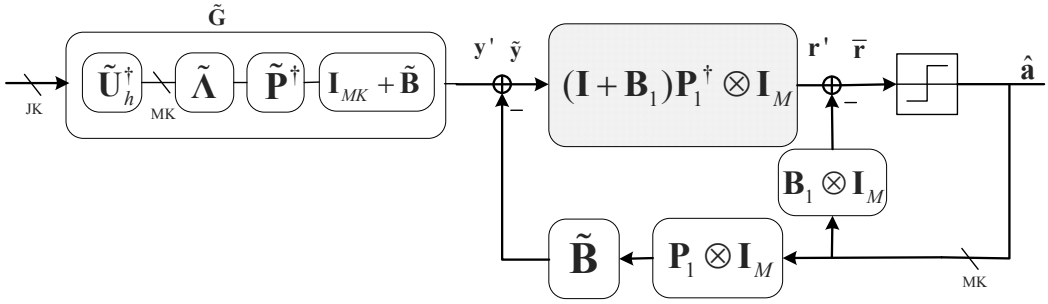


Figure 5.5: The receiver of the ST-GTD transceiver

Note that the conventional GMD-based system is the special case of GMDM-TP with $\mathbf{W} = \mathbf{I}_K$. So, its MSE vector can be written as

$$\mathbf{v}_I = [\nu_I(0), \nu_I(1), \ldots, \nu_I(K-1)]^T, \tag{5.27}$$

where $\nu_I(k)$ is obtained by (5.25) by letting $\mathbf{W} = \mathbf{I}_K$. From (5.25), one can verify that

$$\mathbf{v}_W = \mathbf{T}\mathbf{v}_I, \tag{5.28}$$

where $T_{i,j} = |W_{j,i}|^2$. The matrix $\mathbf{T}$ is a doubly stochastic matrix. By Theorem 1.1, we have

$$\mathbf{v}_W \prec_+ \mathbf{v}_I. \tag{5.29}$$

This implies that any unitary temporal precoder can equalize the MSEs of the blocks in one ST-block, so that $\mathbf{v}_W \prec_+ \mathbf{v}_I$ is satisfied.

If $\mathbf{W}$ is chosen to be the unitary matrix $\mathbf{Z}$ in which

$$|\mathbf{Z}_{ij}| = \frac{1}{\sqrt{K}},\ 0 \le i, j \le K-1, \tag{5.30}$$

one can show that the error covariance matrix has equal diagonal elements [6], [7]. Let $\nu_Z(k)$ be the MSE in the kth block of a ST-block of the GMDM-TP with the temporal precoder $\mathbf{Z}$. From (5.28) and (5.30), one can show that $\nu_Z(k)$ are equal and given by

$$\nu_Z(k) = \frac{\sigma_a^2}{K} \sum_{i=0}^{K-1} \frac{1}{\sigma_i^2}, \tag{5.31}$$

for all k. Hence, $\mathbf{R_{ee}}(\mathbf{Z})$ has equal diagonal elements by (5.24) and (5.25). By Lemma 1.1, one can further show that

$$\mathbf{v}_Z \prec_+ \mathbf{v}_W \prec_+ \mathbf{v}_I, \tag{5.32}$$

where $\mathbf{v}_Z = [\nu_Z(0), \ldots, \nu_Z(K-1)]^T$. The unitary matrix $\mathbf{Z}$ satisfying (5.30) is an optimal precoder among the class of unitary matrices in the sense that it equalizes $[\mathbf{R_{ee}}]_{i,i}$ such that $\mathbf{v}_Z \prec_+ \mathbf{v}_W$. The total MSE over one ST-block for GMDM-TP is independent of the choice of $\mathbf{W}$ and can be shown to be

$$\xi_{gmdm-tp} = Tr\big(\mathbf{R_{ee}}(\mathbf{W})\big) = M\sigma_a^2 \sum_{k=0}^{K-1} \frac{1}{\sigma_k^2}. \tag{5.33}$$

In fact, the total MSE of the conventional GMD-based system $\xi_{gmdm} = \xi_{gmdm-tp}$. However, the BER performance is different since the diagonal entries of $\mathbf{R_{ee}}(\mathbf{W})$ of GMD-TP are redistributed. The BER discussion is relegated to section 5.7.

5.5 Space-Time GTD MMSE Transceivers

In this section, we propose the ST-GTD MMSE DFE transceiver based on the ST-GTD introduced in Chapter 4. The proposed precoder and receiver of ST-GTD transceiver are shown in Fig. 5.4 and Fig. 5.5, respectively. It is assumed that the transmitter could predict the channels $\mathbf{H}(k)$ for $0 \leq k \leq K-1$ before sending a ST-block $\tilde{\mathbf{a}}$ and the receiver can perfectly track the channels. There are well-studied methods [37], [29] which we can exploit here for the channel prediction. In Chapter 4, we used Wiener filters to do channel prediction in Jakes' channel model where $\mathbf{H}(k)$ are correlated for different k. Before the precoding of an ST-block, the Wiener prediction filters predict $\mathbf{H}(1), \ldots, \mathbf{H}(K-1)$ based upon previous P channel matrices $\mathbf{H}(0), \ldots, \mathbf{H}(-P+1)$. In the following subsections, we will introduce the design procedures for the ST-GTD MMSE transceiver based on

perfect and imperfect channel prediction.

5.5.1 Perfect Channel Prediction

Here, we consider the design of the ST-GTD transceiver based on perfect channel prediction and use it as a theoretical foundation for the development of the ST-GTD transceivers with imperfect channel prediction or even without channel prediction in the later section. In either case, this system serves as a benchmark for performance comparisons.

Since the transmitter and receiver have perfect CSI of $\tilde{\mathbf{H}}$, the matrix $\tilde{\mathbf{V}}_h$ in spatial precoder $\tilde{\mathbf{F}}$ and $\tilde{\mathbf{U}}_h^\dagger$ in feedforward matrix $\tilde{\mathbf{G}}$ can be obtained from the block SVD of $\tilde{\mathbf{H}}$ as in (5.16). The power loading matrix $\tilde{\Sigma}_f$ in GMDM-TP only uses the instantaneous CSI, i.e., $\mathbf{H}(k)$ at time k, and allocates power on a per-block basis. Here, since $\mathbf{H}(k)$ for $0 \leq k \leq K-1$ are available by channel prediction, a more sophisticated power allocation algorithm over space and time can be used. The power loading matrix $\tilde{\mathbf{\Sigma}}_f$ takes the form of (5.17) and the diagonal elements $\sigma_{f,i}^2(k)$ are given by the space-time water filling formula,

$$\sigma_{f,i}^2(k) = \begin{cases} \frac{1}{\lambda} - \frac{\sigma_w^2}{\sigma_a^2 \sigma_{h,i}^2(k)}, & 0 \leq i \leq L(k)-1 \\ 0, & L(k) \leq i \leq M-1 \end{cases}, \tag{5.34}$$

where $L(k)-1$ is the largest number such that $1/\lambda - \sigma_w^2/(\sigma_a^2 \sigma_{h,L(k)-1}^2(k))$ is positive. The threshold $1/\lambda$ is determined by the power constraint over an ST-block instead of a block, i.e.,

$$KP_0 = \sigma_a^2 Tr(\tilde{\mathbf{F}}^\dagger \tilde{\mathbf{F}}) = \sigma_a^2 \sum_{k=0}^{K-1} \sum_{i=0}^{M-1} \sigma_{f,i}^2(k). \tag{5.35}$$

Substituting (5.34) into (5.35), we have

$$\frac{1}{\lambda} = \frac{KP_0 + \sum_{k=0}^{K-1} \sum_{i=0}^{L(k)-1} \sigma_w^2/\sigma_{h,i}^2(k)}{\sigma_a^2 \sum_{k=0}^{K-1} L(k)}. \tag{5.36}$$

Since $\tilde{\mathbf{\Sigma}}_f$ can be obtained from (5.34), $\tilde{\Lambda}$ can then be calculated from (5.19).

Moreover, ST-GTD [38] can be applied to get $\tilde{\mathbf{P}}$, $\mathbf{P}_1$, $\tilde{\mathbf{B}}$ and $\mathbf{B}_1$ for the precoder and the receiver as follows. Firstly, we construct the block diagonal matrix

$$\tilde{\mathbf{\Theta}} = diag\big(\mathbf{\Theta}(0), \mathbf{\Theta}(1), \ldots, \mathbf{\Theta}(K-1)\big), \tag{5.37}$$

where $\mathbf{\Theta}(k)$ is given by (5.6). Applying the spatial GMD to $\mathbf{\Theta}(k)$ for all k, we have $\mathbf{\Theta}(k) = \mathbf{Q}(k)\mathbf{R}(k)\mathbf{P}^{\dagger}(k)$ as in (5.7). One can obtain the spatial precoder $\tilde{\mathbf{P}}$ from the block GMD of $\tilde{\mathbf{\Theta}}$ as

$$\tilde{\mathbf{\Theta}} = \tilde{\mathbf{Q}}\tilde{\mathbf{R}}\tilde{\mathbf{P}}^{\dagger}, \tag{5.38}$$

where $\tilde{\mathbf{P}}$, $\tilde{\mathbf{R}}$ and $\tilde{\mathbf{Q}}$ are block diagonal matrices consisting of $\mathbf{P}(k)$, $\mathbf{R}(k)$ and $\mathbf{Q}(k)$ on the diagonals, respectively. Let

$$\mathbf{d} = [\sigma_0, \ldots, \sigma_{K-1}]^T, \tag{5.39}$$

where σ_i is defined in (5.8). $\tilde{\mathbf{R}}$ can be factored as

$$\tilde{\mathbf{R}} = \mathbf{D}_{\mathbf{R}}(\mathbf{I}_{KM} + \tilde{\mathbf{B}}) \tag{5.40}$$

where $\mathbf{D}_{\mathbf{R}} = \mathbf{\Sigma} \otimes \mathbf{I}_M$, $\mathbf{\Sigma} = diag(\mathbf{d})$, and the spatial feedback matrix $\tilde{\mathbf{B}}$ is a block diagonal matrix with $\mathbf{B}(k)$ on the diagonal. $\mathbf{B}(k)$ are strictly upper triangular $M \times M$ matrices given by (5.9).

Secondly, the temporal GTD of $\mathbf{\Sigma}$ gives the temporal precoder $\mathbf{P}_1$. We decompose $\mathbf{\Sigma}$ as

$$\mathbf{\Sigma} = \mathbf{Q}_1\mathbf{R}_1\mathbf{P}_1^{\dagger}, \tag{5.41}$$

where $\mathbf{P}_1$ and $\mathbf{Q}_1$ are $K \times K$ unitary matrices, and $\mathbf{R}_1$ is a $K \times K$ upper triangular matrix. The necessary and sufficient condition for the GTD in (5.41) to be possible is

$$\mathbf{r}_1 \prec_{\times} \mathbf{d}, \tag{5.42}$$

in which $\mathbf{r}_1$ is a $K \times 1$ vector consisting of diagonal elements of $\mathbf{R}_1$ and $\mathbf{d}$ is given by (5.39). The temporal domain feedback matrix $\mathbf{B}_1$ is given by

$$\mathbf{B}_1 = \mathbf{D}_{\mathbf{R}_1}^{-1}\mathbf{R}_1 - \mathbf{I}_K, \tag{5.43}$$

where $\mathbf{D}_{\mathbf{R}1} = diag(\mathbf{r}_1)$. So, the ST-GTD decomposes $\tilde{\mathbf{\Theta}}$ as

$$\tilde{\mathbf{\Theta}} = \tilde{\mathbf{Q}}\Big((\mathbf{Q}_1^{\dagger}\mathbf{D}_{\mathbf{R}1}(\mathbf{I}_K + \mathbf{B}_1)\mathbf{P}_1) \otimes \mathbf{I}_M)(\mathbf{I}_{KM} + \tilde{\mathbf{B}})\Big)\tilde{\mathbf{P}}^{\dagger}, \tag{5.44}$$

which is entirely different from the direct GMD on $\tilde{\mathbf{\Theta}}$. For the design of the ST-GMD MMSE

transceiver, the temporal GTD of $\mathbf{\Sigma}$ in (5.41) is replaced by the temporal GMD in which all the entries of $\mathbf{r}_1$ in (5.42) equal $(\prod_{k=0}^{K-1}\sigma_k)^{1/K}$.

5.5.2 Imperfect Channel Prediction

In practice, perfect channel prediction is not possible. The mismatch between the ST-GTD transceiver designed by the predicted CSI and the actual channel may result in performance degradation. To alleviate the mismatch caused by channel prediction error, we modify the design procedure.

At the transmitter end, we use the Wiener prediction filter to predict the channel matrix $\tilde{\mathbf{H}}$ as described in Sec. 4.4.2 at the beginning of each ST-block, and denote the predicted channel matrix as $\hat{\mathbf{H}}$. $\tilde{\mathbf{V}}_h$ can be implemented with no channel prediction, since it is block diagonal with $\mathbf{V}_h(k)$ on its diagonal. And the computation of $\mathbf{V}_h(k)$ only relies on the instantaneous CSI $\mathbf{H}(k)$. As for the power loading matrix $\tilde{\mathbf{\Sigma}}_f$, the space-time water filling formula in (5.34) requires all the singular values $\sigma_{h,i}(k)$ of $\tilde{\mathbf{H}}$, so the predicted singular values $\hat{\sigma}_{h,i}(k)$ obtained from $\hat{\mathbf{H}}$ are used instead. The spatial precoder $\tilde{\mathbf{P}}$ is a block diagonal matrix with $\mathbf{P}(k)$ on its diagonal, so we can implement the precoder without predicted CSI. Letting $\mathbf{a}' = (\mathbf{P}_1 \otimes \mathbf{I}_M)\tilde{\mathbf{a}}$, the precoded block at time k is given by

$$[\mathbf{t}]_{kM:kM+M-1} = \mathbf{P}(k)[\mathbf{a}']_{kM:kM+M-1}. \tag{5.45}$$

One can see that only $\mathbf{P}(k)$ is required at block time k. To get the spatial precoder $\mathbf{P}(k)$, we construct $\mathbf{\Theta}(k)$ in (5.6) using $\sigma_{h,i}(k)$ and $\tilde{\mathbf{\Sigma}}_f$, and then do the GMD decomposition as in (5.7). Using the predicted $\hat{\sigma}_{h,i}(k)$, we construct $\tilde{\mathbf{\Theta}}$ as in (5.37) and the temporal precoder $\mathbf{P}_1$ is obtained from the ST-GTD of $\tilde{\mathbf{\Theta}}$ in (5.44).

At the receiver end, $\tilde{\mathbf{H}}$ is available without channel prediction, because the receiver can store the signal until $\tilde{\mathbf{H}}$ is estimated. Using the predicted channel matrix $\hat{\mathbf{H}}$, the receiver first designs the power loading matrix $\tilde{\mathbf{\Sigma}}_f$ and temporal precoder $\mathbf{P}_1$ as at the transmitter. $\tilde{\mathbf{U}}_h$, $\tilde{\mathbf{\Lambda}}$, $\tilde{\mathbf{P}}$, $\tilde{\mathbf{B}}$ and $\mathbf{\Sigma}$ are obtained in a similar fashion as in the last subsection. $\mathbf{R}_1$ are obtained from the QR decomposition of $\mathbf{\Sigma}\mathbf{P}_1$ as

$$\mathbf{\Sigma}\mathbf{P}_1 = \mathbf{Q}_1\mathbf{R}_1. \tag{5.46}$$

The temporal feedback matrix $\mathbf{B}_1$ is given by (5.43).

One can also verify that both the design procedures in this and the last subsection lead to the

same ST-GTD or ST-GMD MMSE transceiver when channel prediction is perfect. For each ST-block time, $\tilde{\mathbf{a}}$ is precoded by the temporal and spatial precoders, transmitted through the channel $\tilde{\mathbf{H}}$ and pre-processed at the receiver by $\tilde{\mathbf{G}}$. The estimation of the transmitted $\tilde{\mathbf{a}}$ at the receiver is done by the successive cancellation algorithm described in Sec. 4.4.3.

5.5.3 Mean Square Error at the Detector

To characterize the performance of the ST-GTD transceiver, we assume perfect channel prediction and calculate the error covariance matrix of the noise component, $\mathbf{e} = \bar{\mathbf{r}} - \tilde{\mathbf{a}}$. Under the assumption of no error propagation in the feedback loop, the error component $\mathbf{e}$ can be expressed as

$$\mathbf{e} = \mathbf{A}[\big(\tilde{\mathbf{G}}\tilde{\mathbf{H}}\tilde{\mathbf{F}} - (\tilde{\mathbf{B}} + \mathbf{I}_M)\big)(\mathbf{P}_1 \otimes \mathbf{I}_M)\mathbf{a} + \tilde{\mathbf{G}}\tilde{\mathbf{w}}], \tag{5.47}$$

where

$$\mathbf{A} = [(\mathbf{I} + \mathbf{B}_1)\mathbf{P}_1^{\dagger} \otimes \mathbf{I}_M]. \tag{5.48}$$

Following a similar derivation procedure as for (5.22), we have the error covariance matrix

$$\begin{aligned} \mathbf{R}_{\mathbf{ee}} &= E(\mathbf{e}\mathbf{e}^{\dagger}) = \sigma_a^2 \mathbf{A}(\mathbf{\Sigma}^{-2} \otimes \mathbf{I}_M)\mathbf{A}^{\dagger} \\ &= \sigma_a^2 [\mathbf{D}_{\mathbf{R}_1}^{-1} \mathbf{D}_{\mathbf{R}_1}^{-\dagger} \otimes \mathbf{I}_M], \end{aligned} \tag{5.49}$$

where $\mathbf{D}_{\mathbf{R}1}$ is as (5.43). The second equality follows from the temporal domain GTD (5.41). The total MSE of the ST-GTD MMSE transceiver over an ST-block is the trace of $\mathbf{R}_{\mathbf{ee}}$, i.e.,

$$\xi_{st-gtdm} = \sigma_a^2 M \sum_{k=0}^{K-1} \frac{1}{|\eta_k|^2} \geq \sigma_a^2 MK \Big(\prod_{k=0}^{K-1} \frac{1}{\sigma_k^2} \Big)^{1/K}, \tag{5.50}$$

where $\eta_k = [\mathbf{r}_1]_k$. The last inequality comes from the AM-GM inequality and $\mathbf{r_1} \prec_{\times} \mathbf{d}$. The lower bound of the total MSE $\xi_{st-gtdm}$ is achieved when the ST-GMD is applied, in which we choose

$$\eta_k = \big(\prod_{i=0}^{K-1} \sigma_i \big)^{1/K}. \tag{5.51}$$

We call this class of ST-GTD MMSE transceiver the ST-GMD MMSE transceiver. The total mean square error of the ST-GMD MMSE transceiver is given by

$$\xi_{st-gmdm} = \sigma_a^2 MK \Big(\prod_{k=0}^{K-1} \frac{1}{\sigma_k^2}\Big)^{1/K}. \tag{5.52}$$

If the conventional spatial water-filling algorithm is used, the power loading matrix $\mathbf{\Sigma}_f$ is the same as $\mathbf{\Sigma}_f$ in (5.17). Then, the total mean square error of the ST-GMD MMSE transceiver is given by

$$\xi_{st-gmdm} = \sigma_a^2 MK \prod_{k=0}^{K-1} \Big(\frac{\sigma_w^2 \lambda(k)}{\sigma_a^2}\Big)^{\frac{L(k)}{MK}} \Big(\prod_{i=0}^{L(k)-1} \frac{1}{\sigma_{h,i}^2(k)}\Big)^{\frac{1}{MK}}.$$

However, the MSE is not minimized since the precoder does not exploit the full knowledge of the channel for power loading. If the space-time power loading algorithm in (5.34) is used, the minimized MSE can be shown to be

$$\xi_{st-gmdm} = \sigma_a^2 MK \Big(\frac{\sigma_w^2 \lambda}{\sigma_a^2}\Big)^t \Big(\prod_{k=0}^{K-1} \prod_{i=0}^{L(k)-1} \frac{1}{\sigma_{h,i}^2(k)}\Big)^{\frac{1}{MK}}, \tag{5.53}$$

where $t = \sum_{k=0}^{K-1} L(k)/(MK)$. This is also the smallest MSE achievable by designing a GMD MMSE transceiver over the big equivalent channel matrix $\tilde{\mathbf{H}}$. The class of ST-GMD MMSE transceivers with space-time power loading is the optimal subclass of ST-GTD MMSE transceivers in terms of total mean square error. Furthermore, the error covariance matrix of a ST-GMD MMSE transceiver has equal diagonal elements. Hence, for every ST-block, the ST-GMD MMSE transceiver minimizes both the arithmetic and geometric MSE, and the average un-coded BER at the high SNR region according to [19].

5.5.4 Complexity

Although the ST-GMD MMSE transceiver and the GMD MMSE transceiver designed for the big channel matrix $\tilde{\mathbf{H}}$ have the same performance, the ST-GMD MMSE transceiver has *lower complexity* both in the operation phase, and the design phase which obtains the precoder and equalizer. In the design phase, both transceivers do the same block SVD as in (5.16), which has complexity $O(KM^3)$. However, ST-GTD requires lower computational complexity, $O(KM^2 + K^2)$, than the complexity of directly applying GTD on $\mathbf{\Theta}$, which is $O(K^2M^2)$ [21], [38]. To evaluate the complexity of op-

erations, we let the systems process one ST-block and compare the number of multiplications and additions. The total complexity of an ST-GTD MMSE transceiver is $O(K^2M + KM^2)$ while that of the GMD MMSE transceiver is $O(K^2M^2 + KM^2)$. More importantly, ST-GTD can be chosen in such a way that channel prediction is not necessary, as we show next.

5.6 Space-Time GTD MMSE Transceivers with Fixed Temporal Precoder

In the previous section, the design of ST-GTD MMSE transceivers relies on channel prediction just as the ST-GTD ZF transceivers in Chapter 4. When the channel prediction becomes unavailable, we can choose $\mathbf{P}_1$ to be a constant DFT or Hadamard matrix just as in Chapter 4 which makes the implementation of the transmitter possible. The conventional spatial water-filling formula in (5.4) is used to obtain power loading matrix $\tilde{\mathbf{\Sigma}}_f$. Now, $\mathbf{R}_1$ is obtained from the QR decomposition of $\mathbf{\Sigma W}$ as

$$\mathbf{\Sigma W} = \mathbf{Q}_1\mathbf{R}_1, \tag{5.54}$$

where $\mathbf{Q}_1$ is also a $K \times K$ unitary matrix and $\mathbf{R}_1$ is a $K \times K$ upper triangular matrix. The temporal feedback matrix $\mathbf{B}_1$ is given by (5.43). The detection at the receiver is exactly the same as in the ST-GTD MMSE transceiver (Fig. 5.5). We call this kind of transceiver the causal ST-GTD MMSE transceiver (CST-GTDM). It is in fact a subclass of the ST-GTD MMSE transceiver. The error covariance matrix of the error component in front of the detector is given by (5.49). And the total mean square error $\xi_{cst-gtdm}$ is given by (5.50).

5.6.1 Comparison of Mean Square Error

In this subsection, we compare the performance of the conventional GMD-based MMSE system [19], the CST-GTD MMSE transceiver, and the ST-GMD MMSE transceiver with perfect channel prediction. The total MSE of the GMD-based MMSE system in one ST-block is

$$\xi_{gmdm} = \sigma_a^2 M \sum_{k=0}^{K-1} \frac{1}{\sigma_k^2}. \tag{5.55}$$

The comparison is presented in Theorem 5.1.

Theorem 5.1: The total mean square errors over one ST-block for the three transceivers are such that

$$\xi_{st-gmdm} \leq \xi_{st-gtdm} \leq \xi_{gmdm}. \tag{5.56}$$

□

Proof: The first inequality follows from (5.50). The proof of the second inequality is similar to that of Theorem 4.2 in Chapter 4. □

The result is independent of the chosen power-loading scheme as long as all three transceivers have the same power loading matrix. For simplicity, we assume that only the ST-GMD transceiver uses the space-time water filling formula in (5.34) and the other two use the spatial power loading formula in (5.4). $\xi_{st-gmdm} \leq \xi_{st-gtdm}$ is still true since space-time water filling gives the minimal MSE. The theorem shows that the class of ST-GTD MMSE transceivers has smaller total MSE per ST-block than the conventional GMD-based MMSE system. In particular, for CST-GTDM, we have

$$\xi_{st-gmdm} \leq \xi_{cst-gtdm} \leq \xi_{gmdm}. \tag{5.57}$$

The temporal precoder $\mathbf{P}_1$ or $\mathbf{W}_1$, and the "nested-feedback-loop" receiver in the ST-GMD or CST-GTD MMSE transceiver redistribute the MSEs of the blocks in each ST-block and also help to reduce the per ST-block arithmetic MSE. This is in contrast to the linear temporal precoder of GMDM-TP which keeps the same per ST-block arithmetic MSE while equalizing the MSEs in blocks.

5.7 Performance Analysis

In this section, we compare the average BER of the conventional GMD-based, GMD-TP, ST-GTD, and ST-GMD MMSE transceivers. For both ST-GTD and ST-GMD transceivers, perfect channel prediction is assumed. We assume N' uses of the time-varying channels. Every K successive uses constitute one ST-block. So the mth ST-block uses the channels:

$$\mathbf{H}(mK + k), \text{ for } 0 \leq k \leq K - 1, \tag{5.58}$$

where $0 \leq m \leq \lceil N'/K \rceil - 1$. N' is assumed to be a large number and a multiple of the ST-block size K. Even numbers of bits, b bits, are allocated for every symbol $a_i(k)$ of each ST-block. For square QAM [94], the BER for each symbol in the kth block of the mth ST-block with bias removal, assuming that there is just one bit error per symbol error, is approximately

$$P_e \approx cQ\Big(A\sqrt{\frac{1}{\nu(mK+k)} - \frac{1}{\sigma_a^2}}\Big), \tag{5.59}$$

where $Q(\cdot)$ is the Q-function defined in [94], $A = \sqrt{(3\sigma_a^2)/(2^b-1)}$, σ_a^2 is the average symbol power, $\nu(mK+k)$ is the per symbol MSE of the kth block in the mth ST-block as in (5.25) and $c = (4/b)(1 - 2^{\frac{-b}{2}})$. Note that the symbol error rate (SER) equals bP_e. The average BER over the entire transmission is hence given by

$$\mathcal{P} = \frac{1}{N'} \sum_{m=0}^{(N'/K)-1} \sum_{k=0}^{K-1} cQ\Big(A\sqrt{\frac{1}{\nu(mK+k)} - \frac{1}{\sigma_a^2}}\Big). \tag{5.60}$$

The function $Q(A\sqrt{1/y - 1/\sigma_a^2})$ for $y \in \mathbb{R}^+$ plays a crucial role in BER analysis. An important property of it is restated as the following lemma.

Lemma 5.1: Assume that $y < \sigma_a^2$. The function $f(y) = Q(A\sqrt{1/y - 1/\sigma_a^2})$ is monotone increasing. It is convex when $y \geq t_2$ or $y \leq t_1$ and concave otherwise, where

$$\begin{aligned} t_1 &= \frac{A^2 + 3\sigma_a^2 - \sqrt{A^4 + 9\sigma_a^4 - 10A^2\sigma_a^2}}{8}, \\ t_2 &= \frac{A^2 + 3\sigma_a^2 + \sqrt{A^4 + 9\sigma_a^4 - 10A^2\sigma_a^2}}{8}. \end{aligned} \tag{5.61}$$

If the constellation is chosen to be BPSK and QPSK, $f(y)$ is convex for all y. □

Proof: See [11]. □

We define the SINR of the nth block as $\Gamma(n)$. The SINR expressions of the nth block for the conventional GMD-based MMSE system, the GMDM-TP with precoder $\mathbf{W}$, and the GMDM-TP

with precoder $\mathbf{Z}$ are defined as

$$\begin{aligned}
&\Gamma_{gmd}(n) = \sigma_a^2/\nu_I(n),\\
&\Gamma_{gmdm-tpw}(n) = \sigma_a^2/\nu_W(n),\\
&\Gamma_{gmdm-tpz}(n) = \sigma_a^2/\nu_Z(n),
\end{aligned} \tag{5.62}$$

where $\nu_I(n)$, $\nu_W(n)$ and $\nu_Z(n)$ are given by (5.27), (5.25) and (5.31), respectively. And the SINR expressions for the ST-GTD and ST-GMD MMSE transceivers are given respectively by

$$\begin{aligned}
&\Gamma_{st-gtdm}(n) = |\eta_n^2|,\\
&\Gamma_{st-gmdm}(n) = \big(\prod_{i=0}^{K-1} \sigma^2(mK+i)\big)^{1/K},
\end{aligned} \tag{5.63}$$

which follow from (5.49) and $m = \lfloor n/K \rfloor$. We also define two SINR regions:
For $b \geq 3$,

$$\mathcal{R}_{cvx} = \big\{\Gamma : \Gamma \geq \frac{\sigma_a^2}{t_1} \cup \Gamma \leq \frac{\sigma_a^2}{t_2}\big\}.$$

For $1 \leq b \leq 2$,

$$\mathcal{R}_{cvx} = \big\{\Gamma : \Gamma \in \mathbb{R}^+\big\}. \tag{5.64}$$

And

$$\mathcal{R}_{ccv} = \big\{\Gamma : \Gamma \in \mathbb{R}^+ \backslash \mathcal{R}_{cvx}\big\}. \tag{5.65}$$

For convenience, we define the average BER in (5.60) as the function of the MSEs of all blocks,

$$\phi(\mathbf{z}) = \frac{c}{N'} \sum_{n=0}^{N'-1} Q\Big(A\sqrt{\frac{1}{z_n} - \frac{1}{\sigma_a^2}}\Big), \tag{5.66}$$

where $\mathbf{z} = [z_0, z_1, \ldots, z_{N'-1}]^T$. By Lemma 5.1 and proposition 3.C.1 in [95], $\phi(\mathbf{z})$ is Schur-convex when $z_n \leq t_1$ or $z_n \geq t_2$ and Schur-concave when $t_1 \leq z_n \leq t_2$ for all n and $b \geq 3$. It is Schur-convex if $1 \leq b \leq 2$.

5.7.1 BER Comparison of the Transceivers

In the following theorem, we compare the BER performance of the conventional GMD-based MMSE system and GMDM-TP and prove that $\mathbf{Z}$ in (5.30) is the optimal temporal unitary precoder for GMDM-TP which minimizes the average BER at the convex region. In what follows, the "convex region" means the SINRs of the transceivers are such that $\Gamma(n) \in \mathcal{R}_{cvx}$ and the "concave region" means $\Gamma(n) \in \mathcal{R}_{ccv}$.

Theorem 5.2: Let $\mathcal{P}_{gmdm}$ be the average BER of the conventional GMD-based system, $\mathcal{P}_{gmdmtp-w}$ be the average BER of the GMDM-TP with the temporal precoder $\mathbf{W}$ and $\mathcal{P}_{gmdmtp-z}$ be that with the temporal precoder $\mathbf{Z}$. Then,

$$\mathcal{P}_{gmdmtp-z} \leq \mathcal{P}_{gmdmtp-w} \leq \mathcal{P}_{gmdm}, \text{ at the convex region,}$$
$$\mathcal{P}_{gmdmtp-z} \geq \mathcal{P}_{gmdmtp-w} \geq \mathcal{P}_{gmdm}, \text{ at the concave region.}$$

□

Proof: The proof is similar to that of Theorem 1 in [36]. □

Next, we compare the BER of the entire class of ST-GTD MMSE transceivers including the ST-GMD and CST-GTD MMSE transceivers with the conventional GMD-based MMSE system. The following lemma is helpful for further analysis:

Lemma 5.2: The function $\Delta(y) = Q(c_1\sqrt{e^{2y}-1})$ is monotone decreasing in y where $c_1 = A/\sigma_a$. Let $\eta = e^y$. For constellations up to 16-QAM ($1 \leq b \leq 4$), $\Delta(y)$ is convex for all $y \in \mathbb{R}^+$. Otherwise, it is convex when $|\eta|^2 \leq t_1'$ or $|\eta|^2 \geq t_2'$ and concave when $t_1' < |\eta|^2 < t_2'$, in which

$$t_1' = \left(K_1 + 1 - \sqrt{K_1^2 - 6K_1 + 1}\right)/(4K_1),$$
$$t_2' = \left(K_1 + 1 + \sqrt{K_1^2 - 6K_1 + 1}\right)/(4K_1), \tag{5.67}$$

and $K_1 = 3/(2^b - 1)$. □

Proof: The proof can be found in [23]. □

For the following discussion, we define two more SINR regions:

For $b \geq 5$,

$$\widehat{\mathcal{R}}_{cvx} = \{\Gamma : \Gamma \geq \frac{1}{t_1'} \cup \Gamma \leq \frac{1}{t_2'}\}.$$

For $1 \leq b \leq 4$,

$$\widehat{\mathcal{R}}_{cvx} = \{\Gamma : \Gamma \in \mathbb{R}^+\}. \tag{5.68}$$

And

$$\widehat{\mathcal{R}}_{ccv} = \{\Gamma : \Gamma \in \mathbb{R}^+ \backslash \widehat{\mathcal{R}}_{cvx}\}. \tag{5.69}$$

In the following theorem, "convex region" and "concave region" are defined by the $\widehat{\mathcal{R}}_{cvx}$ and $\widehat{\mathcal{R}}_{ccv}$, respectively.

Theorem 5.3: Let $\mathcal{P}_{st-gtdm}$ be the BER of the ST-GTD MMSE transceiver and $\mathcal{P}_{st-gmdm}$ be that of ST-GMD MMSE transceiver. Then,

$$\mathcal{P}_{st-gmdm} \leq \mathcal{P}_{st-gtdm} \leq \mathcal{P}_{gmdm}, \text{ in the convex region}$$
$$\mathcal{P}_{st-gmdm} \geq \mathcal{P}_{st-gtdm} \geq \mathcal{P}_{gmdm}, \text{ in the concave region.}$$

□

Proof: Using Lemma 5.2 and following similar procedures to those in the proof of Theorem 4.3 in Chapter 4, we can obtain the result. □

From Theorem 5.3, it also follows that $\mathcal{P}_{st-gmdm} \leq \mathcal{P}_{cst-gtdm} \leq \mathcal{P}_{gmdm}$ in the convex region, and the other way around for the concave region.

In the following discussion, let $\mathbf{x} = [x_0, x_1, \ldots, x_{N'-1}]^T$, where

$$x_n = \frac{\sigma_a^2}{|\eta_n|^2}, \tag{5.70}$$

for $n = mK + k$. The average BER of the ST-GTD MMSE transceiver can be expressed as

$$\mathcal{P}_{st-gtdm} = \phi(\mathbf{x}), \tag{5.71}$$

where $\phi(\cdot)$ is defined in (5.66). Now, we define a quantity as

$$\mathcal{P}^*_{st-gtdm} = \phi(\mathbf{y}), \tag{5.72}$$

in which $\mathbf{y} = [y_0, y_1, \ldots, y_{N'-1}]^T$ and $y_{mK+k} = 1/K \sum_{i=0}^{K-1} x_{mK+i}$ for $0 \leq k \leq K-1$. By the similar argument as the proof of Theorem 5.2, one can verify that $\mathcal{P}^*_{st-gtdm}$ is the lower bound of $\mathcal{P}_{st-gtdm}$ in convex region, i.e., $\mathbf{\Gamma}_{st-gtdm}(n) \in \mathcal{R}_{cvx}$ and $\mathbf{\Gamma}^*_{st-gtdm}(n) \in \mathcal{R}_{cvx}$, where $\mathbf{\Gamma}^*_{st-gtdm}(n) = \sigma_a^2/y_n$. At the concave region, $\mathcal{P}^*_{st-gtdm} \geq \mathcal{P}_{st-gtdm}$. Notice that the two bounds are achieved for the class of ST-GMD MMSE transceivers since $x_{mK} = x_{mK+1} = \ldots = x_{mK+K-1}$. That is $\mathcal{P}^*_{st-gmdm} = \mathcal{P}_{st-gmdm}$. In the following, we compare the performance of GMDM-TP and ST-GMD MMSE transceiver. The theorem also gives us some hints about the BER performance of CST-GTDM.

Theorem 5.4: Let $\mathcal{P}_{gmdmtp-z}$ be the average BER of the GMDM-TP with temporal precoder $\mathbf{Z}$, and $\mathcal{P}_{st-gmdm}$ be the BER of the ST-GMD MMSE transceiver. Then,

$$\mathcal{P}_{st-gmdm} \leq \mathcal{P}^*_{st-gtdm} \leq \mathcal{P}_{gmdmtp-z}. \tag{5.73}$$

□

Proof: To prove the second inequality, first observe that the BER of GMDM-TP is given by

$$\mathcal{P}_{gmdmtp-z} = \phi(\mathbf{v}), \tag{5.74}$$

where $\phi(\cdot)$ is defined in (5.66), $\mathbf{v} = [v_0, \ldots, v_{N'-1}]^T$. For $0 \leq k \leq K-1$, v_{mK+k} is given by

$$v_{mK+k} = \frac{\sigma_a^2}{K} \sum_{i=0}^{K-1} \frac{1}{\sigma^2_{mK+i}} = \frac{\xi_{gmdm}(m)}{MK}. \tag{5.75}$$

Notice that $\mathcal{P}^*_{st-gtdm}$ is given by (5.72) and

$$y_{mK+k} = \frac{\sigma_a^2}{K} \sum_{i=0}^{K-1} \frac{1}{|\eta_{mK+i}|^2} = \frac{\xi_{st-gtdm}(m)}{MK}. \tag{5.76}$$

Since Theorem 5.1 shows that $\xi_{st-gtdm}(m) \leq \xi_{gmdm}(m)$, so $y_{mK+k} \leq v_{mK+k}$. As $f(y) = Q(A\sqrt{1/y - 1/\sigma_a^2})$ is a nondecreasing function of y by Lemma 5.1, we can conclude that $\mathcal{P}^*_{st-gtdm} \leq \mathcal{P}_{gmdmtp-z}$.

The BER of the ST-GMD MMSE transceiver is given by

$$\mathcal{P}_{st-gmdm} = \phi(\mathbf{v}'), \tag{5.77}$$

where $\mathbf{v}' = [v'_0, \ldots, v'_{N'-1}]^T$, and by (5.52)

$$v'_{mK+k} = \sigma_a^2 (\prod_{k=0}^{K-1} \frac{1}{\sigma^2_{mK+k}})^{1/K} = \frac{\xi_{st-gmdm}(m)}{MK}. \tag{5.78}$$

Again from Theorem 5.1, we have $v'_{mK+k} \leq y_{mK+k}$. Similarly, we can conclude that $\mathcal{P}_{st-gmdm} \leq \mathcal{P}^*_{st-gtdm}$. □

For the convex region, very low SINR or high SINR, if $\mathcal{P}^*_{cst-gtdm}$ is a tight lower bound of $\mathcal{P}_{cst-gtdm}$, it is very likely, from Theorem 5.4, that $\mathcal{P}_{cst-gtdm} \leq \mathcal{P}_{gmdmtp-z}$; then we have

$$\mathcal{P}_{st-gmdm} \leq \mathcal{P}_{cst-gtdm} \leq \mathcal{P}_{gmdmtp-z} \leq \mathcal{P}_{gmdm}.$$

The simulations in 5.8 also support this conjecture.

5.7.2 Block size and the BER Performance

In this section, the relationship between the size of the ST-block and the BER performance is explored. We first consider the GMDM-TP in Theorem 5.5. In the following theorem, "convex region" and "concave region" are defined by $\mathcal{R}_{cvx}$ and $\mathcal{R}_{ccv}$, respectively.

Theorem 5.5: Let $\mathcal{P}^{(K)}_{gmdmtp-z}$ denote the average BER of the GMDM-TP with ST-block size K over N' channel uses and $\mathcal{P}^{(qK)}_{gmdmtp-z}$ denote that with ST-block size qK, for $q, K \in \mathbb{N}$. Then,

$$\mathcal{P}^{(qK)}_{gmdmtp-z} \begin{cases} \leq \mathcal{P}^{(K)}_{gmdmtp-z} & \text{, in the convex region,} \\ \geq \mathcal{P}^{(K)}_{gmdmtp-z} & \text{, in the concave region.} \end{cases} \tag{5.79}$$

□

Proof: Express $\mathcal{P}^{(K)}_{gmdmtp-z}$ and $\mathcal{P}^{(qK)}_{gmdmtp-z}$ in terms of $\phi(\cdot)$ in (5.66) and follow analog steps in the proof of Theorem 4.4. □

In Theorem 5.6, "convex region" and "concave region" are defined by $\widehat{\mathcal{R}}_{cvx}$ and $\widehat{\mathcal{R}}_{ccv}$, respec-

tively.

Theorem 5.6: Let $\mathcal{P}_{st-gmdm}^{(K)}$ denote the BER of the ST-GMD MMSE transceiver with ST-block size K and $\mathcal{P}_{st-gmdm}^{(qK)}$ denote that with ST-block size qK, for $q, K \in \mathbb{N}$. Then,

$$\mathcal{P}_{st-gmdm}^{(qK)} \begin{cases} \leq \mathcal{P}_{st-gmdm}^{(K)} & \text{, in the convex region,} \\ \geq \mathcal{P}_{st-gmdm}^{(K)} & \text{, in the concave region.} \end{cases} \tag{5.80}$$

□

Proof: Similar to the proof of Theorem 4.4. □

In the convex region, from Theorems 5.5 and 5.6, we can conclude that both $\mathcal{P}_{gmdmtp-z}^{(qK)}$ and $\mathcal{P}_{st-gmdm}^{(qK)}$ are non-increasing functions of q. As the ST-block size gets larger, the BER performance of the GMDM-TP and ST-GMD MMSE transceivers improves monotonically. Larger ST-block size is more favorable because it gains more diversity from the time-varying channels. In the concave region, the relationship is the other way around, so it is better to have small ST-block size.

5.7.3 MMSE v.s. Zero-forcing

In this section, we show explicitly that the ST-GMD MMSE transceiver has better performance than the zero-forcing case in terms of mean square error and bit error rate.

Theorem 5.7: The bias removed per symbol MSE of ST-GMD MMSE transceiver with space-time power loading is smaller than that of the zero-forcing ST-GMD transceiver in Chapter 4, i.e.,

$$\overline{\xi}_{st-gmdm-br} \leq \overline{\xi}_{stgmd_{zf}}, \tag{5.81}$$

where

$$\begin{aligned} \frac{\sigma_a^2}{\overline{\xi}_{st-gmdm-br}} &= \frac{\sigma_a^2}{\overline{\xi}_{st-gmdm}} - 1, \\ \overline{\xi}_{st-gmdm} &= \frac{\xi_{st-gmdm}}{MK}, \end{aligned} \tag{5.82}$$

and from Chapter 4, $\overline{\xi}_{stgmd_{zf}}$ is given by

$$\overline{\xi}_{stgmd_{zf}} = \frac{\sigma_a^2 \sigma_w^2 M}{P_0} \Big(\prod_{k=0}^{K-1} \prod_{i=0}^{M-1} \frac{1}{\sigma_{h,i}^2(k)} \Big)^{1/MK}. \tag{5.83}$$

□

Proof: See Appendix. □

Theorem 5.8: Let $\mathcal{P}_{stgmd-zf}$ be the average bit error probability of the zero-forcing ST-GMD transceiver in Chapter 4. Then,

$$\mathcal{P}_{st-gmdm} \leq \mathcal{P}_{stgmd-zf}. \tag{5.84}$$

□

Proof:

$$\begin{aligned} \mathcal{P}_{st-gmdm} &= \frac{K}{N'} \sum_{m=0}^{(N'/K)-1} cQ\Big(A \sqrt{\frac{1}{\sigma_a^2} \Big(\frac{\sigma_a^2}{\overline{\xi}_{st-gmdm}^{(m)}} - 1 \Big)} \Big) \\ &= \frac{K}{N'} \sum_{m=0}^{(N'/K)-1} cQ\Big(A \sqrt{\frac{1}{\sigma_a^2} \Big(\frac{\sigma_a^2}{\overline{\xi}_{stgmd-br}^{(m)}} \Big)} \Big), \end{aligned} \tag{5.85}$$

and

$$\mathcal{P}_{stgmd-zf} = \frac{K}{N'} \sum_{m=0}^{(N'/K)-1} cQ\Big(A \sqrt{\frac{1}{\sigma_a^2} \Big(\frac{\sigma_a^2}{\overline{\xi}_{stgmd_{zf}}^{(m)}} \Big)} \Big), \tag{5.86}$$

where $\overline{\xi}_{st-gmdm}^{(m)}$ denotes the per symbol MSE of the mth ST-block of the ST-GMD MMSE transceiver, $\overline{\xi}_{stgmdm-br}^{(m)}$ is the bias removed per symbol MSE of the mth ST-block and $\overline{\xi}_{stgmd_{zf}}^{(m)}$ is that of the zero-forcing ST-GMD. The second equality of (5.85) follows from (5.82). By Lemma 4.1 in Chapter 4, $Q(A\sqrt{1/y})$ is a non-increasing function of y. And Theorem 5.7 says $\overline{\xi}_{stgmdm-br}^{(m)} \leq \overline{\xi}_{stgmd_{zf}}^{(m)}$. Hence, we can conclude that $\mathcal{P}_{st-gmdm} \leq \mathcal{P}_{stgmd-zf}$. □

5.7.4 Capacity Performance

In this subsection, we discuss the performance of the ST-GMD MMSE transceiver in terms of Gaussian mutual information. We will show that ST-GMD MMSE transceiver is mutual information lossless. This means the transceiver maximizes the ensemble average of the Gaussian mutual information between the input signal $\tilde{\mathbf{a}}$ and the signal $\bar{\mathbf{r}}$ in front of the detector (see Fig. 5.4 and Fig. 5.5) with respect to $\tilde{\mathbf{H}}$, and the maximized average mutual information is equal to the ergodic channel capacity of $\tilde{\mathbf{H}}$ [96]. The fact that the mutual information is lossless also implies that each substream $[\tilde{\mathbf{a}}]_i$ can use an independent optimal coding scheme to achieve channel capacity.

Under the short-term power constraint, in which the power associated with each channel realization must equal the average power P_0K, the ergodic capacity [96] of the channel $\tilde{\mathbf{H}}$ is

$$\begin{aligned} C &= E_{\tilde{\mathbf{H}}}\Big[\max_{\mathbf{R_x}:Tr(\mathbf{R_x})=KP_0} \log\big(\det(\mathbf{I}+\tilde{\mathbf{H}}^{\dagger}\mathbf{R_x}\tilde{\mathbf{H}}/\sigma_w^2)\big)\Big] \\ &= E_{\tilde{\mathbf{H}}}\Big[\sum_{k=0}^{K-1}\sum_{i=0}^{M-1} \log\big(1+\frac{\sigma_a^2\sigma_{f,i}^2(k)\sigma_{h,i}^2(k)}{\sigma_w^2}\big)\Big] \\ &= E_{\tilde{\mathbf{H}}}\Big[MK\log\Big((\frac{\sigma_a^2}{\lambda\sigma_w^2})^t\big(\prod_{k=0}^{K-1}\prod_{i=0}^{L(k)-1}\sigma_{h,i}^2(k)\big)^{\frac{1}{MK}}\Big], \end{aligned} \tag{5.87}$$

where $\sigma_{f,i}^2(k)$ is given by the space-time water-filling formula (5.34) and t is defined in (5.53). The third equality is obtained by direction substitution with (5.34). In the ST-GMD MMSE transceiver, the channel $\tilde{\mathbf{H}}$ is converted into equivalent parallel subchannels with equal MSEs. After the bias removal at the detector, the MSE of each subchannel is given by $\overline{\xi}_{st-gmdm-br}$ in (5.82). Therefore, the average mutual information of the input and output signals of the equivalent subchannels is

$$\begin{aligned} I_{st-gmdm} &= E_{\tilde{\mathbf{H}}}\Big[MK\log\big(1+\frac{\sigma_a^2}{\overline{\xi}_{st-gmdm-br}}\big)\Big] \\ &= E_{\tilde{\mathbf{H}}}\Big[MK\log\big(\frac{\sigma_a^2}{\overline{\xi}_{st-gmdm}}\big)\Big]. \end{aligned} \tag{5.88}$$

Substituting for $\overline{\xi}_{st-gmdm}$ from (5.82) and (5.53), we have

$$C = I_{st-gmdm}.$$

Hence, we can conclude that the ST-GMD MMSE transceiver is mutual information lossless.

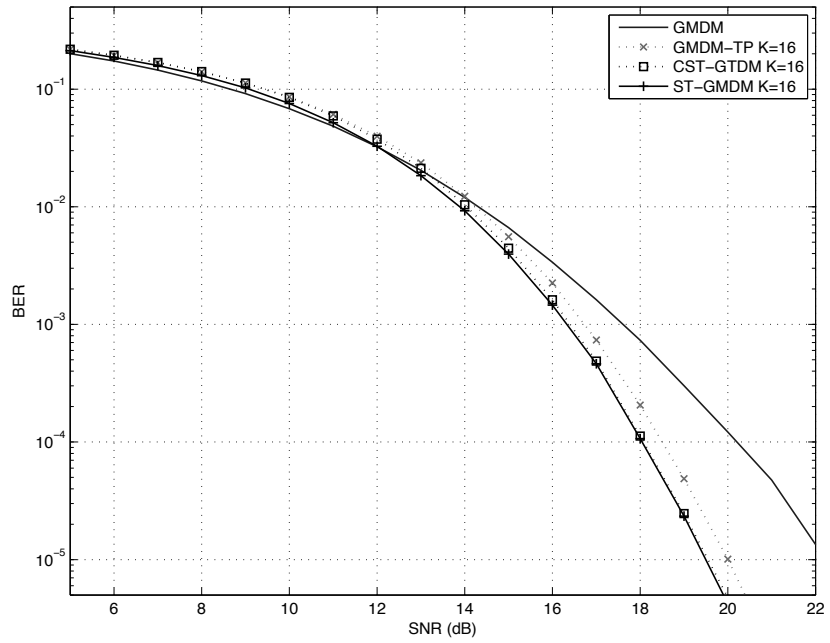


Figure 5.6: BER performance of GMDM, GMDM-TP, ST-GMDM and CST-GTDM.

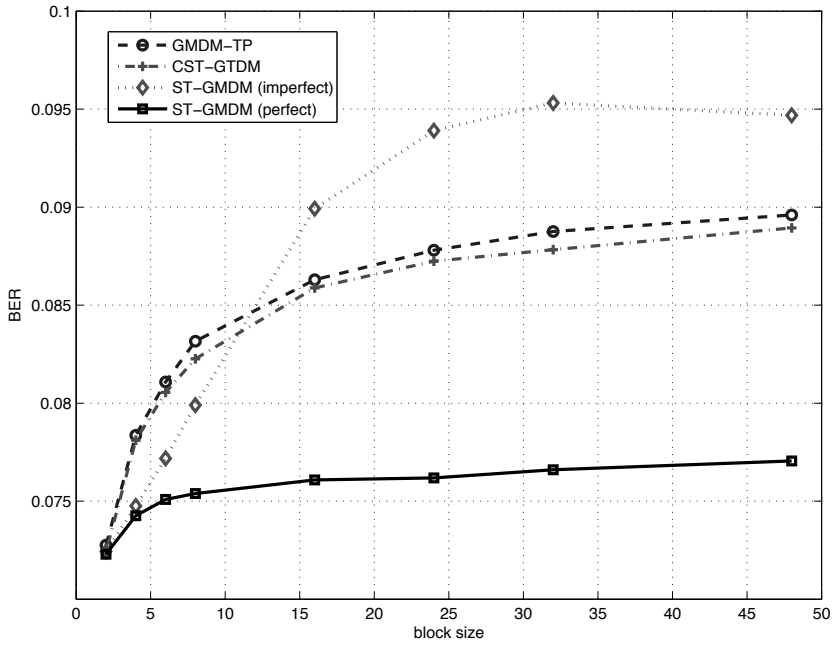


Figure 5.7: BER performance v.s. block size at SNR=10 dB.

5.8 Numerical Results

In this section, simulation results on the average BERs of the GMD-based MMSE system, GMDM-TP, ST-GMD and CST-GTD MMSE transceivers are presented. We also demonstrate how the ST-block size affects the BER performance. Comparisons between the proposed MMSE transceivers and the zero-forcing transceivers in Chapter 4 are also given. The channel model and noise are the same as in Sec. 4.3.1. The channel matrices $\mathbf{H}(k)$ are 3×3 complex Gaussian random matrices. The elements of $\mathbf{H}(k)$ are i.i.d. with zero mean and unit variance. For Examples 1, 2 and 4, Jakes' channel model is adopted. The modulation scheme is 16-QAM. Both transmitter and receiver have perfect CSI at current time k. We assume that perfect channel prediction is available only for the MMSE and zero-forcing ST-GMD transceivers. The temporal precoding matrices of the GMDM-TP and CST-GTDM are $K \times K$ DFT matrices. $N' = 2^{20}$ data blocks are sent through the channels for BER performance evaluation.

Example 1: The ST-block size is $K = 16$. The product $f_d N_c T_s$ which appears in Jakes' model (4.2) is 0.1. Fig. 5.6 shows the BER performance of the conventional GMD-based MMSE system (GMDM) , GMDM-TP, CST-GTDM and ST-GMD MMSE transceivers for different SNRs. For 16-QAM, the entire SNR region is convex. Fig. 5.6 shows $\mathcal{P}_{st-gmdm} < \mathcal{P}_{cst-gtdm} < \mathcal{P}_{gmdmtp-z}$, which verifies Theorem 5.4 and substantiates our conjecture that $\mathcal{P}^*_{cst-gtdm} \approx \mathcal{P}_{cst-gtdm}$. The performance of the CST-GTDM is close to the ST-GMD MMSE transceiver. At BER 10^{-5}, the SNR gain of the ST-GMD MMSE transceiver over the CST-GTDM transceiver is about 0.05 dB. At the high SNR region, $\mathcal{P}_{st-gmdm}$, $\mathcal{P}_{cst-gtdm}$ and $\mathcal{P}_{gmdmtp-z}$ are smaller than $\mathcal{P}_{gmdm}$. This conforms to the nature predicted by Theorem 5.2 and Theorem 5.3. However, at the low SNR region, $\mathcal{P}_{st-gmdm}$ and $\mathcal{P}_{cst-gtdm}$ are greater than $\mathcal{P}_{gmd}$ which contradicts Theorem 5.3. This is not surprising since Theorem 5.3 neglects the effect of error propagation in the decision feedback loop. For ST-GMD and CST-GTD MMSE transceivers, the errors might propagate through the entire ST-block, i.e., K blocks, which is much more severe than what is seen in the conventional GMD-based MMSE transceiver. Hence, in the absence of error propagation which corresponds to the high SNR region, the ST-GMD MMSE transceiver has the best BER performance. At BER 10^{-4}, the SNR gains of the ST-GMD, the CST-GTD and the GMD-TP MMSE transceiver over the GMD-based MMSE system are 2.16 dB, 2.12 dB and 1.7 dB, respectively.

Example 2: In this example, various choices of ST-block size are compared and the performance of the ST-GMD MMSE transceiver with imperfect channel prediction is demonstrated. The order

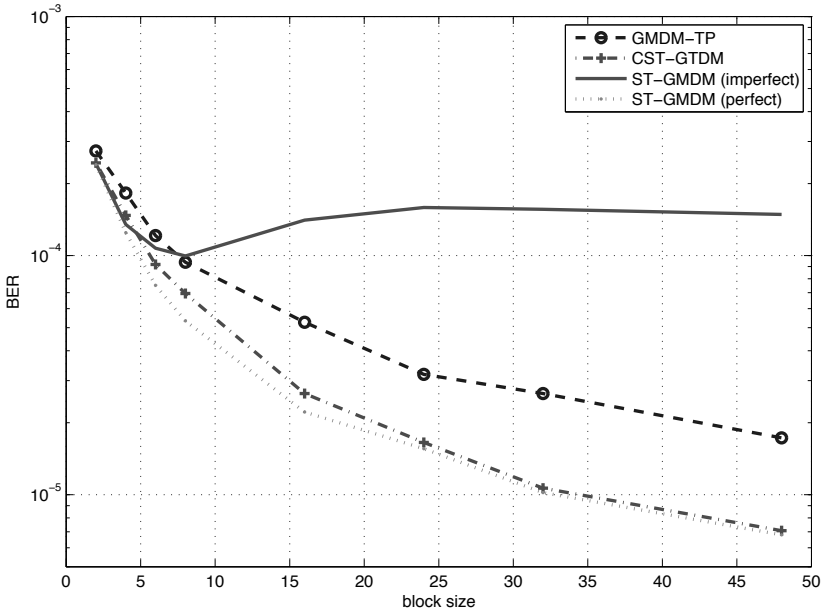


Figure 5.8: BER performance v.s. block size at SNR=19 dB.

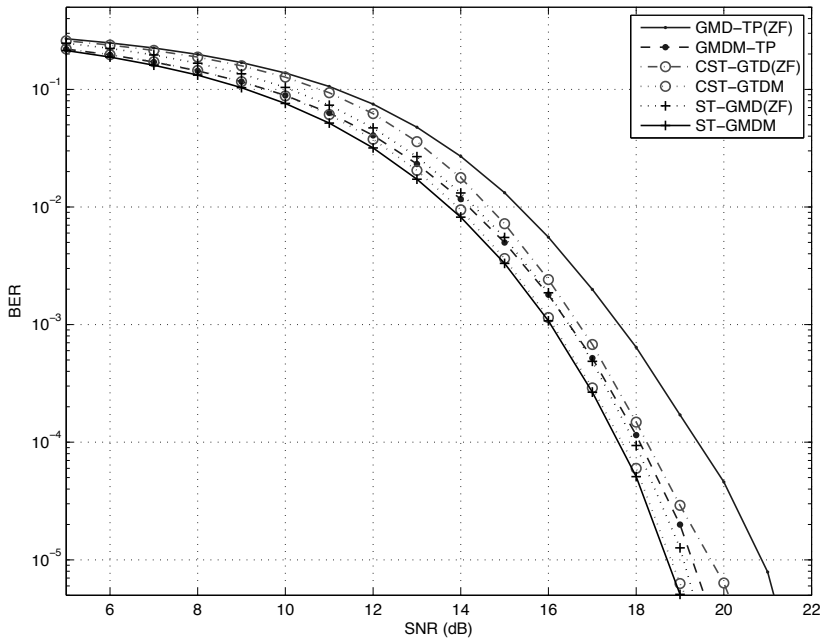


Figure 5.9: BER performance of zero-forcing and MMSE transceivers.

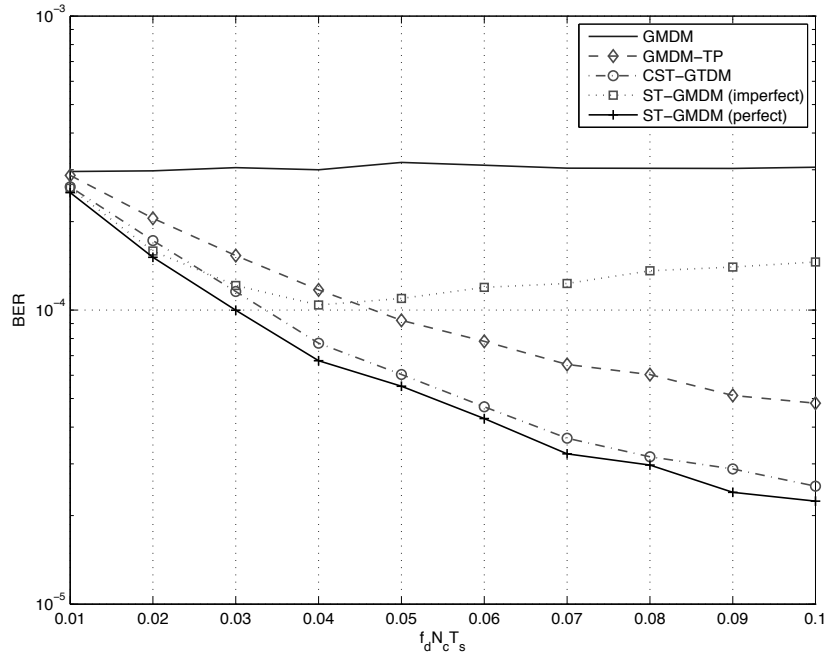


Figure 5.10: BER performance v.s. $f_dN_cT_s$ at $K = 16$ and $SNR = 19$ dB.

of the Wiener prediction filter is set to be equal to ST-block size K and $f_dN_cT_s = 0.1$. Fig. 5.7 and Fig. 5.8 show $\mathcal{P}_{st-gmdm}$, $\mathcal{P}_{cst-gtdm}$ and $\mathcal{P}_{gmdmtp-z}$. At the low SNR region, $\mathcal{P}_{st-gmdm}$, $\mathcal{P}_{cst-gtdm}$ and $\mathcal{P}_{gmdmtp-z}$ increase with respect to K. This is because the larger the ST-block size the more severe the error propagation in low SNR. For GMDM-TP, the phenomenon also results from the concavity of the BER function in the low SNR region. At the high SNR region, $\mathcal{P}_{st-gmdm}$, $\mathcal{P}_{cst-gtdm}$ and $\mathcal{P}_{gmdmtp-z}$ decrease with respect to K, which is best illustrated by Fig. 5.8. These results verify Theorem 5.5 and Theorem 5.6. Notice that the BER performance of the CST-GTD and ST-GMD MMSE transceiver are almost the same when $K = 32$. Fig. 5.7 and Fig. 5.8 also show that the performance of the ST-GMD transceiver with imperfect channel prediction is as good as the ideal ST-GMD transceiver when K is small. This is because prediction error is almost negligible when K is small. However, the ST-GMD transceiver has worse BER than the CST-GTD transceiver when K is slightly larger since degradation caused by the prediction error is too severe.

Example 3: In this example, we compare the average BER performance of the proposed GMD-TP, ST-GMD, and CST-GTD MMSE transceivers, with the corresponding zero-forcing transceivers in [36] and Chapter 4. The ST-block size K is 16. $\mathbf{H}(k)$ for different k are independent. The results

are shown in Fig. 5.9. At BER 10^{-4}, the SNR gains of GMD-TP, ST-GMD, and CST-GTD MMSE transceivers over their zero-forcing counterparts are 1.32 dB, 0.37 dB and 0.57 dB, respectively.

Example 4: The BER performances of the transceivers are evaluated using Jakes' channel model. Fig. 5.10 shows the average BER performances of the conventional GMD-based system, CST-GTD and ST-GMD MMSE transceivers with perfect and imperfect channel prediction for different values of the product $f_d N_c T_s$, which appears in (4.2). Larger $f_d N_c T_s$ implies the channels are changing at faster rates, and $\mathbf{H}(k)$ for different k are more uncorrelated. Hence, as $f_d N_c T_s$ increases, the average BERs of CST-GTD and ideal ST-GMD transceivers drop quickly due to the rich temporal diversity offered by the time-varying channels. However, channel prediction error increases when the channel is changing at a faster rate [29]. For $f_d N_c T_s > 0.026$, the ST-GMDM transceiver with imperfect channel prediction has poorer performance than CST-GTDM. For the system considered in [29], where the carrier frequency $f_c = 2$ GHz, the symbol period $T_s = 2.5$ μs and the terminal speed $\nu = 30$ m/s; if we set $N_c = 50$, which is also suggested in [29], the product $f_d N_c T_s$ will be 0.025. This means that ST-GMDM with imperfect channel prediction is useful for the system in [29] when the terminal is traveling at a speed slower than 30 m/s (i.e., 67 mph).

5.9 Conclusions

Three MIMO MMSE decision feedback transceivers for narrowband MIMO time-varying channels were proposed. The GMDM-TP is a cost efficient structure which equalizes the MSE in both temporal and spatial domains to minimize the average BER. Under the perfect channel prediction, the ST-GMD MMSE transceiver is shown to be the optimal in terms of arithmetic MSE and average BER in the convex region. Moreover, it was shown to be mutual information lossless. The CST-GTDM has the same asymptotic BER performance as the ST-GMD MMSE transceiver, and has smaller arithmetic MSE than both the conventional GMD-based MMSE systems and the GMDM-TP. The dependency of average BER on the ST-block size has also been analyzed. We also showed the advantage of the ST-GMD MMSE transceiver over the ST-GMD ZF transceiver in terms of arithmetic MSE and average BER.

5.10 Proof of Theorem 5.7

Recall that the space-time water-filling formula is given in (5.34). $L(k)$ is determined from

$$\frac{1}{\lambda} - \frac{\sigma_w^2}{\sigma_a^2 \sigma_{h,i}^2(k)} \leq 0, \text{ for } L(k) \leq i \leq M-1. \tag{5.89}$$

In other words,

$$\sigma_{h,i}^2(k) \leq \frac{\lambda \sigma_w^2}{\sigma_a^2}, \text{ for } L(k) \leq i \leq M-1. \tag{5.90}$$

Hence, it is necessary and sufficient to prove the theorem for

$$\sigma_{h,L(k)}^2(k) = \ldots = \sigma_{h,M-1}^2(k) = \frac{\lambda \sigma_w^2}{\sigma_a^2} = \sigma_h^2, \tag{5.91}$$

in which the value is independent of k. Using (5.91), one can reformulate $\overline{\xi}_{st-gmdm}$ as

$$\overline{\xi}_{st-gmdm} = \sigma_a^2 \sigma_h^2 \Big(\prod_{k=0}^{K-1} \prod_{i=0}^{M-1} \frac{1}{\sigma_{h,i}^2(k)} \Big)^{\frac{1}{MK}}. \tag{5.92}$$

So,

$$\frac{\sigma_a^2}{\overline{\xi}_{st-gmdm-br}} = \frac{1}{\sigma_h^2} \Big(\prod_{k=0}^{K-1} \prod_{i=0}^{M-1} \sigma_{h,i}^2(k) \Big)^{\frac{1}{MK}} - 1. \tag{5.93}$$

From the power constraint,

$$KP_0 = \sigma_a^2 \sum_{k=0}^{K-1} \sum_{i=0}^{L(k)-1} \big(\frac{1}{\lambda} - \frac{\sigma_w^2}{\sigma_a^2 \sigma_{h,i}^2(k)} \big), \tag{5.94}$$

we can write

$$\frac{P_0}{\sigma_w^2} = \frac{M}{\sigma_h^2} - \frac{1}{K} \sum_{k=0}^{K-1} \sum_{i=0}^{M-1} \frac{1}{\sigma_{h,i}^2(k)}. \tag{5.95}$$

Plugging this into (5.83), we have

$$\frac{\sigma_a^2}{\overline{\xi}_{stgmd_{zf}}} = \frac{1}{\sigma_h^2} \big(\prod_{k=0}^{K-1} \prod_{i=0}^{M-1} \sigma_{h,i}^2(k) \big)^{\frac{1}{MK}} - \frac{\frac{1}{MK} \sum_{k,i} \frac{1}{\sigma_{h,i}^2(k)}}{(\prod_{k,i} \frac{1}{\sigma_{h,i}^2(k)})^{1/(MK)}}.$$

The second term on the right hand side of the equality is greater than one by the AM-GM inequality. Hence,

$$\frac{\sigma_a^2}{\overline{\xi}_{stgmd_{zf}}} \leq \frac{\sigma_a^2}{\overline{\xi}_{st-gmdm-br}}, \tag{5.96}$$

and the proof is complete.

Chapter 6

MIMO Broadcast DFE Transceivers with QoS Constraints: Min-Power and Max-Rate Solutions

This chapter considers two joint design problems of linear precoder, decision feedback equalizer (DFE) and bit allocation for multi-input multi-output (MIMO) broadcast (BC) channels. The first problem is a power minimization problem (min-power) with a total bitrate constraint and per data stream symbol error rate (SER) specifications. The second problem is a rate maximization problem (max-rate) with a total transmit power constraint and per data stream SER specifications. For a given broadcast DFE transceiver, optimal bit allocation formulas for both problems are derived. A particular class of joint triangularization (JT) is applied to obtain the optimal broadcast DFE transceivers for the min-power and max-rate *QoS* problems, namely the minimum power JT broadcast DFE transceiver (MPJT) and the maximum rate JT broadcast DFE transceiver (MRJT), respectively. Two suboptimal broadcast DFE transceivers, the power minimized QR broadcast DFE transceiver (PMQR) and the rate maximized QR broadcast DFE transceiver (RMQR), are also proposed for the min-power and max-rate *QoS* problems, respectively. The proposed suboptimal designs apply QR decompositions instead of the particular class of JT. Moreover, integer bit allocation problems for both *QoS* problems are addressed. This work also shows the duality of the proposed MPJT and MRJT transceivers. At the end, numerical results are presented to demonstrate the performance of the proposed MPJT, MRJT, PMQR and RMQR transceivers under different *QoS* constraints, and verify the duality of the proposed MPJT and MRJT transceivers. Most of the results in this chapter have been reported in our recently submitted journal paper [65].

6.1 Introduction

Multi-input multi-output (MIMO) broadcast (BC) channels have aroused much interest in recent years. Capacity regions have been established for several scenarios [40]-[43]. In the common-message (multicast) scenario, the base station has only one precoder to encode a common message while the users at the receiver ends can decode the message independently. This prevents one from adapting singular value decomposition (SVD) to the design, since a SVD-based transceiver requires one channel dependent precoder for each MIMO channel. Moreover, the QR-based transceiver fails to achieve channel capacity with scalar Gaussian code [97], [44]. For the two-user case, the joint equal diagonal triangularization (JET) is proposed in [44], [45] for the design of a broadcast decision feedback equalizer (DFE) transceiver which can achieve common message MIMO broadcast channel capacity with scalar Gaussian codes.

In the context of transceiver design for point to point (PTP) MIMO systems, the power minimization (min-power) and the rate maximization (max-rate) problems with quality of service (*QoS*) constraints have been studied extensively in [51]-[58]. In [51] and [58], linear optimal zero-forcing (ZF) transceivers with bit allocation that minimize the total transmit power for a given target rate and equal per substream symbol error rate (SER) constraints are developed. Optimal ZF solutions, with the aim of minimizing the total transmit power under unequal per substream SER or bit error rate (BER) constraints, are considered in [52]-[53] to support multiple data streams, such as voice, data and video. Each of the data streams in general has different BER or SER requirements [52]-[53], [93]. In [54], linear MMSE transceivers are designed by minimizing the total transmit power with possibly different per substream *QoS* constraints, in terms of mean square error (MSE) or BER. Linear transceiver designs with bit allocation which maximize total bit rate with power and SER constraints are considered in [55], [56] and [58]. In [57], the authors consider both the min-power and max-rate design problems with possibly unequal SER constraints for DFE transceivers.

In this work, we investigate two DFE transceiver *QoS* design problems for common-message MIMO broadcasting channels, which are generalizations or extensions of the point to point MIMO min-power or max-rate problems in [51]-[58]. To the best of our knowledge, this is the first work addressing the min-power and max-rate *QoS* problems for common-message MIMO broadcast channels. The JET broadcast DFE transceiver proposed in [44] can not provide solutions for the *QoS* problems. The first problem addressed in this chapter is a min-power problem with a total bit rate constraint and per data substream symbol error rate (SER) specifications. The second problem is

a max-rate problem with a total transmit power constraint and per data substream SER specifications. For both problems, we jointly optimize the bit allocation, the precoder, feed-forward and feedback equalizers. The optimal solutions are obtained in two steps. For a given broadcast DFE transceiver, we design the optimal bit allocation, and then the transceivers are designed to minimize the transmit power or maximize the transmission rate based on the optimal bit allocation. For given *QoS* specifications, one can obtain the optimal solutions for the min-power and max-rate problems, namely the minimum power joint triangularization (JT) broadcast DFE transceiver (MPJT) and the maximum rate JT DFE transceiver (MRJT), respectively, by exploiting a particular class of joint triangularization satisfying the *dominance condition* (to be described later). For the two-user case, if the SER specifications satisfy the *admissible condition* (to be described later), MPJT and MRJT can always be designed. For a more than two-user case, finding a MPJT or MRJT could be difficult or impossible even though the SER specifications satisfy the *admissible condition*, since there is no theory that guarantees the existence of joint triangularization for more than two matrices such that the *dominance condition* is satisfied.

A power minimized QR broadcast DFE transceiver (PMQR) and a rate maximized QR broadcast DFE transceiver (RMQR) are also proposed. They can be easily designed for an arbitrary number of users. Instead of using the joint triangularization satisfying *dominance condition*, QR decompositions are exploited. Therefore, PMQR and RMQR are suboptimal solutions for the min-power and max-rate *QoS* problems, respectively. This chapter also addresses the integer bit allocation problem for the proposed broadcast DFE transceivers by introducing suboptimal bit allocation procedures.

For point to point MIMO channels, the authors in [58] showed that if the linear transceiver is optimal for the power-minimizing problem, it is also optimal for the rate-maximizing problem. And the converse is true. In this chapter, we shall show the duality of the MPJT and MRJT transceivers for MIMO broadcast channels. Given that an MPJT is an optimal solution for a min-power problem, it is also a solution for a max-rate problem with the power constraint set to be the transmit power of the MPJT. Likewise, given that a MRJT is optimal for a max-rate problem, it is also optimal for a min-power problem with the target rate set to be the rate of the MRJT. In the simulations, the optimality of the MPJT transceiver is demonstrated against the suboptimal PMQR transceiver for the min-power problem. Comparisons have also been made between the MRJT and RMQR transceivers for the max-rate problem. The effects of integer bit allocation for the transceivers are shown. Numerical examples are also provided to verify the duality of the MPJT and MRJT

transceivers.

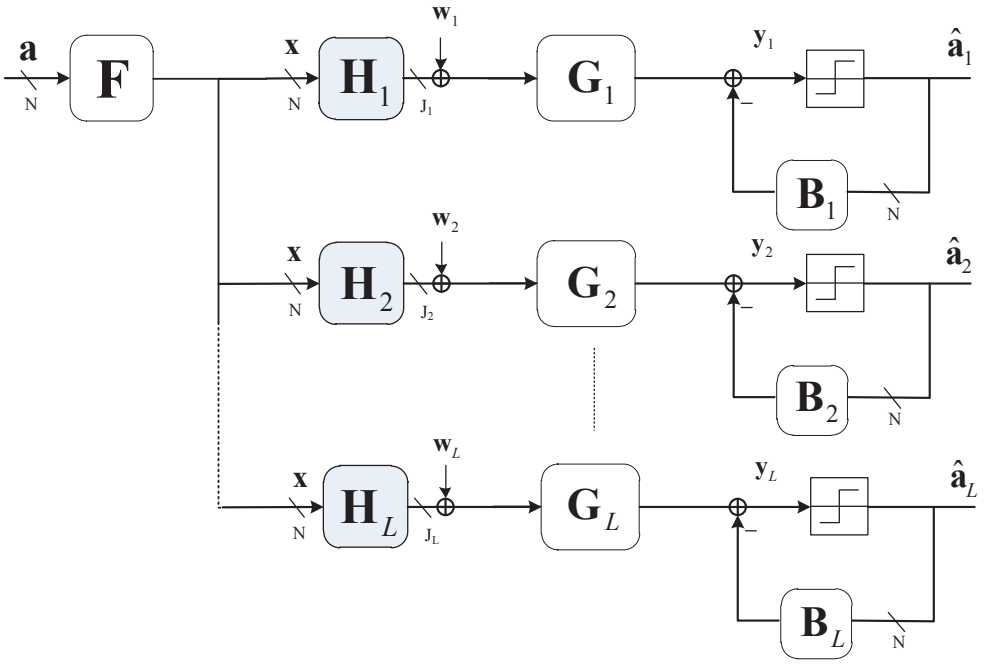


Figure 6.1: The DFE transceiver for broadcast channel

6.2 Outline

The sections in this chapter are structured as follows. We start by reviewing the concept of generalized singular value [46] and the joint triangularization [45], and introduce the structure of a DFE transceiver for MIMO broadcast channels in Section 6.3. In Section 6.4, we discuss the joint design of a broadcast DFE transceiver and bit allocation for the min-power *QoS* problem. The max-rate *QoS* problem of a broadcast DFE transceiver is investigated in Section 6.5. Section 6.6 connects the proposed MPJT and MRJT transceivers by showing their duality in the min-power and max-rate *QoS* problems. Numerical examples for the designs are provided in Section 6.7. Concluding remarks are given in Section 6.8.

6.3 Preliminaries and System Model

The generalized singular values [46] and the joint triangularization [45] are two important concepts used in this work. We review them first:

Definition 1: The *generalized singular values* (GSVs) of the ordered matrix pair $(\mathbf{A}_1, \mathbf{A}_2)$ are elements of the set

$$S = \{s \in \mathcal{R}^+ : \det(\mathbf{A}_1^\dagger \mathbf{A}_1 - s^2 \mathbf{A}_2^\dagger \mathbf{A}_2) = 0\}, \tag{6.1}$$

where $\mathbf{A}_1 \in \mathcal{C}^{J_1 \times N}$, $\mathbf{A}_2 \in \mathcal{C}^{J_2 \times N}$ and $J_1 \geq N$. $\mathbf{u}(\mathbf{A}_1, \mathbf{A}_2)$ will denote a vector composed of all GSVs.

Theorem 6.1:(The joint triangularization (JT) [45]). Let $\mathbf{r}$ be an $N \times 1$ vector consisting of positive real elements, r_n, and $\mathbf{A}_1 \in \mathcal{C}^{J_1 \times N}$ and $\mathbf{A}_2 \in \mathcal{C}^{J_2 \times N}$ be two matrices of rank N where $J_1, J_2 \geq N$. There exists upper triangular matrices $\mathbf{R}_1, \mathbf{R}_2 \in \mathcal{C}^{N \times N}$, a unitary matrix $\mathbf{P} \in \mathcal{C}^{N \times N}$, and *semi-unitary matrices* $\mathbf{Q}_1 \in \mathcal{C}^{J_1 \times N}$ and $\mathbf{Q}_2 \in \mathcal{C}^{J_2 \times N}$, such that

$$\mathbf{A}_1 = \mathbf{Q}_1 \mathbf{R}_1 \mathbf{P}^\dagger, \ \mathbf{A}_2 = \mathbf{Q}_2 \mathbf{R}_2 \mathbf{P}^\dagger, \tag{6.2}$$

and $[\mathbf{R}_1]_{n,n} / [\mathbf{R}_2]_{n,n} = r_n$ for $n = 0, 1, \ldots, N-1$, if and only if

$$\mathbf{r} \prec_\times \mathbf{u}(\mathbf{A}_1, \mathbf{A}_2). \tag{6.3}$$

□

Without loss of generality, we can make the diagonal entries of $\mathbf{R}_1$ and $\mathbf{R}_2$ real and positive.

In this chapter, we consider the scenario that a base station broadcasts *common messages* to several subscribed users through MIMO narrowband channels. The broadcast channel is modeled as

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{x} + \mathbf{w}_i, \ i = 1, 2 \ldots, L, \tag{6.4}$$

where $\mathbf{y}_i$ is a $J_i \times 1$ received signal vector of the ith user, $\mathbf{x}$ is an $N \times 1$ transmit signal vector, $\mathbf{H}_i$ is a $J_i \times N$ ($J_i \geq N$) channel matrix for the ith user, and $\mathbf{w}_i$ is a $J_i \times 1$ circular symmetric additive Gaussian noise vector for the ith user with its statistics given by $E(\mathbf{w}_i) = \mathbf{0}$, $E(\mathbf{w}_i \mathbf{w}_i^\dagger) = \mathbf{I}_{J_i}$.

The *common message* broadcast channel capacity is determined by the worst channel and has the expression:

$$C_{BC} = \max_{\mathbf{C_{xx}}} \min_{i=1,\ldots,L} \log_2 \det\Big(\mathbf{I}_{J_i} + \mathbf{H}_i \mathbf{C_{xx}} \mathbf{H}_i^{\dagger}\Big), \tag{6.5}$$

where the maximization is over all positive definite matrices $\mathbf{C_{xx}}$ subject to the power constraint Tr $(\mathbf{C_{xx}}) \leq P_0$ [41].

A generic DFE transceiver for a common message broadcast MIMO channel is shown in Fig. 6.1. The precoder $\mathbf{F}$ is of size $N \times N$, the feed forward receiver matrices $\mathbf{G}_i$ are of size $N \times J_i$ and the feedback matrices $\mathbf{B}_i$ are strictly upper triangular matrices of size $N \times N$. It is assumed that the channel state information $\mathbf{H}_i$ is available both at the receivers and the transmitter. The data symbol, a_n, is a b_n-bit carrying square QAM symbol from the nth substream and is transmitted through L subchannels to L users. The total number of bits per data block, which consists of N QAM symbols, is

$$\sum_{n=0}^{N-1} b_n.$$

The data symbols a_n are assumed to be zero mean and uncorrelated, i.e., $E(\mathbf{a}) = \mathbf{0},\ E(\mathbf{a}\mathbf{a}^{\dagger}) = \mathbf{I}_N$. Furthermore, $\mathbf{a}$ is uncorrelated with the channel noise $\mathbf{w}_i$, i.e., $E(\mathbf{a}\mathbf{w}_i^{\dagger}) = \mathbf{0}$.

The error component prior to the detector of the ith user is

$$\mathbf{e}_i = \mathbf{y}_i - \mathbf{a},$$

and the error covariance matrix is given by

$$\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i} = E(\mathbf{e}_i \mathbf{e}_i^{\dagger}). \tag{6.6}$$

Denote the mean square error (MSE) of the nth data stream for the ith user as

$$\sigma_{i,n}^2 = [\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i}]_{n,n}. \tag{6.7}$$

Under the high bitrate and low SER assumptions, $b_n \gg 1$, the SER of the QAM detector for the nth

substream from user i can be approximated [94] as

$$SER_{i,n} \approx 4Q\Big(\sqrt{\frac{3}{2^{b_n}-1}\gamma_{i,n}}\Big), \tag{6.8}$$

where $\gamma_{i,n}$ is the SINR of the nth substream from the ith user and can be approximated [25] by $\gamma_{i,n} = 1/\sigma^2_{i,n}$. The above SER expression is good for Gaussian interference. For non-Gaussian impulse noise [49], the SER may be higher than (6.8).

6.4 Power Minimized Broadcast DFE Transceivers with QoS Constraints

In this section, we discuss the joint design of a MIMO broadcast DFE transceiver for a power minimization (min-power) problem under the target rate and per data substream SER constraints. Under the total bit rate and SER constraints, we optimize the bit allocation $\{b_n\}$, the precoder $\mathbf{F}$, the receiver matrices $\mathbf{G}_i$ and the feedback matrices $\mathbf{B}_i$ such that the total transmit power is minimized. To avoid a complicated integer programming problem, we relax $b_n \in \mathcal{Z}^+$ to $b_n \in \mathcal{R}$. The problem can be formulated as

$$\begin{aligned} \min_{\mathbf{F},\mathbf{G}_i,\mathbf{B}_i,\{b_n\}} \quad & \text{Tr}\,(\mathbf{F}\mathbf{F}^\dagger) \\ s.t. \quad (a) \quad & SER_{i,n} \le \eta_{i,n} \\ (b) \quad & \sum_{n=0}^{N-1} b_n \ge B_0,\ b_n \in \mathcal{R}, \end{aligned} \tag{6.9}$$

where B_0 denotes the target total bit rate and $\eta_{i,n}$ is the SER constraint of the nth substream for the ith user. The optimization problem in (6.9) reduces to the min-power problem in [57] when there is only one subscriber and the zero-forcing constraint is applied. It shall further reduce to the min-power problem in [51] and [58] if we only consider linear transceivers. Note that per substream SER constraints have been widely used in the min-power and max-rate problems for PTP MIMO channels [51], [53], [55]-[58], even in the cases where the bit allocation variables, $\{b_n\}$, are optimization variables [51], [55]-[58]. It is also shown in [55] and [25] that SER constraints are good approximations of BER constraints for moderately low BERs and Gray coded bit mapping. Here, a careful treatment of the min-power problem with per substream BER constraints is available later in Sec. 6.4.6.

Using (6.8) and assuming $b_n >> 1$, the SER constraints in (6.9) can be re-expressed as

$$2^{b_n}\sigma_{i,n}^2\Gamma_{i,n} \leq 1, \tag{6.10}$$

where the *QoS coefficients*

$$\Gamma_{i,n} = (Q^{-1}(\eta_{i,n}/4))^2/3.$$

Reducing the SER constraints in (6.10) to a single constraint, we can translate the problem in (6.9) into

$$\begin{aligned} \min_{\mathbf{F},\mathbf{G}_i,\mathbf{B}_i,\{b_n\}} \quad & \text{Tr}\,(\mathbf{F}\mathbf{F}^\dagger) \\ s.t. \quad (a) \quad & \max_{i\in S_1, n\in S_2} 2^{b_n}\sigma_{i,n}^2\Gamma_{i,n} \leq 1 \\ (b) \quad & \sum_{n=0}^{N-1} b_n \geq B_0,\ b_n \in \mathcal{R}, \end{aligned} \tag{6.11}$$

where $S_1 = \{1, 2, \ldots, L\}$ and $S_2 \in \{0, 1, \ldots, N-1\}$. The following lemma is useful in solving the above optimization problem.

Lemma 6.1: Suppose $\mathcal{L}_0 = \{\mathbf{F}, \mathbf{B}_i, \{b_n\}, \mathbf{G}_i\}$ is an optimal solution set for (6.11), then $\mathcal{L}_1 = \{\mathbf{F}, \mathbf{B}_i, \{b_n\}, \mathbf{G}_i^*\}$ is also an optimal solution set, where the MMSE feedforward matrices are given by

$$\mathbf{G}_i^* = (\mathbf{I}_N + \mathbf{B}_i)\mathbf{F}^\dagger\mathbf{H}_i^\dagger(\mathbf{I}_{J_i} + \mathbf{H}_i\mathbf{F}\mathbf{F}^\dagger\mathbf{H}_i^\dagger)^{-1}. \tag{6.12}$$

□

Proof: For given $\mathbf{F}$ and $\mathbf{B}_i$, [25] shows that the nth row of $\mathbf{G}_i^*$ independently minimizes the mean square error of the nth subchannel, $\sigma_{i,n}^2$ in (6.7). Hence,

$$2^{b_n}\Gamma_{i,n}\sigma_{i,n}^2(\mathbf{G}_i^*) \leq 2^{b_n}\Gamma_{i,n}\sigma_{i,n}^2(\mathbf{G}_i) \leq 1.$$

The first constraint in (6.11) is still valid if one replaces $\mathbf{G}_i$ in $\mathcal{L}_0$ with $\mathbf{G}_i^*$. □

By Lemma 6.1, there always exists a set of optimal feed forward matrices $\mathbf{G}_i$ in the form of (6.12) if an optimal solution set exists for (6.11). So, in this work, we set $\mathbf{G}_i$ to be in the form of (6.12). The

error covariance matrices in (6.6) can be shown to be

$$\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i}(\mathbf{G}_i^*) = (\mathbf{I}_N + \mathbf{B}_i)(\mathbf{I}_N + \mathbf{F}^\dagger \mathbf{H}_i^\dagger \mathbf{H}_i \mathbf{F})^{-1}(\mathbf{I}_N + \mathbf{B}_i)^\dagger. \tag{6.13}$$

6.4.1 Optimal Bit Allocation

When the optimal $\mathbf{G}_i^*$ are applied, the optimization problem (6.11) becomes

$$\begin{aligned} \min_{\mathbf{F},\mathbf{B}_i,\{b_n\}} \quad & \text{Tr}\,(\mathbf{F}\mathbf{F}^\dagger) \\ s.t. \quad (a) \quad & \max_{i\in S_1, n\in S_2} 2^{b_n}\sigma_{i,n}^2\Gamma_{i,n} \le 1 \\ (b) \quad & \sum_{n=0}^{N-1} b_n \ge B_0,\ b_n \in \mathcal{R}. \end{aligned} \tag{6.14}$$

The optimal bit allocation is obtained by considering the weighted MSE constraint in (6.14). We first review a lemma of vector norm from [108].

Lemma 6.2: If $\mathbf{x} \in \mathcal{R}^N$, then $||\mathbf{x}||_\infty \ge ||\mathbf{x}||_1/N$. □

Proof: Please refer to [108]. □

By Lemma 6.2, the maximal weighted MSEs in (6.14) have an arithmetic mean (AM) lower bound such that

$$\max_{n\in S_2, i\in S_1}\{\sigma_{i,n}^2\Gamma_{i,n}2^{b_n}\} \ge \frac{1}{N}\Big(\sum_{n\in S_2} 2^{b_n}\max_{i\in S_1}\{\sigma_{i,n}^2\Gamma_{i,n}\}\Big). \tag{6.15}$$

Applying the AM-GM inequality to the right hand side of (6.15), we obtain a smaller geometric mean (GM) lower bound

$$\max_{n\in S_2, i\in S_1}\{\sigma_{i,n}^2\Gamma_{i,n}2^{b_n}\} \ge 2^b\Big(\prod_{n\in S_2}\max_{i\in S_1}\{\sigma_{i,n}^2\Gamma_{i,n}\}\Big)^{\frac{1}{N}}, \tag{6.16}$$

where $b = (1/N)\sum_{n\in S_2} b_n$. The equality in (6.16) holds if and only if one chooses b_n such that

$$2^{b_n}\max_{i\in S_1}\{\Gamma_{i,n}\sigma_{i,n}^2\} = c \text{ (constant)}. \tag{6.17}$$

Taking logarithms on both sides, we have

$$b_n = D - \log_2\Big(\max_{i\in S_1}\{\Gamma_{i,n}\sigma_{i,n}^2\}\Big), \tag{6.18}$$

where D is a constant.

In Theorem 6.2 below, it is shown that the maximal weighted MSE achieves the GM lower bound with the optimal $\{b_n\}$. Before we reach that, we introduce the following lemma which is useful for the later discussions.

Lemma 6.3: Let $\mathbf{F} = \alpha \mathbf{F}_1$, where $\alpha > 0$ and $\mathbf{F}_1 \in \mathcal{C}^{N \times N}$. Then, $[\mathbf{R}_{\mathbf{e}_i \mathbf{e}_i}(\mathbf{G}_i^*)]_{n,n} = \sigma_{i,n}^2(\mathbf{G}_i^*)$ in (6.13) is a non-increasing function of α. □

Proof: See appendix. □

Theorem 6.2: Given that $\mathbf{F}$ and $\mathbf{B}_i$ are optimal, the *optimal bit allocation formula* for problem (6.14) is

$$b_n = D - \log_2\Big(\max_{i \in S_1}\{\Gamma_{i,n}\sigma_{i,n}^2\}\Big), \tag{6.19}$$

where D is a constant chosen such that $\sum_{n=0}^{N-1} b_n = B_0$. □

Proof: In the following, we shall first prove that the optimal bit allocation has the form as in (6.19). Using the first part of the result, we shall be able to show that the constant D in (6.19) should be chosen such that the total bit rate is equal to B_0.

Suppose that $(\mathbf{F}, \mathbf{B}_i, \{\tilde{b}_n\})$ is the optimal solution for (6.14) and the equality in (6.17) does not hold, i.e.

$$2^{\tilde{b}_k} \max_i \Gamma_{i,n}\sigma_{i,n}^2 \neq c, \tag{6.20}$$

for all n. The total rate is

$$B_1 = \sum_{n=0}^{N-1} \tilde{b}_n,$$

and it satisfies the total bit rate constraint in (6.14), $B_1 \geq B_0$. Let $\{b_k^*\}$ be the bit allocation given by (6.19) and D be the constant such that $\sum_{n=0}^{N} b_n^* = B_1$. By (6.16), one can get

$$\max_{n \in S_2, i \in S_1} \sigma_{i,n}^2 \Gamma_{i,n} 2^{b_n^*} = 2^{\frac{B_1}{N}} \Big(\prod_{n=0}^{N-1} \max_i \sigma_{i,n}^2 \Gamma_{i,n}\Big)^{\frac{1}{N}}, \tag{6.21}$$

where the right hand side term is also the lower bound of $\max_{n\in S_2, i\in S_1} \sigma_{i,n}^2 \Gamma_{i,n} 2^{\tilde{b}_k}$. Therefore,

$$\max_{n\in S_2, i\in S_1} \sigma_{i,n}^2 \Gamma_{i,n} 2^{b_n^*} < \max_{n\in S_2, i\in S_1} \sigma_{i,n}^2 \Gamma_{i,n} 2^{\tilde{b}_n} \leq 1. \tag{6.22}$$

The first inequality follows from (6.20) and the last inequality is from the MSE constraints in (6.14). Suppose the optimal precoder $\mathbf{F}$ is scaled by a factor of α where $\alpha > 0$, i.e., $\alpha\mathbf{F}$. By Lemma 6.3, $\max_{n\in S_2, i\in S_1} \sigma_{i,n}^2 \Gamma_{i,n} 2^{b_k}$ is a non-increasing function of α. Therefore, one can scale down $\mathbf{F}$ as $\alpha\mathbf{F}$ ($0 < \alpha < 1$) such that

$$\max_{n\in S_2, i\in S_1} \sigma_{i,n}^2 \Gamma_{i,n} 2^{b_k^*} \leq 1.$$

The scaling leads to lower transmit power, $\alpha^2 \text{Tr}\,(\mathbf{F}\mathbf{F}^\dagger)$ and a contradiction arises on the optimality of $(\mathbf{F}, \mathbf{B}_i, \{\tilde{b}_n\})$.

Now, we are ready to prove the second result. Suppose that $(\mathbf{F}, \mathbf{B}_i, \{b_n^*\})$ is optimal and $\{b_k^*\}$ is given by (6.19) where the D is chosen such that

$$\sum_{n=0}^{N-1} b_n^* = B_1 > B_0.$$

Let $\{\bar{b}_n\}$ be given by (6.19) but D be such that $\sum_{n=S_2} \bar{b}_n = B_0$. Since $B_0 < B_1$, the maximal weighted MSE is given by

$$\begin{aligned} \max_{n\in S_2, i\in S_1} \sigma_{i,n}^2 \Gamma_{i,n} 2^{\bar{b}_k} &= 2^{\frac{B_0}{N}} \Big(\prod_{n=0}^{N-1} \max_i \sigma_{i,n}^2 \Gamma_{i,n} \Big)^{\frac{1}{N}} \\ &< 2^{\frac{B_1}{N}} \Big(\prod_{n=0}^{N-1} \max_i \sigma_{i,n}^2 \Gamma_{i,n} \Big)^{\frac{1}{N}} \leq 1. \end{aligned} \tag{6.23}$$

By Lemma 6.3, $\big(\prod_{n=0}^{N-1} \max_i \sigma_{i,n}^2 \Gamma_{i,n}\big)^{\frac{1}{N}}$ is a continuous non-increasing function of α. Using the same arguments as the first part of the proof, we can conclude the proof.

□

6.4.2 Optimal Feedback Matrices and Precoder

The remaining problem is optimizing the precoding matrix and feedback matrices. Substituting the optimal b_n from (6.19) into (6.14) and setting $\mathbf{F} = \tilde{\mathbf{F}}\mathbf{P}$ where $\mathbf{P}$ is an $M \times M$ unitary matrix, we

reduce the optimization problem (6.14) to

$$\begin{aligned} \min_{\mathbf{P},\tilde{\mathbf{F}},\mathbf{B}_i} \quad & \text{Tr}\,(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^{\dagger}) \\ s.t. \quad & \prod_{n\in S_2} \max_{i\in S_1}\Big(\sigma_{i,n}^2\Gamma_{i,n}\Big) \leq 2^{-B_0}. \end{aligned} \tag{6.24}$$

To solve the problem above, the following lemma is introduced.

Lemma 6.4: For $a_{i,n} \in \mathcal{R}^{++}$ ($i \in S_1$ and $n \in S_2$), a strictly increasing function $f : \mathcal{R}^{++} \mapsto \mathcal{R}^{++}$ and a strictly decreasing function $g : \mathcal{R}^{++} \mapsto \mathcal{R}^{++}$,

$$\prod_{n\in S_2} \max_{i\in S_1}\{f(a_{i,n})\} \geq \max_{i\in S_1}\Big\{\prod_{n\in S_2} f(a_{i,n})\Big\}, \tag{6.25}$$

$$\prod_{n\in S_2} \min_{i\in S_1}\{g(a_{i,n})\} \leq \min_{i\in S_1}\Big\{\prod_{n\in S_2} g(a_{i,n})\Big\}. \tag{6.26}$$

The equalities hold if and only if there exists a $j \in S_1$ such that

$$a_{j,n} \geq a_{i,n}, \text{ for } i \in S_1, n \in S_2. \tag{6.27}$$

□

Proof: See appendix. □

Using Lemma 6.4, we obtain the lower bound for the geometric mean of weighted MSEs in (6.24)

$$\prod_{n\in S_2} \max_{i\in S_1}\{\Gamma_{i,n}\sigma_{i,n}^2\} \geq \max_{i\in S_1}\{\prod_{n\in S_2} \Gamma_{i,n}\sigma_{i,n}^2\}, \tag{6.28}$$

and the equality holds if and only if the *dominance conditions* hold, i.e., there exists $j \in S_1$ such that

$$\sigma_{j,n}^2/\sigma_{i,n}^2 \geq \Gamma_{i,n}/\Gamma_{j,n}, \text{ for } i \neq j \text{ and } n \in S_2. \tag{6.29}$$

The dominance condition implies that there exists one user who has the largest weighted MSEs for all of his data substreams when compared to other users. Let us now assume that the *dominance condition* is satisfied and relegate the discussion of the condition to the next section, Sec. 6.4.3. The

problem in (6.24) becomes

$$\begin{aligned} \min_{\mathbf{P},\tilde{\mathbf{F}},\mathbf{B}_i} \quad & \text{Tr}\,(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^{\dagger}) \\ s.t. \quad & \max_{i\in S_1}\{\prod_{n\in S_2}\Gamma_{i,n}\sigma_{i,n}^2\} \leq 2^{-B_0}. \end{aligned} \tag{6.30}$$

To optimize $\mathbf{B}_i$ and $\mathbf{P}$, we consider the left term of the inequality constraint in (6.30),

$$\max_{i\in S_1}\Big\{\tilde{\Gamma}_i\prod_{n\in S_2}\sigma_{i,n}^2\Big\} \geq \max_{i\in S1}\Big\{\tilde{\Gamma}_i\det\big(\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i}(\mathbf{G}_i^*)\big)\Big\}, \tag{6.31}$$

where $\tilde{\Gamma}_i = \prod_{n=0}^{N-1}\Gamma_{i,n}$. The inequality follows from Hadamard's inequality and the equality holds if and only if $\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i}(\mathbf{G}_i^*)$ is diagonal. The following theorem shows that the optimal $\mathbf{P}$ and $\mathbf{B}_i$ shall make the equality in (6.31) hold. Before that, an useful lemma is introduced.

Lemma 6.5: The function $g(\alpha) = \det\big(\mathbf{I}_N + \alpha^2\tilde{\mathbf{F}}^{\dagger}\mathbf{H}_i^{\dagger}\mathbf{H}_i\tilde{\mathbf{F}}\big)$ is a continuous strictly increasing function of α for $\alpha > 0$ and $\mathbf{H}_i\tilde{\mathbf{F}} \neq \mathbf{0}$. □

Proof: See appendix. □

Theorem 6.3: Suppose $(\tilde{\mathbf{F}}, \mathbf{P}, \mathbf{B}_i)$ is the optimal solution for (6.30). The optimal $\mathbf{P}$ and $\mathbf{B}_i$ are such that $\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i}(\mathbf{G}_i^*)$ in (6.13) is diagonal (The diagonalization is always possible as shown later in this subsection). □

Proof: Suppose $(\tilde{\mathbf{F}}, \tilde{\mathbf{P}}, \tilde{\mathbf{B}}_i)$ is optimal for (6.30) and $(\tilde{\mathbf{P}}, \tilde{\mathbf{B}}_i)$ does not make $\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i}(\mathbf{G}_i^*)$ diagonal. Applying the Hadamard's inequality from (6.31) and substituting $\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i}(\mathbf{G}_i^*)$ from (6.13), we have

$$\begin{aligned} 2^{-B_0} &\geq \max_{i\in S_1}\{\tilde{\Gamma}_i\prod_{n\in S_2}\sigma_{i,n}^2(\mathbf{G}_i^*, \tilde{\mathbf{P}}, \tilde{\mathbf{B}}_i)\} \\ &> \max_{i\in S_1}\Big\{\tilde{\Gamma}_i\det\big(\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i}(\mathbf{G}_i^*)\big)\Big\} = \max_{i\in S_1}\frac{\tilde{\Gamma}_i}{\det\big(\mathbf{I}_N + \tilde{\mathbf{F}}^{\dagger}\mathbf{H}_i^{\dagger}\mathbf{H}_i\tilde{\mathbf{F}}\big)}. \end{aligned} \tag{6.32}$$

Using the fact from Lemma 6.5 that $\det\big(\mathbf{I}_N + \alpha^2\tilde{\mathbf{F}}^{\dagger}\mathbf{H}_i^{\dagger}\mathbf{H}_i\tilde{\mathbf{F}}\big)$ is a continuous increasing function of α, one can scale $\tilde{\mathbf{F}}$ by a factor of $\alpha = 1 - \epsilon$ (ϵ is an arbitrary small positive number) such that

$$\max_{i\in S_1}\frac{\tilde{\Gamma}_i}{\det\big(\mathbf{I}_N + \alpha^2\tilde{\mathbf{F}}^{\dagger}\mathbf{H}_i^{\dagger}\mathbf{H}_i\tilde{\mathbf{F}}\big)} \leq 2^{-B_0}, \tag{6.33}$$

and construct a new feasible solution for (6.30), i.e. $(\alpha\tilde{\mathbf{F}}, \ddot{\mathbf{P}}, \ddot{\mathbf{B}}_i)$, where $\ddot{\mathbf{P}}$ and $\ddot{\mathbf{B}}_i$ diagonalize $\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i}(\mathbf{G}_i^*)$. By (6.31) and (6.33), the MSE constraint of (6.30) is satisfied, i.e.,

$$\begin{aligned}&\max_{i\in S_1}\{\tilde{\Gamma}_i \prod_{n\in S_2} \sigma_{i,n}^2(\mathbf{G}_i^*, \ddot{\mathbf{P}}, \ddot{\mathbf{B}}_i)\}\\ &= \max_{i\in S_1} \frac{\tilde{\Gamma}_i}{\det\big(\mathbf{I}_N + \alpha^2\tilde{\mathbf{F}}^\dagger\mathbf{H}_i^\dagger\mathbf{H}_i\tilde{\mathbf{F}}\big)} \le 2^{-B_0}.\end{aligned} \tag{6.34}$$

The new feasible solution requires less power than $(\tilde{\mathbf{F}}, \tilde{\mathbf{P}}, \tilde{\mathbf{B}}_i)$, i.e.,

$$\alpha^2 \text{Tr}\,(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^\dagger) < \text{Tr}\,(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^\dagger),$$

which leads to a contradiction on the optimality of $(\tilde{\mathbf{F}}, \tilde{\mathbf{P}}, \tilde{\mathbf{B}}_i)$. □

To make $\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i}(\mathbf{G}_i^*)$ diagonal, we construct auxiliary matrices and apply *joint triangularization* on them as follows:

$$\mathbf{A}_i = \begin{bmatrix} \mathbf{H}_i\tilde{\mathbf{F}} \\ \mathbf{I}_N \end{bmatrix} = \mathbf{Q}_i\mathbf{R}_i\mathbf{P}^\dagger = \mathbf{Q}_i\mathbf{D}_{\mathbf{R}_i}(\mathbf{I}_N + \mathbf{B}_i)\mathbf{P}^\dagger, \tag{6.35}$$

where $\mathbf{Q}_i$ is a $(J_i + N) \times N$ semi-unitary matrix, $\mathbf{R}_i$ is an $N \times N$ upper triangular matrix, $\mathbf{P}$ is an $N \times N$ unitary matrix, $\mathbf{D}_{\mathbf{R}_i}$ is a diagonal matrix composed of $[\mathbf{R}_i]_{n,n}$ on its diagonal and $\mathbf{B}_i$ is an $N \times N$ strictly upper triangular matrix. Given that the $\mathbf{P}$ and $\mathbf{B}_i$ are obtained from (6.35), one can verify that

$$\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i}(\mathbf{G}_i^*) = \mathbf{D}_{\mathbf{R}_i}^{-2}. \tag{6.36}$$

Note that here, joint triangularization for any number of $\mathbf{A}_i$ is possible since one can fix $\mathbf{P}$ and perform QR decomposition on $\mathbf{A}_i\mathbf{P}$. If $\mathbf{P} = \mathbf{I}_N$, the operations are just ordinary QR decompositions.

Applying the optimal $\mathbf{P}$ and $\mathbf{B}_i$ obtained from the triangularization in (6.35) to (6.30), it follows from (6.31) and (6.13) that (6.30) can be reduced to

$$\begin{aligned}\min_{\tilde{\mathbf{F}}} \quad & \text{Tr}\,(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^\dagger) \\ s.t. \quad & \min_{i\in S_1} \log_2\Big(\frac{\det\big(\mathbf{I}_{J_i} + \mathbf{H}_i\tilde{\mathbf{F}}\tilde{\mathbf{F}}^\dagger\mathbf{H}_i^\dagger\big)}{\tilde{\Gamma}_i}\Big) \ge B_0.\end{aligned} \tag{6.37}$$

The problem is the dual problem of the determinant maximization problem with linear matrix inequality [47]. The optimal solution $\tilde{\mathbf{F}}^*$ can be solved numerically and has the following property.

Property 1 The optimal solution $\tilde{\mathbf{F}}^*$ for (6.37) is such that

$$\min_{i\in S_1} \log_2\Big(\frac{\det\big(\mathbf{I}_{J_i} + \mathbf{H}_i\tilde{\mathbf{F}}^*\tilde{\mathbf{F}}^{*\dagger}\mathbf{H}_i^\dagger\big)}{\tilde{\Gamma}_i}\Big) = B_0. \tag{6.38}$$

Proof: See appendix. □

6.4.3 Dominance Condition

Recall that to convert (6.24) to (6.30), the *dominance condition* (6.29) should be satisfied. By (6.36), the *dominance condition* is equivalent to

$$\frac{r_{i,n}}{r_{j,n}} \geq \sqrt{\frac{\Gamma_{i,n}}{\Gamma_{j,n}}}, \tag{6.39}$$

for some $j \in S_1$ and $i \neq j$ where $r_{i,n} = [\mathbf{R}_i]_{n,n}$ and $\mathbf{R}_i$ is given by (6.35). In other words, only a particular class of JT which satisfies (6.39) yields optimal $\mathbf{P}$ and $\mathbf{B}_i$ for problem (6.24). Before further discussions, we introduce the following two lemmas.

Lemma 6.6: For $\mathbf{w} \in \mathcal{R}_{++}^N$ and $\mathbf{d} \in \mathcal{R}_{++}^N$, if $w_n \leq d_n$, then $w_{[n]} \leq d_{[n]}$, where $[n]$ denotes the index of the nth largest element in a vector. □

Proof: See appendix. □

Lemma 6.7: For $\mathbf{w} \in \mathcal{R}_{++}^N$ and $\mathbf{u} \in \mathcal{R}_{++}^N$, there exists a $\mathbf{d} \in \mathcal{R}_{++}^N$ such that

$$\mathbf{d} \prec_\times \mathbf{u}, \text{ and } w_n \leq d_n, \tag{6.40}$$

for $n \in S_2$ if and only if

$$\mathbf{w} \prec_{\times w} \mathbf{u}. \tag{6.41}$$

□

Proof: See appendix. □

For given channel state information, by Theorem 6.1 and Lemma 6.7, one can verify that the necessary condition for the existence of a joint triangularization satisfying the *dominance condition* is

$$\begin{aligned}\mathbf{\Gamma}(i,j) &= \Big[\sqrt{\frac{\Gamma_{i,0}}{\Gamma_{j,0}}}, \sqrt{\frac{\Gamma_{i,1}}{\Gamma_{j,1}}}, \ldots, \sqrt{\frac{\Gamma_{i,N-1}}{\Gamma_{j,N-1}}}\Big]^T \\ &\quad \prec_{\times w} \mathbf{u}(\mathbf{A}_i, \mathbf{A}_j),\end{aligned} \tag{6.42}$$

for some j and $i \neq j$ where $\Gamma_{i,n}$ are the predetermined *QoS coefficients* in (6.10) and $\mathbf{u}(\mathbf{A}_i, \mathbf{A}_j)$ is defined in Definition 1. Note that $\mathbf{u}(\mathbf{A}_i, \mathbf{A}_j)$ is determined by the channels, $\mathbf{H}_i$. For convenience, we name the necessary condition in particular the *admissibility condition*. By Lemma 6.7, the *admissibility condition* implies that there exists an $\mathbf{r} \in \mathcal{R}_{++}^N$ such that

$$\mathbf{r} \prec_{\times} \mathbf{u}(\mathbf{A}_i, \mathbf{A}_j), \text{ and } \mathbf{\Gamma}(i,j) \leq \mathbf{r}.$$

For the two-user case, it follows from Theorem 6.1 that $\Gamma_{i,n}$ fulfilling the *admissibility condition* also guarantees the existence of a joint triangularization satisfying the *dominance condition* in (6.39). For the case with more than two users, the existence of such a joint triangularization is not guaranteed even when the *admissibility condition* is fulfilled. This is mainly because there is no theoretical result which supports more than two matrices and guarantees the existence of a joint triangularization satisfying the *dominance condition* in (6.39), such as Theorem 6.1 for the two-user case.

In summary, to solve the power minimization problem in (6.9), we first solve for $\tilde{\mathbf{F}}$ from (6.37) numerically. Next, we construct the auxiliary design matrices $\mathbf{A}_i$ as in (6.35), and perform joint triangularizations on $\mathbf{A}_i$ to obtain $\mathbf{P}$ and $\mathbf{B}_i$. The optimal $\mathbf{G}_i$ and optimal bit allocation are given by (6.12) and (6.19), respectively. Suppose that there exists a joint triangularization satisfying the *dominance condition* in (6.39) for the given *QoS* coefficients, $\Gamma_{i,n}$, we can obtain an optimal design for the transmit power minimizing problem formulated in (6.9), namely the optimal minimum power JT broadcast DFE transceiver (MPJT).

6.4.4 Power Minimized QR (PMQR) Broadcast DFE Transceiver

For a more than two-user case, finding an optimal MPJT may be difficult or impossible even when the *QoS* coefficients, $\Gamma_{i,n}$, fulfill the *admissibility condition*. In this section, we propose a suboptimal design for the power minimization problem in (6.9). For an arbitrary number of users, we can set $\tilde{\mathbf{F}} = \alpha\tilde{\mathbf{F}}^*$, where $\alpha > 0$ and $\tilde{\mathbf{F}}^*$ is obtained by solving (6.37). QR decompositions are performed on the auxiliary matrices in (6.35) to design $\mathbf{B}_i$ and $\mathbf{P}$ (which is $\mathbf{I}_N$). The feedforward equalizer $\mathbf{G}$ and bit allocation $\{b_n\}$ are given by (6.12) and (6.19), respectively. Here, the *dominance condition* in (6.29) is not guaranteed to be fulfilled. We need to check if the weighted MSE constraint in (6.24) is satisfied or not. In (6.24), $\prod_{n=0}^{N-1} \max_{i\in S_1}\left(\sigma_{i,n}^2\Gamma_{i,n}\right)$ is a non-increasing function of α by Lemma 6.3. To make the power in (6.24) as small as possible and fulfill the weighted MSE constraint, we iteratively adjust α and reiterate the transceiver design such that α is the smallest number satisfying

$$\prod_{n=0}^{N-1} \max_{i\in S_1}\left(\sigma_{i,n}^2\Gamma_{i,n}\right) \leq 2^{-B_0}.$$

The choice only gives us a suboptimal solution, which is referred to as the power minimized QR broadcast DFE transceiver (PMQR), for the original design problem (6.9).

The idea of iteratively designing the precoding matrices and equalizers at receivers is widely used in the context of multi-user MIMO system [59]-[63], where independent data streams for different users are encoded with different precoding matrices at the transmitter. In our case, common data streams for all the subscribed users are encoded with a single precoding matrix, and the bit allocation is included in the joint optimization. It is possible to exploit some ideas from [59]-[63], to further improve the iterative algorithm here in this section.

6.4.5 Integer Bit Allocation

Recall that in Sec. 6.4.1 we have relaxed the bit allocation, b_n, to real numbers. However, the square QAM symbols can only carry an even number of bits. Hence, a suboptimal strategy is required for practical implementation. We modify the bit allocation formula in (6.19) to

$$\bar{b}_n = 2\left\lfloor 0.5\Big(D - \log_2\big(\max_{i\in S_1}\{\Gamma_{i,n}\sigma_{i,n}^2\}\big)\Big)\right\rfloor, \tag{6.43}$$

which truncates b_n in (6.19) to the nearest even number smaller than or equal to b_n. Since we are truncating b_n, the constraint (a) in (6.11) is always satisfied while the total bit rate might not meet

the target rate B_0. Therefore, we propose the following algorithm.

1. Set $B = B_0$.
2. Solve optimization problem (6.11) and compute $\bar{b}_n$ by (6.43).
3. If $\sum_{n=0}^{N-1} \bar{b}_n \geq B$ then stop, else set $B_0 = B_0 + 1$ and go to 2).

There are at most $2N$ iterations for the above algorithm.

6.4.6 BER Constraints

In this section, we consider the min-power problem with per substream SER constraints. It is sometimes desirable to directly apply per substream BER constraints in (6.9), i.e.,

$$BER_{i,n} \leq \eta_{i,n}. \tag{6.44}$$

When $b_n \geq 2$ and the BER is smaller than 10^{-3}, it is shown in [64] that the BER of the QAM detector for the nth substream from user i is well approximated by

$$BER_{i,n} \approx 0.2 \exp\Big(\frac{-1.6\gamma_{i,n}}{2^{b_n}-1}\Big), \tag{6.45}$$

where $\gamma_{i,n}$ is given in (6.8). Since $BER_{i,n}$ is a decreasing and continuous function of $\gamma_{i,n}/(2^{b_n}-1)$, assuming $b_n >> 1$, one can re-express the BER constraints in (6.44) as in (6.10), except that the *QoS* coefficients $\Gamma_{i,n}$ are replaced by

$$\Gamma_{i,n} = \frac{\log(5\eta_{i,n})^{-1}}{1.6}. \tag{6.46}$$

Hence, the min-power problem with BER constraints can be treated as in (6.11). All the results in this section shall still follow.

6.5 Rate Maximized Broadcast DFE Transceivers with QoS Constraints

In applications where the data rate is the main concern, formulating data rate as the objective function directly gives the maximum rate design. In this section, we design the broadcast DFE

transceiver in Fig. 6.1 by maximizing the total bit rate under the total transmit power constraint and the per substream SER *QoS* specifications. To avoid a complicated integer programming problem, $b_n \in \mathcal{Z}^+$ is relaxed to $b_n \in \mathcal{R}$. The problem can be formulated as

$$\begin{aligned} \max_{\mathbf{F},\mathbf{G}_i,\mathbf{B}_i,\{b_n\}} \quad & \sum_{n\in S_2} b_n \\ s.t. \quad (a) \quad & \text{Tr}\,(\mathbf{F}\mathbf{F}^\dagger) \le P_0 \\ (b) \quad & SER_{i,n} \le \eta_{i,n},\ i \in S_1,\ n \in S_2, \end{aligned} \tag{6.47}$$

where P_0 denotes the total transmit power constraint, and $\eta_{i,n}$ is the nth substream SER constraint for the ith user, $S_1 = \{1, 2, \ldots, L\}$ and $S_2 = \{0, 1, \ldots, N-1\}$. Note that a careful treatment of the max-rate problem with per substream BER constraints is available later in Sec. 6.5.4. Also, the optimization problem in (6.47) reduces to the max-rate problem in [57] when there is only one subscriber and the zero-forcing constraint is applied. It further reduces to the max-rate problem in [56] and [58] if we only consider linear transceivers.

6.5.1 Optimal Bit Allocation

To solve (6.47), we shall first derive the *optimal bit allocation* for a given broadcast DFE transceiver. Using the SER expression as in (6.8) and the fact that $Q(\sqrt{x})$ is a monotonic decreasing function of x, the SER constraints in (6.47) can be re-expressed as

$$(2^{b_n} - 1)\sigma_{i,n}^2 \Gamma_{i,n} \le 1, \tag{6.48}$$

where the *QoS coefficient* $\Gamma_{i,n}$ is given by (6.10). Reducing L SER constraints in (6.48) to a single constraint, (6.47) becomes

$$\begin{aligned} \max_{\mathbf{F},\mathbf{G}_i,\mathbf{B}_i,\{b_n\}} \quad & \sum_{n\in S_2} b_n \\ s.t. \quad (a) \quad & \text{Tr}\,(\mathbf{F}\mathbf{F}^\dagger) \le P_0 \\ (b) \quad & (2^{b_n} - 1) \max_{i\in S_1}\{\sigma_{i,n}^2 \Gamma_{i,n}\} \le 1. \end{aligned} \tag{6.49}$$

One can claim that the equality in the second constraint of (6.49) holds for the optimal solution $\{b_n^*\}$. Otherwise, the total bit rate can be made even higher by increasing $\{b_n^*\}$ until the equalities are met, which leads to a contradiction. Without loss of optimality, we can replace the inequality

constraint (b) in (6.49) with

$$(2^{b_n}-1)\max_{i\in S_1}\{\sigma_{i,n}^2\Gamma_{i,n}\}=1,$$

and hence the *optimal bit allocation* is given by

$$b_n^*=\min_{i\in S_1}\log_2\Big(1+\frac{1}{\sigma_{i,n}^2\Gamma_{i,n}}\Big). \tag{6.50}$$

Applying the *optimal bit allocation* and setting $\mathbf{F}=\tilde{\mathbf{F}}\mathbf{P}$ where $\mathbf{P}$ is an $N\times N$ unitary matrix and $\tilde{\mathbf{F}}\in\mathcal{C}^{N\times N}$, (6.49) can be reduced to

$$\begin{aligned}\max_{\tilde{\mathbf{F}},\mathbf{P},\mathbf{G}_i,\mathbf{B}_i}\quad & \sum_{n\in S_2}\min_{i\in S_1}\log_2\Big(1+\frac{1}{\sigma_{i,n}^2\Gamma_{i,n}}\Big)\\ s.t.\quad & \text{Tr}\,(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^\dagger)\le P_0.\end{aligned} \tag{6.51}$$

In the above discussions, we have relaxed b_n to be real numbers. However, the square QAM symbols only carry an even number of bits. In practice, one can modify the *optimal bit allocation formula* in (6.50) as

$$\bar{b}_n=2\Big\lfloor 0.5\min_{i\in S_1}\log_2\Big(1+\frac{1}{\sigma_{i,n}^2\Gamma_{i,n}}\Big)\Big\rfloor, \tag{6.52}$$

which truncates b_n^* in (6.50) to the nearest even number smaller than or equal to b_n^*. The new integer bit allocation, $\{\bar{b}_n\}$, is still a valid but suboptimal solution (i.e., even integer) for (6.47) since the SER constraints in (6.48) are still satisfied.

For fixed $\tilde{\mathbf{F}}$, the remaining optimization problem of (6.51) is

$$\max_{\mathbf{P},\mathbf{G}_i,\mathbf{B}_i}\sum_{n\in S_2}\min_{i\in S_1}\Big\{\log_2\big(1+\frac{1}{\sigma_{i,n}^2\Gamma_{i,n}}\big)\Big\}. \tag{6.53}$$

By taking logs on both sides of (6.26) in Lemma 6.4, we obtain the upper bound for (6.53)

$$\begin{aligned}&\max_{\mathbf{P},\mathbf{G}_i,\mathbf{B}_i}\sum_{n\in S_2}\min_{i\in S_1}\Big\{\log_2\Big(1+\frac{1}{\sigma_{i,n}^2\Gamma_{i,n}}\Big)\Big\}\\ &\le\max_{\mathbf{P},\mathbf{G}_i,\mathbf{B}_i}\min_{i\in S_1}\Big\{\sum_{n\in S_2}\log_2\Big(1+\frac{1}{\sigma_{i,n}^2\Gamma_{i,n}}\Big)\Big\};\end{aligned} \tag{6.54}$$

the equality holds if and only if the *dominance condition* as in (6.29) is satisfied. At high SNR, the

upper bound in (6.54) can be approximated as

$$\max_{\mathbf{P},\mathbf{G}_i,\mathbf{B}_i} \min_{i\in S_1}\Big\{\log_2\Big(\frac{1}{\tilde{\Gamma}_i \prod_{n\in S_2}\sigma_{i,n}^2}\Big)\Big\}, \tag{6.55}$$

where $\tilde{\Gamma}_i = \prod_{n\in S_2}\Gamma_{i,n}$. Applying the *max-min inequality* [104] on the upper bound, we have

$$\max_{\mathbf{P},\mathbf{G}_i,\mathbf{B}_i} \min_{i\in S_1} f_i(\mathbf{P},\mathbf{G}_i,\mathbf{B}_i) \le \min_{i\in S_1}\max_{\mathbf{P},\mathbf{G}_i,\mathbf{B}_i} f_i(\mathbf{P},\mathbf{G}_i,\mathbf{B}_i), \tag{6.56}$$

where

$$f_i(\mathbf{P},\mathbf{G}_i,\mathbf{B}_i) = \log_2\Big(\frac{1}{\tilde{\Gamma}_i \prod_{n\in S_2}\sigma_{i,n}^2}\Big).$$

The equality holds when the function f satisfies the *saddle point property*, i.e., there exists a $j \in S_1$ such that

$$f_j(\mathbf{P},\mathbf{G}_j,\mathbf{B}_j) \le f_i(\mathbf{P},\mathbf{G}_i,\mathbf{B}_i), \tag{6.57}$$

for every $i \in S_1$. One can verify by substitution that the *dominance condition* as in (6.29) is sufficient for (6.57) to hold. Suppose that the *dominance condition* holds, the optimization problem in (6.51) can be translated to

$$\begin{aligned} \max_{\tilde{\mathbf{F}}} \quad & \min_{i\in S_1}\max_{\mathbf{P},\mathbf{G}_i,\mathbf{B}_i} \Big\{\log_2\Big(\frac{1}{\tilde{\Gamma}_i \prod_{n\in S_2}\sigma_{i,n}^2}\Big)\Big\} \\ s.t. \quad & \text{Tr}\,(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^\dagger) \le P_0. \end{aligned} \tag{6.58}$$

6.5.2 Optimal Feed-Forward, Feedback Matrices and Precoder

Since the objective function in (6.58) is a monotonic decreasing function of the MSEs, $\sigma_{i,n}^2$, [11], [25], the optimal feed-forward matrices $\mathbf{G}_i^*$ for given precoding matrices, $\tilde{\mathbf{F}}$ and $\mathbf{P}$, and feedback matrices $\mathbf{B}_i$, have the same form as in (6.12). Applying the optimal $\mathbf{G}^*$, the error covariance matrices, $\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i}$, in (6.6) can be shown to be in the form of (6.13).

To optimize $\mathbf{B}_i$ and $\mathbf{P}$, we consider the MSE product in (6.58)

$$\tilde{\Gamma}_i \prod_{n\in S_2}\sigma_{i,n}^2 = \tilde{\Gamma}_i \prod_{n\in S_2}[\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i}]_{n,n} \ge \tilde{\Gamma}_i \det(\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i}), \tag{6.59}$$

where the inequality follows from Hadamard's inequality and the equality holds if and only if $\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i}$ is diagonal. The lower bound,

$$\tilde{\Gamma}_i \det(\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i}) = \frac{\tilde{\Gamma}_i}{\det(\mathbf{I}_{J_i} + \mathbf{H}_i\tilde{\mathbf{F}}\tilde{\mathbf{F}}^\dagger\mathbf{H}_i^\dagger)}, \tag{6.60}$$

is independent of the choice of $\mathbf{B}_i$ and $\mathbf{P}$. On account of the objective function, $\log_2(1/(\tilde{\Gamma}_i \prod_{n\in S_2} \sigma_{i,n}^2))$, being a monotonic decreasing function of $\tilde{\Gamma}_i \prod_{n\in S_2} \sigma_{i,n}^2$, the optimal $\mathbf{B}_i$ and $\mathbf{P}$ of (6.58) for given $\tilde{\mathbf{F}}$ and $\mathbf{G}_i^*$ are such that $\tilde{\Gamma}_i \prod_{n\in S_2} \sigma_{i,n}^2$ is minimized. By (6.59) and (6.60), one can see that the optimal $\mathbf{P}$ and $\mathbf{B}_i$ will make $\mathbf{R}_{\mathbf{e}_i\mathbf{e}_i}$ in (6.13) diagonal, and can be obtained from the joint triangularizations as in (6.35).

Applying the optimal $\mathbf{G}_i^*$, $\mathbf{P}$ and $\mathbf{B}_i$ to (6.58), it follows from (6.59) and (6.60) that (6.58) can be reduced to

$$\begin{aligned} \max_{\tilde{\mathbf{F}}} \quad & \min_{i\in S_1} \log_2\Big(\frac{\det\big(\mathbf{I}_{J_i} + \mathbf{H}_i\tilde{\mathbf{F}}\tilde{\mathbf{F}}^\dagger\mathbf{H}_i^\dagger\big)}{\tilde{\Gamma}_i}\Big) \\ s.t. \quad & \text{Tr}\,(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^\dagger) \le P_0. \end{aligned} \tag{6.61}$$

The problem falls into the category of a determinant maximization problem with linear matrix inequality [47]. The optimal solution $\tilde{\mathbf{F}}^*$ can be solved numerically and has the following property.

Property 2 The optimal solution $\tilde{\mathbf{F}}^*$ for (6.61) is such that

$$Tr\,(\tilde{\mathbf{F}}^*\tilde{\mathbf{F}}^{*\dagger}) = P_0. \tag{6.62}$$

Proof: See appendix. □

In summary, to obtain the solution for the rate maximization problem in (6.47), we first solve for the optimal $\tilde{\mathbf{F}}$ from (6.61) numerically, construct the auxiliary design matrices $\mathbf{A}_i$ as in (6.35), and perform joint triangularizations on $\mathbf{A}_i$ to obtain $\mathbf{P}$ and $\mathbf{B}_i$. The optimal $\mathbf{G}_i$ and optimal bit allocation are given by (6.12) and (6.50), respectively. Recall that to convert (6.51) to (6.58), the *dominance condition* as in (6.39) should be satisfied. If $\mathbf{B}_i$ and $\mathbf{P}$ are obtained from the joint triangularization satisfying the *dominance condition* for the given *QoS* coefficients , the whole design $(\tilde{\mathbf{F}}, \mathbf{P}, \mathbf{B}_i, \mathbf{G}_i, \{b_n\})$, namely the optimal maximum rate JT broadcast DFE transceiver (MRJT), is optimal for the rate maximization problem (6.47). The existence of such a joint triangularization was elaborated in Sec. 6.4.3.

6.5.3 Rate Maximized QR (RMQR) Broadcast DFE Transceiver

For an arbitrary number of users, we can first solve for the optimal $\tilde{\mathbf{F}}$ from (6.61) numerically and perform QR decompositions on the auxiliary matrices in (6.35) to design $\mathbf{B}_i$ and $\mathbf{P}$ (which is $\mathbf{I}_N$). The $\mathbf{G}_i$ and bit allocation are given by (6.12) and (6.50), respectively. QR decompositions only give us a suboptimal solution since the *dominance condition* is not guaranteed to hold. So, (6.51) does not translate to (6.58) in general. We refer to the design for the rate maximization problem (6.47) as the bit rate maximized QR broadcast DFE transceiver (RMQR).

6.5.4 BER Constraints

As in Sec. 6.4.6, it is sometimes desirable to directly apply per substream BER constraints in (6.47). Using the BER expression as in (6.45) and the fact that $\exp(-x)$ is a monotonic decreasing function of x, the BER constraints in (6.44) can be re-expressed as in (6.48), except that the *QoS* coefficients $\Gamma_{i,n}$ are given by (6.46). Hence, the max-rate problem with BER constraints can be treated as (6.49). All the results in this section shall still follow.

6.6 The Duality of MPJT and MRJT

In the previous sections, we have shown that the minimum power JT broadcast DFE transceiver (MPJT) is optimal for the power minimization problem in (6.9), and the maximum rate JT broadcast DFE transceiver (MRJT) is optimal for the rate maximization problem in (6.47), if the *dominance conditions* are satisfied. Although MPJT and MRJT are optimal solutions for two different problems, respectively, they are well connected. While a MPJT is an optimal solution for the power minimization in (6.9), it is also an optimal solution for the rate maximization in (6.47) with the power constraint set to the minimal power obtained in (6.9). Similarly, the MRJT is optimal for the power minimization problem in (6.9) with target rate B_0 being the maximal rate from (6.47). The results are formally stated by the following two theorems.

Theorem 6.4: Suppose the MPJT, $\chi_1 = (\mathbf{F}^*, \mathbf{B}_i^*, \mathbf{G}_i^*, \{b_n^*\})$, is optimal for the power minimization problem in (6.9) with target bitrate B_0 and transmit power P_0. Then, χ_1 is also an optimal solution for the rate maximization problem in (6.47) with the transmit power constraint, P_0, given such that

the *dominance condition* as in (6.39) holds for

$$\tilde{\mathbf{F}} = \alpha \tilde{\mathbf{F}}^*, \tag{6.63}$$

where $\tilde{\mathbf{F}}^*$ is an optimal solution for (6.61), $\alpha = 1$ or $1 - \epsilon$, and ϵ is an arbitrary small positive number. □

Proof: Suppose the MPJT, $\chi_1 = (\mathbf{F}^*, \mathbf{B}_i^*, \mathbf{G}_i^*, \{b_n^*\})$, is optimal for the power minimization problem in (6.9) under the total bitrate constraint, $\sum_{n \in S_2} b_n^* \geq B_0$, and the SER constraints, $SER_{i,n} \leq \eta_{i,n}$. The minimal transmit power is Tr $(\mathbf{F}^* \mathbf{F}^{*\dagger}) = P_0$. By Theorem 6.2, the total bitrate is

$$\sum_{n \in S_2} b_n^* = B_0.$$

Take P_0 as the power constraint and consider the rate maximization problem in (6.47) under the same SER constraints. Given that the *dominance condition* holds, there exists an MRJT, $\chi_2 = (\overline{\mathbf{F}}, \overline{\mathbf{B}}_i, \overline{\mathbf{G}}_i, \{\bar{b}_n\})$, which is an optimal solution for the rate maximization problem and the maximal rate is

$$B_1 = \sum_{n \in S_2} \bar{b}_n = \min_{i \in S_1} \log_2 \Big(\frac{\det\big(\mathbf{I}_{J_i} + \mathbf{H}_i \overline{\mathbf{F}\mathbf{F}}^\dagger \mathbf{H}_i^\dagger\big)}{\tilde{\Gamma}_i} \Big), \tag{6.64}$$

where the last equality follows from (6.61). By *Property 2*, the transmit power is

$$\text{Tr } (\overline{\mathbf{F}\mathbf{F}}^\dagger) = P_0.$$

The rate $B_1 \geq B_0$, otherwise, it will lead to a contradiction on the optimality of χ_2 since χ_1 is a feasible solution for the rate maximization problem with a higher rate if $B_0 > B_1$. If $B_1 = B_0$, χ_1 achieves the same rate as χ_2 and hence is also an optimal solution for the rate maximization problem. For the case where $B_1 > B_0$, we shall show in the following that it leads to a contradiction. For $B_1 > B_0$, we can scale $\overline{\mathbf{F}}$ by a factor of $\alpha = 1 - \epsilon$ (ϵ is a small positive number) and re-design an MRJT such that it requires less power than P_0 to achieve the rate B_0 since B_1 in (6.64) is a continuous strictly increasing function of α by Lemma 6.5. This leads to a contradiction for χ_1 being optimal for (6.9). □

Theorem 6.5:

Suppose the MRJT, $\chi_1 = (\mathbf{F}^*, \mathbf{B}_i^*, \mathbf{G}_i^*, \{b_n^*\})$, is optimal for the rate maximization problem in (6.47)

under the power constraint, P_0, and the maximum rate is B_0. Then, χ_1 is also an optimal solution for the power minimization problem in (6.9) with the target rate, B_0, given that the *dominance condition* in (6.39) holds for $\tilde{\mathbf{F}} = \alpha\tilde{\mathbf{F}}^*$, where $\tilde{\mathbf{F}}^*$ is an optimal solution for (6.37), $\alpha = 1$ or $1+\epsilon$, and ϵ is an arbitrary small positive number. □

Proof: Suppose the MRJT, $\chi_1 = (\mathbf{F}^*, \mathbf{B}_i^*, \mathbf{G}_i^*, \{b_n^*\})$, is optimal for the rate maximization problem in (6.47) under the power constraint, Tr $(\mathbf{F}\mathbf{F}^\dagger) \leq P_0$. The maximum rate is assumed to be B_0. By Property 2, the transmit power is

$$\text{Tr } (\mathbf{F}^*\mathbf{F}^{*\dagger}) = P_0.$$

Take B_0 as the rate constraint and consider the power minimization problem in (6.9) under the same SER constraints. Given that the *dominance condition* holds, there exists an MPJT, $\chi_2 = (\overline{\mathbf{F}}, \overline{\mathbf{B}}_i, \overline{\mathbf{G}}_i, \{\bar{b}_n\})$, which is an optimal solution, and the minimal power is Tr $(\overline{\mathbf{F}}\overline{\mathbf{F}}^\dagger) = P_1$. The power $P_1 \leq P_0$. The case where $P_1 > P_0$ leads to a contradiction on the optimality of χ_2 for (6.9) since χ_1 is a feasible solution of (6.9) and requires less power than χ_2. If $P_1 = P_0$, χ_1 achieves the same power as χ_2 and hence is also the solution for the power minimization problem. For the case where $P_1 < P_0$, a contradiction shall arise as shown in the following. For the power minimization problem in (6.9), the optimal bit allocation is given by (6.19) and the total rate is

$$B_0 = \min_{i \in S_1} \log_2\Big(\frac{\det\big(\mathbf{I}_{J_i} + \mathbf{H}_i\overline{\mathbf{F}}\overline{\mathbf{F}}^\dagger\mathbf{H}_i^\dagger\big)}{\tilde{\Gamma}_i}\Big), \tag{6.65}$$

where the equality follows from Property 1. One can scale $\overline{\mathbf{F}}$ by a factor of $\alpha = 1+\epsilon$ (ϵ is an small positive number) and re-design the MPJT such that the new transceiver achieves a rate higher than B_0 with the transmit power less than or equal to P_0 since B_0 in (6.65) is a continuous strictly increasing function of α by Lemma 6.5. This is a contradiction with χ_1 being optimal for (6.47). □

Following the same techniques as in Theorem 6.4 and Theorem 6.5, one can establish similar duality results for the min-power and max-rate problems with per substream BER constraints. Furthermore, suppose that the optimal $\mathbf{G}_i$ for min-power problem (6.9) is given by (6.12), one can follow similar procedures as the above to show that the optimal solution for (6.9) is optimal for (6.47), and the optimal solution for (6.47) is also optimal for (6.9) even when the dominance condition is not satisfied.

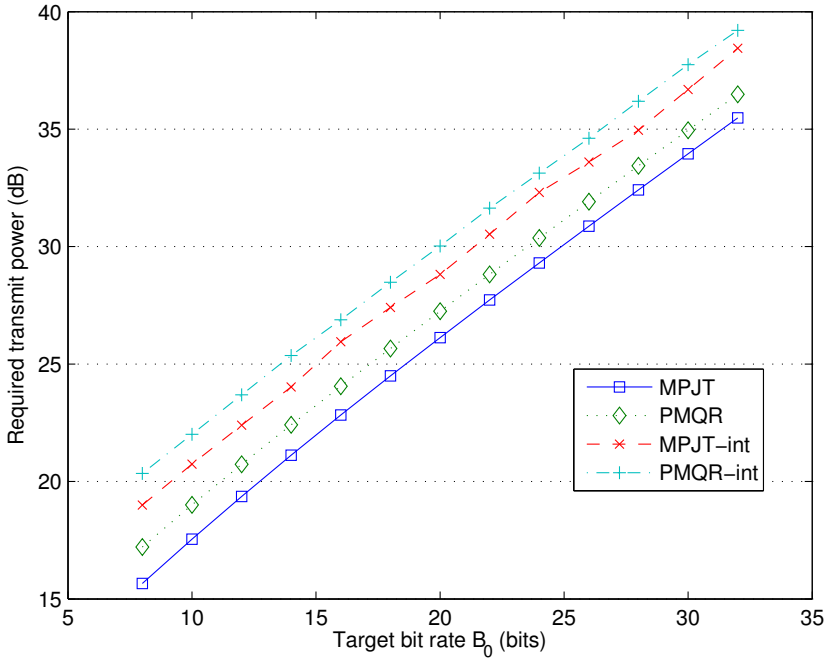


Figure 6.2: Average minimal transmit power versus target bit rate B_0

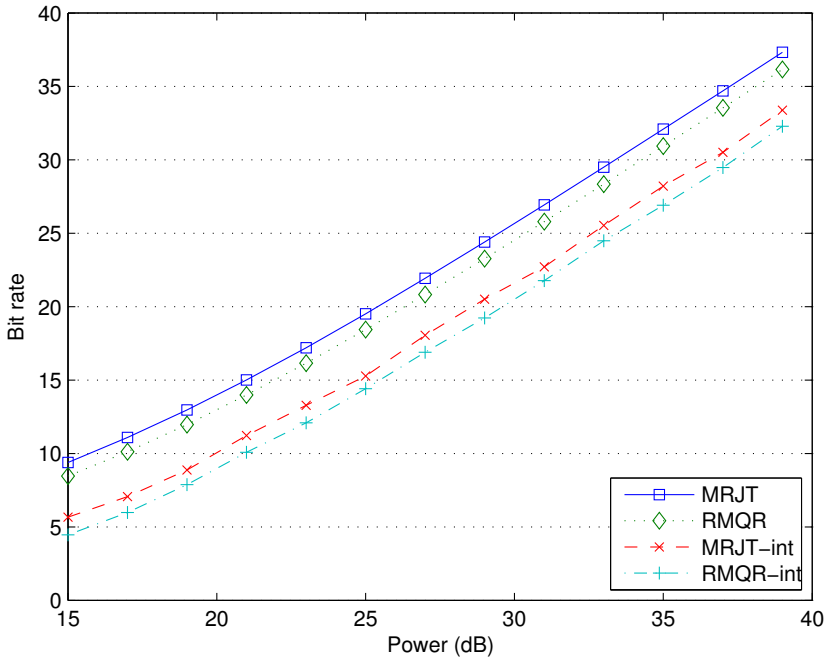


Figure 6.3: Average total bit rate versus transmission power P_0

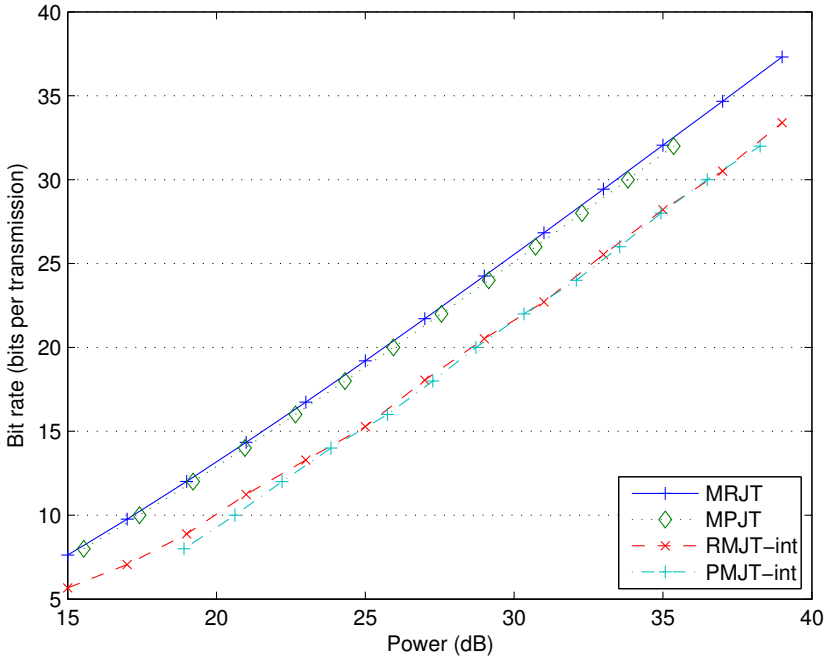


Figure 6.4: Duality of max-rate and min-power problem

6.7 Numerical Results

In this section, we present several numerical results on the design of the proposed MPJT, PMQR, RMJT and RMQR transceivers. A two-user scheme is considered and the SER constraints in (6.9) and (6.47) are set such that $\eta_{i,n} = 10^{-3}$. Choosing the *QoS* parameters, $\Gamma_{i,n}$, in this way guarantees that the *admissibility condition* in (6.42) can always be fulfilled because either

$$\mathbf{\Gamma}(1,2) = \mathbf{1} \prec_{\times w} \mathbf{u}(\mathbf{A}_1, \mathbf{A}_2), \text{ or}$$

$$\mathbf{\Gamma}(2,1) = \mathbf{1} \prec_{\times w} \mathbf{u}(\mathbf{A}_2, \mathbf{A}_1).$$

The optimal solutions for (6.9) and (6.47), the MPJT and MRJT, respectively, are therefore obtainable. The channels for both users are 4×4 MIMO Rayleigh channels $\mathbf{H}_i$. The elements of $\mathbf{H}_i$ are complex Gaussian random variables whose real and imaginary parts are independent with zero mean and variance $1/2$. The channel noise is described in (6.4). Both the transmitter and receiver have perfect channel state information.

Example 1: In this example, we solve the power minimization problem in (6.9) for each of 500 channel realizations. For different target bit rates, we compute the minimal transmit power for both MPJT and PMQR transceivers, and the power is averaged over all channel realizations. For integer bit allocation, we use the suboptimal strategy introduced in Sec. 6.4.5 which gives even $\bar{b}_n$. The results are shown in Fig. 6.2 where the suffix "int" indicates the use of suboptimal solutions given in Sec. 6.4.5 for MPJT and PMQR transceivers. The figure shows that both the MPJT transceivers with optimal bit allocation and suboptimal integer bit allocation always require less power than the PMQR transceivers with optimal bit allocation and suboptimal integer bit allocation, respectively. For both the optimal and integer allocation, the MPJT transceivers have $1.3dB$ gain over the PMQR transceivers. We also observe that each transceiver requires approximately $2.9dB$ more power doing even number bit allocation than the optimal bit allocation.

Example 2: In this example, we design the proposed MRJT and RMQR transceivers by maximizing the bit rate for different transmission power constraints. The designs are done for 500 randomly generated channel realizations, and for each realization we compute bit rate via (6.50) and (6.52). The total bit rates are averaged over 500 channel realizations and shown in Fig 6.3. The suffix "int" indicates the use of (6.52), the even number bit allocation. For both cases of optimal bit allocation and even number bit allocation, the results show that the MRJT broadcast DFE transceivers achieve higher bit rate than the RMQR broadcast DFE transceiver. For the same bit rate, the MRJT broadcast DFE transceiver has 1.3dB power advantage over the RMQR broadcast DFE transceiver. We also observe that each transceiver requires approximately $2.9dB$ more power doing even number bit allocation than doing the optimal bit allocation.

The numerical results of the proposed MRJT and MPJT broadcast DFE transceivers are plotted together in Figure 6.4 for comparison. The lines of MPJT and MRJT broadcast DFE transceivers overlap almost everywhere. This implies that when both transceivers operate at the same transmit power, their bit rates are the same. It suggests the duality of the proposed MRJT broadcast DFE transceiver and the MPJT broadcast DFE transceiver as discussed in Sec. 6.6.

6.8 Conclusions

In this chapter, we solved the min-power and max-rate *QoS* design problems for MIMO broadcast DFE transceivers. By applying a particular class of joint triangularization which satisfies the *dominance condition*, we proposed MPJT and MRJT transceivers, which are optimal solutions for

the min-power and max-rate *QoS* problems, respectively. Two corresponding suboptimal broadcast DFE transceivers, MPQR and MRQR transceivers, which exploit simple QR decompositions, were also proposed for the min-power and max-rate *QoS* problems, respectively. Numerical results show that the suboptimal designs require 1.3dB more power than the optimal designs. Integer bit allocation algorithms were also developed for the proposed systems. We then showed the duality of the proposed MPJT and MRJT transceivers for the min-power and max-rate *QoS* problems and the results were verified by numerical examples.

6.9 Appendix

6.9.1 Proof of Lemma 6.3

Substituting $\mathbf{F} = \alpha\mathbf{F}_1$ into (6.13), one gets

$$\sigma_{i,n}^2 = [(\mathbf{I}_N + \mathbf{B}_i)(\mathbf{I}_N + \alpha^2\mathbf{C}_i^\dagger\mathbf{C}_i)^{-1}(\mathbf{I}_N + \mathbf{B}_i)^\dagger]_{n,n},$$

where $\mathbf{C}_i = \mathbf{H}_i\mathbf{F}_1$. The first derivative of $\sigma_{i,n}^2$ with respect to α is

$$\frac{\partial\sigma_{i,n}^2}{\partial\alpha} = [(\mathbf{I} + \mathbf{B}_i)\frac{\partial\mathbf{D}_i}{\partial\alpha}(\mathbf{I} + \mathbf{B}_i)^\dagger]_{n,n},$$

where $\mathbf{D}_i = (\mathbf{I} + \alpha^2\mathbf{F}_1^\dagger\mathbf{H}_i^\dagger\mathbf{H}_i\mathbf{F})^{-1}$. The derivative of $\mathbf{D}_i$ with respect to α is given by [109]

$$\frac{\partial\mathbf{D}_i}{\partial\alpha} = -\mathbf{D}_i\frac{\partial{\mathbf{D}_i}^{-1}}{\partial\alpha}\mathbf{D}_i^\dagger = -2\alpha\mathbf{D}_i\mathbf{C}_i^\dagger\mathbf{C}_i\mathbf{D}_i^\dagger. \tag{6.66}$$

Hence, one can conclude that

$$\frac{\partial\sigma_{i,n}^2}{\partial\alpha} = -2\alpha[\mathbf{A}_i^\dagger\mathbf{A}_i]_{n,n} \leq 0,$$

where $\mathbf{A}_i = \mathbf{C}_i\mathbf{D}_i^\dagger(\mathbf{I} + \mathbf{B}_i)^\dagger$ and $\sigma_{i,n}^2$ is a non-increasing function of α.

6.9.2 Proof of Lemma 6.4

The inequality (6.25) follows immediately from the fact that

$$\prod_{n\in S_2}\max_{i\in S_1}\{f(a_{i,n})\} \geq \prod_{n\in S_2} f(a_{i,n}).$$

We shall first prove the sufficient condition of the equality. Suppose there exists a $j \in S_1$ such that

$$a_{j,n} \geq a_{i,n},$$

for all $i \in S_1$ and $n \in S_2$. Then, we have

$$\begin{aligned}\prod_{n\in S_2} \max_{i\in S_1}\{f(a_{i,n})\} &= \prod_{n\in S_2} f(a_{j,n}) \\ &= \max_{i\in S_1}\Big\{ \prod_{n\in S_2} f(a_{i,n})\Big\}.\end{aligned}$$

The necessary part shall be proved by contradiction. Assume that the equality in (6.25) holds, and for every $j \in S_1$, there exists $i \neq j$ such that

$$a_{j,n} < a_{i,n}$$

for some n. This shall lead to

$$\prod_{n\in S_2} f(a_{j,n}) < \prod_{n\in S_2} \max_{i\in S_1}\{f(a_{i,n})\}. \tag{6.67}$$

Taking $\max_{j\in S_1}$ on both sides, (6.67) becomes

$$\max_{j\in S_1}\Big\{ \prod_{n\in S_2} f(a_{j,n})\Big\} < \prod_{n\in S_2} \max_{i\in S_1}\{f(a_{i,n})\},$$

which contradicts with the original assumption. For (6.26), the proof is similar.

6.9.3 Proof of Lemma 6.5

By Jacobi's formula [110],

$$\begin{aligned}\frac{\partial \det(\mathbf{A})}{\partial \alpha} &= \det(\mathbf{A})\text{Tr}\,(\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial \alpha}) \\ &= 2\alpha \det(\mathbf{A})\text{Tr}\,(\mathbf{H}_i\tilde{\mathbf{F}}\mathbf{A}^{-1}\tilde{\mathbf{F}}^\dagger\mathbf{H}_i^\dagger) > 0,\end{aligned} \tag{6.68}$$

where $\mathbf{A} = \big(\mathbf{I}_N + \alpha^2\tilde{\mathbf{F}}^\dagger\mathbf{H}_i^\dagger\mathbf{H}_i\tilde{\mathbf{F}}\big)$. The last inequality follows from $\alpha > 0$ and $\mathbf{A}$ being positive definite. Hence, $g(\alpha)$ is a continuous strictly increasing function of α.

6.9.4 Proof of Property 1

Suppose that $\tilde{\mathbf{F}}$ in (6.37) is scaled by a positive factor of α. By Lemma 6.5, the function

$$\min_{i\in S_1} \log_2\Big(\frac{\det\big(\mathbf{I}_{J_i} + \alpha^2 \mathbf{H}_i \tilde{\mathbf{F}}\tilde{\mathbf{F}}^\dagger \mathbf{H}_i^\dagger\big)}{\tilde{\Gamma}_i}\Big) \tag{6.69}$$

is a strictly increasing function of α. Let us assume that the optimal solution for (6.37) is such that

$$\min_{i\in S_1} \log_2\Big(\frac{\det\big(\mathbf{I}_{J_i} + \mathbf{H}_i \tilde{\mathbf{F}}^*\tilde{\mathbf{F}}^{*\dagger} \mathbf{H}_i^\dagger\big)}{\tilde{\Gamma}_i}\Big) > B_0,$$

and the minimal power is $P_0 = \text{Tr}\,(\tilde{\mathbf{F}}^*\tilde{\mathbf{F}}^{*\dagger})$. One can scale $\tilde{\mathbf{F}}^*$ by $\alpha = 1-\epsilon$, where ϵ is an arbitrary small positive number, such that

$$\min_{i\in S_1} \log_2\Big(\frac{\det\big(\mathbf{I}_{J_i} + \alpha^2 \mathbf{H}_i \tilde{\mathbf{F}}^*\tilde{\mathbf{F}}^{*\dagger} \mathbf{H}_i^\dagger\big)}{\tilde{\Gamma}_i}\Big) \geq B_0.$$

The adjusted power

$$\alpha^2 \text{Tr}\,(\tilde{\mathbf{F}}^*\tilde{\mathbf{F}}^{*\dagger}) < P_0,$$

which contradicts with the optimality of $\tilde{\mathbf{F}}^*$. Hence, we can conclude the proof.

6.9.5 Proof of Lemma 6.6

Let us define a permutation of a vector, $\mathbf{x} \in \mathcal{R}^N$, by a permutation function $m = g(n)$, where $0 \leq n, m \leq N-1$. The permutation takes the mth element of $\mathbf{x}$ and copies it to the nth entry of the permuted vector. The permuted vector defined by $g(n)$ is given by

$$\mathbf{x}_g = [x_{g(0)}, x_{g(1)}, \ldots, x_{g(N-1)}]^T.$$

Let $g_1(n)$ denote the index of the nth largest element in $\mathbf{w}$. Given that

$$\mathbf{w} \leq \mathbf{d}, \tag{6.70}$$

where "$\leq$" denotes element-wise inequality, the permutation defined by $g_1(n)$ preserves the ordering, i.e.,

$$\mathbf{w}_{g_1} \leq \mathbf{d}_{g_1}. \tag{6.71}$$

Suppose the permutations, $g_1(n)$ and $g(n)$, preserve the ordering in (6.70), i.e.,

$$\mathbf{w}_{g_1} \leq \mathbf{d}_{g}. \tag{6.72}$$

If the elements of $\mathbf{d}_g$ are not in decreasing order, there exist n_1 and n_2 such that $d_{g(n_1)} < d_{g(n_2)}$ where $0 \leq n_1 < n_2 \leq N-1$. We can define a new permutation function

$$g_2(n) = \begin{cases} g(n_2), n = n_1 \\ g(n_1), n = n_2 \\ g(n), n \neq n_1, n_2. \end{cases}$$

The new permuted vector $\mathbf{d}_{g_2}$ is the same as $\mathbf{d}_g$ except that $d_{g(n_1)}$ and $d_{g(n_2)}$ are swapped. Using the fact that $w_{g_1(n_1)} \geq w_{g_1(n_2)}$ and (6.72), we have

$$w_{g_1(n_1)} \leq d_{g(n_1)} < d_{g(n_2)} = d_{g_2(n_1)},$$
$$w_{g_1(n_2)} \leq w_{g_1(n_1)} \leq d_{g(n_1)} = d_{g_2(n_2)}.$$

Hence, it shows that the permutation $g_2(n)$, preserves the ordering in (6.72), i.e.,

$$\mathbf{w}_{g_1} \leq \mathbf{d}_{g_2}. \tag{6.73}$$

Taking $\mathbf{w}_{g_1}$ and $\mathbf{d}_{g_1}$ from (6.71), one can iteratively apply $g_2(n)$ like permutations on $\mathbf{d}_{g_1}$ such that the elements of $\mathbf{d}_{g_1}$ are in descending order and the ordering is preserved as in (6.73). Finally, we have

$$w_{[n]} \leq d_{[n]}.$$

6.9.6 Proof of Lemma 6.7

The sufficient part of the proof is similar to Proposition 5.A.9 in [95]. The necessary part can be shown via the definition of weakly majorization. Suppose $\mathbf{d} \prec_{\times} \mathbf{u}$ and $w_n \leq d_n$. Since $w_n \leq d_n$, Lemma 6.6 shows that

$$w_{[n]} \leq d_{[n]}. \tag{6.74}$$

By the definition of multiplicative majorization, we have

$$\prod_{i=0}^{n} u_{[i]} \geq \prod_{i=0}^{n} d_{[i]} \geq \prod_{i=0}^{n} w_{[i]}, \tag{6.75}$$

where $0 \leq n \leq N-1$ and the second inequality follows from (6.74). Hence, by definition $\mathbf{w} \prec_{\times w} \mathbf{u}$.

6.9.7 Proof of Property 2

Let us assume that the optimal solution for (6.61) is such that Tr $(\tilde{\mathbf{F}}^{*}\tilde{\mathbf{F}}^{*\dagger}) < P_0$ and the maximum rate is

$$B_0 = \min_{i \in S_1} \log_2\Big(\frac{\det\big(\mathbf{I}_{J_i} + \mathbf{H}_i\tilde{\mathbf{F}}^{*}\tilde{\mathbf{F}}^{*\dagger}\mathbf{H}_i^{\dagger}\big)}{\tilde{\Gamma}_i}\Big).$$

One can scale $\tilde{\mathbf{F}}^{*}$ by $\alpha = 1+\epsilon$ (ϵ is an arbitrary small positive number) such that α^2Tr $(\tilde{\mathbf{F}}^{*}\tilde{\mathbf{F}}^{*\dagger}) \leq P_0$. Since the objective rate function as in (6.69) is a strictly increasing function of α, the modified rate

$$\min_{i \in S_1} \log_2\Big(\frac{\det\big(\mathbf{I}_{J_i} + \alpha^2\mathbf{H}_i\tilde{\mathbf{F}}^{*}\tilde{\mathbf{F}}^{*\dagger}\mathbf{H}_i^{\dagger}\big)}{\tilde{\Gamma}_i}\Big) > B_0,$$

which contradicts with the optimality of $\tilde{\mathbf{F}}^{*}$. Therefore, Tr $(\tilde{\mathbf{F}}^{*}\tilde{\mathbf{F}}^{*\dagger}) = P_0$.

Chapter 7

Optimized DFT-FB Transceivers over Linear Time-Varying Channels

In this chapter, we consider the optimization of a DFT modulated filterbank transceiver (DFT-FBT) over linear time varying (LTV) channels. The DFT-FBT is a generalization of the Affine Fourier transform based OFDM (Affine OFDM) and the chirped OFDM, which are suggested in recent literature for transmission over LTV channels. For both known LTV channels and unknown wide sense stationary uncorrelated scattering (WSSUS) statistical channels, we show how to optimize the transmitting and receiving prototypes of DFT-FBT such that the signal to interference and noise ratio (SINR) at the receiver is maximized. After the optimization, the channel dependent part, like OFDM, is a set of scalar multipliers at the receiver end that adapts to the equivalent memoryless channel on a block basis. Simulation results show that the bit error rate (BER) performance of the optimized DFT-FBT over LTV channels is superior to the Affine OFDM. Most of the results in this chapter have been reported in our paper [66].

7.1 Introduction

In recent years, orthogonal frequency division multiplexing (OFDM) systems have found many applications in wideband communications. One of the advantages of OFDM systems is their ability to combat ISI induced by the transmission over frequency selective channels efficiently. By the insertion of cyclic prefix (CP) of length L at the transmitter, the inverse discrete Fourier transform (IDFT) and discrete Fourier transform (DFT) operations at the transmitter and receiver can convert any frequency selective channel of order L into parallel one-tap subchannels. And the equalizer at the receiver is just a set of scalar multipliers. Due to its many fascinating features, OFDM tech-

nology has been adopted for wireless wideband communications systems, like DVB-T [71] and the evolving IEEE 802.16e standard for WiMAX [72]. In these applications, the wireless channels are typically LTV multipath channels, which are also called doubly selective channels. The time variation may be caused by Doppler shift due to the mobility of the transmitter or receiver, carrier frequency offset, or phase noise [99], [73]. This leads us to study the case where the channel changes continuously within one block time. Windowed Fourier functions in $\mathbf{L}^2(\mathbb{R}^2)$ [101] have been reported in [74], [75], serving as good approximate eigenfunctions of practical LTV channels. They are in the form of $f(n-lN)\exp(j2\pi kn/M)$, and the transceiver based on these eigenfunctions can be represented as a DFT-FBT as in Fig. 7.1.

Moreover, the results reported in [68], [69] also confirm this approximation. They show that if the support of the spreading function $S(\upsilon,\tau)$ of the LTV channels is maximally concentrated on a line in delay-Doppler plane, the chirp modulated Fourier basis is a set of eigenfunctions of these channels. The chirped OFDM schemes in [67], [68] are generalized by [69] as Affine OFDM. However, they all fail to diagonalize general LTV channels in which the support is not necessarily on a line in the delay-Doppler plane. The authors propose algorithms to optimize the chirp rate parameters at the transmitter and receiver to minimize MSE or maximize SIR at the receiver. But their algorithms are highly nonlinear and the perfect knowledge of LTV channels is required.

In this chapter, we consider the DFT-FBT whose prototypes are arbitrary and not limited to the form of chirp waveforms as in Affine OFDM. Our design goal is to optimize DFT-FBT over LTV channels such that the SINR is maximized at the receiving detectors. With the treatment of inter-symbol interference (ISI) and inter-carrier interference (ICI) as Gaussian noise as in multiuser detection for CDMA systems [100], the equivalent channel is memoryless within one block time. Standard channel estimation techniques of OFDM can be applied directly to design the FEQ [70] to correct scalar ambiguity from block to block. Firstly, the case in which the transmitter and receiver have perfect knowledge of the channel is considered. We extend the SIR maximization algorithm for DFT-FBT over LTI channels [76] to our case where the channel is LTV. We show how to optimize the transmitter and the receiver such that the SIR at the receiver is maximized. In addition, a new iterative algorithm that allows us to maximize the SINR of the DFT-FBT is proposed. Secondly, WSSUS stochastic channels [99] are considered. Based on the statistics of the channel, we propose an algorithm that optimizes the prototypes for the maximizations of average SIR/SINR. In this case, both prototypes are stationary for fixed statistics.

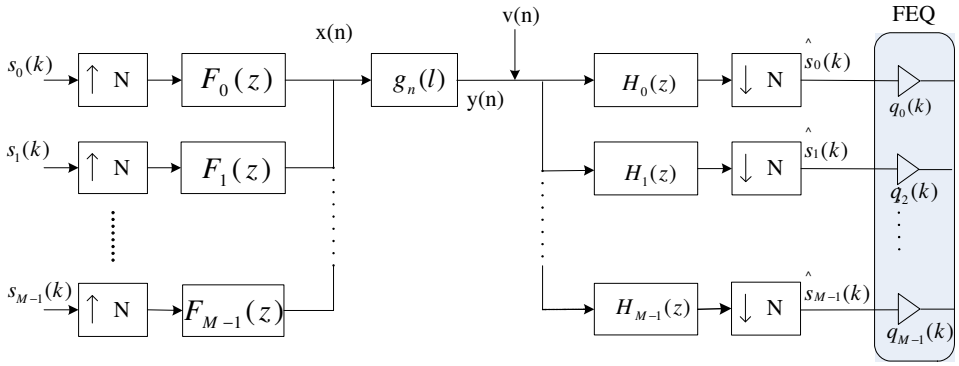


Figure 7.1: General filter bank transceiver

7.2 Outline

The sections in this chapter are structured as follows. We start by introducing a filter bank transceiver for LTV channels, and derive the input/output description for such system in Section 7.3. In Section 7.4, we address the DFT-FBT optimization problem by proposing an iterative algorithm which maximizes the SIR at the receiver, based on the full knowledge of channel state information. Section 7.5 considers the SIR maximization problem of a DFT-FBT for WSSUS channels in which only channel statistics is available. In Section 7.6, the SINR maximization of the DFT-FBT is addressed for both cases of channels. Numerical examples for the designs are provided in Section 7.7. Concluding remarks are given in Section 7.8.

7.3 The FBT for LTV Channels

The block diagram of a FB transceiver is shown in Fig. 7.1. There are M subchannels and the decimation ratio is N. We assume that $L = N - M$ redundant samples are added, where $N > M$. The LTV channel is characterized by the channel response $g_n(l)$, for $0 \leq l \leq L$. In the following discussions, we will derive the input-output (I/O) description of a FBT over LTV channels. Assume that all transmitting filters, $f_i(n)$, and and receiving filters, $h_i(n)$, are of order n_f and n_h, respectively. Let

$$\tilde{g}_i^{(k)}(n) = \sum_{l=0}^{L} g_n(l) f_i[n - kN - l], \tag{7.1}$$

which has support on $n \in kN, kN+1, \ldots, kN+n_f+L$. The output signal of the channel is

$$y(n) = \sum_{k=-\infty}^{\infty} \sum_{i=0}^{M-1} s_i(k) \tilde{g}_i^{(k)}(n). \tag{7.2}$$

The downsampled signal after the m-th subband filter $H_m(z)$ can be expressed in Z-domain as

$$\hat{S}_m(z) = \sum_{k=-\infty}^{\infty} \sum_{i=0}^{M-1} s_i(k) [H_m(z) \tilde{G}_i^{(k)}(z)]_{\downarrow N}, \tag{7.3}$$

in which $\tilde{G}_i^{(k)}(z) = \sum_{n=-\infty}^{\infty} \tilde{g}_i^{(k)}(n) z^{-n}$. Define

$$g_i^{(k)}(n) = \tilde{g}_i^{(k)}(n+kN), \text{for } n \in 0, 1, \ldots, n_f + L.$$

With $z_{k,i}(n)$ denoting the contribution of $s_i(k)$ to $\hat{S}_m(z)$, the advanced version $z_{k,i}(n+k)$ has Z-transform

$$s_i(k) [H_m(z) G_i^{(k)}(z)]_{\downarrow N}.$$

To facilitate the derivation of the I/O description, we define

$$\left[H_m(z) G_i^{(k)}(z) \right]_{\downarrow N} = \sum_n b_{i,m}^{(k)}(n) z^{-n}. \tag{7.4}$$

So, the output signal $\hat{s}_m(k)$ can be expressed as

$$\begin{aligned} \hat{s}_m(k) &= s_m(k) b_{m,m}^{(k)}(0) + \sum_{n \neq 0} s_m(k-n) b_{m,m}^{(k-n)}(n) \\ &\quad + \sum_{i \neq m} \sum_n s_i(k-n) b_{i,m}^{(k-n)}(n). \end{aligned}$$

From the above expression, we have the time-varying I/O description for the FBT from the i-th input to the m-th output at the k-th block time as

$$T_{i,m}^{(k)}(z) = b_{i,m}^{(k)}(0) + \sum_{n \neq 0} b_{i,m}^{(k-n)}(n) z^{-n}. \tag{7.5}$$

For the m-th output, the ISI and ICI come from $\sum_{n \neq 0} b_{m,m}^{(k-n)}(n) z^{-n}$ and $\sum_n b_{i,m}^{(k-n)}(n) z^{-n}$, respectively, for all $i \neq m$. Note that $T_{i,m}^{(k)}(z)$ is a useful notation, although there is no physical system with this transfer function. Unfortunately, the ISI-free solution for LTV channels is unknown. To

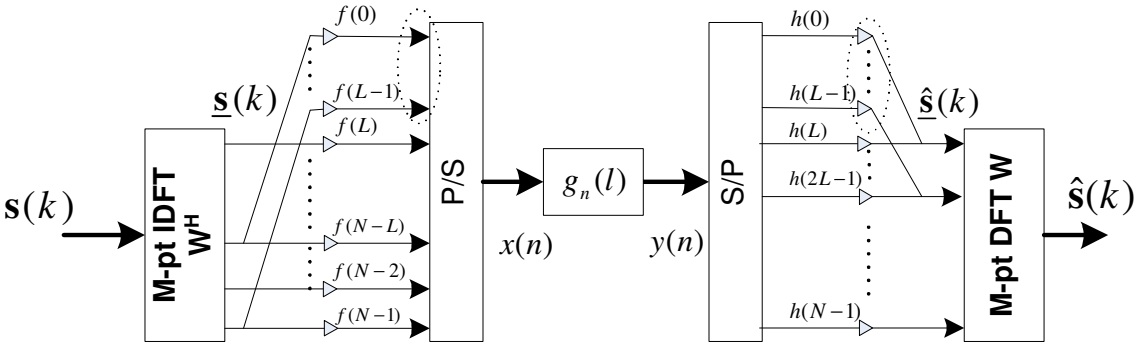


Figure 7.2: DFT-FBT transceiver with $n_h = n_f = N - 1$

have near ISI-free property, the transmitting filters $F_i(z)$ and the receiving filters $H_i(z)$ could be optimized such that these terms are as small as possible.

For efficient implementation, the class of DFT-FBT is often considered as a good candidate. The transmitting filters $F_i(z)$ and the receiving filters $H_i(z)$, for $0 \leq i \leq M-1$, are modulated versions of prototype filters $h(n)$ and $f(n)$ given by

$$\begin{aligned} F_i(z) &= W_M^{iL} \sum_{n=0}^{n_f} f(n) W_M^{-in} z^{-n} \\ H_i(z) &= W_M^{-iL} \sum_{n=0}^{n_h} h(n) W_M^{in} z^{n}, \end{aligned} \tag{7.6}$$

where $W_M = \exp(-j2\pi/M)$.

The special case of DFT-FBT with $n_f = n_h = N-1$ is as in Fig. 7.2. It becomes the Affine OFDM if $h(0) = \ldots = h(L-1) = 0$, and

$$\begin{aligned} f(n) &= e^{j2\pi c_1 (n-L)^2}, \\ h(n) &= e^{j2\pi (c_0 (n-L) - c_1 (n-L)^2)}, \end{aligned} \tag{7.7}$$

in which c_0 and c_1 are constant. The authors in [69] show that if the LTV channel takes the form of

$$g_n(l) = v(l) e^{-j2\pi c_1 l^2} e^{j2\pi (2c_1 l - c_0) n}, \tag{7.8}$$

where $v(l)$ is a function of integer l, the coefficients $h(n)$ and $f(n)$ would make the equivalent channel matrix from $\underline{\mathbf{s}}(k)$ to $\hat{\underline{\mathbf{s}}}(k)$ a circulant matrix. Hence, the LTV channel is diagonalizable by

$\mathbf{W}$ and $\mathbf{W}^\dagger$. This is called the linear delay/Doppler spreading channel in [68].

7.4 SIR Optimization of DFT-FBT for Known LTV Channels

For the class of linear delay/Doppler spreading channel, the channel is always diagonalized, if the prototypes of a DFT-FBT take the form of (7.7). However, LTV channels in general are not linear delay/Doppler spreading, i.e., the support of $S(v,l)$ is not necessarily on a line. Unfortunately, there is no ISI-free solution in terms of $F_i(z)$ and $H_i(z)$ for general LTV channels. In the following, we consider the class of DFT-FBT and show how to design the prototype filters so that SIR at the receiver is maximized.

In the following discussions, the input signals $s_i(n)$ are assumed to be zero mean WSS and white random process with the same variance, i.e,

$$\mathbb{E}\{s_i(n')s_j^*(n)\} = \sigma_s^2\delta(i-j)\delta(n'-n). \tag{7.9}$$

Assume that the prototypes adapt to the channel on a per block basis. Let $\mathbf{f}^{(k)} = \left[f^{(k)}(0)f^{(k)}(1)\dots f^{(k)}(n_f)\right]^T$ and $\mathbf{h}^{(k)} = \left[h^{(k)}(0)h^{(k)}(1)\dots h^{(k)}(n_h)\right]^T$ be the transmitting and receiving prototypes, respectively, at the k-th block time. Also, let n_h and n_f be no greater than $N-1$.

Define a $1\times(n_h+1)$ row vector and a $1\times(n_f+L+1)$ row vector as

$$\begin{aligned}
\mathbf{b}_{i,f}^{(k)}(n) &= \left[g_i^{(k)}(nN) \quad \dots \quad g_i^{(k)}(n_h+nN)\right] \\
\mathbf{b}_{i,h}^{(k)}(n) &= \left[h_i^{(k)}(-nN) \quad \dots \quad h_i^{(k)}(n_f+L-1-nN)\right]
\end{aligned}$$

where $h_i^{(k)}(\cdot)$ is the i-th receiving filter at the k-th block time. So, $b_{i,m}^{(k)}(n)$ can be expressed as the product of matrices as

$$b_{i,m}^{(k)}(n) = \mathbf{b}_{i,f}^{(k)}(n)\mathbf{D}_m\mathbf{h}^{(k)} = \mathbf{b}_{m,h}^{(k)}(n)\mathbf{G}^{(k)}\mathbf{D'}_i\mathbf{f}^{(k)}, \tag{7.10}$$

where

$$\begin{aligned}
\mathbf{D}_i &= diag(1, W_M^i, \dots, W_M^{in_h})W_M^{-iL}, \\
\mathbf{D'}_i &= diag(1, W_M^{-i}, \dots, W_M^{-in_f})W_M^{iL},
\end{aligned}$$

$$\mathbf{G}^{(k)} = \begin{bmatrix} g_{kN}(0) & 0 & \cdots & 0 \\ \vdots & g_{kN+1}(0) & \ddots & \vdots \\ g_{kN+L}(L) & \vdots & \ddots & 0 \\ 0 & g_{kN+L+1}(L) & & g_{kN+nf}(0) \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & g_{kN+L+nf}(L) \end{bmatrix}.$$

The power gains for the desired signal, the ISI and the ICI of the m-th output $\hat{s}_m(k)$ can be shown to be

$$\begin{aligned} \frac{P_{sig}^{(k)}(m)}{\sigma_s^2} &= \left| b_{m,m}^{(k)}(0) \right|^2 \\ &= \mathbf{h}^{(k)\dagger} \mathbf{D}_m^\dagger \mathbf{b}_{m,f}^{(k)\dagger}(0) \mathbf{b}_{m,f}^{(k)}(0) \mathbf{D}_m \mathbf{h}^{(k)}, \\ \frac{P_{isi}^{(k)}(m)}{\sigma_s^2} &= \sum_{n \neq 0} \left| b_{m,m}^{(k-n)}(n) \right|^2 \\ &= \sum_{n \neq 0} \mathbf{h}^{(k-n)\dagger} \mathbf{D}_m^\dagger \mathbf{b}_{m,f}^{(k-n)\dagger}(n) \mathbf{b}_{m,f}^{(k-n)}(n) \mathbf{D}_m \mathbf{h}^{(k-n)}, \\ \frac{P_{ici}^{(k)}(m)}{\sigma_s^2} &= \sum_{n, i \neq m} \left| b_{i,m}^{(k-n)}(n) \right|^2 \\ &= \sum_{i \neq m, n} \mathbf{h}^{(k-n)\dagger} \mathbf{D}_m^\dagger \mathbf{b}_{i,f}^{(k-n)\dagger}(n) \mathbf{b}_{i,f}^{(k-n)}(n) \mathbf{D}_m \mathbf{h}^{(k-n)}. \end{aligned} \tag{7.11}$$

Therefore, the SIR at the k-th block time can be expressed as

$$SIR_{\mathbf{h}}^{(k)} = \frac{\mathbf{h}^{(k)\dagger} \mathbf{Q}_{0,f}^{(k)} \mathbf{h}^{(k)}}{\mathbf{h}^{(k)\dagger} \mathbf{Q}_{1,f}^{(k)} \mathbf{h}^{(k)} + \sum_{n \neq 0} \mathbf{h}^{(k-n)\dagger} \mathbf{Q}_{2,f}^{(k-n)}(n) \mathbf{h}^{(k-n)}}$$

where $\mathbf{Q}_{0,f}^{(k)}$, $\mathbf{Q}_{1,f}^{(k-n)}$ and $\mathbf{Q}_{2,f}^{(k-n)}$ are positive definite matrices given by

$$\begin{aligned} \mathbf{Q}_{0,f}^{(k)} &= \sigma_s^2 \sum_m \mathbf{D}_m^\dagger \mathbf{b}_{m,f}^{(k)\dagger}(0) \mathbf{b}_{m,f}^{(k)}(0) \mathbf{D}_m, \\ \mathbf{Q}_{1,f}^{(k)} &= \sigma_s^2 \sum_{m, i \neq m} \mathbf{D}_m^\dagger \mathbf{b}_{i,f}^{(k)\dagger}(0) \mathbf{b}_{i,f}^{(k)}(0) \mathbf{D}_m, \\ \mathbf{Q}_{2,f}^{(k)}(n) &= \sigma_s^2 \sum_{m,i} \mathbf{D}_m^\dagger \mathbf{b}_{i,f}^{(k)\dagger}(n) \mathbf{b}_{i,f}^{(k)}(n) \mathbf{D}_m. \end{aligned}$$

Similarly we can express the same SIR in terms of $\mathbf{f}^{(k)}$ as

$$SIR_{\mathbf{f}}^{(k)} = \frac{\mathbf{f}^{(k)\dagger}\mathbf{Q}_{0,h}^{(k)}\mathbf{f}^{(k)}}{\mathbf{f}^{(k)\dagger}\mathbf{Q}_{1,h}^{(k)}\mathbf{f}^{(k)} + \sum_{n\neq 0}\mathbf{f}^{(k-n)\dagger}\mathbf{Q}_{2,h}^{(k-n)}(n)\mathbf{f}^{(k-n)}},$$

where $\mathbf{Q}_{0,h}^{(k)}$, $\mathbf{Q}_{1,h}^{(k)}$ and $\mathbf{Q}_{2,h}^{(k)}(n)$ are positive definite matrices similar to $\mathbf{Q}_{0,f}^{(k)}$, $\mathbf{Q}_{1,f}^{(k)}$ and $\mathbf{Q}_{2,f}^{(k)}(n)$ by replacing $\mathbf{b}_{m,f}^{(k)}(n)$ and $\mathbf{D}_m$ with $\mathbf{b}_{m,h}^{(k)}(n)$ and $\mathbf{G}^{(k)}\mathbf{D'}_m$, respectively.

$SIR^{(k)}$ depends on the prototypes prior to block k. It is not easy to optimize $SIR^{(k)}$ with respect to $\mathbf{h}^{(k)}$ or $\mathbf{f}^{(k)}$. Observe that if the first L coefficients of the $\mathbf{h}^{(k)}$'s are set to zero, the inter-block interference (IBI) $\sum_{n\neq 0}\mathbf{h}^{(k-n)\dagger}\mathbf{Q}_{2,f}^{(k-n)}(n)\mathbf{h}^{(k-n)}$ is zero. Hence, we assume that all the blocks are friendly blocks such that they do not allow too much IBI to leak to the other blocks. Suppose that the IBI is much smaller than the ICI within the block itself, i.e.,

$$\begin{aligned}\mathbf{h}^{(k)\dagger}\mathbf{Q}_{1,f}^{(k)}\mathbf{h}^{(k)} &\gg \sum_{n\neq 0}\mathbf{h}^{(k-n)\dagger}\mathbf{Q}_{2,f}^{(k-n)}(n)\mathbf{h}^{(k-n)},\\ \mathbf{h}^{(k)\dagger}\mathbf{Q}_{1,f}^{(k)}\mathbf{h}^{(k)} &\gg \sum_{n\neq 0}\mathbf{h}^{(k)\dagger}\mathbf{Q}_{2,f}^{(k)}(n)\mathbf{h}^{(k)}.\end{aligned} \tag{7.12}$$

So,

$$SIR_{\mathbf{h}}^{(k)} \approx \frac{\mathbf{h}^{(k)\dagger}\mathbf{Q}_{0,f}^{(k)}\mathbf{h}^{(k)}}{\mathbf{h}^{(k)\dagger}(\mathbf{Q}_{1,f}^{(k)} + \sum_{n\neq 0}\mathbf{Q}_{2,f}^{(k)}(n))\mathbf{h}^{(k)}} = \widehat{SIR}_{\mathbf{h}}^{(k)}.$$

The SIR maximization problem becomes

$$\max_{\mathbf{h}^{(k)}} \frac{\mathbf{h}^{(k)\dagger}\mathbf{Q}_{0,f}^{(k)}\mathbf{h}^{(k)}}{\mathbf{h}^{(k)\dagger}(\mathbf{Q}_{1,f}^{(k)} + \sum_{n\neq 0}\mathbf{Q}_{2,f}^{(k)}(n))\mathbf{h}^{(k)}}, \text{ for all } k. \tag{7.13}$$

Solving (7.13) is much simpler since $\mathbf{h}^{(k)}$ could be optimized on a per block basis. The computer experiments of the optimization problem (7.13) for various LTV channels also confirm the friendly block assumptions in (7.12). Furthermore, if the assumptions are guaranteed to hold, the actual SIR at the receiver is approximately equal to the optimal SIR value in (7.13).

Similarly, given $\mathbf{h}^{(k)}$ and the channel response, and under similar assumptions, the optimal $\mathbf{f}^{(k)}$ can be obtained by solving the optimization problem,

$$\max_{\mathbf{f}^{(k)}} \frac{\mathbf{f}^{(k)\dagger}\mathbf{Q}_{0,h}^{(k)}\mathbf{f}^{(k)}}{\mathbf{f}^{(k)\dagger}(\mathbf{Q}_{1,h}^{(k)} + \sum_{n\neq 0}\mathbf{Q}_{2,h}^{(k)}(n))\mathbf{f}^{(k)}}, \text{ for all } k. \tag{7.14}$$

Given one of the two prototype filters, we can optimize the other prototype filter so that the SIR

is maximized. The optimal $\mathbf{h}^{(k)}$ and $\mathbf{f}^{(k)}$ can be obtained by solving the optimization problems, involving the Rayleigh-Ritz ratio, in (7.13) and (7.14), respectively. By solving the optimization problems alternatively and iteratively, the SIR will increase monotonically.

7.5 SIR Optimization for Unknown WSSUS Channels

For wideband wireless communications, the motion of the mobile terminal, and the variation of the surrounding objects makes channels time-varying. In the design of a wideband wireless communications system over LTV channels, channel estimation is a big challenge since the channel may change due to the Doppler effect within one block time. Instead of real-time LTV channel estimation, we can design a statistically optimized DFT-FBT according to the channel statistics. In the statistically optimized DFT-FBT, the transmitting and receiving filters are stationary if the channel statistics do not change. The only part that adapts from block to block is the FEQ which is designed based on the equivalent memoryless channel.

In the following, we consider $g_n(l)$ as a wide sense stationary uncorrelated scattering (WSSUS) channel [99]. The cross-correlation of the zero-mean stochastic processes $g_n(l)$ is given by

$$\mathbb{E}\{g_{n_1}(l_1)g^*_{n_2}(l_2)\} = R_{l_2}(n_1 - n_2)\delta(l_1 - l_2), \tag{7.15}$$

where $R_{l_2}(n)$ is any arbitrary function. We will propose an algorithm to optimize the prototypes $\mathbf{h}$ and $\mathbf{f}$ of the DFT-FBT such that the ratio of the average signal power to the average interference power is maximized. For convenience, we assume that $n_h = n_f$ and n_h could be greater than $N-1$ since the transmitting and receiving filters are stationary. The average signal, ISI and ICI power are calculated by taking the expectation on (7.11) over the random process $g_n(l)$. From (7.4) and (7.6), one can verify that

$$\begin{aligned} b^{(k)}_{i,m}(n) &= \sum_{n_1=0}^{n_h}\sum_{l=0}^{L} g_{n_1+(n+k)N}(l) f(n_1 - l + nN) \\ &\quad \cdot h(n_1) W_M^{-i(n_1+nN-l-L)+m(n_1-L)}. \end{aligned}$$

Based on WSSUS assumption, we can express $\mathbb{E}_g|b^{(k)}_{i,m}(n)|^2$ in terms of $\mathbf{h}$ as

$$\mathbb{E}_g \left| b^{(k)}_{i,m}(n) \right|^2 = \mathbf{h}^\dagger \mathbf{\Lambda}^\dagger_{m-i} \mathbf{J}_f(n) \mathbf{\Lambda}_{m-i} \mathbf{h}, \tag{7.16}$$

where

$$\begin{aligned}
\mathbf{J}_f(n) &= \sum_{l=0}^{L} \mathbf{\Delta}_{f,l,n}^{\dagger} \mathbf{J}_l \mathbf{\Delta}_{f,l,n} \\
\mathbf{\Lambda}_m &= diag(1, W_M^m, \ldots, W_M^{mn_h}) \\
\mathbf{\Delta}_{f,l,n} &= diag(f(-l+nN), \ldots, f(n_h - l + nN)),
\end{aligned}$$

and $\mathbf{J}_l$ is an $n_h + 1$ by $n_h + 1$ matrix whose (n_2, n_1)-th element is $R_l(n_1 - n_2)$.

Similarly, we can express $\mathbb{E}_g |b_{i,m}^{(k)}(n)|^2$ in terms of $\mathbf{f}$ as

$$\mathbb{E}_g \left| b_{i,m}^{(k)}(n) \right|^2 = \mathbf{f}^{\dagger} \mathbf{\Lambda}_{m-i}^{\dagger} \mathbf{J}_h(n) \mathbf{\Lambda}_{m-i} \mathbf{f}, \tag{7.17}$$

where

$$\begin{aligned}
\mathbf{J}_h(n) &= \sum_{l=0}^{L} \mathbf{S}_{-l+nN}^{\dagger} \mathbf{\Delta}_h^{\dagger} \mathbf{J}_l \mathbf{\Delta}_h \mathbf{S}_{-l+nN} \\
\mathbf{\Delta}_h &= diag(h(0), h(1), \ldots, h(n_h)),
\end{aligned}$$

and $\mathbf{S}_m$ is an n_h+1 by n_h+1 shifting matrix that shifts a vector up by m. Applying these expressions, the average signal, ISI, and ICI power can be reformulated in terms of $\mathbf{f}$ or $\mathbf{h}$ as

$$\begin{aligned}
\bar{P}_{sig}^{(k)}(m) &= \sigma_s^2 \mathbf{h}^{\dagger} \mathbf{J}_f(0) \mathbf{h} = \sigma_s^2 \mathbf{f}^{\dagger} \mathbf{J}_h(0) \mathbf{f}, \\
\bar{P}_{isi}^{(k)}(m) &= \sigma_s^2 \mathbf{h}^{\dagger} \sum_{n \neq 0} \mathbf{J}_f(n) \mathbf{h} = \sigma_s^2 \mathbf{f}^{\dagger} \sum_{n \neq 0} \mathbf{J}_h(n) \mathbf{f}, \\
\bar{P}_{ici}^{(k)}(m) &= \sigma_s^2 \mathbf{h}^{\dagger} (\sum_{j=1}^{M-1} \mathbf{\Lambda}_j^{\dagger} \sum_n \mathbf{J}_f(n) \mathbf{\Lambda}_j) \mathbf{h} \\
&= \sigma_s^2 \mathbf{f}^{\dagger} (\sum_{j=1}^{M-1} \mathbf{\Lambda}_j^{\dagger} \sum_n \mathbf{J}_h(n) \mathbf{\Lambda}_j) \mathbf{f}.
\end{aligned}$$

Note from the above expressions that the average signal, ISI, and ICI power are constant over all subbands and block time. The power is independent of the block index since both $\mathbf{h}$ and $\mathbf{f}$ are stationary over the blocks and the channel is WSSUS. The independence of power from the subchannel index comes from the nature of DFT-FBT and the averaging over channels. So, we have average SIR in terms of $\mathbf{h}$ as

$$\overline{SIR}_{\mathbf{h}} = \frac{\mathbf{h}^{\dagger} \overline{\mathbf{Q}}_{0,f} \mathbf{h}}{\mathbf{h}^{\dagger} \overline{\mathbf{Q}}_{1,f} \mathbf{h}}, \tag{7.18}$$

where $\overline{\mathbf{Q}}_{0,f}$, $\overline{\mathbf{Q}}_{1,f}$ of dimension n_h+1 by n_h+1 are positive definite matrices

$$\begin{aligned}\overline{\mathbf{Q}}_{0,f} &= \sigma_s^2 \mathbf{J}_f(0), \\ \overline{\mathbf{Q}}_{1,f} &= \sigma_s^2 (\sum_{m=0}^{M-1} \mathbf{\Lambda}_m^\dagger \sum_n \mathbf{J}_f(n) \mathbf{\Lambda}_m - \mathbf{J}_f(0)).\end{aligned} \tag{7.19}$$

Similarly, in terms of $\mathbf{f}$,

$$\overline{SIR}_{\mathbf{f}} = \frac{\mathbf{f}^\dagger \overline{\mathbf{Q}}_{0,h} \mathbf{f}}{\mathbf{f}^\dagger \overline{\mathbf{Q}}_{1,h} \mathbf{f}}, \tag{7.20}$$

where $\overline{\mathbf{Q}}_{0,h}$, $\overline{\mathbf{Q}}_{1,h}$ can be obtained from (7.19) by replacing $\mathbf{J}_f(n)$ by $\mathbf{J}_h(n)$. The optimal $\mathbf{f}$ and $\mathbf{h}$ can be obtained by solving the eigen problem involving the Rayleigh-Ritz ratio of (7.20) and (7.18). We can use the iterative algorithm mentioned in the previous section to get satisfactory SIR.

7.6 The SINR Optimization

The BER of communications systems depends on the receiving SINR. In light of this dependence, we propose algorithms to maximize the receiving SINR in this section. Assume that the noise $v(n)$ is AWGN with zero mean and variance σ_v^2. The power of the received noise after the receiving filter $H_k(z)$ can be shown to be

$$\sigma_v^2 \sum_{n=0}^{n_h} |h(n)|^2.$$

So, we can formulate the average SINR for the unknown channel case in terms of $\mathbf{h}$ as

$$\overline{SINR}_{\mathbf{h}} = \frac{\mathbf{h}^\dagger \overline{\mathbf{Q}}_{0,f} \mathbf{h}}{\mathbf{h}^\dagger (\overline{\mathbf{Q}}_{1,f} + \sigma_v^2 \mathbf{I}) \mathbf{h}}. \tag{7.21}$$

From the last section, we know that all the subchannels have the same the average $\overline{SINR}_{\mathbf{h}}$. So, the maximization of $\overline{SINR}_{\mathbf{h}}$ is equivalent to the minimization of the BER of DFT-FBT provided that the BER is a monotonic decreasing function of SINR. Given $\mathbf{f}$, we can solve for the optimal $\mathbf{h}$ such that SINR is maximized just as in SIR maximization. However, given $\mathbf{h}$, the SINR optimization problem is

$$\max_{\mathbf{f}} \frac{\mathbf{f}^\dagger \overline{\mathbf{Q}}_{0,h} \mathbf{f}}{\mathbf{f}^\dagger \overline{\mathbf{Q}}_{1,h} \mathbf{f} + \sigma_v^2 \mathbf{h}^\dagger \mathbf{h}} \text{ s.t. } \|\mathbf{f}\| \leq 1. \tag{7.22}$$

This is not a standard SIR maximization problem and cannot be solved directly. However, by Lemma 7.1 below, it is equivalent to

$$\max_{\mathbf{f}} \frac{\mathbf{f}^\dagger \overline{\mathbf{Q}}_{0,h} \mathbf{f}}{\mathbf{f}^\dagger (\overline{\mathbf{Q}}_{1,h} + \sigma_v^2 \mathbf{h}^\dagger \mathbf{h}) \mathbf{f}} \text{ s.t. } \|\mathbf{f}\| = 1, \tag{7.23}$$

where $\overline{\mathbf{Q}}_{0,h}$ and $\overline{\mathbf{Q}}_{1,h} + \sigma_v^2 \mathbf{h}^\dagger \mathbf{h}$ are positive definite matrices. A similar technique has been used in [77]. We recognize this is the same as the SIR maximization problem in the previous section. Therefore, we can solve $\mathbf{h}$ and $\mathbf{f}$ iteratively to get a satisfactory SINR value. The SINR maximization for DFT-FBT over known LTV channels can be solved similarly.

Lemma 7.1: The optimal solution of the optimization problem (7.22) is $\mathbf{f}^{**} = \mathbf{f}^*/\|\mathbf{f}^*\|$, where $\mathbf{f}^*$ is the optimal solution to the unconstrained optimization problem

$$\max_{\mathbf{f}} \frac{\mathbf{f}^\dagger \overline{\mathbf{Q}}_{0,h} \mathbf{f}}{\mathbf{f}^\dagger (\overline{\mathbf{Q}}_{1,h} + \sigma_v^2 \mathbf{h}^\dagger \mathbf{h}) \mathbf{f}}. \tag{7.24}$$

□

Proof: For any nontrivial $\|\mathbf{f}\| \leq 1$, let $\mathbf{f}_3 = \mathbf{f}/\|\mathbf{f}\|$ and $\eta = \sigma_v^2 \mathbf{h}^\dagger \mathbf{h} > 0$. The ratio

$$\frac{\mathbf{f}^\dagger \overline{\mathbf{Q}}_{0,h} \mathbf{f}}{\mathbf{f}^\dagger \overline{\mathbf{Q}}_{1,h} \mathbf{f} + \eta} = \frac{\mathbf{f_3}^\dagger \overline{\mathbf{Q}}_{0,h} \mathbf{f_3}}{\mathbf{f_3}^\dagger \overline{\mathbf{Q}}_{1,h} \mathbf{f_3} + \eta/\|\mathbf{f}\|^2} \leq \frac{\mathbf{f_3}^\dagger \overline{\mathbf{Q}}_{0,h} \mathbf{f_3}}{\mathbf{f_3}^\dagger \overline{\mathbf{Q}}_{1,h} \mathbf{f_3} + \eta}.$$

The last inequality holds if and only if $\|\mathbf{f}\| = 1$. So, for the optimization problem (7.22), we just need to consider the feasible set $\|\mathbf{f}\| = 1$, and this is equivalent to (7.23). And (7.23) is equivalent to the unconstrained problem in (7.24) since if $\mathbf{f}^*$ maximizes (7.24), so does $\mathbf{f}^*/\|\mathbf{f}^*\|$. □

7.7 Design Examples

In this section, we design DFT-FBT over WSSUS channels and compare the performance with Affine OFDM and traditional OFDM. The the channel response $g_n(l)$ is a wide-sense stationary stochastic process given by

$$g_n(l) = \sum_{i=0}^{K-1} \alpha_l(n) \delta(l-i), \tag{7.25}$$

where K is the number of multipaths and the path gains $\alpha_l(n)$ have zero mean and follow Clark's model [99]

$$\mathbb{E}[\alpha_l(n) \alpha^*(n_1)] = P_l J_0(2\pi f_D T_s (n - n_1)). \tag{7.26}$$

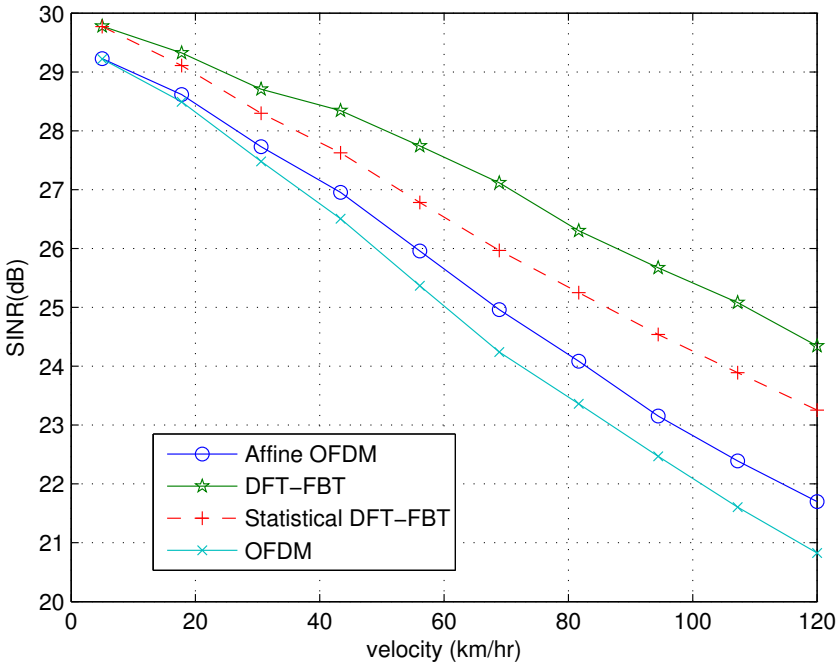


Figure 7.3: Optimized SINR versus speed

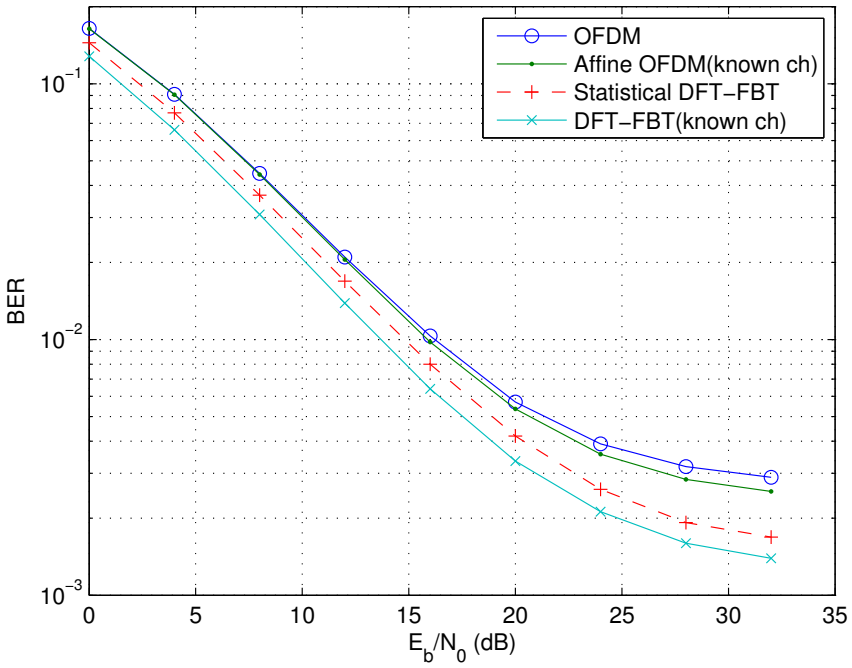


Figure 7.4: Bit error rate versus E_b/N_0

f_D is the maximum Doppler shift depending on the mobile agent velocity v and is given by $f_D = f_c v/c$, where c is the speed of light and f_c is the carrier frequency. P_l denotes the average power of the l-th path and T_s is the sampling period. Here, f_c is set to 5 GHz and T_s to $2\mu s$. The average power gain for the l-th path is $P_l = 2^{-l}$. The channel gains are normalized such that the sum of the power from all multipaths is one. The number of subbands is $M = 64$, $K = 9$, and the size of a block $N = 80$. The length of prototype filters $\mathbf{h}$ and $\mathbf{f}$ are 80. For the known channel case, Affine OFDM and DFT-FBT are designed based on $g_n(l)$. The Chirp rate parameters c_1 and c_0 in (7.7) of the Affine OFDM are designed and updated on a per block basis by the algorithm in [69]. We design our DFT-FBT by alternatively optimizing $\mathbf{h}^{(k)}$ and $\mathbf{f}^{(k)}$ to maximize the SINR. The prototypes are also updated on a per block basis. For the unknown channel case, we design the statistical optimized DFT-FBT according to the channel statistics (Sec. 7.5), P_l and f_D. For the optimization of $\mathbf{h}$ and $\mathbf{f}$, the objective function is the average SINR given in (7.18) and (7.20).

The plot of SINR at the receiver versus the mobile terminal velocity is shown in Fig. 7.3. The variance of AWGN noise is 0.001. The SINR is the average SINR over 500 realizations of channels. We observe that the increase of velocity, hence the Doppler frequency, lowers the SINR at the receiver. The DFT-FBT with known channel maintains really good SINR even at high Doppler frequency, while the SINRs of the traditional OFDM and the Affine OFDM drop tremendously. Furthermore, the statistical DFT-FBT (Sec. 7.5) outperforms the Affine OFDM even though it only has the statistics of the LTV channels and the prototypes are stationary over all transmission blocks.

Fig. 7.4 shows the bit error rate performance of OFDM, Affine OFDM, DFT-FBT and statistical DFT-FBT. FEQ as in Fig. 7.1 is used. 4-QAM symbols are transmitted in each subband. The BER is the average probability of error over 500 realizations of channels. The velocity of the mobile terminal is 100 km/hr. The Affine OFDM outperforms traditional OFDM only when E_b/N_0 is greater than 15 dB. The gain of DFT-FBT and statistical DFT-FBT over Affine OFDM increases as E_b/N_0 gets larger. When BER equals to $3 \cdot 10^{-3}$, the gains are approximately 7 dB and 4 dB, respectively.

7.8 Conclusions

We have proposed methods of designing DFT-FBT over known and unknown LTV channels. By formulating the SINR as a Rayleigh-Ritz ratio, we were able to optimize the prototypes of DFT-FBT such that the SINR at the receiver is maximized. Even when only channel statistics are available, the statistical DFT-FBT performs better than the Affine OFDM which requires exact channel knowl-

edge. In general, the optimized DFT-FBT provides substantial performance gain over the Affine OFDM in LTV channels.

Chapter 8

Channel Estimation for OFDM Systems with Co-Pilots

In Chapter 7, orthogonal frequency division multiplexing (OFDM) transceivers are designed based on channel state information (CSI). Channel estimation is an essential part of an OFDM system which provides updated CSI. In this chapter, a new pilot-aided subspace channel estimation algorithm for OFDM systems is proposed. The channel model is a sparse L-multipath channel with gains subject to Rayleigh fading. The new algorithm constructs M^2 virtual pilot tones from M physical pilot tones, and employs the subspace method, MUSIC, to estimate the multipath time delays. With the delay information, a minimum mean square error (MMSE) estimator is derived to estimate the frequency response of OFDM systems. The proposed algorithm can identify up to $O(M^2)$ multipath delays by using M pilot tones. It is also demonstrated that the proposed channel estimation algorithm can help the OFDM system to keep relatively good bit error rate (BER) performance, compared to the conventional subspace channel estimation method with uniform pilots, even when the number of multipaths is greater than the number of pilots. Moreover, it could potentially improve the spectral efficiency by reducing the number of pilots for given L.

8.1 Introduction

One of the advantages of OFDM systems is their ability to combat ISI induced by the transmission over frequency selective channels efficiently [106]. A good channel estimation method which tracks the time-varying channel parameters is essential in performing coherent demodulation for doubly selective fading channels. The continuous time impulse response of a doubly selective wireless channel for high speed data transmission is often modeled as the superposition of sparse multi-

paths and each of them is associated with some time delay [111]. The multipath time delays are slowly time varying while the gain of each multipath is relatively fast time varying and is subject to Rayleigh fading [79]. It is advisable to place pilot subcarriers in each OFDM symbol [78], [83].

The traditional channel estimation approach consists of two steps. First, the least square (LS) estimates are obtained over the pilot tones. Then these preliminary estimates are interpolated over the frequency grids. In [80] and [82], the MMSE channel estimators for OFDM system are derived based on the correlation in the frequency domain or time domain and the preliminary LS estimates at the pilot tones. The correlation is not easily available at the receiver, and may require complicated pilot placement and large overhead [86]. In [83], the frequency response is interpolated with a linear/spline function based on the LS channel estimates at pilot tones. However, the method may require a large number of pilot tones. The Fourier transform based methods in [81] and [85] first estimate the constrained temporal channel response under MMSE or LS criterion, and interpolate the frequency response with the Fourier transform. The FT-based methods have low complexity and may require a smaller number of pilot tones than the frequency interpolation methods in [83]. However, they require knowledge of the channel support in the temporal domain for good performance, and require the normalized multipath delays to be on integer grids. To give a comparison benchmark for the class of FT-based methods, we construct an FT-based MMSE estimator with a genie in Sec. 8.6. The genie has knowledge of the normalized channel delays rounded to the nearest integers.

The methods above do not exploit the nature of the parametric channel model in which the channel can be characterized by delays and complex amplitudes. In practice, the delays present much slower variations in time than the amplitudes. This may allow the two types of parameters to be handled independently. In [84], the authors use the subspace method which estimates the multipath delays of the parametric channel model from uniform pilot tones, and apply the delay information to design an MMSE estimator for the channel gains. The multipath delay estimation works at a much slower rate than the MMSE amplitude estimation due to the quasi-stationarity of the multipath delays. The subspace method allows the normalized multipath delays to be non-integers which can potentially improve the performance of the channel estimator over the FT-based channel estimators. However, the subspace channel estimator can never identify more multipaths than the number of pilots minus one. To get good channel estimation when the number of multipaths increases, one needs to add more pilots per OFDM symbol which sacrifices the spectral

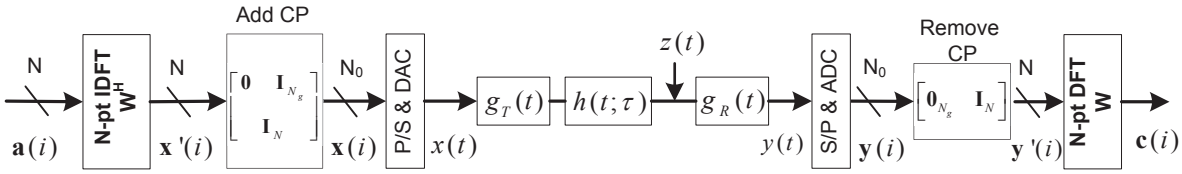


Figure 8.1: The OFDM system.

efficiency of the OFDM systems. Inspired by the work on co-array in [89], we introduce the alternating pilot placement in OFDM systems which allows us to construct M^2 virtual pilot tones from M physical pilot tones. The subspace method, MUSIC [90], can still be employed to estimate the multipath time delays. With acquired multipath delays, an MMSE estimator is designed to estimate the channel frequency response from the pilot tones in two consecutive OFDM symbols. Theoretically speaking, the proposed algorithm can identify up to $O(M^2)$ multipath delays by using M pilot tones. For a fixed number of multipath delays, the new algorithm can potentially increase the spectral efficiency of OFDM systems by using a smaller number of physical pilots.

8.2 Outline

The sections in this chapter are structured as follows. In Sec. 8.3, we introduce the channel models, OFDM systems and the concept of the nested array. An overview of the subspace channel estimation method with uniform pilots is presented in Sec. 8.4. In Sec. 8.5, we introduce the idea of co-pilots and propose a subspace channel estimation algorithm. Numerical examples for the designs are provided in Sec. 8.6. Concluding remarks are given in Sec. 8.7.

8.3 Preliminaries

In this section, we shall first describe the OFDM system model considered in this chapter. Sec. 8.3.2 reviews the concept of a nested array which plays a key role in the development of our new channel estimation algorithm.

8.3.1 System Model

Fig. 8.1 displays the block diagram of the OFDM system used in this paper. The system has N subcarriers, and for each transmission, a size N OFDM symbol $\mathbf{a}(i) = [a_0(i), a_1(i), \ldots a_{N-1}(i)]^T$ is transmitted through the subcarriers, where i represents the OFDM symbol index. Suppose the transmission bandwidth is $1/T_s$, so the sampling period for the OFDM signals is T_s. The transmit filter $g_T(t)$ is applied to limit the transmission bandwidth to $1/T_s$. It is assumed that the channel $h(t;\tau)$ is a multipath linear time-varying Rayleigh fading channel, characterized by

$$h(t;\tau) = \sum_{l=0}^{L-1} h_l(t)\delta(\tau - \tau_l(t)), \tag{8.1}$$

where $h_l(t)$ is the channel gain for the lth path, L is the total number of the multipaths, and $\tau_l(t)$ is the time delay for the lth path and is confined in the time interval $\tau_l \in [0, \tau_{max}]$. The channel gains, $\{h_l(t)\}$, are independent wide sense stationary complex Gaussian processes and their auto-correlations are characterized by the Jakes' model, in which

$$E[h_l(t+T)h_l^*(t)] = \sigma_{h_l}^2 J_0(2\pi f_{d,l} T), \tag{8.2}$$

where $J_0(\cdot)$ is the zero-th order Bessel function of first kind, $f_{d,l}$ is the Doppler spread of the lth path and $\sigma_{h_l}^2$ is the power of the lth path. The channel is time-varying as a result of the *slow* variations of $\tau_l(t)$ and the *fast* variation variations of the amplitudes $h_l(t)$ due to Doppler shift. To remove the intersymbol interference (ISI) caused by the multipaths, a sufficiently long cyclic prefix (CP) of length N_g is added in the front of each OFDM symbol at the transmitter end which brings the total length of an OFDM symbol to N_0 ($N_0 = N_g + N$). The CP is then removed at the receiver. We assume that the channel $h(t;\tau)$ is quasi-stationary, i.e., the channel does not change within one OFDM symbol time, $N_0 T_s$, but can vary from one OFDM symbol to another OFDM symbol. Moreover, the delays, $\tau_l(t)$, are assumed to be stationary for N_I (assumed to be even) OFDM symbols. In practice, each delay can be considered as a constant in N_I OFDM symbols provided that their variations in $N_I N_0 T_s$ are much smaller than the temporal resolution T_s of the system. To quantify, for a relative radial movement between a transmitter and a receiver with velocity ν, the variation of the delay $\tau_l(t)$ in $N_I N_0 T_s$ s is $\Delta\tau = \nu N_I N_0 T_s / c$ (where $c = 3 \times 10^5$ km/s). The condition $\Delta\tau << T_s$ leads to $N_I << (c/\nu)/N_0$. Therefore, for $\nu = 30$ km/h, we have $N_I << 3.6 \cdot 10^7/(N_0)$, meaning that the delays, $\tau_l(t)$, can be treated as a constant (τ_l) for a large number of OFDM symbols, N_I.

Under perfect synchronization, for the ith OFDM symbol time, the output of the receiver is an $N \times 1$ vector, $\mathbf{c}(i) = [c_0(i), c_1(i), \ldots, c_{N-1}(i)]^T$, and

$$c_n(i) = \frac{1}{T_s} G_n^T G_n^R H_n(i) a_n(i) + z_n(i), \tag{8.3}$$

where G_n^T and G_n^R are the values of the Fourier transform of $g_T(t)$ and $g_R(t)$ evaluated at frequency $w = \frac{2\pi n}{NT_s}$, $E\{|a_n(i)|^2\} = \sigma_a^2$, and $z_n(i)$ is i.i.d. white complex Gaussian noise with zero mean and variance σ_z^2. The channel frequency response $H_n(i)$ is given by

$$H_n(i) = \sum_{l=0}^{L-1} h_l(iN_0T_s) e^{-\frac{j2\pi n \tau_l}{NT_s}}. \tag{8.4}$$

In this paper, we assume $G_R(jw)$ is flat for the entire OFDM signal bandwidth $1/T_s$, and G_n^R and G_n^T are equal to $\sqrt{T_s}$.

8.3.2 Nested Array

In this subsection, we shall review the concept of difference co-array and the nested array proposed by [89].

Definition 8.1: *(Difference Co-Array):* Consider an array of of N sensors, with the positions given by the integer set $\mathcal{I}$ defined as

$$\mathcal{I} = \{n_i \in \mathcal{N}, i = 0, 1, \ldots, N-1\}.$$

Define the set

$$\mathcal{S}_{diff} = \{n_i - n_j,\ n_i \in \mathcal{I},\ 0 \leq i, j \leq N-1\}.$$

In our definition of the set $\mathcal{S}_{diff}$, we allow repetition of its elements. We define the set $\mathcal{S}_{du}$, consisting of the distinct elements of the set $\mathcal{S}_{diff}$. Then, the *difference co-array* of the given array is defined as the array which has sensors located at positions given by the set $\mathcal{S}_{du}$. □

The number of elements in the difference co-array, given by the set $\mathcal{S}_{du}$, directly decides the distinct values of the cross correlation terms in the covariance matrix of the signal received by an antenna array with sensor positions defined by $\mathcal{I}$. Using these distinct cross correlation terms in different

ways, one can substantially increase the number of sources that can be detected by the array. Each such technique uses a part or whole of the resulting difference co-array, instead of the original array, to perform direction of arrival (DOA) estimation.

The two-level nested array [89] is a concatenation of two uniform linear arrays (ULAs): inner and outer. The sensor positions of the inner ULA are given by the set

$$\mathcal{S}_{inner} = \{m_0 d_1 | m_0 = 1, 2, \ldots, M_0\}, \tag{8.5}$$

and those of the outer ULA are defined by the set

$$\mathcal{S}_{outer} = \{m_1(M_0 + 1)d_1 | m_1 = 1, 2, \ldots, M_1\}. \tag{8.6}$$

The difference co-array of this nested array is a filled ULA with $2M_1(M_0 + 1) - 1$ elements whose positions are given by the set

$$\mathcal{S}_{ca} = \{nd_1, n = -M, \ldots, M\}, \tag{8.7}$$

where $M = M_1(M_0 + 1) - 1$. The two-level nested array can attain $2M_1(M_0 + 1) - 1$ degrees of freedom in the co-array using only $M_0 + M_1$ elements.

8.4 Subspace Channel Estimation with Uniform Pilots

In [84] and [87], the authors proposed pilot-aided channel estimation methods based on subspace methods. The algorithms have two stages: the channel model parameter acquisition stage and tracking stage. The acquisition stage includes the detection of the number of multipaths, L, via the MDL criterion, and the estimation of multipath delays, τ_l, using subspace methods like ESPRIT or MUSIC. In the tracking stage, an MMSE estimator is applied to estimate the frequency response of the channel. It is assumed that M_s pilots are evenly inserted into N subcarriers with spacing $d_s \leq NT_s/\tau_{max}$ to avoid aliasing (according to sampling theorem). The pilot tone set

$$\mathcal{S} = \big\{s(m) \big| s(m) = md_s,\ m = 0, \ldots, M_s - 1\big\} \tag{8.8}$$

denotes all the pilot positions. The least square (LS) estimates of $H_{s(m)}(i)$ on the pilot tones are obtained by inverting the $c_{s(m)}(i)$ in (8.3) by the known pilot symbols, $a_{s(m)}(i)$, as

$$\mathbf{b}^{\mathcal{S}}(i) = \mathbf{W}_{\mathcal{S}}\mathbf{h}(i) + \mathbf{z}'(i), \tag{8.9}$$

where $\mathbf{b}^{\mathcal{S}}(i) \in \mathcal{C}^{M_s \times 1}$, $b_m^{\mathcal{S}}(i)$ is the least square estimate of $H_{s(m)}(i)$, $\mathbf{W}_{\mathcal{S}} \in \mathcal{C}^{M_s \times L}$, $[\mathbf{W}_{\mathcal{S}}]_{m,l} = \exp\{\frac{-j2\pi m d_s \tau_l}{NT_s}\}$, $\mathbf{h}(i) = [h_0(iN_0T_s) \dots h_{L-1}(iN_0T_s)]^T$, $\mathbf{z}'(i) \in \mathcal{C}^{M_s \times 1}$, $z'_m(i) = z_{s(m)}(i)/a_{s(m)}(i)$, $E\{\mathbf{z}'(i)\mathbf{z}'^{\dagger}(i)\} = \sigma_{z'}^2\mathbf{I}_{M_s}$ and $\sigma_{z'}^2 = E\{1/|a_{s(m)}(i)|^2\}\sigma_z^2$. In the acquisition stage, the covariance matrix for the LS estimated signals is estimated via

$$\mathbf{R} = \frac{1}{N_I}\sum_{i=0}^{N_I-1} \mathbf{b}^{\mathcal{S}}(i)\mathbf{b}^{\mathcal{S}}(i)^{\dagger}. \tag{8.10}$$

When N_I goes large, the matrix $\mathbf{R}$ converges to

$$\lim_{N_I \to \infty} \mathbf{R} = \mathbf{W}_{\mathcal{S}} E\{\mathbf{h}(i)\mathbf{h}^{\dagger}(i)\}\mathbf{W}_{\mathcal{S}}^{\dagger} + \sigma_{z'}^2\mathbf{I}_{M_s}. \tag{8.11}$$

Standard MDL method in [84] and MUSIC can be used to estimate L and τ_l from $\mathbf{R}$, and reconstruct $\mathbf{W}_{\mathcal{S}}$ in (8.9). After that, the algorithms switch to the tracking stage, which estimates frequency response $H_n(i)$ by the MMSE estimator given by

$$\tilde{\mathbf{h}}(i) = \mathbf{W}_{\mathbf{H}}\Big(\frac{\beta}{SNR}\mathbf{R}_{hh}^{-1} + \mathbf{W}_{\mathcal{S}}^{\dagger}\mathbf{W}_{\mathcal{S}}\Big)^{-1}\mathbf{W}_{\mathcal{S}}^{\dagger}\mathbf{b}^{\mathcal{S}}(i), \tag{8.12}$$

where $\tilde{\mathbf{h}}(i) = [\hat{H}_0(i), \dots, \hat{H}_{N-1}(i)]^T$, $\mathbf{R}_{hh} = diag(\sigma_{h_0}^2, \dots, \sigma_{h_{L-1}}^2)$, $\beta = E\{|a_n(i)|^2|\}E\{|1/a_n(i)|^2\}$, $SNR = E\{|a_n(i)|^2|\}/\sigma_z^2$, and $\mathbf{W}_{\mathbf{H}}$ is an $N \times L$ Vandermode matrix with $[\mathbf{W}_{\mathbf{H}}]_{n,l} = \exp\{\frac{-j2\pi\tau_l n}{NT_s}\}$.

8.5 Subspace Channel Estimation with Co-pilots

The subspace channel estimator in Sec. 8.4 has two fundamental limitations. It can only identify up to $M_s - 1$ multipath delays given that the number of pilots is M_s. In [88], the author suggests that a large aperture helps to improve the identifiability of MUSIC. So, the identifiability of the subspace method in Sec. 8.4 could be limited by the small pilot aperture, $M_s d_s$, when $M_s d_s$ is much smaller than N. To improve identifiability, one may try to increase M_s. However, doing so would harm the spectral efficiency of the OFDM system, which is given by $\eta = (N - M_s)/(N + N_g)$.

In this section, we apply different placements of pilots in even and odd OFDM symbols. M_0 pilots are inserted in even OFDM symbols while M_1 pilots are inserted in odd OFDM symbols. M_0 and M_1 are such that $M_0 M_1 < N$. By computing the cross-correlation of the even and odd pilot signals, it is possible to construct a set of virtual pilots (co-pilots) which have more degrees of freedom and bigger apertures than the uniform pilots in Sec. 8.4. For the even OFDM symbols, the pilot tone indices are chosen from the set

$$\mathcal{S}_0 = \big\{ s_0(m_0) = (m_0 + 1) M_1 d \,\big|\, m_0 = 0, \ldots, M_0 - 1 \big\}, \tag{8.13}$$

where $d = \lfloor \min((N-1)/(M_0 M_1), N T_s/\tau_{max}) \rfloor$. Similarly, the pilot tone indices of the odd OFDM symbols are chosen from

$$\mathcal{S}_1 = \big\{ s_1(m_1) = (m_1 + 1) d \,\big|\, m_1 = 0, \ldots, M_1 - 1 \big\}. \tag{8.14}$$

We call the placement, in particular, the *alternating nested pilot placement*. The dense pilots in odd OFDM symbols correspond to the inner ULA of a nested array while the sparse pilots in even OFDM symbols correspond to the outer ULA. Instead of having the *nested structure* [89] pilots in a single OFDM symbol, the dense and sparse pilots are inserted in the odd and even OFDM symbols, respectively, which leads to better spectral efficiency.

From $\mathcal{S}_0$ and $\mathcal{S}_1$, we can define a difference set $\mathcal{S}_{dif}$ as

$$\begin{aligned} \mathcal{S}_{dif} &= \Big\{ x_{m_0} - y_{m_1} \,\Big|\, x_{m_0} \in \mathcal{S}_0,\ y_{m_1} \in \mathcal{S}_1 \Big\} \\ &= \Big\{ md \,\Big|\, 0 \le m \le M_0 M_1 - 1 \Big\}. \end{aligned} \tag{8.15}$$

The cardinality (with no repetition) of the difference set, which is $|S_{dif}|$, corresponds to the number of distinct cross correlation values obtained from the pilot signals. Using these distinct cross correlation values judiciously in different ways, one can substantially increase the number of multipath delays that can be estimated and the number is upper bounded by $|S_{dif}|$ [89]. Here, we have $|S_{dif}| = M_0 M_1$, which is the maximum degree of freedom a difference set constructed from a size M_0 and a size M_1 set can achieve. For a fixed total number of pilots in an even OFDM symbol and an odd OFDM symbol, $2M_s = M_0 + M_1$ (assumed to be even here), $|S_{dif}|$ is maximized by dividing $2M_s$ pilots evenly to the odd and even symbols, i.e. $M_0 = M_1 = M_s$. Therefore, we choose both M_0 and M_1 to be M_s.

As in (8.9), the least square estimates of $H_{s_0(m_0)}(2i)$ for an even OFDM symbol is given by

$$\mathbf{b}^{(0)}(2i) = \mathbf{W}_{\mathcal{S}_0}\mathbf{h}^{(0)}(2i) + \mathbf{z}^{(0)}(2i), \tag{8.16}$$

where $\mathbf{b}^{(0)}(2i) \in \mathcal{C}^{M_0\times 1}$, $b_m^{(0)}(2i)$ is the least square estimate of $H_{s_0(m)}(2i)$, $\mathbf{W}_{\mathcal{S}_0} \in \mathcal{C}^{M_0\times L}$, $[\mathbf{W}_{\mathcal{S}_0}]_{m_0,l} = W_l^{(m_0+1)M_1}$, $W_l = \exp\{\frac{-j2\pi d\tau_l}{NT_s}\}$, $\mathbf{z}^{(0)}(2i) \in \mathcal{C}^{M_0\times 1}$, $\mathbf{h}^{(0)}(2i) = \mathbf{h}(2i)$ and $E\{\mathbf{z}^{(0)}(2i)\mathbf{z}^{(0)\dagger}(2i)\} = \sigma_{z'}^2\mathbf{I}_{M_0}$. Similarly, the least square estimates of $H_{s_1(m_1)}(2i+1)$ for an odd OFDM symbol is

$$\mathbf{b}^{(1)}(2i+1) = \mathbf{W}_{\mathcal{S}_1}\mathbf{h}^{(1)}(2i+1) + \mathbf{z}^{(1)}(2i+1), \tag{8.17}$$

where $\mathbf{b}^{(1)}(2i+1) \in \mathcal{C}^{M_1\times 1}$, $b_m^{(1)}(2i+1)$ is the least square estimate of $H_{s_1(m)}(2i+1)$, $\mathbf{W}_{\mathcal{S}_1} \in \mathcal{C}^{M_1\times L}$, $[\mathbf{W}_{\mathcal{S}_1}]_{m_1,l} = W_l^{M_1-m_1}$, $\mathbf{z}^{(1)}(2i+1) \in \mathcal{C}^{M_1\times 1}$, and $E\{\mathbf{z}^{(1)}(2i+1)\mathbf{z}^{(1)\dagger}(2i+1)\} = \sigma_{z'}^2\mathbf{I}_{M_1}$. Taking the LS estimates, $\mathbf{b}^{(0)}$ and $\mathbf{b}^{(1)}$, from N_I consecutive OFDM symbols, one can compute the correlation matrix

$$\mathbf{R}_{N_I} = \frac{2}{N_I}\sum_{i=0}^{N_I/2-1}\mathbf{b}^{(1)}(2i+1)\mathbf{b}^{(0)\dagger}(2i), \tag{8.18}$$

where N_I is the OFDM symbol time in which the multipath delay time is stationary. When $N_I \to \infty$, we have

$$\begin{aligned}\mathbf{R}_{dif} &= \lim_{N_I\to\infty}\mathbf{R}_{N_I} = E\{\mathbf{b}^{(1)}\mathbf{b}^{(0)\dagger}\} = \mathbf{W}_{\mathcal{S}_1}\mathbf{R}_{\mathbf{h}^{(1)}\mathbf{h}^{(0)}}\mathbf{W}_{\mathcal{S}_0}^{\dagger} \\ &+ \mathbf{W}_{\mathcal{S}_1}E\{\mathbf{h}^{(1)}\mathbf{z}^{(0)\dagger}\} + E\{\mathbf{z}^{(1)}\mathbf{h}^{(0)}\}\mathbf{W}_{\mathcal{S}_0}^{\dagger} + E\{\mathbf{z}^{(1)}\mathbf{z}^{(0)\dagger}\},\end{aligned} \tag{8.19}$$

where $\mathbf{R}_{\mathbf{h}^{(1)}\mathbf{h}^{(0)}} = J_0(2\pi f_d N_0 T_s)diag(\sigma_{h_0}^2,\ldots,\sigma_{h_{L-1}}^2)$. The last three terms are zeros since the channel gains and the noise have zero means and are uncorrelated, and the noises $\mathbf{z}^{(0)}$ and $\mathbf{z}^{(1)}$ are uncorrelated. Performing a vectorization on $\mathbf{R}_{dif}$, we get

$$\mathbf{q} = vec(\mathbf{R}_{dif}) = \underbrace{\left(\mathbf{W}_{\mathcal{S}_0}^{*} \odot \mathbf{W}_{\mathcal{S}_1}\right)}_{\mathbf{A}}\mathbf{p}, \tag{8.20}$$

where $\mathbf{q} \in \mathcal{C}^{M_0M_1\times 1}$, $[\mathbf{A}]_{n,l} = W_l^{-n}$, and $\mathbf{p} = J_0(2\pi f_d N_0 T_s)\cdot[\sigma_{h_0}^2,\ldots,\sigma_{h_{L-1}}^2]^T$. Comparing the above equation to (8.9), one can identify that $\mathbf{A}$ has the same Vandermonde form as $\mathbf{W}_{\mathcal{S}}$ with M_0M_1 pilots, $\mathbf{p}$ corresponds to $\mathbf{h}(i)$, and the noise term is zero. So, we have constructed a set of *virtual pilots* with M_0M_1 distinct tones in the co-array domain. We call the *virtual pilots*, in particular,

the *co-pilots*, which suggests that they are constructed from the concept of co-array.

Also, we can compute another correlation matrix

$$\mathbf{R}'_{N_I} = \frac{2}{N_I} \sum_{i=0}^{N_I/2-1} \mathbf{b}^{(1)*}(2i+1)\mathbf{b}^{(0)T}(2i), \tag{8.21}$$

and it converges to

$$\mathbf{R}'_{dif} = \lim_{N_I \to \infty} \mathbf{R}'_{N_I} = \mathbf{W}^*_{\mathcal{S}_1} \mathbf{R}_{\mathbf{h}^{(1)}\mathbf{h}^{(0)}} \mathbf{W}^T_{\mathcal{S}_0}. \tag{8.22}$$

The vectorization $\mathbf{R}'_{dif}$ yields

$$\mathbf{q}' = vec(\mathbf{R}'_{dif}) = \underbrace{\left(\mathbf{W}_{\mathcal{S}_0} \odot \mathbf{W}^*_{\mathcal{S}_1}\right)}_{\mathbf{A}'} \mathbf{p}, \tag{8.23}$$

where $\mathbf{q}' \in \mathcal{C}^{M_0M_1 \times 1}$ and $\mathbf{p}$ is given by (8.20). Taking the 2nd to the last element of $\mathbf{q}'$ and stacking it with $\mathbf{q}$, we have

$$\tilde{\mathbf{q}} = \begin{bmatrix} \mathbf{J}\mathbf{q}'_{2:end} \\ \mathbf{q} \end{bmatrix} = \tilde{\mathbf{A}}\mathbf{p}, \tag{8.24}$$

where $\mathbf{J} \in \mathcal{R}^{(M_0M_1-1)\times(M_0M_1-1)}$ with ones on its anti-diagonal and zeros elsewhere, and the lth column of $\tilde{\mathbf{A}}$ is given by

$$[\tilde{\mathbf{A}}]_{:,l} = \begin{bmatrix} W_l^{M_0M_1-1} \dots W_l^0 \dots W_l^{-M_0M_1+1} \end{bmatrix}^T.$$

Extracting the $(M_0M_1 - i)$th to $(2M_0M_1 - 1 - i)$th row of $\tilde{\mathbf{q}}$, we have the vector

$$\tilde{\mathbf{q}}^{(i)} = \mathbf{A}\mathbf{D}_p\mathbf{v}_i,$$

where $\mathbf{A}$ is given by (8.20), $\mathbf{D}_p = diag(\mathbf{p})$ and $\mathbf{v}_i = [W_0^i, W_1^i, \dots, W_{L-1}^i]^T$. Using $\tilde{\mathbf{q}}^{(i)}$, we can construct the matrix

$$\mathbf{G} = \begin{bmatrix} \tilde{\mathbf{q}}^{(0)} & \tilde{\mathbf{q}}^{(1)} & \dots & \tilde{\mathbf{q}}^{(M_0M_1-1)} \end{bmatrix} = \mathbf{A}\mathbf{D}_p\mathbf{A}^\dagger. \tag{8.25}$$

The matrix $\mathbf{G}$ has the same form as the covariance matrix of the LS estimated signals from a set of M_0M_1 uniform pilots as in (8.11), except that the noise term is gone. We can take $\mathbf{R}_{N_I}$ and $\mathbf{R}'_{N_I}$

as the approximations for $\mathbf{R}_{dif}$ and $\mathbf{R}'_{dif}$, respectively, and construct the estimated $\mathbf{G}$ accordingly. Then, the MDL method and the MUSIC algorithm are applied on $\mathbf{G}$ to estimate to multipath delays. Note that applying the subspace methods on the covariance matrix $\mathbf{G}$ can identify up to $M_0M_1 - 1$ multipaths, theoretically.

After the estimation of the number of multipaths and time delays, τ_l, the channel gains, $\mathbf{h}(i)$, are estimated via an MMSE estimator. Stacking $\mathbf{b}(2i)$ in (8.16) and $\mathbf{b}(2i+1)$ in (8.17) (we omit the superscripts (0) and (1) here), we have

$$\mathbf{b}'(2i) = [\mathbf{b}^T(2i)\ \mathbf{b}^T(2i+1)]^T = \mathbf{W}\mathbf{h}'(2i) + \overline{\mathbf{z}}(2i), \tag{8.26}$$

where $\mathbf{W} = diag(\mathbf{W}_{\mathcal{S}_0}, \mathbf{W}_{\mathcal{S}_1})$, $\mathbf{h}'(2i) = [\mathbf{h}^T(2i)\ \mathbf{h}^T(2i+1)]^T$, and $\overline{\mathbf{z}}(2i) = [\mathbf{z}^T(2i)\ \mathbf{z}^T(2i+1)]^T$. Then, the MMSE estimate of $\mathbf{h}'(2i)$ is given by

$$\hat{\mathbf{h}}'(2i) = \left(\sigma^2_{z'}\mathbf{R}^{-1}_{h'h'} + \mathbf{W}^\dagger\mathbf{W}\right)^{-1}\mathbf{W}^\dagger\mathbf{b}'(2i), \tag{8.27}$$

where $\hat{\mathbf{h}}'(2i) = [\hat{\mathbf{h}}^T(2i)\ \hat{\mathbf{h}}^T(2i+1)]^T$, and $\hat{\mathbf{h}}(i)$ is the MMSE estimate of $\mathbf{h}(i)$. The MMSE estimator of the frequency response $H_n(i)$ is given by

$$\tilde{\mathbf{h}}(i) = \mathbf{W}_{\mathbf{H}}\hat{\mathbf{h}}(i), \tag{8.28}$$

where $\tilde{\mathbf{h}}(i)$ and $\mathbf{W}_{\mathbf{H}}$ are defined in (8.12).

8.6 Numerical Results

In this section, we present numerical examples of channel estimation for the OFDM system in Sec. 8.3, using the subspace channel estimator with uniform pilots in Sec. 8.4, the proposed co-pilot method in Sec. 8.5, uniformly spaced comb-type pilots with linear interpolation [83], and a Fourier transform (FT) based MMSE estimator with genie aided temporal constraint. For the genie aided FT based MMSE estimator, the number of multipaths and the normalized delays which are rounded to the nearest integers, i.e., $\tau'_l = \text{round}(\tau_l/T_s)$, are provided by a genie as the prior information. The pilots are inserted uniformly into both even and odd OFDM symbols. A similar MMSE estimator as in (8.27) is used to estimate the temporal channel response $\hat{\mathbf{h}}'(2i)$. Here, the matrix $\mathbf{W} = \mathbf{I}_2 \otimes \mathbf{W}_{\mathbf{g}}$, where $\mathbf{W}_g \in \mathcal{C}^{M_g \times L}$, $[\mathbf{W}_g]_{m,l} = \exp(-2\pi m\tau'_l/M_g)$ and M_g is the number of pilots per OFDM

symbol. The frequency response is estimated with an FT interpolator as in (8.28). The genie aided FT based MMSE estimator provides the performance benchmark for the class of classical FT based MMSE estimators in which the multipath delays fall on the integer grids of $nT_s, n \in \mathcal{N}$ [81], [85]. The number of subcarriers, N, is 64. The length of cyclic prefix is $N_g = 16$. So, the size of an OFDM symbol is $N_0 = 80$. The maximum delay for the multipath channel in (8.1), τ_{max}, is $16T_s$. The normalized multipath delays are given by $\tau_l/T_s = 1 + \text{round}(l(\tau_{max}/T_s - 2)/(L-1)) + \mu_l$, where μ_l are assumed to be independent, uniformly distributed between $-\sigma_\tau/2$ and $\sigma_\tau/2$, and stationary for $N_I = 2000$ consecutive OFDM symbols. The term μ_l perturbs τ_l/T_s from an integer grid. For $\sigma_\tau = 0$, τ_l/T_s takes an integer value. The power profile of the lth delay in (8.2) is given by $\sigma_{h_l}^2 = C \cdot \alpha^l$, where $\alpha = 0.8$ and C is chosen such that $\sum_{l=0}^{L-1} \sigma_{h_l}^2 = 1$. The normalized Doppler frequency in (8.2) is $f_{d,l}T = f_{d,l}N_0T_s = 0.1$. The data and pilot symbols are QPSK symbols. For all the channel estimation methods considered here, the number of pilots per OFDM symbol is set to be 5, i.e., $M_s = M_0 = M_1 = M_g = 5$. In the delay acquisition stage of the uniform pilot and the proposed co-pilot method, both algorithms take $N_I = 2000$ OFDM symbols, construct their own covariance matrices, and run the MDL and MUSIC algorithm to estimate τ_l. In the tracking stage, MMSE estimators in the forms of (8.27) and (8.28) are applied every two OFDM symbols to estimate $H_n(i)$. For the uniform pilot method, $\mathbf{W}$ and $\mathbf{b}'(2i)$ in the MMSE estimator are given by $\mathbf{W} = diag(\mathbf{W}_\mathcal{S}, \mathbf{W}_\mathcal{S})$ and $\mathbf{b}'^T(2i) = [(\mathbf{b}^\mathcal{S}(2i))^T\ (\mathbf{b}^\mathcal{S}(2i+1))^T]$, instead.

Example 1: In this example, we evaluate the average normalized MSE, $1/N \sum_{n=0}^{N-1} E\{|\hat{H}_n(i) - H_n(i)|^2/|H_n(i)|^2\}$, of the above mentioned channel estimators for different numbers of multipaths (L). One million OFDM symbols are transmitted and the channel varies according to the parameters described above. The SNR is

$$SNR = 10\log_{10}\Big(\frac{\sigma_a^2(N-M_s)}{N_0\sigma_z^2}\Big) = 30 \text{ dB}.$$

The results are shown in Fig. 8.2 and Fig. 8.3 for $\sigma_\tau = 0.1$ and $\sigma_\tau = 0.5$, respectively. The proposed co-pilot method performs better than the uniform pilot method when $L \geq 4$. The result conforms to the justification that the co-pilot method can identify more than 4 multipaths, which is the fundamental limit of the uniform pilot method. For $L = 1, 2$, the co-pilot method also outperforms the uniform pilot method since it has a larger pilot aperture. The FT based MMSE estimator with a genie has lower MSE for the case with $\sigma_\tau = 0.1$, in which the normalized multipath delays (τ_l/T_s) are more likely to fall on integer grids. For $L < 7$ and $\sigma_\tau = 0.5$, the proposed co-pilot method out-

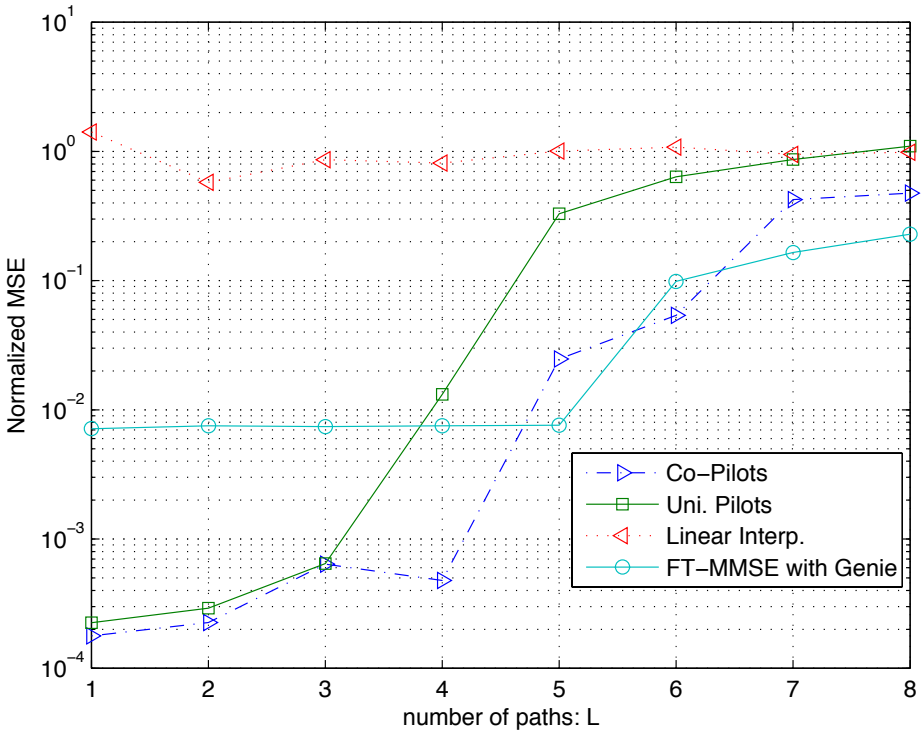


Figure 8.2: Average MSE of the estimated frequency response with $\sigma_\tau = 0.1$.

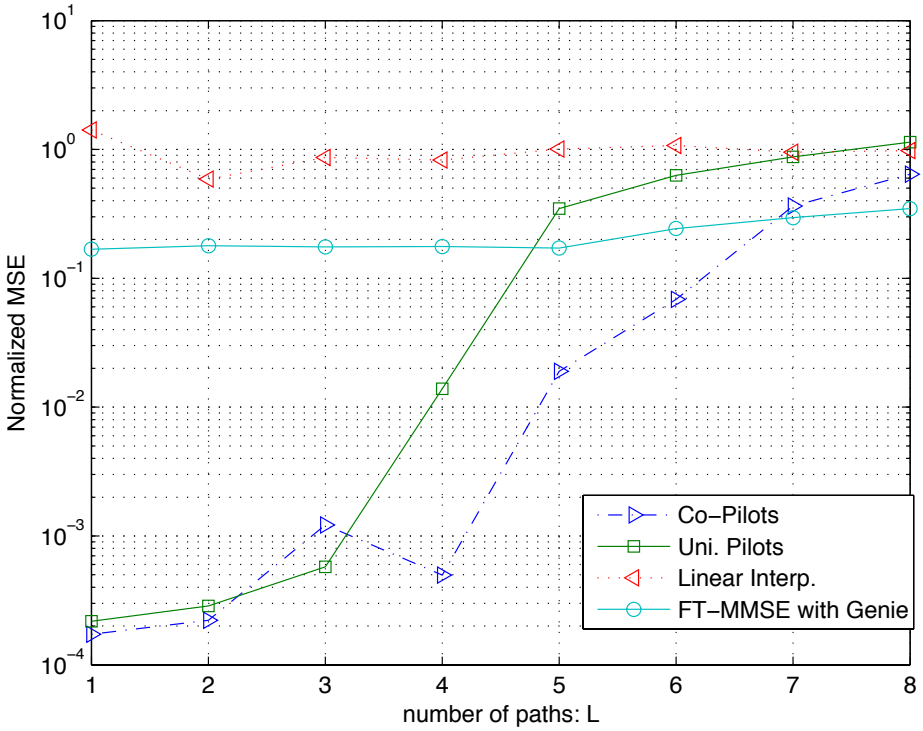


Figure 8.3: Average MSE of the estimated frequency response with $\sigma_\tau = 0.5$.

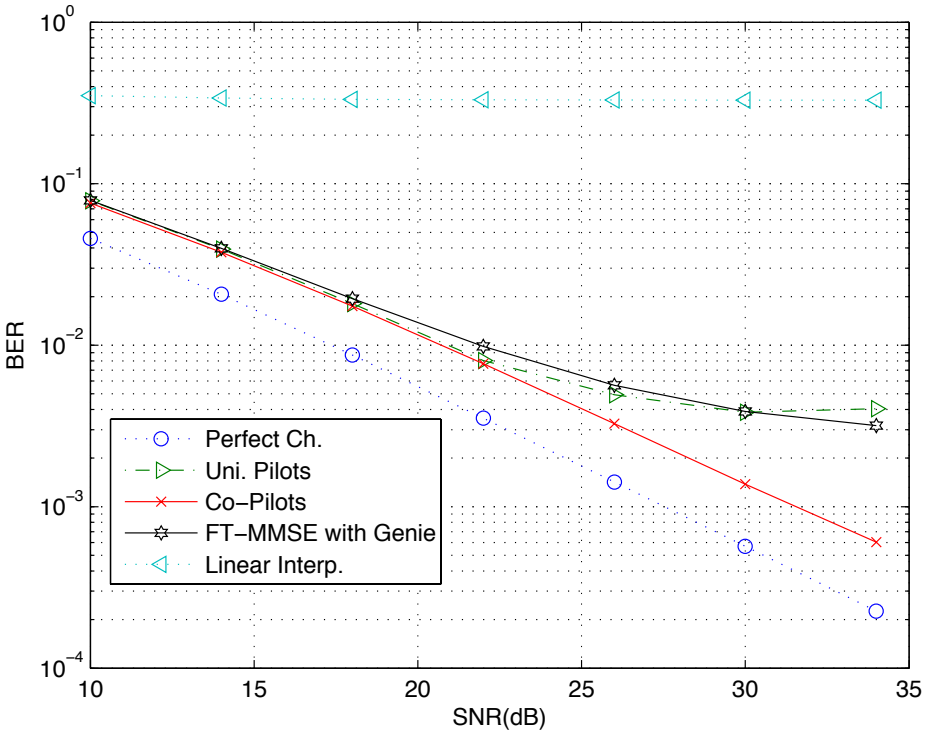


Figure 8.4: Average BER of the OFDM system with $\sigma_\tau = 0.1$

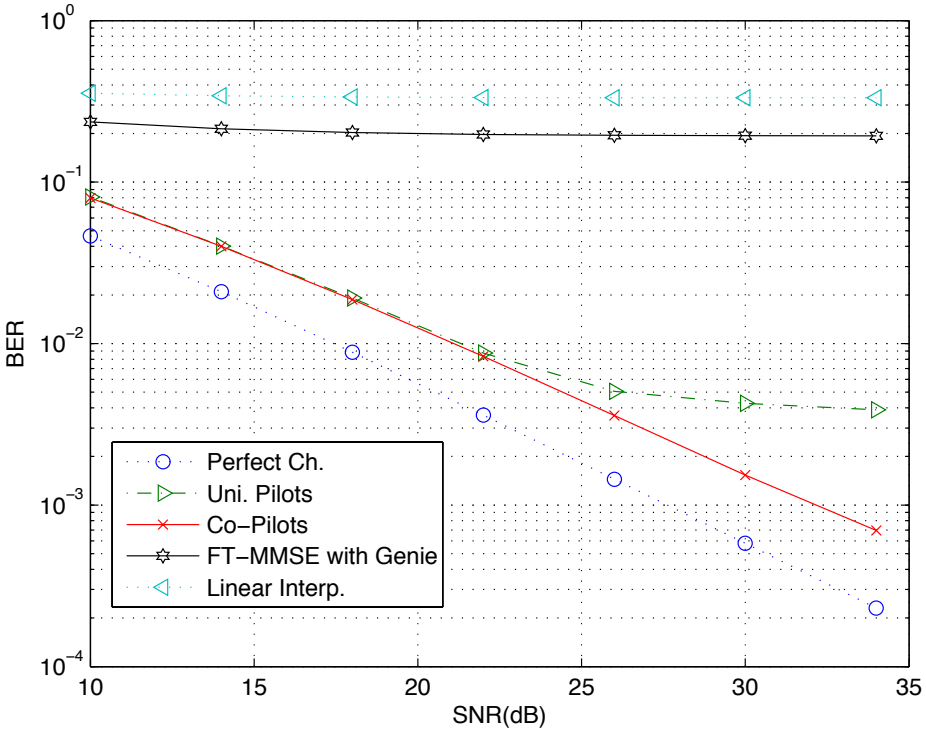


Figure 8.5: Average BER of the OFDM system with $\sigma_\tau = 1$

performs the FT based MMSE estimator. However, the FT based MMSE estimator benefits from the genie and has lower MSEs than all the other three methods when $L > 7$ and $\sigma_\tau = 0.5$. The comb-type pilot method with linear interpolation has bad MSE performance since it can not capture the variation of $H_n(i)$ in the frequency domain.

Example 2: We compare the BER performance of the OFDM systems equipped with the above mentioned channel estimators for different SNRs. The number of multipaths is $L = 4$, $\sigma_\tau = 0.1$ or $\sigma_\tau = 1$ and other system parameters are described as above. The BERs are the average BERs computed over 10^6 OFDM symbols. The BERs of the OFDM system with perfect channel knowledge are simulated and serve as the benchmark for all other systems with channel estimators. The results are given in Fig. 8.4 and Fig. 8.5. Among the four OFDM systems with channel estimators, the co-pilot system has the best BER performance, then the uniform pilot system, then the FT based MMSE system, and then the linear interpolation system. For the case with $\sigma_\tau = 0.1$, the FT based MMSE system with the genie and the uniform pilot system have almost the same BERs for all the SNRs. When $\sigma_\tau = 1$, the BER performance of the FT based system degrades tremendously. This suggests that the FT-based MMSE systems will perform poorly if the normalized multipath delays (τ_l/T_s) are not integers.

8.7 Conclusions

In this chapter, we have proposed a new subspace channel estimation algorithm which can identify up to $O(M^2)$ multipath delays by using M physical pilot tones. For doubly selective channels, it was shown that the new algorithm is capable of improving the spectral efficiency of OFDM systems by using a lesser number of pilots for a given number of multipath delays.

Chapter 9

Conclusions

In this thesis, we have studied transceiver design problems with various kinds of channels. By exploiting the majorization theory, the proposed generalized geometric mean decomposition (GGMD), the proposed space-time generalized triangular decomposition (ST-GTD) and joint triangularization (JT), we were able to analytically solve the transceiver optimization problems with LTI flat MIMO channels, LTV flat MIMO channels, and broadcast flat MIMO channels. Novel transceiver structures and detection algorithms were developed. Performance analysis was done to confirm the improvements over the conventional design, and relevant theoretical issues were studied. In addition to MIMO transceiver designs, new transceiver design and channel estimation algorithms were proposed for the multi-carrier systems.

The first part of the thesis focuses on decision feedback equalizer (DFE) transceiver design with LTI flat MIMO channels. In Chapter 2, based on CSIT and CSIR, a novel butterfly structured MIMO DFE transceiver was constructed by using the newly proposed GGMD. The proposed transceiver works for general LTI flat MIMO channels, and is found to be most useful for cyclic prefix (CP) systems since its *design* and *implementation complexity* are much less than that of a conventional GMD MMSE transceiver. In Chapter 3, performance analyses for the proposed GGMD DFE transceiver were derived. We showed that the proposed design minimizes the average MSE, minimizes the average SER, and maximizes the Gaussian mutual information under the total power constraint. For the application in CP systems, the SER analysis shows that the GGMD DFE transceiver outperforms the SC-CP MMSE system for all SINRs, and the OFDM MMSE system in the moderate high SINR region.

In the second part of the thesis, the transceiver designs for slowly LTV flat MIMO channels were studied. In Chapter 4, we developed the ST-GTD based on the generalized triangular decomposi-

tion (GTD). ST-GTD can be used to decompose the slowly LTV flat MIMO channels and construct an ST-GMD DFE transceiver with a zero-forcing (ZF) constraint. Under perfect channel prediction, the system minimizes both the average MSE at the detector in each ST-block, and the average per ST-block BER in the moderate high SNR region. For practical applications where channel prediction is not available, a causal ST-GTD ZF transceiver (CST-GTD), which has the same asymptotic BER performance as the ST-GMD ZF DFE transceiver, was also proposed. In the analysis, we showed that the proposed transceivers have smaller MSEs and BERs in the moderate high SNR region than the conventional GMD ZF DFE transceiver. In Chapter 5, three transceivers were proposed under the MMSE criterion. Firstly, a channel independent temporal precoder is super-imposed on the conventional block-wise GMD MMSE DFE transceiver to take advantage of the temporal diversity. Secondly, ST-GTD is applied to design the ST-GMD MMSE transceiver under the MMSE criterion (ST-GMDM) which is optimal in terms of arithmetic MSE, Gaussian mutual information and average BER for the ST-block transmission scheme. Finally, we proposed a causal ST-GTD MMSE DFE transceiver (CS-GTDM) which does not require channel prediction and has the same asymptotic BER performance as ST-GMDM. In the analysis, we verified the performance advantage of the proposed systems over the conventional GMD MMSE DFE transceiver and compared the proposed systems in terms of MSE and BER. Also, we analyzed the dependency of the average BERs of the proposed systems on ST-block size, and showed the advantage of ST-GMDMs over ST-GMD ZF transceivers (proposed in Chapter 4).

Chapter 6 focuses on the DFE transceiver design for MIMO broadcast channels. Two joint design problems of the linear precoder, DFE and bit allocation are considered. The first problem is a *min-power QoS* problem with a total bitrate constraint and per stream SER/BER constraints, and the second problem is a *max-rate QoS* problem with a total power constraint and per stream SER/BER constraints. By applying a particular class of joint triangularization which satisfies the *dominance condition*, we propose the minimum power JT broadcast DFE transceiver (MPJT) and the maximum rate JT broadcast DFE transceiver (MRJT), which are optimal solutions for the *min-power* and *max-rate QoS* problems, respectively. Two corresponding suboptimal broadcast DFE transceivers, MPQR and MRQR transceivers which exploit simple QR decompositions, are also proposed for the min-power and max-rate *QoS* problems, respectively. Integer bit allocation algorithms are also developed for the proposed systems. Moreover, the theoretical results show the duality of the proposed MPJT and MRJT transceivers for the min-power and max-rate *QoS* problems and the results

are verified by numerical examples.

In Chapter 7, the discrete Fourier transform modulated filterbank transceiver (DFT-FBT) design problem with LTV scalar channels was studied. For the case with perfect CSI and that with channel statistics, we proposed iterative algorithms which jointly optimize the prototype filters such that the SINR at the receiver is maximized. Simulation results reveal that the optimized DFT-FBTs provide substantial gains over the conventional Affine OFDM which requires exact CSI at both ends. In Chapter 8, a new pilot-aided subspace channel estimation algorithm for OFDM systems with sparse L-multipath LTV channels was proposed. The new algorithm can identify up to $O(M^2)$ multipath delays with M physical pilot tones and the normalized delays are not limited to be on integer grids. The simulation results reveal that the new algorithm can help the OFDM system to keep relatively good BER performance, compared to the conventional subspace channel estimator, even when the number of multipaths is greater than the number of physical pilots.

There are various topics worthy of future research. In Chapters 2 and 3, the proposed GGMD enables us to construct an implementation-efficient GGMD DFE transceiver which can convert MIMO channels into parallel subchannels with equal SINRs. In *QoS* problems, one may have different SINR requirements for the substreams of an implementation efficient DFE transceiver. It is thus essential to develop a new decomposition which can help us to design such a transceiver. Moreover, the results are obtained based on the assumption of perfect CSIT and CSIR. Having a robust design with respect to the channel uncertainty, which arises in practical communication systems, is desirable. In Chapters 4 and 5, channel independent temporal precoders are super-imposed on the DFE transceivers to exploit the temporal diversity without the need of channel prediction. The design of a channel dependent temporal precoder based on the channel statistics is not known and worthy of pursuing. In Chapter 6, elegant analytical solutions for the *min-power* and *max-rate QoS* problems are derived by exploiting a particular class of joint triangularization. There is a theory that guarantees the existence of such a JT decomposition for two-user cases. For more than two users, relevant theoretical results are still unavailable. It will be useful if one could develop similar results for the JT decomposition of more than two matrices.

Bibliography

[1] K.-H. and D. P. Petersen, "Optimum linear coding for vector channels," *IEEE Trans. on Commun.*, vol. 24, no. 12, pp. 1283-1290, Dec. 1976.

[2] J. Saltz, "Digital transmission over cross-coupled linear channels," *AT&T Tech. J.*, vol. 64, no. 6, pp. 1147-1159, July-Aug. 1985.

[3] H. S. Malvar and D. H. Staelin, "Optimal pre- and postfilters for multichannel signal processing," *IEEE Trans. Acoust., Speech. & Sig. Proc.*, vol. 36, no. 2, pp. 287-289, Feb. 1988.

[4] X-G. Xia, "New precoding for intersymbolinterference cancellation using nonmaximally decimated multirate filter banks with ideal FIR equalizers," *IEEE Trans. Sig. Proc.*, vol. 45, no. 10, pp. 2431-2441, Oct. 1997.

[5] T. Li and Z. Ding, "Joint transmitter-receiver optimization for partial response channels based on nonmaximally decimated filter bank precoding technique," *IEEE Trans. Sig. Proc.*, vol. 47, no. 9, pp. 2407-2414, Sep. 1999.

[6] Y. Ding, T. Davidson, Z. Luo, K. Wong, "Minimum BER block precoders for zero-forcing equalization," *IEEE Trans. Sig. Proc.*, vol. 51, no. 9, pp. 2410-2423, Sep. 2003.

[7] Y. P. Lin and S. M. Phoong, "BER minimized OFDM systems with channel independent precoders," *IEEE Trans. Sig. Proc.*, vol. 51, no. 9, pp.2369-2380, Sep. 2003.

[8] L. J. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiple access," *IEEE Trans. Commun.*, vol. 30, pp. 665-675, July 1985.

[9] J. S. Chow, J. C. Tu, and J. M. Cioffi, "A discrete multitone transceiver system for HDSL applications," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 895-908, Aug. 1991.

[10] H. Sari, G. Karam, and I. Jeanclaude, "Transmission techniques for digital terrestrial broadcasting," *IEEE Commun. Mag.*, vol. 33, no. 2, pp. 100-109, Feb. 1995.

[11] D. P. Palomar, J. M. Cioffi and M. A. Lagunas, "Joint Tx-Rx beamformering design for multicarrier MIMO channels: a unified framework for convex optimization," *IEEE Trans. Sig. Proc.*, vol. 51, no. 9, Sep. 2003.

[12] J. Yang and S. Roy, "Joint transmitter-receiver optimization for multi-input multi-output systems with decision feedback," *IEEE Trans. Info. Theory*, vol. 40, no. 5, pp.1334-1347, Sep. 1994.

[13] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Tech. Rep., AT&T Bell Labs.*, 1995.

[14] J. B. Anderson, "Array gain and capacity for known random channels with multiple elements arrays at both ends," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 11, pp.2172-2178, Nov. 2000.

[15] G. J. Foschinini, "Layered space-time architecture for wireless communications in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp.41-59, 1996.

[16] J. Zhang, A. Kavcic, and K. M. Wong, "Equal-diagonal QR decomposition and its application to precoder design for successive-cancellation detection," *IEEE Trans. Inform. Theory*, vol. 51, no. 1, pp. 154-172, Jan. 2005.

[17] Y. Jiang, J. Li, and W. W. Hager, "Joint transceiver design for MIMO communications using geometric mean decomposition, " *IEEE Trans. Sig. Proc.*, pp.3791-3803, Oct. 2005.

[18] Y. Jiang, J. Li, and W. W. Hager, "Uniform channel decomposition for MIMO communications, " *IEEE Trans. Sig. Proc.*, pp.4283-4294, Nov. 2005.

[19] F. Xu, T. N. Davidson, J. K. Zhang and K. M. Wong, "Design of block transceivers with decision feedback detection," *IEEE Trans. Sig. Proc.*, vol. 54, no. 3, pp. 965-978, Mar. 2006.

[20] Y. Jiang, W. W. Hager, and J. Li, "The geometric mean decomposition," *Linear Algebra and Its Applications*, vol. 396, pp. 373-384, Feb. 2005.

[21] Y. Jiang, W. W. Hager, and J. Li, "Generalized triangular decomposition," *Mathematics of computation.,* vol. 77, no. 262, pp. 1037-1056, Apr. 2008.

[22] Y. Jiang, D. P. Palomar, and M. K. Varanasi, "Precoder optimization for nonlinear MIMO transceiver based on arbitrary cost function," *Proc. Conf. on Information Sciences and Systems,* pp. 119-124, Mar. 2007.

[23] M. B. Shenouda and T. N. Davidson, "A framework for designing MIMO systems with decision feedback equalization or Tomlinson-Harashima precoding," *IEEE J. Sel. Areas Commun.,* vol. 26, no. 2, pp. 401-411, Feb. 2008.

[24] S. Lin, W. W. L. Ho and Y.-C. Liang, "Block diagonal geometric mean decomposition (BD-GMD) for MIMO broadcast channels," *IEEE Trans. Sig. Proc.,* vol. 7, no. 7, pp. 2778-2789, Jul. 2008.

[25] S. Bergman, D. P. Palomar, and B. Ottersten, "Joint bit allocation and precoding for MIMO Systems with decision feedback detection," *IEEE Trans. Sig. Proc.,* vol. 57, no. 11, pp. 4509-4521, Nov. 2009.

[26] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney, "MMSE decision-feedback equalizers and coding-part I: equalization results, " *IEEE Trans. Commun.,* vol. 43, no. 10, pp.2582-2594, Oct. 1995.

[27] C. H. Liu and P. P. Vaidyanathan, "Generalized geometric mean decomposition and DFE transceiver design-part I: design and complexity," *IEEE Trans. Sig. Processing,* vol. 60, no. 6, pp. 3124-3133, June 2012.

[28] C. H. Liu and P. P. Vaidyanathan, "Generalized geometric mean decomposition and DFE MMSE transceiver design-part II: performance analysis," *IEEE Trans. Sig. Processing,* vol. 60, no. 6, pp. 3124-3133, June 2012.

[29] S. Zhou, G. B. Giannakis, "How accurate channel prediction needs to be for transmit-beamforming with adaptive modulation over Rayleigh MIMO channels?," *IEEE Trans. Wireless Commun.,* vol. 3, no. 3, pp.1285-1294, July 2004.

[30] F. Dietrich, R. Hunger, M. Joham and W. Utschick, "Linear precoding over time-varying channels in TDD systems," *Proc. ICASSP'03,* vol. 5, pp.117-120, Hong Kong, Apr. 2003.

[31] S. K. Jayaweera and H. V. Poor, "Capacity of multiple-antenna systems with both receiver and transmiitter channel state information," *IEEE Trans. Info. Theory,* vol. 49, no. 10, pp.2697-2709, Oct. 2003.

[32] G. Lebrun, J. Gao, and M. Faulkner, "MIMO transmission over a time-varying channel using SVD," *IEEE Trans. Wireless Commun.,* vol. 4, no. 2, pp.757-764, Mar. 2005.

[33] D. Samardzija and N. Mandayam, "Impact of pilot design on achievable data rates in multiple antenna multiuser TDD systems," *IEEE J. Sel. Areas Commun.,* vol. 25, no. 7, pp.1370-1379, Sep. 2007.

[34] P. Komulainen, A. Tolli, M. Latva-aho, and M. Juntti, "Downlink assisted uplink zero forcing for TDD multiuser MIMO systems," *Eurasip J. Wireless Commun. and Networking,* vol. 2009, no. 8, Mar. 2009.

[35] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication, " *IEEE Trans. Commun.,* vol.46, no.3, pp.357-366, Mar. 1998.

[36] C. H. Liu and P. P. Vaidyanathan, "ZF-DFE transceiver for time-varying MIMO channels with channel-independent temporal precoder," *Proc. ISCAS'10,* pp.3525-3528, Paris, May 2010.

[37] A. Duel-Hallen, S. Hu, and H. Hallen, "Long-range prediction of fading signals," *IEEE Signal Precessiong Mag.,* vol. 17, pp.62-75, May 2000.

[38] C. H. Liu and P. P. Vaidyanathan, "Zero-forcing DFE transceiver design over slowly time-varying MIMO channels using ST-GTD," *IEEE Trans. Sig. Proc.* vol. 58, no. 11, Nov. 2010.

[39] C. H. Liu and P. P. Vaidyanathan, "MMSE DFE transceiver design over slowly time-varying MIMO channels using ST-GTD," *IEEE Trans. Sig. Proc.,* vol. 59, no. 1, pp. 277-289, Jan. 2011.

[40] W. Yu and J. Cioffi, "Sum capacity of a Gaussian vector broadcast channel," *IEEE Trans. Info. Theory,* vol. 50, no. 9, pp. 1875-1892, Sep. 2004.

[41] N. Jindal and Z. Q. Luo, "Capacity limits of multiple antenna multicast," in *ISIT'06,* Seattle, USA, Jul. 2006, pp. 1841-1845.

[42] H. Weingarten, Y. Steinber, and S. Shamai, "On the capacity region of the multi-antenna broadcast channel with common messages," in *ISIT'06,* Seattle, USA, Jul. 2006, pp.2195-2199.

[43] D. Blackwell, L. Breiman, and A. J. Thomasian, "The capacity of a class of channels," *The Annals of Mathematical Statistics,* vol. 30, pp. 1229-1241, Dec. 1959.

[44] A. Khina, Y. Kochman, and U. Erez, "Decomposing the MIMO broadcast channel," in *58th Annual Allerton Conference*, Illinois, USA, Sep. 2010, pp. 102-106.

[45] A. Khina, Y. Kochman, and U. Erez,"Joint unitary triangularization for MIMO networks" *IEEE Trans. on Sig. Proc.,* vol. 60, no. 1, pp. 326-336, Jan. 2012.

[46] C. F. Van Loan, "Generalizing the singular value decomposition," *SIAM J. Numer. Analy.,* vol. 13, no. 1, pp. 76-83, Mar. 1976.

[47] L. Vandenberghe, S. Boyd, and S.-P. Wu, "Determinant maximization with linear matrix inequality constraints," *SIAM J. Matrix Anal. Applic.*, vol. 19, pp. 499-533, 1998.

[48] C. H. Liu and P. P. Vaidyanathan, "MIMO broadcast DFE transceiver design with bit allocation under QoS constraints," in the Proc. IEEE Statistical Sig. Proc. Workshop, Ann Arbor, USA, Aug. 2012.

[49] J.-S. Seo, S.-J. Cho, and K. Feher, "Impact of non-Gaussian impulsive noise on the performance of high-level QAM," *IEEE Trans. Electromagnetic Compatibility,* vol. 31, no. 2, pp. 177-180, May 1989.

[50] Y.-P. Lin and S.-M. Phoong, "Perfect discrete multitone modulation with optimal transceivers," *IEEE Trans. Signal Processing,* vol. 48, no. 6, pp. 1702-1711, June 2000.

[51] Y.-P. Lin and S.-M. Phoong, "Optimal ISI-Free DMT transceivers for distorted channels with colored noise," *IEEE Trans. Signal Processing,* vol. 49, no. 11, pp. 2702-2712, Nov. 2001.

[52] A. Yasotharan, "Multirate zero-forcing Tx-Rx design for MIMO channel under BER constraints," *IEEE Trans. Signal Processing,* vol. 54, no. 6, pp. 2288-2300, June 2006.

[53] S. Dasgupta and A. Pandharipande, "Optimum multiflow biorthogonal DMT with unequal subchannel assignment," *IEEE Trans. Signal Processing,* vol. 53, no. 9, pp. 3572-3582, Sep. 2005.

[54] D. P. Palomar, M. A. Lagunas, and J. M. Cioffi, "Optimum linear joint transmit-receive processing for MIMO channels with QoS Constraints," *IEEE Trans. Signal Processing*, vol. 52, no. 5, pp. 1179-1197, May 2004.

[55] D. P. Palomar and S. Barbarossa, "Designing MIMO communication systems: constellation choice and linear transceiver design," *IEEE Trans. Signal Processing*, vol. 53, no. 10, pp. 3804-3818, Oct. 2005.

[56] C.-C. Li, Y.-P. Lin, S.-H. Tsai, and P. P. Vaidyanathan "Optimization of transceiver with bit allocation to maximize bit rate for MIMO transmission," *IEEE Trans. Signal Processing*, vol. 57, no. 12, pp. 3556-3560, Dec. 2009.

[57] C. C. Weng, C. Y. Chen and P. P. Vaidyanathan, "MIMO transceivers with decision feedback, bit loading and limited feedback: therory and optimization," *IEEE Trans. on Sig. Proc.*, vol. 58, no. 3, pp. 1334-1346, Mar. 2010.

[58] C. C. Li and Y. P. Lin, "On the duality of MIMO transceiver designs with bit allocation," *IEEE Trans. on Sig. Proc.*, vol. 59, no. 8, pp. 3775-3787, Aug. 2011.

[59] S. Serbetli and A. Yener, "Transceiver optimization for multiuser MIMO systems," *IEEE Trans. Sig. Proc.*, vol. 52, no. 1, pp. 214-226, Jan. 2004.

[60] J. Zhang, S. Zhou and J. Wang, "Joint linear transmitter and receiver design for the downlink of multiuser MIMO systems," *IEEE Comm. Letter*, vol. 9, no. 11, pp. 991-993, Nov. 2005.

[61] J.-H. Chang, L. Tassiulas, and F. Rashid-Farrokhi, "Joint transmitter receiver diversity for efficient space division multiaccess," *IEEE Trans. Wireless Comm.*, vol. 1, no. 1, pp. 16-26, Jan. 2002.

[62] Q. H. Spencer, A. L. Swindlehurst and M. Haardt, "Fast power minimization with QoS contraints in multi-user MIMO downlinks," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 4, pp. 816-819, Apr. 2003.

[63] C. Hellings, M. Joham and W. Utschick, "Gradient-based power minimization in MIMO broadcast channels with linear precoding," *IEEE Trans. Sig. Proc.*, vol. 60, no. 2, pp. 877-890, Feb. 2012.

[64] S. T. Chung and A. J. Goldsmith, "Degrees of freedom in adaptive modulation: a unified view," *IEEE Trans. Commun.*, vol. 49, no. 9, pp. 1561-1571, Sep. 2001.

[65] C.-H. Liu and P. P. Vaidyanathan, "MIMO broadcast DFE transceivers with QoS constraints: min-power and max-rate solutions," *submitted to IEEE Trans. Signal Processing*.

[66] C.-H. Liu and P. P. Vaidyanathan, "Optimized DFT-FB transceivers over LTV channels," *Signals, Systems and Computers, 2009 Conference Record of the Forty-Third Asilomar Conference on*, pp.83-87, Nov. 2009.

[67] M. Martone, "A multicarrier system based on the fractional Fourier transform for time-frequency-selective channels," *Proc. IEEE ICASSP*, May 2001.

[68] S. Barbarossa, R. Torti, "Chirped-OFDM for transmissions over time-varying channels with linear delay/Doppler spreading," *IEEE Trans. Communications*, Vol. 49, June 2001.

[69] T. Erseghe, N. Laurenti and V. Cellini, "A multicarrier architecture based upon the Affine Fourier transform," *IEEE Trans. Communications*, Vol. 53, May 2005.

[70] S. Coleri, M. Ergen, A. Puri and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Trans. Broadcasting*, Vol. 48, Sep. 2002.

[71] U. Reimers, "DVB-T: the COFDM-based system for terrestrial television," Electronics and Communication Engineering Journal," Vol. 9, Feb. 1997.

[72] IEEE Std 802.16e, "Amendment for physical and medium access control layers for combined fixed and mobile operation in licensed bands," Sep. 2005.

[73] T. Pollet, M. V. Bladel and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Trans. Communications*, Vol. 43, April 1995.

[74] K. Liu, Y. Kadous, and A. M. Sayeed, "Orthogonal time-frequency signaling over doubly dispersive channels," *IEEE Trans. Information Theory*, Vol. 50, Nov. 2004.

[75] W. Kozek and A. F. Molishch, "Nonorthogonal pulse shapes for multicarrier communications in doubly dispersive channels," *IEEE J. Select. Areas Communications*, Vol. 16, Oct. 1998.

[76] S. M. Phoong, Y. B. Chang, and C. Y. Chen, "DFT modulated filter bank transceivers for multipath fading channels," *IEEE Trans. Signal Processing,* Vol. 53, No. 1, Jan. 2005.

[77] C. Y. Chen and P. P. Vaidyanathan, "MIMO radar waveform optimization with prior information of the extended clutter and target," *IEEE Trans. Signal Processing*, vol. 58, no. 9, pp. 3533-3544, Sep. 2009.

[78] M. Hsieh and C. Wei, "Channel estimation for OFDM systems based on comb-type pilot arrangement in frequency selective fading channels," *IEEE Trans. Consumer Electron.*, vol. 44, no. 2, pp. 217-225, Feb. 1998.

[79] M. C. Vanderveen, A-J. V. D. Veen, and A. Paulraj, "Estimation of multipath parameters in wireless communication," *IEEE Trans. Comm.*, vol. 46, no. 3, pp.682-690, Mar. 1998.

[80] P. Hoeher, S. Kaiser, and I. Robertson, "Two-dimensional pilot-symbol-aided channel estimation by wiener filtering," in *Proc. ICASSP'97,* April 1997.

[81] V. K. Jones and G. G. Raileigh, "Channel estimation for wireless OFDM systems," in *Proc. IEEE GLOBECOM'98,* pp. 980-985, 1998.

[82] O. Edfors, M. Sandell, J.-J V. D. Beek, and S. K. Wilson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Comm.*, vol. 46, no. 7, pp.931-939, July 1998.

[83] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Trans. Broadcasting*, vol. 48, no. 3, pp.223-229, Sep. 2002.

[84] B. Yang, K. B. Letaief, R. S. Cheng, and Z. Cao, "Channel estimation for OFDM transmission in multipath fading channels based on parametric channel modeling," *IEEE Trans. Comm.*, vol. 49, no. 3, pp. 467-479, Mar. 2001.

[85] M. Morelli and U. Mengali, "A Comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Processing*, vol. 49, no. 12, pp. 3065-3073, Dec. 2001.

[86] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Pilot-based channel estimatiom for OFDM systems by tracking the delay-subspace," *IEEE Trans. Wireless Comm.*, vol. 3, no. 1, pp. 315-325, Jan. 2004.

[87] I. C. Wong and B. L. Evans, "Sinusoidal modeling and adapitive channel prediction in mobile OFDM systems," *IEEE Trans. Sig. Proc.*, vol. 56, no. 4, pp. 1601-1615, Apr. 2008.

[88] P. P. Vaidyanathan and P. Pal, "Direct-MUSIC on sparse arrays," Signal Processing and Communications (SPCOM), 2012 International Conference on, July 2012.

[89] P. Pal and P. P. Vaidyanathan, "Nested arrays: a novel approach to array processing with enhanced degrees of freedom," *IEEE Trans. Sig. Proc.,* vol. 58, no. 8, pp. 4167-4181, Aug. 2010.

[90] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276-280, Mar. 1986.

[91] P. P. Vaidyanathan, S. M. Phoong and Y. P. Lin, *Signal processing and optimization for transceiver systems.* Cambridge, U.K.: Cambridge Univ. Press, 2010.

[92] P. P. Vaidyanathan, *Multirate systems and filter banks.* Englewood Cliffs, NJ: Prentice-Hall, 1993.

[93] J. D. Gibson, Ed., *Multimedia communications-directions and innovations,* Academic, 2001.

[94] J. G. Proakis, *Digital communications,* McGraw-Hill, 2001.

[95] A. W. Marshall, I. Olkin and B. C. Arnold, *Inequalities: theory of majorization and its applications,* Springer, 2011.

[96] A. Goldsmith, *Wireless communications,* Cambridge Univ. Press, 2005.

[97] D. P. Palomar and Y. Jiang, "MIMO transceiver design via majorization theory," *Foundations and Trends in Communications and Information Theory,* vol. 3, issue 4, pp. 331-551, Nov. 2006.

[98] D. Tse and P. Viswanath, *Fundamentals of wireless communication,* Cambridge Univ. Press, 2005.

[99] S. Haykin and M. Moher, *Modern wireless communications,* Prentice Hall 2005.

[100] S. Verdu, *Multiuser detection*, Cambridge 1998.

[101] S. Mallat, *A wavelet tour of signal processing,* Academic Press, 2009.

[102] L. N. Trefethen and D. Bau, *Numerical linear algebra,* Siam, 1997.

[103] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms, third edition,* The MIT Press, 2009.

[104] S. Boyd and L. Vandenberghe, *Convex optimization,* Cambridge, U.K.: Cambridge Univ. Press, 2004.

[105] R. A. Horn and C. R. Johnson, *Topics in matrix analysis,* Cambridge Univ. Press, 1991.

[106] Y. P. Lin, S. M. Phoong and P. P. Vaidyanathan, *Filterbank transceivers for OFDM and DMT systems,* Cambridge Univ. Press, 2011.

[107] D. Knuth, *The art of computer programming, volume 3: sorting and searching, third edition,* Addison-Wesley, 1997.

[108] G. H. Golub and C. F. Van Loan, *Matrix computations,* Johns Hopkins Univ. Press, 1996.

[109] T. K. Moon and W. C. Stirling, *Mathematical methods and algorithms for signal processing,* Prentice-Hall, 2000.

[110] J. R. Magnus and H. Neudecker, *Matrix differential calculus with applications in statistics,* Wiley, 1999.

[111] G. L. Stuber, *Principles of mobile communications, 2nd Ed,* Kluwer Academic, 2001.