國 立 交 通 大 學
電信工程學系
博 士 論 文

於通道偏差下單載波及多載波區塊式傳輸
系統之強健式接收機設計

Robust Receiver Design for Single- and Multi-Carrier Block Transmission Systems Under Channel Mismatch

研 究 生：林 志 遠

指導教授：李 大 嵩 博士

中 華 民 國 九十六 年 五 月

於通道偏差下單載波及多載波區塊式傳輸
系統之強健式接收機設計
Robust Receiver Design for Single- and Multi-Carrier Block Transmission Systems Under Channel Mismatch

研 究 生: 林 志 遠　　　　　　　Student: Chih-Yuan Lin

指導教授: 李 大 嵩 博士　　　　　Advisor: Dr. Ta-Sung Lee

國 立 交 通 大 學
電 信 工 程 學 系
博 士 論 文

A Dissertation
Submitted to Institute of Communication Engineering
College of Electrical and Computer Engineering
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy
in
Communication Engineering

May 2007
Hsinchu, Taiwan, R.O.C.

中 華 民 國 九 十 六 年 五 月

國 立 交 通 大 學
博碩士論文全文電子檔著作權授權書

本授權書所授權之學位論文，為本人於國立交通大學 電信工程 系所 系統 組， 95 學年度第 二 學期取得博士學位之論文。

論文題目：於通道偏差下單載波及多載波區塊式傳輸系統之強健式接收機設 計

指導教授：李大嵩 教授

■ 同意　□不同意

本人茲將本著作，以非專屬、無償授權國立交通大學與台灣聯合大學系統圖書館：基於推動讀者間「資源共享、互惠合作」之理念，與回饋社會與學術研究之目的，國立交通大學及台灣聯合大學系統圖書館得不限地域、時間與次數，以紙本、光碟或數位化等各種方法收錄、重製與利用；於著作權法合理使用範圍內，讀者得進行線上檢索、閱覽、下載或列印。

<div align="center">論文全文上載網路公開之範圍及時間：</div>

本校及台灣聯合大學系統區域網路	■ 中華民國 96 年 7 月 30 日公開
校外網際網路	■ 中華民國 96 年 7 月 30 日公開

授 權 人：林志遠

親筆簽名：　林志遠

中華民國　96　年　6　月　26　日

國 立 交 通 大 學

博碩士紙本論文著作權授權書

本授權書所授權之學位論文，為本人於國立交通大學 電信工程 系所 系統 組， 95 學年度第 二 學期取得博士學位之論文。

論文題目：於通道偏差下單載波及多載波區塊式傳輸系統之強健式接收機設計

指導教授：李大嵩 教授

■ 同意

本人茲將本著作，以非專屬、無償授權國立交通大學，基於推動讀者間「資源共享、互惠合作」之理念，與回饋社會與學術研究之目的，國立交通大學圖書館得以紙本收錄、重製與利用；於著作權法合理使用範圍內，讀者得進行閱覽或列印。

本論文為本人向經濟部智慧局申請專利(未申請者本條款請不予理會)的附件之一，申請文號為：＿＿＿＿＿＿＿＿＿＿＿＿＿，請將論文延至＿＿年＿＿月＿＿日再公開。

授 權 人：林志遠

親筆簽名：＿林志遠＿＿＿

中華民國 96 年 6 月 26 日

國家圖書館博碩士論文電子檔案上網授權書

ID:9013530

本授權書所授權之論文為授權人在國立交通大學_電機_學院_電信工程_系所_ _系統_組_95_學年度第_二_學期取得博士學位之論文。

論文題目：於通道偏差下單載波及多載波區塊式傳輸系統之強健式接收機設計

指導教授：李大嵩 教授

茲同意將授權人擁有著作權之上列論文全文（含摘要），非專屬、無償授權國家圖書館，不限地域、時間與次數，以微縮、光碟或其他各種數位化方式將上列論文重製，並得將數位化之上列論文及論文電子檔以上載網路方式，提供讀者基於個人非營利性質之線上檢索、閱覽、下載或列印。

授權人：林志遠

親筆簽名：_林志遠_

民國 96 年 6 月 26 日

推 薦 函

一、事由：本校電信研究所博士班研究生 林志遠 提出論文以參加國立交通大學

博士班論文口試。

二、說明：本校電信研究所博士班研究生 林志遠 已完成本校電信研究所規定之

學科課程及論文研究之訓練。

有關學科部分，林君已修滿三十六學分之規定（請查閱學籍資料）並通過資

格考試。

有關論文部分，林君已完成其論文初稿，相關之論文亦發表於國際期刊（請

查閱附件）並滿足論文計點之要求。

總而言之，林君已具備國立交通大學電信研究所應有之教育及訓練水準，因

此特推薦

林君參加國立交通大學電信工程學系博士班論文口試。

交通大學電信工程學系教授　　李大嵩

國 立 交 通 大 學

論 文 口 試 委 員 審 定 書

本校 ___電信工程___ 學系博士班 ___林志遠___ 君

所提論文 於通道偏差下單載波及多載波區塊式傳輸系統之強健式接
收機設計
Robust Receiver Design for Single- and Multi-Carrier Block
Transmission Systems Under Channel Mismatch

合於博士資格水準、業經本委員會評審認可。

口試委員：

指導教授：
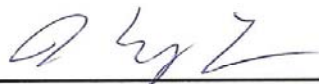
系 主 任： 教授

中 華 民 國 九 十 六 年 五 月 十 五 日

Department of Communication Engineering
National Chiao Tung University
Hsinchu, Taiwan, R.O.C.

Date: May 15, 2007

We have carefully read the dissertation entitled

Robust Receiver Design for Single- and Multi-Carrier Block

Transmission Systems Under Channel Mismatch

submitted by ___Chih-Yuan Lin___ in partial fulfillment of the requirements of the degree of DOCTOR OF PHILOSOPHY and recommend its acceptance.
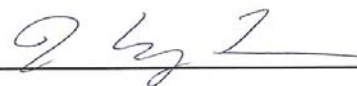
Thesis Advisor: _____

Chairman
Department of Communication Engineering: _____

於通道偏差下單載波及多載波區塊式傳輸系統之
強健式接收機設計

學生：林志遠　　　　　　　指導教授：李大嵩 博士

國立交通大學電信工程學系

摘要

多輸入多輸出正交分頻多工（MIMO-OFDM）系統及多輸入多輸出單載波迴旋前綴（MIMO SC-CP）系統能非常有效地補償頻率選擇性衰減及支援高資料傳輸率，因此已獲得許多系統設計者的注意，在本篇論文中，吾人將分別針對上述兩種系統在特定的通訊環境中設計接收機架構。在 MIMO-OFDM 系統部分，吾人考慮 CP 長度比通道階數（Channel Order）短的通訊環境，首先針對單輸入多出輸出（SIMO）模式進行設計，吾人聯合利用接收端的空間及頻率資源提出一個能有效地消除 CP 不足所造成之內符碼干擾（ISI）及內載波干擾（ICI）的強制最佳化（Constrained Optimization）線性等化器，其最佳化問題在等效的無強制（Unconstrained）廣義旁波帶消除（GSC）機制下進行求解，之後吾人再將所提出之 GSC 等化器架構推廣至 MIMO 模式。除此之外，吾人進一步假設接收端無法得知精確之通道參數而需採用最小平方（LS）技術進行估測，並提出利用擾動分析（Perturbation Analysis）技術將通道估計錯誤之效應明確地併入 GSC 系統模型，這使得 LS 通道誤差之特性能被應用，藉以推導出一個能對抗通道估計錯誤的封閉（Closed-Form）強健解。吾人亦推導出此強健式等化器的近似輸出訊號干擾雜訊比（SINR），由此結果可看出相較於非強解健解的一些優點。

在 MIMO SC-CP 系統部分，吾人考慮一個時變通道環境，同時假設通道參數亦採用 LS 技術估測而得。因為通道在時間上的變化會破壞頻域訊號間的正交性，使得低複雜度之各頻（Per-Tone）等化機制無法實現，所以在頻域處理訊號將不再具有優勢。因此，吾人提出直接於時域中處理訊號，於時域中，吾人發現訊號特徵矩陣能夠被分為數個擁有正交元素的群組，能自然地被用來設計群組式（Group-Wise）訊號偵測技術，為了實現此特性吾人提出一個 GSC 接收機，其能同時抑制通道時間變化與通道估計錯誤所產生之通道偏差效應。由驗證顯示吾人於 MIMO-OFDM 系統中針對 GSC 所設計的擾動分析數學架構亦能將本部分所考慮的通道偏差併入系統方程式中，基於通道時間變化與通道估計錯誤的統計假設吾人亦可推導出一個封閉（Closed-Form）強健解。透過一些數值範例可證實吾人所提出之接收機架構在所考慮的通訊環境中效能明顯優於現存的方法。

Robust Receiver Design for Single- and Multi-Carrier Block Transmission Systems Under Channel Mismatch

Student: Chih-Yuan Lin   **Advisor: Dr. Ta-Sung Lee**

Department of Communication Engineering
National Chiao Tung University

Abstract

MIMO orthogonal frequency division multiplexing (MIMO-OFDM) and MIMO single-carrier with cyclic prefix (MIMO SC-CP) have drawn a lot of attention since they can effectively compensate frequency selective channels and can support high data rates. In this dissertation, we will design receiver architectures for both of them, each under a specific communication environment. For MIMO-OFDM systems, we consider a scenario that the adopted CP length is shorter than the channel order. By jointly exploiting the receiver spatial and frequency resources, we first propose a constrained optimization based linear equalizer, which can mitigate the resultant inter-symbol interference and inter-carrier interference (ICI) incurred by the insufficient CP insertion, for the SIMO case. The optimization problem is solved under an equivalent unconstrained generalized sidelobe canceller (GSC) setup. Then the proposed GSC-based equalization framework is generalized to the MIMO case. Moreover, in this case we further assume that the channel parameters are not exactly known but are estimated using the least-squares (LS) training technique. We propose to apply the perturbation analysis technique to explicitly incorporate the channel parameter error into the GSC system model; this allows us to exploit the presumed LS channel error properties for deriving a closed-form robust solution against the net detrimental effects caused by the channel estimation errors. A closed-form approximate output SINR expression of the proposed robust equalizer is also derived, based on which some appealing advantages over the non-robust counterpart can be inferred.

For MIMO SC-CP systems, we consider a communication environment that the channel is time-varying, under the assumption that the channel parameters are also estimated via the LS training technique. Since the channel temporal variation destroys the orthogonality between signal components in frequency domain, low-complexity per-tone based equalizations can no longer be realized. As a result, there are no specific advantages of processing signals in frequency domain,

and we propose to directly process signals in time domain. By this way, it is observed that in time domain the signal signatures can be arranged into groups of orthogonal components, leading to a very natural group-wise symbol recovery scheme. To realize this figure of merit, we propose a GSC-based receiver, which also takes into account the mitigation of channel mismatch effects caused by the channel temporal variation and the imperfect estimation. It is shown that the proposed perturbation analysis framework as well enables us to model the channel mismatch effects into the system equation and, in turn, to further exploit the statistical assumptions on the channel temporal variation and the estimation error for deriving a closed-form robust solution. By some numerical examples, it is confirmed that the proposed receiver architectures outperform the existing methods under the respectively considered communication environments.

Acknowledgement

Foremost, I would like to thank my advisor Prof. Ta-Sung Lee for providing me with the opportunity to complete my Ph.D. dissertation at National Chiao Tung University. I learned much from him to improve my working and research styles. I especially want to thank Dr. Jwo-Yuh Wu, whose support made my dissertation work possible. I am very grateful for his motivation, enthusiasm, and immense knowledge in matrix algebra.

I also want to thank the former members of the Communication System Design and Signal Processing (CSDSP) Lab: Dr. Juinn-Horng Deng, Dr. Gau-Joe Lin, and Dr. Chung-Lien Ho for giving me an introduction into the CSDSP Lab and for sharing their knowledge. Thank also goes to my present student colleagues: Fang-Shuo Tseng and Wen-Fang Yang for their inspiring discussions.

Finally, I thank my mother Yueh-Yun Lin, younger sister Bei-Yun Lin, younger brother Chih-Sheng Lin, and girl friend Kan-Wen Lin for their faith and support.

Contents

List of Figures

List of Tables

List of Notations

$\mathbb{C}^m$	set of m-dimensional complex vectors
$\mathbb{C}^{m \times n}$	set of $m \times n$ complex matrices
$\mathbf{X}^T$	transpose of matrix $\mathbf{X}$
$\mathbf{X}^*$	complex conjugate of $\mathbf{X}$
$\mathbf{X}^H$	complex conjugate transpose of $\mathbf{X}$
$\mathbf{X}_{i,j}$	(i,j) th entry of the matrix $\mathbf{X}$
$\|\mathbf{x}\|$	vector two-norm
$\|\mathbf{X}\|$	matrix Frobenius norm
$Tr(\mathbf{X})$	trace of the matrix $\mathbf{X}$
$\odot$	matrix Hadamard product
$\otimes$	matrix Kronecker product
$E\{y\}$	expected value of the random variable y
$\mathbf{I}_m$	$m \times m$ identity matrix
$\mathbf{0}_m$	$m \times m$ zero matrix
$\mathbf{0}_{m \times n}$	$m \times n$ zero matrix
$diag\{\mathbf{x}\}$	$m \times m$ diagonal matrix with the elements of the vector $\mathbf{x}$ on the main diagonal
$Diag\{\mathbf{C}_1,...,\mathbf{C}_m\}$	block diagonal matrix with the diagonal block submatrices $\mathbf{C}_p$, $1 \le p \le m$

List of Acronyms

AP   access point

BER   bit error rate

V-BLAST   vertical Bell Laboratories layered Space-Time

CP   cyclic prefix

DL   diagonal loading

FDE   frequency-domain equalization

FFT   fast Fourier transform

GSC   generalized sidelobe canceller

ICI   intercarrier interference

IFFT   inverse fast Fourier transform

ISI   inter-symbol interference

LS   least-squares

LSFE   layered space-frequency equalization

MCM   multicarrier modulation

MIMO   multi-input multi-output

ML   maximum likelihood

MMSE   minimum mean square error

MU   mobile unit

OFDM   orthogonal frequency division multiplexing

PA   partial adaptivity

PAR   peak-to-average ratio

PIC   parallel interference cancellation

pre-MRC   pre-maximal ratio combining

PTEQ   per-tone equalization

RHS   right-hand-side

RR   reduced-rank

SC-CP   single-carrier with cyclic prefix

SIC   successive interference cancellation

SIMO   single-input multi-output

SIR   signal-to-interference ratio

SINR   signal to interference-plus-noise ratio

SM   spatial multiplexing

STC   space-time coding

TEQ   time-domain equalization

TIR   target impulse response

ZF   zero-forcing

Chapter 1

Introduction

In this introductory chapter, some background materials about block transmission schemes and multi-antenna systems are presented. What follow up are the literature survey, dissertation contributions and an overview of this dissertation.

1.1 Basics of OFDM and SC-CP

Orthogonal frequency division multiplexing (OFDM) [49], [71] is a kind of multicarrier modulation (MCM) scheme. Unlike the classical MCM, in which guard bands are required to separate the different subcarriers (see Figure 1.1(a)), the orthogonality nature of the OFDM subcarriers allows their sidebands to overlap (see Figure 1.2(b)). This makes OFDM the most spectrally efficient among the MCM schemes.
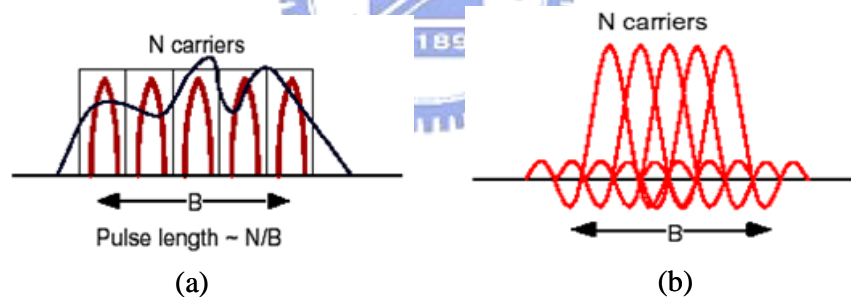
Figure 1.1. (a) Classical MCM, and (b) OFDM modulation.

In practical applications of the OFDM modulation, the transmitter segments the signal streams into blocks, which are block-wise fed into the IFFT device. Prior to transmission, a cyclic prefix (CP) is inserted in front of each post-IFFT signal block [71]. At the receiver the CP portion should be discarded before any signal processing starts. As long as CP length is longer than the channel order, the inter-symbol interference (ISI) coming from the previous block is removed, and, moreover, the delayed replicas of an OFDM block are ensured to always have an integer number of cycles within the FFT interval, such that orthogonality between subcarriers is preserved. The

receiver then can only use the FFT device followed by the per-tone one-tap equalizers to recover the transmitted OFDM block [71] (the overall transceiver architecture is shown in Figure 1.2). Due to its simplicity and robustness against frequency selective channels, the OFDM modulation recently has been adopted in a lot of commercial communication systems, such as DAB, DVB and 802.11a/g WLAN. Furthermore, the lately established IEEE 802.16 standard [66], [81] also adopts OFDM in its PHY layer. The main advantage of OFDM comes from the insertion of CP, but, however, the presence of CP could greatly decrease the effective data rate; for example, in the IEEE 802.11a standard, 25% spectral resource is wasted due to the CP insertion. In addition, another fundamental drawback of OFDM is high peak-to-average ratio (PAR) of the transmitted signal, which would incur high clipping rate: each signal sample that is beyond the saturation limit of the power amplifier suffers clipping to this limit value. This leads to non-linear signal distortion and creates additional bit errors at the receive end [49].
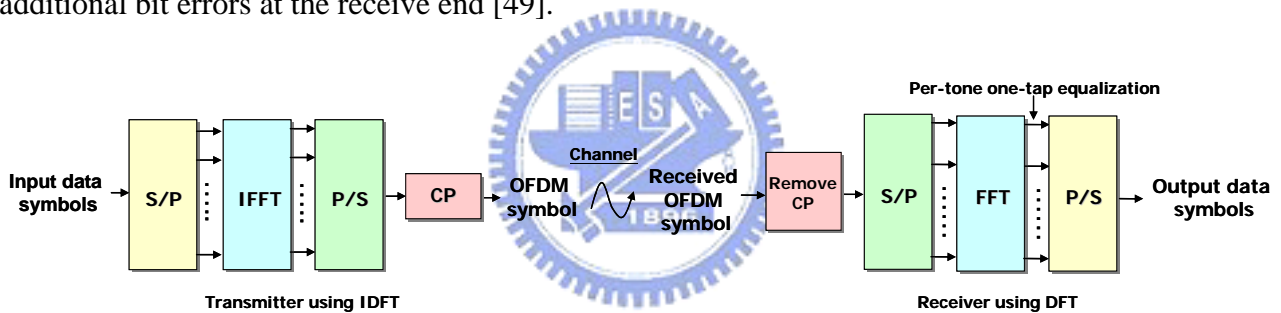
Figure 1.2. Block diagram of the OFDM transceiver.

An alternative block transmission scheme is called as single-carrier with cyclic prefix (SC-CP) [19], [20] (see Figure 1.3). With the CP length longer than the channel delay spread, it allows to use the low-complexity one-tap frequency-domain equalization (FDE) to compensate the negative effects caused by frequency-selective channels. Furthermore, since SC-CP system alleviates the IFFT operation (and hence avoids the coherent superposition of the signal samples) at the transmitter, the PAR of SC-CP signals is essentially much lower than that of OFDM signals. In summary, SC-CP has similar performance, bandwidth efficiency, and low signal processing complexity as OFDM, but with much lower signal PAR. The next-generation wireless communication standard, like 3GPP-LTE [17], [45], has considered the coexistence of OFDM and SC-CP in a system: OFDM is used in the downlink transmission and SC-CP accounts for the uplink

counterpart (see Figure 1.4) [45]. Such a deployment leads to two advantages: *1)* shifts most signal processing complexity to the base station (two IFFTs and one FFT at the base station, but only one FFT at the mobile station), and *2)* allows the power amplifier at the battery-driven mobile station to work more efficiently (since SC-CP has relatively lower signal PAR).
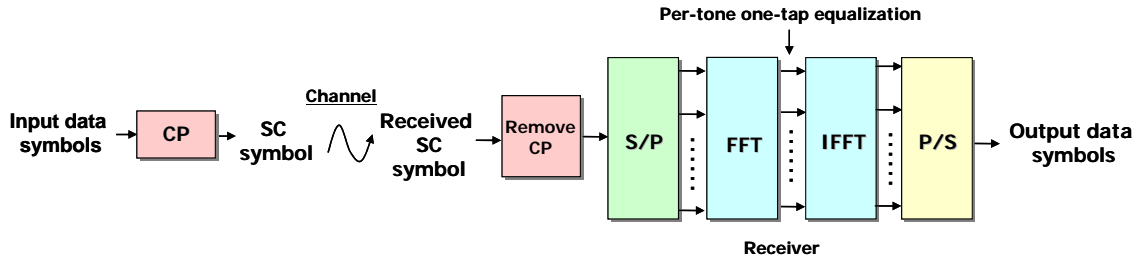
Figure 1.3. Block diagram of the SC-CP transceiver.
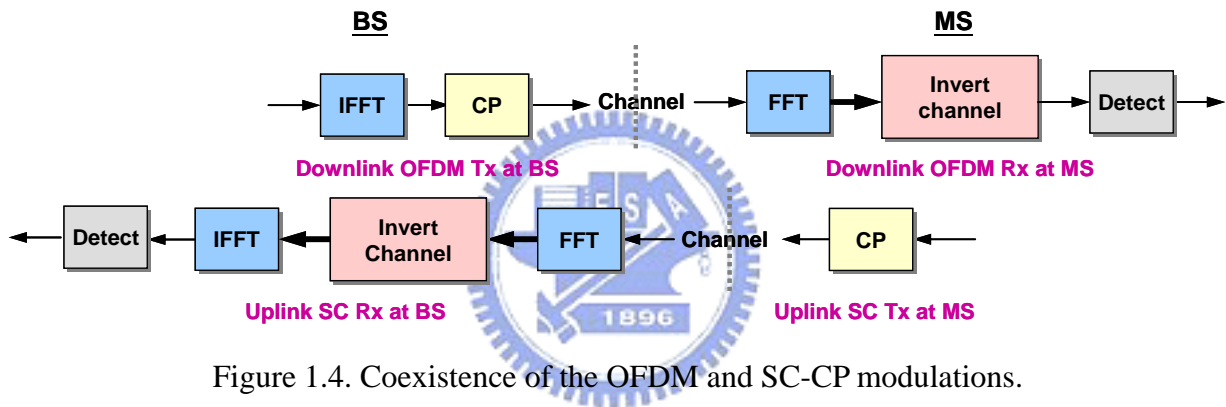
Figure 1.4. Coexistence of the OFDM and SC-CP modulations.

1.2 Basics of Multi-Antenna Systems

Multiple antennas, as is widely known, can offer an additional antenna gain for various specific tasks. If a multiuser system is equipped with an antenna array at the base station, user terminals can transmit and receive with a smaller gain, leading to lower system interference level and thus larger system capacity. Moreover, multi-antenna array can also offer a much larger degrees-of-freedom for interference suppression. This certainly results in a better system output SINR performance compared with the single-antenna system. Recent theoretical results show that if transmit and receive ends are respectively with M and N antennas, a.k.a., multiple-input multiple-output (MIMO) systems (see Figure 1.5), the point-to-point system capacity can linearly increase with $\min(M, N)$ [10]. As a result, the MIMO transmission method has been recognized as one promising solution to achieve high spectral efficiency and link quality, which are believed to be

two major requirements of the next-generation wireless networks. The MIMO research topics are generally categorized into two types: *1)* spatial multiplexing (SM) and *2)* diversity techniques. A well established SM scheme is known as Bell Laboratories layered Space-Time (BLAST) [82]. At the transmitter multiple independent data streams are simultaneously transmitted via different antenna branches and are detected at the receiver based on their unique spatial signatures. Intuitively, SM technique creates several parallel spatial transmission links so as to increase data rate. Such a system setup can significantly increase system spectral efficiency without the need of increasing transmit power and system bandwidth. On the other hand, the well-known diversity techniques are the space-time coding (STC) schemes [4], [29], [70], in which the transmitted symbols are appropriately mapped into multiple transmit antennas. The receiver then can exploit the artificially induced signal redundancy to obtain diversity gain. Recently, MIMO schemes are usually combined with the OFDM or the SC-CP modulation, a.k.a., MIMO-OFDM [51] or MIMO SC-CP [15], [85], such that the system has a two-fold advantage (coming from MIMO and OFDM (or SC-CP)). MIMO-OFDM and MIMO SC-CP are widely believed to play important roles in the next-generation wireless communication systems, since they can efficiently combat frequency-selective channels and can support high data rates. The next-generation wireless standards, like IEEE 802.16 [66], [81] and 3GPP-LTE [17], [45], have also respectively included them into the specifications.
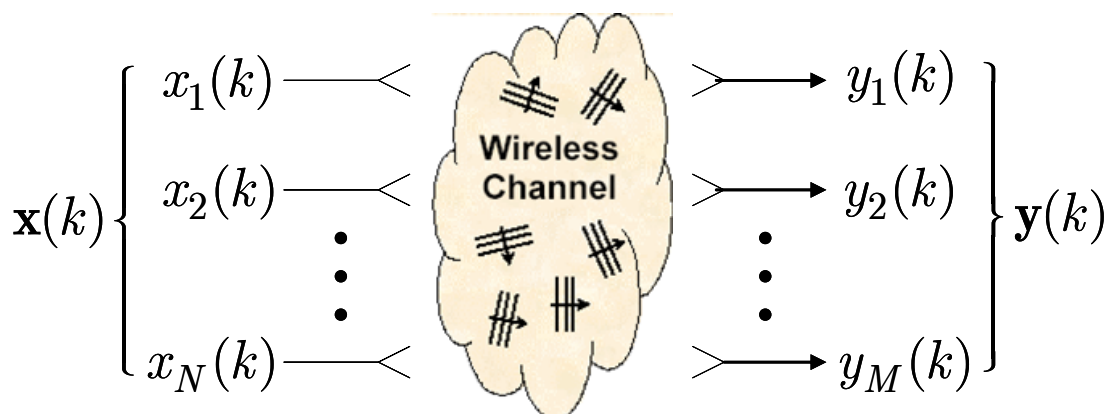
Figure 1.5. Illustration of a MIMO system.

1.3 Related Literature Review

The conventional OFDM systems rely on the insertion of CP to remove the channel-induced ISI, leading to a very simple one-tap equalization implementation [71]. To further boost the effective data rate, a natural strategy is to reduce CP overhead and design a receiver to compensate the resultant negative effects. The time-domain equalization (TEQ) techniques have been extensively studied for discrete multi-tone systems. In these approaches, a time-domain equalizer is employed to properly squeeze the channel impulse response (CIR) to be a prescribed target impulse response (TIR) such that the CP length can be reduced to be the TIR order. In [2] and [65], minimum mean square error (MMSE) based TEQ are proposed to obtain a composite CIR, which is best close to the desired TIR. In [5] and [46], the maximum shortening-SNR method is proposed to maximize the ratio of the energies of the CIR inside and outside a target window; by properly choosing the window size the undesired ISI can be largely suppressed. A lot of researchers are also based on the so-called bit rate maximization criterion to design TEQ [39], [76]. However, such methods should usually exploit the nonlinear optimization techniques to search the optimal TEQ solution, and thus a globally optimal solution is not always guaranteed. In [37], a semiblind TEQ method, which does not require the CIR and channel noise statistics, is proposed for VDSL systems to maximize the frequency-domain signal-to-interference ratio (SIR). The above described methods are usually based on the offline designs: using the derived equalizer coefficients at the beginning of the communication link once and for all during the whole transmission period. However, such approaches are unable to work well if the channel varies with time. In [77], several adaptive TEQ methods are proposed to track the channel variation, so as to be able to maximize the bit rate as high as possible all the time. For MIMO systems, [1] proposes to design TEQ based on minimizing the average energy of the error sequence between the equalized MIMO CIR and a desired MIMO TIR with shorter memory. As an alternative to the TEQ approach, the per-tone equalization (PTEQ) [33] transfers the equalization problem from time domain to frequency domain and has better performance than the TEQ methods at the expense of computation complexity.

One particularly attractive feature unique to the block transmission systems is the

low-complexity per-tone FDE scheme. Such a figure of merit, however, hinges crucially on the time-invariant assumption on the background channels. When the channel is otherwise subject to fast temporal variation, the orthogonality among the signal components in frequency domain will no longer be preserved, and tone-by-tone signal recovery is then rendered impossible. To design communication architectures for block transmission systems under time-varying channels is a hot research topic in recent years. In [59] and [86], a method called intercarrier interference (ICI) self cancellation scheme is proposed for OFDM systems. Based on the observation that the ICI components at the adjacent subcarriers are highly similar, at the transmitter one data symbol is assigned, with judiciously designed weights, to multiple subcarriers, and then at the receiver by linear combination of the signals at those subcarriers the ICI components in the received signal can be mutually cancelled. Such approaches are quite simple but, however, the spectral efficiency is low since multiple subcarriers only carry one data symbol. In [55], a simple MMSE-based linear receiver is proposed, in which a banded approximation to the frequency-domain channel matrix is adopted such that the computation of the MMSE weight matrix can be substantially conserved. The linear receiver can also be combined with the parallel interference cancellation (PIC) mechanism to further improve the system performance [28]. In [68], a two-stage receiver is proposed for OFDM systems: in the first stage a max-SINR filter is exploited to suppress ICI and in the second stage an iterative MMSE detector is used to estimate transmitted source symbols; the two-stage receiver architecture also can be further extended for SC-CP systems [67]. It is noted that, to the best of our knowledge, the receiver design under time-varying channels for MIMO block transmission systems remains scarce in the literature.

Receiver designs are in general based on a crucial assumption that the channel state information is perfectly known at the receiver end. However, it is impossible to achieve this in practice, under either time-invariant channels or time-varying channels. As a result, when channel estimation error appears, the derived solutions are rendered suboptimal. Robust linear receiver design against channel estimation errors has been addressed in the context of multiuser communication [52], [53], [54], [84]. By modeling the channel mismatch as a random variable with

known statistical characteristics, the probability-constrained optimization approach [52], [53] exploits the Gaussianality assumption on the estimation error, and the solution is obtained through linear [52] or nonlinear [53] programming. In [54] and [84], the channel error is, on the other hand, treated as a "deterministic" perturbation. Based on a min-max type formulation, the optimization problem in [54] is solved by using the convex programming technique; the solution in [86] admits a diagonal loading form, with the optimal loading factor determined via an iterative procedure. We note that the deterministic formulation of model error is also used in robust beamformer design [35], [42], [64], [75], in which exact statistical characterization of model uncertainty due to, e.g., array calibration error, unknown antenna coupling effect, etc., is difficult (or even impossible) to track.

1.4 Main Contributions

The contributions of this dissertation are summarized as follows:

1. A generalized sidelobe canceller (GSC) based channel equalization framework is proposed for MIMO-OFDM systems with insufficient CP. By jointly exploiting spatial and frequency resources and based on the block signal representation, the desired signal and the ISI/ICI components in the received signal are respectively characterized by the specific signatures, which allows us to leverage the constrained-optimization technique to suppress the interference part and then extract the desired one. It is shown that the proposed method is able to completely suppress the channel-induced distortion via the space-frequency processing, leading to comparable BER performance to the idea case (with sufficient CP insertion).

2. By exploiting the perturbation analysis technique, we incorporate the detrimental effect caused by channel estimation error into the proposed GSC-based MIMO-OFDM architecture to derive a closed-form robust version against channel mismatch. It is shown that the proposed framework merely requires the knowledge of the first- and second-order error statistics but, unlike [52] and [53], is free from any priori assumptions on the error distributions. An analytical expression of the SINR performance for the robust method is derived, based on which the proposed robust method is theoretically proven to yield higher output SINR than the non-robust counterpart.

3. A new constrained-optimization based receiver architecture for MIMO SC-CP systems under time-varying channels is proposed. By exploiting the imbedded cyclic-shift structure of the channel matrix, we can realize a low-complexity group-wise symbol detection setup. Moreover, based on the proposed perturbation analysis framework the channel mismatch caused by the channel temporal variation and the estimation error is also considered into the solution equation, leading to a closed-form robust solution. In comparison with the existing methods, it is shown that the proposed method has better BER performance, but only with comparable computation complexity.

1.5 Organization of Dissertation

The remaining of this dissertation is organized as follows.

In Chapter 2, a GSC based equalizer for ISI/ICI suppression is proposed for uplink SIMO OFDM systems without CP. Based on the block representation of the CP-free OFDM system, there is a natural formulation of the ISI/ICI suppression problem under the GSC framework. By further exploiting the signal and ISI/ICI signature matrix structures, a computationally efficient partially adaptive (PA) implementation of the GSC-based equalizer is proposed for complexity reduction. The proposed scheme can be extended for the design of a pre-equalizer, which pre-suppresses the ISI/ICI and realizes CP-free downlink transmission to ease the computational burden of the mobile unit, which is assumed to be equipped with only one antenna.

Chapter 3 generalizes the ISI/ICI suppression scheme proposed in Chapter 2 for MIMO-OFDM systems with variable CP length (shorter than the channel order) and derives an associated robust version against channel estimation error, assuming that the channel parameters are estimated via the commonly used least-squares (LS) training technique. The channel parameter error is explicitly incorporated into the constraint-free GSC system model through the perturbation analysis technique; this allows us to exploit the presumed LS channel error properties for deriving a closed-form solution, and can also facilitate an associated analytic performance analysis. A closed-form approximate SINR expression for the proposed robust scheme is given, and an appealing formula of the achievable SINR improvement over the non-robust counterpart is further specified. Our analytic

results bring out several intrinsic features of the proposed solution.

In Chapter 4, we consider MIMO SC-CP block transmissions over time-varying multipath channels, under the assumption that the channel parameters are not exactly known but are estimated via the LS training technique. While the channel temporal variation is known to negate the tone-by-tone frequency-domain equalization facility, it is otherwise shown that in the time domain the signal signatures can be arranged into groups of orthogonal components, leading to a very natural yet efficient group-by-group symbol recovery scheme. To realize this figure of merit we propose a GSC based receiver which also takes into account the mitigation of channel mismatch effects caused by the channel temporal variation and the imperfect estimation. Based on the robustness analysis framework proposed in Chapter 3, the channel mismatch is directly modeled into the system equations. Then a closed-form solution can be obtained by further exploiting the statistical assumptions on the channel mismatch. Within the considered framework the proposed robust equalizer can be further combined with the successive interference cancellation (SIC) mechanism for further performance enhancement.

Finally, Chapter 5 concludes this dissertation and discusses future extensions of this research.

Chapter 2

Low-Complexity Transceiver for CP-Free Multi-Antenna OFDM Systems

2.1 Overview

Conventional OFDM transmission relies on the insertion of CP to remove the channel induced ISI and ICI, leading to a very simple one-tap equalization implementation [71]. The presence of CP, however, greatly decreases the effective data rate; for example, in the IEEE 802.11a standard, 25 % spectral resource is wasted due to the CP insertion. In recently proposed WLAN systems, multiple antennas are suggested to be employed at the access point (AP) to provide the receive diversity or beamforming gain for enhancing the link quality. Multiple receive antennas, as is widely known, can offer an additional antenna gain for interference suppression. As a result, a natural approach to improving the spectral efficiency in OFDM systems would be to consider CP-free transmission and exploit the multi-antenna resource for combating the channel-induced distortion.

In this chapter, we will propose a multi-antenna transceiver architecture for the CP-free OFDM systems to jointly suppress ISI and ICI. We consider that multiple antennas are equipped at the AP and a single antenna at the mobile unit (MU). In the uplink stage, the degrees-of-freedom offered by multiple receive antennas and their associated subcarriers are exploited for ISI and ICI suppression based on the constrained-optimization technique. The optimization problem is reformulated into an unconstrained one via the GSC principle. The proposed CP-free GSC-based solution can compensate the channel distortion and enhance the data rate, at the expense of increased receiver complexity as compared with the conventional CP-based systems. To mend this penalty, we propose a partial adaptivity (PA) GSC scheme, which involves a series of reduced rank processing [22], [25], for complexity reduction. In time division duplex OFDM systems over slowly time-varying channel, the AP can pre-equalize the ISI/ICI based on the channel knowledge estimated in the uplink stage so as to relieve the MU's computational burden. Since the downlink stage is the dual operation of the

uplink stage, the GSC method is again a natural candidate for the pre-equalizer design based on the block transmission system model. The proposed GSC-based pre-equalizer, which enjoys the same figure of merits as in the uplink counterpart, is shown to effectively pre-suppress the ISI/ICI.

2.2 Uplink Signal Model

We consider the discrete-time baseband model of a multi-antenna CP-free OFDM system with Q subcarriers and M receive antennas. The source symbol sequence $s(n)$ is segmented to obtain consecutive $Q \times 1$ symbol blocks. The kth one among them is $s(k) = \left[s_0(k), ..., s_{Q-1}(k) \right]^T$, where $s_j(k) = s(kQ + j)$, $0 \leq j \leq Q - 1$. After the inverse fast Fourier transform (IFFT), we obtain the $Q \times 1$ OFDM symbol

$$\mathbf{x}(k) = \mathbf{F}^{-1}\mathbf{s}(k), \tag{2.1}$$

where $\mathbf{x}(k) = \left[x_0(k), ..., x_{Q-1}(k) \right]^T$, with $x_j(k) = x(kQ + j)$, $0 \leq j \leq Q - 1$, and $\mathbf{F}$ is the Q-point FFT matrix [79]. The signal $\mathbf{x}(k)$ is then transmitted through the composite channel $h^{(m)}(l)$, including the response of the cascade connection of the transmitter filter, the channel, and the receive filter, to the mth receive antenna, $1 \leq m \leq M$. We assume that the channel orders of $h^{(m)}(l)$'s are all L, which is reasonable since we can always choose L as the largest one among them. Assuming that the transmitter is synchronized with the receiver, the $Q \times 1$ received data vector of the mth antenna branch, in terms of the block data model [79], is

$$\mathbf{r}^{(m)}(k) = \mathbf{H}_0^{(m)}\mathbf{F}^{-1}\mathbf{s}(k) + \mathbf{H}_1^{(m)}\mathbf{F}^{-1}\mathbf{s}(k-1) + \mathbf{n}^{(m)}(k), \tag{2.2}$$

in which $\mathbf{H}_0^{(m)} \in \mathbb{C}^{Q \times Q}$ is a lower triangular Toeplitz matrix with

$$\left[h^{(m)}(0), ..., h^{(m)}(L), 0, ..., 0 \right]^T \tag{2.3}$$

as its first column, $\mathbf{H}_1^{(m)} \in \mathbb{C}^{Q \times Q}$ is an upper triangular Toeplitz matrix with

$$\left[0, ..., 0, h^{(m)}(L), ..., h^{(m)}(1) \right] \tag{2.4}$$

as its first row, and $\mathbf{n}^{(m)}(k)$ is the noise vector. The following assumptions are made in the sequel:
1. The source symbols $s(n)$ are zero mean, uncorrelated, and $E\left\{ s(n_1)s^*(n_2) \right\} = \delta\left(n_1 - n_2 \right)$,

where $E\{y\}$ denotes the expectation of the random variable y, and $\delta(.)$ is the Kronecker delta function.

2. The channel taps $h^{(m)}(l)$ are modelded as i.i.d. zero-mean complex Gaussian random variables with variance σ_l^2 over m and l.

3. The elements of $\mathbf{n}^{(m)}(k)$'s are modeled as AWGN, for $1 \leq m \leq M$, with the same power σ_v^2.

4. Perfect channel knowledge is available at the AP.

2.3 Channel-Induced Distortion

Based on the block system model, this section characterizes the channel-induced interference for a CP-free OFDM system. From (2.3) and (2.4), it is easy to verify that the matrix $\mathbf{H}_0^{(m)} + \mathbf{H}_1^{(m)} \in \mathbb{C}^{Q \times Q}$ is a circulant matrix with the first row equal to $\left[h^{(m)}(0), ..., 0, h^{(m)}(L), ..., h^{(m)}(1) \right]$. This suggests that we can rewrite the block signal model (2.2) as

$$\mathbf{r}^{(m)}(k) = \left(\mathbf{H}_0^{(m)} + \mathbf{H}_1^{(m)} \right) \mathbf{F}^{-1} \mathbf{s}(k) + \mathbf{H}_1^{(m)} \mathbf{F}^{-1} \left(\mathbf{s}(k-1) - \mathbf{s}(k) \right) + \mathbf{n}^{(m)}(k). \tag{2.5}$$

The first term on the right-hand-side of (2.5), which depends entirely on the currently transmitted OFDM symbol, has the appealing property that the signature matrix is circulant. The second term, on the other hand, shows a penalty of removing CP: it accounts for the multipath induced distortion and can cause loss of the orthogonality among subcarriers. From (2.5), we directly obtain the corresponding frequency-domain description

$$\mathbf{z}^{(m)}(k) = \mathbf{F}\mathbf{r}^{(m)}(k) = \mathbf{F}(\mathbf{H}_0^{(m)} + \mathbf{H}_1^{(m)})\mathbf{F}^{-1}\mathbf{s}(k) + \mathbf{F}\mathbf{H}_1^{(m)}\mathbf{F}^{-1}(\mathbf{s}(k-1) - \mathbf{s}(k)) + \mathbf{F}\mathbf{n}^{(m)}(k). \tag{2.6}$$

Since $\mathbf{H}_0^{(m)} + \mathbf{H}_1^{(m)}$ is circulant, the model (2.6) reduces to

$$\mathbf{z}^{(m)}(k) = \mathbf{D}^{(m)}\mathbf{s}(k) + \mathbf{F}\mathbf{H}_1^{(m)}\mathbf{F}^{-1}\mathbf{e}(k) + \mathbf{F}\mathbf{n}^{(m)}(k), \tag{2.7}$$

where $\mathbf{D}^{(m)} \in \mathbb{C}^{Q \times Q}$ is the diagonal matrix with the qth diagonal entry being the channel frequency response evaluated at the qth subcarrier of the mth receive antenna and $\mathbf{e}(k) = \mathbf{s}(k-1) - \mathbf{s}(k)$ induces the ISI/ICI perturbation. Compared with the time-domain expression (2.5) (or (2.2)), the

frequency-domain model (2.7) has two appealing properties. First, the desired signals are decoupled. In particular, the diagonal nature of the signal signature $\mathbf{D}^{(m)}$ will impose certain structure that can simplify the computation of the GSC equalizer parameters. Second, the frequency-domain perturbation is specified by the interference signature matrix $\mathbf{F}\mathbf{H}_1^{(m)}\mathbf{F}^{-1}$ (see (2.7)). This in turn implies that the interference subspace is spanned by some columns of the FFT matrix $\mathbf{F}$. This fact can be further exploited for deriving a closed-form PA implementation of the proposed GSC equalizer for complexity reduction. Hence, instead of relying on the time-domain counterpart, we will focus on the frequency-domain model (2.7) in the following discussions.

2.4 Proposed GSC-Based Equalizer

In this section, we collect the data from multiple receive antennas to form a multi-channel system model. By exploiting a special structure inherent in the interference signature matrix, we propose a GSC-based equalizer scheme as well as an associated reduced complexity implementation

2.4.1 Multi-Channel Signal Model

Stacking $\mathbf{z}^{(m)}(k)$ (see (2.7)), $1 \leq m \leq M$, from all the receive antennas gives the following $MQ \times 1$ post-FFT multi-channel model

$$\mathbf{z}(k) = \left[\mathbf{z}^{(1)T}, ..., \mathbf{z}^{(M)T} \right]^T = \mathbf{D}\mathbf{s}(k) + \underbrace{\mathbf{F}_M \mathbf{H}_1 \mathbf{F}^{-1}}_{\mathbf{H}_I} \mathbf{e}(k) + \mathbf{v}(k) = \begin{bmatrix} \mathbf{D} & \mathbf{H}_I \end{bmatrix} \begin{bmatrix} \mathbf{s}(k) \\ \mathbf{e}(k) \end{bmatrix} + \mathbf{v}(k), \tag{2.8}$$

in which

$$\mathbf{D} = \left[\mathbf{D}^{(1)T} \cdots \mathbf{D}^{(M)T} \right]^T \in \mathbb{C}^{MQ \times Q} \tag{2.9}$$

and

$$\mathbf{H}_1 = \left[\mathbf{H}_1^{(1)T} \cdots \mathbf{H}_1^{(M)T} \right]^T \in \mathbb{C}^{MQ \times Q} \tag{2.10}$$

are respectively the concatenated multi-channel signal signature matrix and ISI/ICI inducing channel matrix, $\mathbf{F}_M$ is an $MQ \times 1$ matrix given by $\mathbf{F}_M = \mathbf{I}_M \otimes \mathbf{F}$, with $\otimes$ denoting the Kronecker product, and $\mathbf{v}(k)$ is the noise component. From (2.4), it can be seen that the nonzero

entries of the upper triangular Toeplitz matrix $\mathbf{H}_1^{(m)}$ occur only in the last L columns. Since $\mathbf{H}_1$ is obtained by stacking $\mathbf{H}_1^{(m)}$, $1 \leq m \leq M$, one on top of another and both $\mathbf{F}_M$ and $\mathbf{F}^{-1}$ are unitary, the block Toeplitz structure of $\mathbf{H}_1$ immediately implies that the overall ISI/ICI signature matrix, namely, $\mathbf{H}_I$, is of rank L. Let $\mathbf{V} \in \mathbb{C}^{MQ \times (MQ-L)}$ be the matrix whose columns form an orthonormal basis of the left null space of $\mathbf{H}_I$. Then, from (2.8) and the fact that $\mathbf{V}^H \mathbf{H}_I = \mathbf{O}$, the $(MQ - L) \times 1$ ISI-free data model is obtained as

$$\mathbf{z}_v(k) = \mathbf{V}^H \mathbf{z}(k) = \mathbf{V}^H \mathbf{D}\mathbf{s}(k) + \mathbf{V}^H \mathbf{v}(k), \tag{2.11}$$

where $(.)^H$ denotes the Hermitian transpose. It is readily observed that if $MQ \geq Q + L$, one can completely extract the Q-dimensional desired signal from (2.11). Since the number of subcarriers is typically much larger than the channel order, i.e., $Q \gg L$, the choice of $M \geq 2$ will satisfy the condition.

2.4.2 Equalization Based on ISI Suppression

Based on the ISI-free data model (2.11), one can use various equalization approaches to recover the desired symbols. Since the columns of the matrix $\mathbf{V}^H \mathbf{D} \in \mathbb{C}^{(MQ-L) \times Q}$ are in general not orthogonal (i.e., $\mathbf{D}^H \mathbf{V}\mathbf{V}^H \mathbf{D} \neq \mathbf{I}_Q$, where $\mathbf{I}_n$ denotes an $n \times n$ identity matrix), the ISI nulling operation in (2.11) induces ICI and so an additional signal separation procedure is required for recovering the transmitted symbol on each subcarrier. To this end, a linear equalizer $\mathbf{W}_v \in \mathbb{C}^{(MQ-L) \times Q}$ of the form $\mathbf{y}_v(k) = \mathbf{W}_v^H \mathbf{z}_v(k)$ can be applied. One typical solution for $\mathbf{W}_v$ is the linear zero-forcing (ZF) [24] or minimum mean square error (MMSE) [60] equalizer with the following $(MQ - L) \times Q$ weight matrix:

$$\mathbf{W}_{zf} = \mathbf{V}^H \mathbf{D}(\mathbf{D}^H \mathbf{V}\mathbf{V}^H \mathbf{D})^{-1}, \tag{2.12}$$

or

$$\mathbf{W}_{mmse} = \left(\mathbf{V}^H \mathbf{D}\mathbf{D}^H \mathbf{V} + \sigma_v^2 \mathbf{I}_{MN-L}\right)^{-1} \mathbf{V}^H \mathbf{D}. \tag{2.13}$$

In case that the channel frequency response undergoes severe fading near some of the FFT grids, and hence some diagonal entries of $\mathbf{D}$ are very close to zero, the ZF solution in (2.12) tends to

amplify the noise effect, even if the SNR is high. The MMSE solution (2.13), on the other hand, can attenuate the noise effect but, as compared with the ZF solution (2.12), involves the larger size matrix inversion $\left(\mathbf{V}^H\mathbf{D}\mathbf{D}^H\mathbf{V} + \sigma_v^2\mathbf{I}_{MN-L}\right)^{-1} \in \mathbb{C}^{(MQ-L)\times(MQ-L)}$. It is noted that this "two-stage" method completely nulls the ISI, but requires extra degrees-of-freedom to suppress the nulling induced ICI. The equalization performance can be further improved if one resorts to other nonlinear method such as the joint maximum likelihood (ML) search [61]. However, the additional performance gain offered by this nonlinear equalizers is obtained at the expense of higher computational complexity. In the next subsection, we propose a GSC-based equalizer for joint ISI and ICI suppression. This method is linear in nature and, by incorporating a PA implementation, the main computations required boil down to inverting an $L \times L$ matrix. Moreover, as one will see, numerical simulations show that it can perform better than the above scheme involving ISI nulling followed by ZF or MMSE ICI suppression.

2.4.3 GSC-Based Equalizer

To exploit the extra degrees-of-freedom in model (2.8) for interference suppression, a commonly used approach is via constrained optimization [18], [56], [72]. Specifically, we will seek for a linear weighting matrix $\mathbf{W} \in \mathbb{C}^{MQ\times Q}$ which satisfies

$$\mathbf{W}^H\mathbf{D} = \mathbf{D}^H\mathbf{D}, \tag{2.14}$$

and minimizes the average output interference-plus-noise power, i.e., $E\left\{\left\|\mathbf{W}^H(\mathbf{H}_I\mathbf{e}(k) + \mathbf{v}(k))\right\|^2\right\}$. With constraint (2.14), the optimal weight $\mathbf{W}$ will linearly combine the desired signal in the maximal-ratio sense (channel matched filtering), and suppress ISI/ICI via minimizing the total output interference energy; the equalized signal will then be approximate to

$$\mathbf{W}^H\mathbf{z}(k) \approx \mathbf{D}^H\mathbf{D}\mathbf{s}(k), \tag{2.15}$$

which can further low-complexity tone-by-tone signal separation. To solve for the optimal $\mathbf{W}$, an efficient approach is to transform the constrained optimization problem into an unconstrained one via the GSC principle [23], [74]. This amounts to decomposing the weighting matrix into

$$\mathbf{W} = \mathbf{D} - \mathbf{B}\mathbf{U}, \tag{2.16}$$

where the signal signature $\mathbf{D}$ represents the non-adaptive component for verifying constraint

(2.14), $\mathbf{B} \in \mathbb{C}^{MQ \times (M-1)Q}$ is the signal blocking matrix with $\mathbf{B}^H \mathbf{D} = \mathbf{O}$, and $\mathbf{U} \in \mathbb{C}^{(M-1)Q \times Q}$ forms the remaining free parameters to be determined. A schematic description of the GSC is shown in Figure 2.1(a).
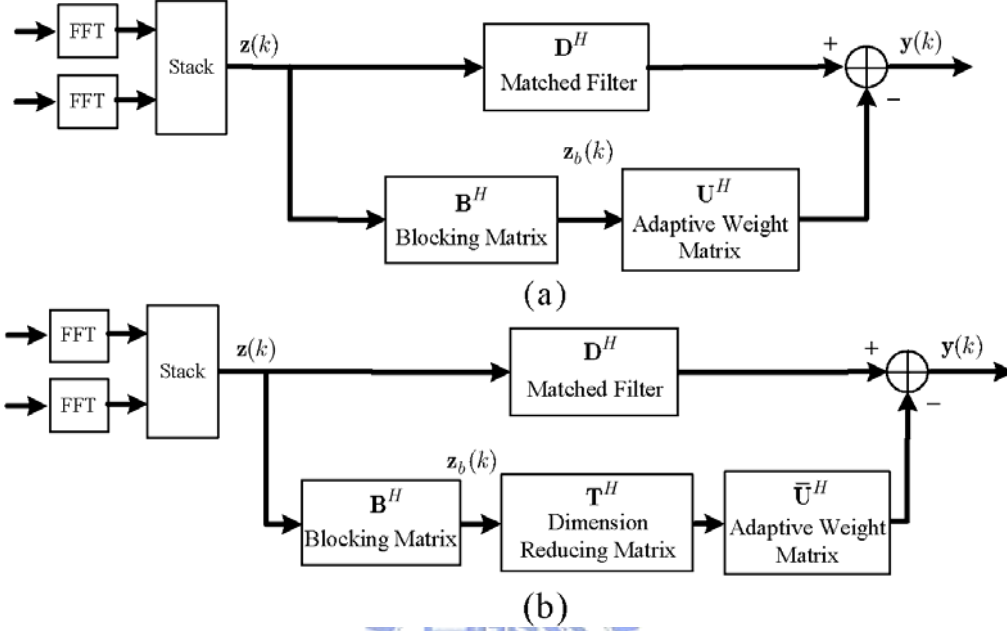
Figure 2.1. Illustration of (a) fully adaptive GSC equalizer (b) partially adaptive GSC equalizer.

With (2.16) and (2.8), the equalized output becomes

$$\mathbf{W}^H \mathbf{z}(k) = \mathbf{z}_d(k) - \mathbf{U}^H \mathbf{z}_b(k), \tag{2.17}$$

where

$$\mathbf{z}_d(k) := \mathbf{D}^H \mathbf{z}(k) = \mathbf{D}^H \mathbf{D}\mathbf{s}(k) + \mathbf{D}^H \mathbf{H}_I \mathbf{e}(k) + \mathbf{D}^H \mathbf{v}(k), \tag{2.18}$$

and

$$\mathbf{z}_b(k) := \mathbf{B}^H \mathbf{z}(k) = \mathbf{B}^H \mathbf{H}_I \mathbf{e}(k) + \mathbf{B}^H \mathbf{v}(k). \tag{2.19}$$

Since the matched filtered branch $\mathbf{z}_d(k)$ is contaminated by

$$\mathbf{i}(k) := \mathbf{D}^H \mathbf{H}_I \mathbf{e}(k) + \mathbf{D}^H \mathbf{v}(k), \tag{2.20}$$

to effectively suppress interference equation (2.17) suggests that the matrix $\mathbf{U}$ should be chosen to render $\mathbf{U}^H \mathbf{z}_b(k)$ as close to $\mathbf{i}(k)$ as possible; more precisely, we can determine $\mathbf{U}$ by minimizing the following cost function

$$J := E\left\{ \left\| \mathbf{i}(k) - \mathbf{U}^H \mathbf{z}_b(k) \right\|^2 \right\}. \tag{2.21}$$

16

With (2.19) and (2.20) and by following the standard procedures [9], the optimal weight, denoted by $\mathbf{U}_{opt}$, can be shown to satisfy the linear equation

$$\left[\mathbf{B}^H \mathbf{R}_{in} \mathbf{B}\right] \mathbf{U} = \mathbf{B}^H \mathbf{R}_{in} \mathbf{D},$$ (2.22)

where

$$\mathbf{R}_{in} := 2\mathbf{H}_I \mathbf{H}_I^H + \sigma_v^2 \mathbf{I}_{MQ} \in \mathbb{C}^{MQ \times MQ}.$$ (2.23)

With (2.22), we have

$$\mathbf{U}_{opt} = \left(\mathbf{B}^H \mathbf{R}_{in} \mathbf{B}\right)^{-1} \mathbf{B}^H \mathbf{R}_{in} \mathbf{D},$$ (2.24)

and the optimal GSC weight is thus

$$\mathbf{W} = \mathbf{D} - \mathbf{B}\mathbf{U}_{opt} = \left(\mathbf{I}_{MQ} - \mathbf{B}\left(\mathbf{B}^H \mathbf{R}_{in} \mathbf{B}\right)^{-1} \mathbf{B}^H \mathbf{R}_{in}\right)\mathbf{D},$$ (2.25)

which is commonly referred to as the fully adaptive (FA) GSC equalizer.

2.5 Reduced-Complexity Equalizer Implementation

To obtain the GSC weight matrix $\mathbf{W}$ in (2.25), one has to compute the blocking matrix $\mathbf{B}$ and to invert the $(M-1)Q \times (M-1)Q$ matrix $\mathbf{B}^H \mathbf{R}_{in} \mathbf{B}$. This section proposes a reduced-complexity implementation of the proposed GSC equalizer. In particular, by respectively exploiting the structures of the signal and ISI/ICI signatures $\mathbf{D}$ and $\mathbf{H}_I$, we will show a very simple approach to computing the blocking matrix $\mathbf{B}$ and deriving a low-complexity PA implementation of $\mathbf{W}$.

2.5.1 Computation of Blocking Matrix B

The columns of the blocking matrix $\mathbf{B}$ form an orthonormal basis for the left null space of the signal signature $\mathbf{D}$. One typical approach for obtaining $\mathbf{B}$ is to perform the SVD on $\mathbf{D}$. Since $\mathbf{D}$ is $MQ \times Q$, the computational burden of the SVD based approach could be prohibitive for a large Q. From (2.9), the matrix $\mathbf{D}$ is a concatenation of M diagonal matrices. There is a very simple way of computing $\mathbf{B}$ if one can further exploit the block diagonal and sparse structures of $\mathbf{D}$. To see this, we denote by $\mathbf{d}_j$ the jth column of $\mathbf{D}$, $1 \leq j \leq Q$. From (2.9), it can be easily

checked that $\mathbf{d}_{j_1}$ and $\mathbf{d}_{j_2}$ are mutually orthogonal whenever $j_1 \neq j_2$. This observation suggests that we can choose $\mathbf{B}$ to be of the form

$$\mathbf{B} = \left[\mathbf{B}^{(1)} \cdots \mathbf{B}^{(M-1)} \right], \tag{2.26}$$

with $\mathbf{B}^{(m')} = \left[\mathbf{B}_1^{(m')T} \quad \cdots \quad \mathbf{B}_M^{(m')T} \right]^T \in \mathbb{C}^{MQ \times Q}$, $1 \leq m' \leq M-1$, where $\mathbf{B}_m^{(m')}$, $1 \leq m \leq M$, is diagonal. Denote by $\mathbf{b}_j^{(m')}$ the jth column of $\mathbf{B}^{(m')}$, $1 \leq j \leq Q$. Then the columns of $\mathbf{B}^{(m')}$ are mutually orthogonal. Moreover, it can be directly checked that $\mathbf{b}_j^{(m_1)H}\mathbf{b}_k^{(m_2)} = 0$, for $m_1 \neq m_2$, and $\mathbf{b}_j^{(m')H}\mathbf{d}_k = 0$ when $j \neq k$. As a result, the $(M-1)Q$ columns of the matrix $\mathbf{B}$ constructed as in (2.26) form an orthonormal basis for the left null space of $\mathbf{D}$ if we choose $\mathbf{b}_j^{(m')}$, with $\left\| \mathbf{b}_j^{(m')} \right\| = 1$, to satisfy the orthogonality condition

$$\begin{cases} \mathbf{b}_j^{(m')H}\mathbf{d}_j = 0 \\ \mathbf{b}_j^{(m_1)H}\mathbf{b}_j^{(m_2)} = 0 \end{cases}, \quad 1 \leq m' \leq M-1, \ 1 \leq m_1 \neq m_2 \leq M-1, \ 1 \leq j \leq Q. \tag{2.27}$$

We note that, for a fixed j, only the $((m-1)Q+j)$-th entries, $1 \leq m \leq M$, in $\mathbf{b}_j^{(m')}$ (and hence $\mathbf{d}_j$) are nonzero. This implies that (2.27) can be simplified as

$$\begin{cases} \sum\limits_{m=1}^{M} \mathbf{b}_j^{(m')*}\left((m-1)Q+j\right)\mathbf{d}_j\left((m-1)Q+j\right) = 0 \\ \sum\limits_{m=1}^{M} \mathbf{b}_j^{(m_1)*}\left((m-1)Q+j\right)\mathbf{b}_j^{(m_2)}\left((m-1)Q+j\right) = 0 \end{cases}, \tag{2.28}$$

with $\sum_{m=1}^{M}\left| \mathbf{b}_j^{(m')}\left((m-1)Q+j\right) \right|^2 = 1$, $1 \leq m' \leq M-1$, $1 \leq m_1 \neq m_2 \leq M-1$, $1 \leq j \leq Q$, where $\mathbf{b}_j^{(m')}(n)$ denotes the nth entry of $\mathbf{b}_j^{(m')}$. Hence, by exploiting the structure of the signal signature matrix $\mathbf{D}$, we can develop $\mathbf{B}$ as in (2.26). Instead of performing SVD on $\mathbf{D}$, all we have to do is to pick up the $M-1$ M-dimensional solutions (i.e., $\{\mathbf{b}_j^{(m')}(j),...,\mathbf{b}_j^{(m')}\left((M-1)Q+j\right)\}$, $1 \leq m' \leq M-1$) based on the orthogonality condition in (2.28) for each $1 \leq j \leq Q$.

2.5.2 PA Implementation

The computation of the adaptive weight matrix $\mathbf{U}_{opt}$ involves the inversion of the $(M-1)Q \times (M-1)Q$ matrix $\mathbf{B}^H\mathbf{R}_{in}\mathbf{B}$. This would lead to a high computational load and poor convergence for real-time implementation whenever Q is large. To resolve this problem, PA GSC

can be incorporated to reduce the size of the adaptive weight matrix as shown as follows. Since $\mathbf{H}_I$ is of rank L (see Section 2.4.1) and $\mathbf{B}$ has orthonormal columns, the post-blocking ISI/ICI signature matrix $\mathbf{B}^H\mathbf{H}_I \in \mathbb{C}^{(M-1)Q \times (M-1)Q}$ is at most of rank L only. This implies that there are certain redundant components in $\mathbf{B}^H\mathbf{H}_I$ that are irrelevant to the post-blocking ISI/ICI information. As a result, it would be plausible to seek for a more compact representation, e.g., through an observation space of dimension L, of the post-blocking ISI/ICI signal. Indeed, if $\mathbf{T} \in \mathbb{C}^{(M-1)Q \times L}$ is a basis for the range space of $\mathbf{B}^H\mathbf{H}_I$ [22], [25], then a "refined" post-blocking ISI/ICI signal can be obtained by passing the blocked output $\mathbf{z}_b(k)$ through $\mathbf{T}^H$, leading to the following L-dimensional data vector $\overline{\mathbf{z}}_b(k) = \mathbf{T}^H\mathbf{B}^H\mathbf{H}_I\mathbf{e}(k) + \mathbf{T}^H\mathbf{B}^H\mathbf{v}(k)$. The reduction in signal dimension via the matrix filtering $\mathbf{T}^H$ can extract the minimal possible information essential for describing the L-dimensional post-blocking interference subspace. The major advantage of such a data refinement is to allow for the use of a reduced size adaptive weight matrix. In fact, based on $\overline{\mathbf{z}}_b(k)$, it only requires to choose an $L \times Q$ weight matrix $\overline{\mathbf{U}}$ to minimize the cost function

$$\min_{\overline{\mathbf{U}}} E\left\{\left\|\mathbf{i}(k) - \overline{\mathbf{U}}^H\mathbf{T}^H\mathbf{B}^H\mathbf{z}(k)\right\|^2\right\},\tag{2.29}$$

which yields the solution

$$\overline{\mathbf{U}} = \left(\mathbf{T}^H\mathbf{B}^H\mathbf{R}_{in}\mathbf{B}\mathbf{T}\right)^{-1}\mathbf{T}^H\mathbf{B}^H\mathbf{R}_{in}\mathbf{D}.\tag{2.30}$$

Based on (2.30), the optimal $MQ \times Q$ GSC weight matrix is thus

$$\overline{\mathbf{W}} = \mathbf{D} - \mathbf{B}\mathbf{T}\overline{\mathbf{U}} = \left[\mathbf{I}_{MQ} - \mathbf{B}\mathbf{T}\left(\mathbf{T}^H\mathbf{B}^H\mathbf{R}_{in}\mathbf{B}\mathbf{T}\right)^{-1}\mathbf{T}^H\mathbf{B}^H\mathbf{R}_{in}\right]\mathbf{D}.\tag{2.31}$$

Compared with (2.25), in which an $(M-1)Q \times (M-1)Q$ matrix inversion is needed, the weight matrix defined in (2.31) involves only an $L \times L$ matrix inversion. This can significantly reduce the computational complexity since, in practice, the channel order L is quite small as compared with the number of subcarriers Q. The $\overline{\mathbf{W}}$ in (2.31) is often referred to as the PA implementation of the GSC weight (see Figure 2.1(b) for a schematic description).

The remaining task is to find the dimension reducing matrix $\mathbf{T}$, whose columns consist of a basis for the range space of $\mathbf{B}^H\mathbf{H}_I$, namely, the interference subspace. The rest of this section will

show that there is a simple closed-form candidate for such a $\mathbf{T}$. To proceed, the term $\mathbf{B}^H \mathbf{H}_I$ is first expanded into

$$\mathbf{B}^H \mathbf{H}_I = \begin{bmatrix} \sum_{m=1}^{M} \mathbf{B}_m^{(1)H} \mathbf{F} \mathbf{H}_1^{(m)} \mathbf{F}^{-1} \\ \vdots \\ \sum_{m=1}^{M} \mathbf{B}_m^{(M-1)H} \mathbf{F} \mathbf{H}_1^{(m)} \mathbf{F}^{-1} \end{bmatrix}, \tag{2.32}$$

where we use (2.8) and (2.26). Define $\mathbf{J}_l \in \mathbb{C}^{Q \times Q}$, $1 \le l \le L$, to be the matrix with zero entries except that those on the $(Q - l + 1)$-th upper diagonal are equal to one. Then we can express $\mathbf{H}_1^{(m)} = \sum_{l=1}^{L} h^{(m)}(l) \mathbf{J}_l$ for $1 \le m \le M$, such that

$$\mathbf{B}^H \mathbf{H}_I = \begin{bmatrix} \hat{\mathbf{f}}_1 \cdots & \hat{\mathbf{f}}_L \end{bmatrix} \begin{bmatrix} \mathbf{f}_N^H \\ \vdots \\ \mathbf{f}_{N-L+1}^H \end{bmatrix} = \mathbf{T}_b \mathbf{F}_L, \tag{2.33}$$

where

$$\hat{\mathbf{f}}_l = \left[\left(\sum_{m=1}^{M} \sum_{j=1}^{L} \left(h^{(m)}(j) \mathbf{B}_m^{(1)H} \mathbf{f}_{j-l+1} \right)^T \right), \dots, \left(\sum_{m=1}^{M} \sum_{j=1}^{L} \left(h^{(m)}(j) \mathbf{B}_m^{(M-1)H} \mathbf{f}_{j-l+1} \right)^T \right) \right]^T \in \mathbb{C}^{MQ \times 1}, \tag{2.34}$$

$\mathbf{f}_q$, $1 \le q \le Q$, are the columns of the FFT matrix $\mathbf{F}$, $\mathbf{T}_b = \begin{bmatrix} \hat{\mathbf{f}}_1 \cdots \hat{\mathbf{f}}_L \end{bmatrix} \in \mathbb{C}^{MQ \times L}$, and $\mathbf{F}_L = \begin{bmatrix} \mathbf{f}_Q \cdots \mathbf{f}_{Q-L+1} \end{bmatrix}^H \in \mathbb{C}^{L \times Q}$. Based on (2.33), $\mathbf{T}_b$ should form a basis for the range space of $\mathbf{B}^H \mathbf{H}_{ISI}$. As a result, a natural and immediate choice for the dimension-reducing matrix $\mathbf{T}$ is $\mathbf{T} = \mathbf{T}_b$.

2.5.3 Computational Complexity

This subsection compares the computational complexity (measured in terms of the number of complex multiplications (CM)) of the FA GSC-based equalizer, PA GSC-based equalizer, and MMSE receiver. To obtain the two GSC-based solutions, one shall first compute the matrix $\mathbf{R}_{in} \in \mathbb{C}^{MQ \times MQ}$ (cf. (2.24) and (2.30)). Based on (2.23) and by definition of $\mathbf{H}_I$ in (2.8), we can express $\mathbf{R}_{in}$ as

$$\mathbf{R}_{in} = 2\mathbf{F}_M \mathbf{H}_1 \mathbf{H}_1^H \mathbf{F}_M^H + \sigma_n^2 \mathbf{I}_{MQ}, \tag{2.35}$$

where $\mathbf{H}_1$ is defined in (2.10). Equation (2.35) implies that the computation of $\mathbf{R}_{in}$ can be implemented using FFT. By further exploiting the sparse structure of $\mathbf{H}_1$ (recalling that the nonzero entries in each submatrix of $\mathbf{H}_1$ all cluster in the upper right corner), the number of flops for computing $\mathbf{R}_{in}$ is determined as $2M^2L^2\log_2^Q + M^2L^3$. Based on (2.25) and (2.31), the total approximate flop counts required to compute FA and PA GSC-based solutions are respectively summarized as below:

$$CM_{FA} = \frac{1}{3}Q^3 + \left(4M^2 + 2M + \frac{5}{2}\right)Q^2 + (M^3L^2 + 2M^2L\log_2^Q)Q, \qquad (2.36)$$

and

$$CM_{PA} = (4M^2 + 2M)Q^2 + \left(\left(M^3 + \frac{5}{2}\right)L^2 + \left(2M^2 + \frac{L}{2} + 1\right)L\log_2^Q\right)Q. \qquad (2.37)$$

For comparison, the total flop count required to obtain the MMSE solution in (2.13) is given as

$$CM_{MMSE} = \frac{1}{3}(MQ - L)^3 + 2Q(MQ - L)^2 + 2MQ(MQ - L). \qquad (2.38)$$

From (2.36), (2.37) and (2.38), it can be seen that the PA scheme can save about one order of computational load over the FA GSC and MMSE solutions.

2.6 Downlink GSC-Based Pre-Equalization

To relieve the MU complexity, one can resort to the pre-equalization technique. Based on the block system model, this section derives a GSC-based pre-equalizer, which serves as the downlink counterpart of the previous work.

2.6.1 Downlink Signal Model

In the downlink stage, the OFDM symbol $\mathbf{s}(k)$ is first weighted by $\mathbf{G}^{(1)},\ldots,\mathbf{G}^{(M)}$, respectively for the M transmit antenna branches, and is then transmitted though the ISI channel. Similar to the synchronous block signal model in (2.8), the $Q \times 1$ frequency-domain noise-free received data at the MU is expressed as

$$\mathbf{z}_d(k) = \begin{bmatrix} \mathbf{D}^{(1)} & \cdots \mathbf{D}^{(M)} \end{bmatrix} \begin{bmatrix} \mathbf{G}^{(1)} \\ \vdots \\ \mathbf{G}^{(M)} \end{bmatrix} \mathbf{s}(k) + \begin{bmatrix} \mathbf{FH}_1^{(1)}\mathbf{F}^{-1} \cdots \mathbf{FH}_1^{(M)}\mathbf{F}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{G}^{(1)} \\ \vdots \\ \mathbf{G}^{(M)} \end{bmatrix} \mathbf{e}(k) = \mathbf{D}_d\mathbf{G}\mathbf{s}(k) + \mathbf{H}_{d,I}\mathbf{G}\mathbf{e}(k),$$

(2.39)

in which $\mathbf{D}_d = \begin{bmatrix} \mathbf{D}^{(1)} & \cdots \mathbf{D}^{(M)} \end{bmatrix} \in \mathbb{C}^{Q \times MQ}$, $\mathbf{G} = \begin{bmatrix} \mathbf{G}^{(1)T} & \cdots \mathbf{G}^{(M)T} \end{bmatrix}^T \in \mathbb{C}^{MQ \times Q}$, and $\mathbf{H}_{d,I} = \begin{bmatrix} \mathbf{FH}_1^{(1)}\mathbf{F}^{-1} \cdots \mathbf{FH}_1^{(M)}\mathbf{F}^{-1} \end{bmatrix} \in \mathbb{C}^{Q \times MQ}$.

2.6.2 GSC-Based Pre-Equalization Weight Matrix

Based on (2.39), this subsection demonstrates how to appropriately choose $\mathbf{G}$ to pre-equalize the ISI/ICI effect characterized by $\mathbf{H}_{d,I}$. To see this, let us take the Hermitian transpose on both side of (2.39) to obtain

$$\mathbf{z}_d^H(k) = \mathbf{s}^H(k)\mathbf{G}^H\mathbf{D}_d^H + \mathbf{e}^H(k)\mathbf{G}^H\mathbf{H}_{d,I}^H. \qquad (2.40)$$

By following the GSC principle, we divide the pre-equalization weight matrix $\mathbf{G}$ into two orthogonal parts, $\mathbf{G} = \mathbf{D}_d^H - \mathbf{B}_d\mathbf{U}_d$, where $\mathbf{B}_d = \begin{bmatrix} \mathbf{B}_d^{(1)}, \ldots, \mathbf{B}_d^{(M-1)} \end{bmatrix} \in \mathbb{C}^{MQ \times (M-1)Q}$, with $\mathbf{B}_d^{(m')} = \begin{bmatrix} \mathbf{B}_{d,1}^{(m')T}, \ldots, \mathbf{B}_{d,M}^{(m')T} \end{bmatrix}^T \in \mathbb{C}^{MQ \times Q}$, $1 \le m' \le M-1$, is a blocking matrix whose columns are chosen from an orthonormal basis for the left null space of $\mathbf{D}_d^H$ and $\mathbf{U}_d$ is the adaptive weight matrix. The schematic description of $\mathbf{G}$ is shown in Figure 2.2(a).
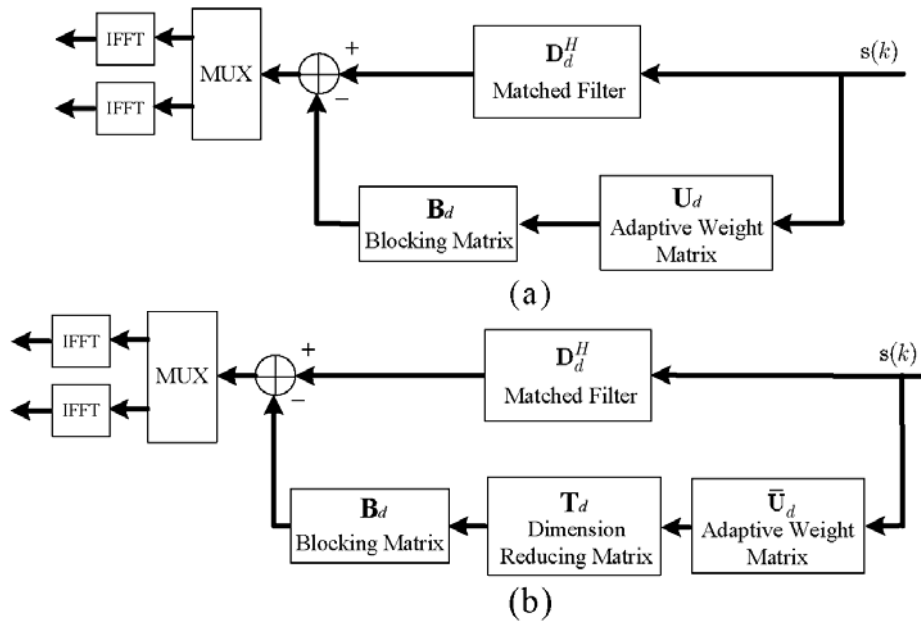
Figure 2.2. Illustration of (a) fully adaptive GSC pre-equalizer (b) partially adaptive GSC

pre-equalizer.

Substituting $\mathbf{G}$ in (2.40), we obtain

$$\mathbf{z}_d^H(k) = \mathbf{s}^H(k)\mathbf{D}_d\mathbf{D}_d^H + \mathbf{e}^H(k)\left[\mathbf{D}_d - \mathbf{U}_d^H\mathbf{B}_d^H\right]\mathbf{H}_{d,I}^H. \tag{2.41}$$

The weight matrix for the downlink stage is determined by minimizing the ISI/ICI power in the received signal at the MU, that is,

$$\min_{\mathbf{U}_d} E\left\{\left\|\mathbf{H}_{d,ISI}\left(\mathbf{D}_d^H - \mathbf{B}_d\mathbf{U}_d\right)\mathbf{e}(k)\right\|^2\right\}. \tag{2.42}$$

After some manipulations, it can be shown that $\mathbf{U}_d$ must satisfy the following linear equation

$$\mathbf{B}_d^H\mathbf{R}_{d,i}\mathbf{B}_d\mathbf{U}_d = \mathbf{B}_d^H\mathbf{R}_{d,i}\mathbf{D}_d^H, \tag{2.43}$$

where $\mathbf{R}_{d,i} = \mathbf{H}_{d,I}^H\mathbf{R}_e\mathbf{H}_{d,I}$. It is easy to check that $\mathbf{B}_d^H\mathbf{R}_{d,i}\mathbf{B}_d \in \mathbb{C}^{(M-1)Q\times(M-1)Q}$ is at most of rank L only (see (2.4) and Section 2.5.2), and the optimal solution cannot be determined from (2.43). A suboptimal solution can instead be obtained based on the diagonal loading technique [11], that is, adding a real value α to the diagonal entries of $\mathbf{B}_d^H\mathbf{R}_{d,i}\mathbf{B}_d$ to improve the matrix condition. In this way, the suboptimal solution of $\mathbf{U}_d$ is immediately computed as

$$\mathbf{U}_d = (\mathbf{B}_d^H\mathbf{R}_{d,in}\mathbf{B}_d)^{-1}\mathbf{B}_d^H\mathbf{R}_{d,i}\mathbf{D}_d^H \in \mathbb{C}^{(M-1)Q\times Q}, \tag{2.44}$$

where $\mathbf{R}_{d,in} = \mathbf{R}_{d,i} + \alpha\mathbf{I}_{MN}$. The FA pre-equalization weight matrix $\mathbf{G} \in \mathbb{C}^{MQ\times Q}$ is thus

$$\mathbf{G} = \left[\mathbf{I}_{MQ} - \mathbf{B}_d(\mathbf{B}_d^H\mathbf{R}_{d,in}\mathbf{B}_d)^{-1}\mathbf{B}_d^H\mathbf{R}_{d,i}\right]\mathbf{D}_d^H. \tag{2.45}$$

If we choose α too large, the ISI/ICI cancellation capability of the pre-equalizer drops, and if too small, the matrix $\mathbf{R}_{d,in}$ remains ill-conditioned [11]. That is, the choice of α is a trade-off between the interference cancellation capability and accuracy of the solution. Referring to the uplink solution (2.25), the diagonal loading value α could be chosen as σ_v^2 (noise power), which can be estimated in the uplink stage

It is necessary to normalize the norm of each column of $\mathbf{G}$ to be 1 in order to keep the instantaneous transmitted power constant. It will be shown in the simulation results that the

pre-equalization performance of the suboptimal weight matrix $\mathbf{G}$ is almost identical to that of the CP-based pre-maximal ratio combining (pre-MRC) scheme [16].

2.6.3 Reduced-Complexity Implementation

To obtain $\mathbf{G}$, we need to compute $\mathbf{B}_d$ and invert the $(M-1)Q \times (M-1)Q$ matrix $\mathbf{B}_d^H \mathbf{R}_{d,in} \mathbf{B}_d$. The proposed low-complexity implementation for the uplink stage suggests an immediate solution for the corresponding downlink part. Specifically, since $\mathbf{D}_d = \left[\mathbf{D}^{(1)} \cdots \mathbf{D}^{(M)} \right]$ is the horizontal concatenation of M diagonal matrices, the matrix $\mathbf{B}_d$, which is the solution to $\mathbf{D}_d \mathbf{B}_d = \mathbf{O}$, can be computed in an analogue way as in Section 2.5.1 without performing an SVD of $\mathbf{D}_d$. On the other hand, to avoid the direct computation of $\left(\mathbf{B}_d^H \mathbf{R}_{d,in} \mathbf{B}_d \right)^{-1}$, we can resort to the PA scheme as in the uplink stage. From (2.41), $\mathbf{B}_d^H \mathbf{H}_{d,I}^H$ appears to be at most of rank L only (see (2.4) and Section 2.5.2), and we can find a dimension reducing matrix $\mathbf{T}_d$ inserted in front of $\mathbf{B}_d$ to reduce the pre-processing dimension. By following the same procedure as in Section 2.5.2, we find that $\mathbf{T}_d$ can be chosen as $\mathbf{T}_d = \left[\mathbf{f}_{d,1} \cdots \mathbf{f}_{d,L} \right] \in \mathbb{C}^{(M-1)Q \times L}$, where

$$\mathbf{f}_{d,l} = \left[\left(\sum_{m=1}^{M} \sum_{j=1}^{L} \left(h^{(m)*}(j) \mathbf{B}_{d,m}^{(1)H} \mathbf{f}_{N+l-j} \right) \right)^T, \cdots, \left(\sum_{m=1}^{M} \sum_{j=1}^{L} \left(h^{(m)*}(j) \mathbf{B}_{d,m}^{(M-1)H} \mathbf{f}_{N+l-j} \right) \right)^T \right]^T \in \mathbb{C}^{MQ \times 1}.$$

Substituting $\mathbf{B}_d \mathbf{T}_d \overline{\mathbf{U}}_d$ for $\mathbf{B}_d \overline{\mathbf{U}}_d$ in (2.42) and after some manipulations, the reduced size weight matrix $\overline{\mathbf{U}}_d$ must satisfy the following linear equation

$$\mathbf{T}_d^H \mathbf{B}_d^H \mathbf{R}_{d,i} \mathbf{B}_d \mathbf{T}_d \overline{\mathbf{U}}_d = \mathbf{T}_d^H \mathbf{B}_d^H \mathbf{R}_{d,i} \mathbf{D}_d^H. \tag{2.46}$$

Similarly, $\overline{\mathbf{U}}_d$ can be obtained as

$$\overline{\mathbf{U}}_d = (\mathbf{T}_d^H \mathbf{B}_d^H \mathbf{R}_{d,i} \mathbf{B}_d \mathbf{T}_d)^{-1} \mathbf{T}_d^H \mathbf{B}_d^H \mathbf{R}_{d,i} \mathbf{D}_d^H. \tag{2.47}$$

Accordingly, the PA pre-equalization weight matrix $\overline{\mathbf{G}} = \mathbf{D}_d^H - \mathbf{B}_d \mathbf{T}_d \overline{\mathbf{U}}_d \in C^{MQ \times Q}$ is obtained as

$$\overline{\mathbf{G}} = \left[\mathbf{I}_{MQ} - \mathbf{B}_d \mathbf{T}_d \left(\mathbf{T}_d^H \mathbf{B}_d^H \mathbf{R}_{d,i} \mathbf{B}_d \mathbf{T}_d \right)^{-1} \mathbf{T}_d^H \mathbf{B}_d^H \mathbf{R}_{d,i} \right] \mathbf{D}_d^H. \tag{2.48}$$

The schematic description is shown in Figure 2.2(b). To maintain equal transmit power on each subcarrier, the norm of each column of $\overline{\mathbf{G}}$ should be normalized to be 1.

24

2.7 Output SINR Performance

This section derives the output signal to interference-plus-noise ratio (SINR) on each subcarrier at the AP end. In the uplink stage, from (2.25), the $MQ \times 1$ equalization weight vector associated with the qth subcarrier is

$$\mathbf{w}_q = \mathbf{d}_q - \mathbf{B}\left(\mathbf{B}^H \mathbf{R}_{in} \mathbf{B}\right)^{-1} \mathbf{B}^H \mathbf{R}_{in} \mathbf{d}_q. \tag{2.49}$$

Then, the equalized output at the qth subcarrier is obtained as

$$\mathbf{y}_q = \mathbf{w}_q^H \mathbf{z}(k) = \mathbf{w}_q^H \mathbf{D}\mathbf{s}(k) + \mathbf{w}_q^H \mathbf{H}_I \mathbf{e}(k) + \mathbf{w}_q^H \mathbf{v}(k) = \mathbf{d}_q^H \mathbf{d}_q s_q(k) + \mathbf{w}_q^H \mathbf{H}_I \mathbf{e}(k) + \mathbf{w}_q^H \mathbf{v}(k), \tag{2.50}$$

where we have used $\mathbf{d}_{j_1}^H \mathbf{d}_{j_2} = 0$ when $j_1 \neq j_2$ (see Section 2.5.1) and $\mathbf{B}^H \mathbf{D} = \mathbf{O}$ (blocking matrix removes the desired signal). The output SINR of the qth subcarrier at the AP end is thus given by

$$SINR_{AP,q} = \frac{(\mathbf{d}_q^H \mathbf{d}_q)^2}{\mathbf{w}_q^H \mathbf{R}_{in} \mathbf{w}_q}. \tag{2.51}$$

Assuming that the GSC equalizer effectively suppress ISI/ICI, i.e., $\mathbf{W}^H \mathbf{H}_I \approx \mathbf{O}$, then by definition of $\mathbf{R}_{in}$ in (2.23), we have

$$SINR_{AP,q} \approx \frac{(\mathbf{d}_q^H \mathbf{d}_q)^2}{\sigma_v^2 \mathbf{w}_q^H \mathbf{w}_q} = \frac{\mathbf{d}_q^H \mathbf{d}_q}{\sigma_v^2 \left(1 + \dfrac{\mathbf{d}_q^H \Lambda_u \mathbf{d}_q}{\mathbf{d}_q^H \mathbf{d}_q}\right)}, \tag{2.52}$$

where the last equality follows from (2.49) with $\Lambda_u = \mathbf{R}_{in} \mathbf{B}\left(\mathbf{B}^H \mathbf{R}_{in} \mathbf{B}\right)^{-2} \mathbf{B}^H \mathbf{R}_{in} \in \mathbb{C}^{MQ \times MQ}$. Since $\mathbf{d}_q^H \mathbf{d}_q = \sum_{m=1}^M \left|\mathbf{d}_q\left((m-1)Q + q\right)\right|^2$, Equation (2.52) roughly confirms that an M-fold diversity gain is achieved. Also, the result in (2.52) tends to indicate that the GSC equalizer first tackles ISI/ICI to yield $\tilde{\mathbf{z}}_q(k) \approx \mathbf{d}_q s_q(k) + \tilde{\mathbf{v}}_q(k)$, where $\tilde{\mathbf{v}}_q(k)$ is the effective noise with variance

$$\tilde{\sigma}_v^2 = \sigma_v^2 (1 + \mathbf{d}_q^H \Lambda_u \mathbf{d}_q / \mathbf{d}_q^H \mathbf{d}_q), \tag{2.53}$$

and then perform a maximum-ratio combining (MRC) over all receive antenna branches to maximize the output SNR, leading to

$$\tilde{\mathbf{y}}_q(k) = \mathbf{d}_q^H \tilde{\mathbf{z}}_q(k) \approx \mathbf{d}_q^H \mathbf{d}_q s_q(k) + \mathbf{d}_q^H \tilde{\mathbf{v}}(k). \tag{2.54}$$

We note that, essentially, the same results can be obtained for the downlink stage.

2.8 Impacts of Channel Correlation

Toward an analytically tractable approach to investigating the channel correlation effect, we will focus on the $M{=}2$ case. We consider the two correlated channels modeled by the following equations [40]

$$\begin{bmatrix} \bar{h}^{(1)}(l) \\ \bar{h}^{(2)}(l) \end{bmatrix} = \begin{bmatrix} 1 & \rho \\ \rho^* & 1 \end{bmatrix}^{1/2} \begin{bmatrix} h^{(1)}(l) \\ h^{(2)}(l) \end{bmatrix}, \tag{2.55}$$

where $h^{(1)}(l)$ and $h^{(2)}(l)$ are modeled as i.i.d. zero-mean Gaussian random variables with variance σ_l^2 and

$$\rho = \frac{E\left\{\bar{h}^{(1)}(l)\bar{h}^{(2)*}(l)\right\}}{\sqrt{E\left\{\left|\bar{h}^{(1)}(l)\right|^2\right\}E\left\{\left|\bar{h}^{(2)}(l)\right|^2\right\}}}.$$

Based on the definition of $\mathbf{D}$ in (2.9), the associated frequency-domain signal signature matrix is

$$\tilde{\mathbf{D}} = \Upsilon^{1/2}\mathbf{D} \in C^{2Q \times Q}, \text{ with } \Upsilon = \begin{bmatrix} \mathbf{I}_Q & \rho\mathbf{I}_Q \\ \rho^*\mathbf{I}_Q & \mathbf{I}_Q \end{bmatrix}. \tag{2.56}$$

From (2.56), the channel frequency responses of the qth subcarriers associated with the two channel branches can be expressed as

$$\begin{bmatrix} \tilde{\mathbf{d}}_q(q) \\ \tilde{\mathbf{d}}_q(Q+q) \end{bmatrix} = \begin{bmatrix} 1 & \rho \\ \rho^* & 1 \end{bmatrix}^{1/2} \begin{bmatrix} \mathbf{d}_q(q) \\ \mathbf{d}_q(Q+q) \end{bmatrix},$$

in which $\tilde{\mathbf{d}}_q(i)$ is the ith entry of the qth column of $\tilde{\mathbf{D}}$. Since the channel taps $h^{(1)}(l)$ and $h^{(2)}(l)$ are i.i.d. zero-mean Gaussian random variables with variance σ_l^2, $0 \le l \le L$, so are the corresponding channel frequency responses $\mathbf{d}_q(q)$ and $\mathbf{d}_q(Q+q)$, $1 \le q \le Q$, with variance $\sigma_d^2 = \sum_{l=0}^L \sigma_l^2$, such that

$$\begin{bmatrix} \tilde{\mathbf{d}}_q(q) \\ \tilde{\mathbf{d}}_q(Q+q) \end{bmatrix} \sim N_c(0, \sigma_d^2\Psi), \text{ with } \Psi = \begin{bmatrix} 1 & \rho \\ \rho^* & 1 \end{bmatrix}, \tag{2.57}$$

i.e., complex Gaussian distribution with mean $\mathbf{0}$ and covariance $\sigma_d^2 \Psi$. To further pin down the impacts of channel correlation on system performance, in the sequel, we will consider the effective system model (2.54). Assume that the source symbols are drawn from the QPSK constellation.

2.8.1 $0 \leq |\rho| < 1$ case

We note from (2.53) that the variance of the effective noise depends on $\mathbf{d}_q$ also. In fact, as Q is large, it can be shown that $\mathbf{d}_q^H \Lambda_u \mathbf{d}_q / \mathbf{d}_q^H \mathbf{d}_q \approx 0$ for a wide range of SNR, and hence $\tilde{\sigma}_v^2 \approx \sigma_v^2$. According to [32], the approximate BER of the qth subcarrier at the AP end, averaged over $\tilde{\mathbf{d}}_q(q)$ and $\tilde{\mathbf{d}}_q(Q+q)$, is upper-bounded by

$$P_{AP,q} = E\left\{ Q\left(\sqrt{\frac{2\left(\left|\tilde{\mathbf{d}}_q(q)\right|^2 + \left|\tilde{\mathbf{d}}_q(Q+q)\right|^2 \right)}{\sigma_v^2}} \right) \right\} \leq G_\rho \left(\frac{1}{\sigma_v^2} \right)^{-r_\rho}, \tag{2.58}$$

where $G_\rho = \prod_{m=1}^{r_\rho} \varepsilon_m^{-1}$, in which the integer r_ρ is the rank of Ψ, ε_m, $1 \leq m \leq r_\rho$, are the nonzero eigenvalues of $\sigma_d^2 \Psi$, and the quantity r_ρ is the available diversity gain. Given Ψ in (2.57), it is easy to show that $G_\rho = \left(\sigma_d^4 \left(1 - |\rho|^2\right) \right)^{-1}$ and $r_\rho = 2$. As a result, even though a two-fold diversity adavantage can be attained for $0 < |\rho| < 1$, the nonzero correlation index ρ will enlarge G_ρ, as compared with the $\rho = 0$ case. The increase in G_ρ tends to shift the BER curve to the right, incurring a larger error level.

2.8.2 $|\rho| = 1$ case

Now, we consider the extreme case of $|\rho| = 1$, i.e., $\rho = e^{j\theta}$, in which θ is a real constant. To evaluate the resultant BER, we shall first examine the overall system model given the complete channel dependency. Substituting $\mathbf{D}$ in (2.56) with its definition in (2.9), the correlated signal signature matrix becomes

$$\tilde{\mathbf{D}} = \begin{bmatrix} \mathbf{I}_N & e^{j\theta}\mathbf{I}_N \\ e^{-j\theta}\mathbf{I}_N & \mathbf{I}_N \end{bmatrix}^{1/2} \begin{bmatrix} \mathbf{D}^{(1)} \\ \mathbf{D}^{(2)} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_N \\ e^{-j\theta}\mathbf{I}_N \end{bmatrix} \mathbf{D}_c, \tag{2.59}$$

where $\mathbf{D}_c = \sqrt{2}\mathbf{D}^{(1)}/2 + e^{j\theta}\sqrt{2}\mathbf{D}^{(2)}/2$. Similarly, from (2.8) and (2.10), we have

$$\tilde{\mathbf{H}}_I = \begin{bmatrix} \mathbf{I}_N \\ e^{-j\theta}\mathbf{I}_N \end{bmatrix} \mathbf{F}\mathbf{H}_{1,c}\mathbf{F}^H, \tag{2.60}$$

where $\mathbf{H}_{1,c} = \sqrt{2}\mathbf{H}_1^{(1)}/2 + e^{j\theta}\sqrt{2}\mathbf{H}_1^{(2)}/2$. Equations (2.59) and (2.60) imply that the column spaces of the signal and interference signature matrices coincide. As a result, the blocking matrix $\overline{\mathbf{B}}$, which by definition must satisfy $\tilde{\mathbf{B}}^H\tilde{\mathbf{D}} = \mathbf{O}$, will also null the ISI/ICI component. The interference cancellation mechanism of the GSC solution thus breaks down in this case since ISI/ICI is no longer retained in the second branch. Substituting $\mathbf{R}_{in}$ in (2.23) into (2.25) and using $\tilde{\mathbf{B}}^H\tilde{\mathbf{H}}_I = \mathbf{O}$, the GSC solution is seen to reduce to $\mathbf{W} = \tilde{\mathbf{D}}$: the receiver only coherently combines the desired signals from the two receive antennas without ISI/ICI suppression. With $\mathbf{w}_q = \tilde{\mathbf{d}}_q$ in (2.50), the output data of the qth subcarrier can be immediately obtained as

$$\mathbf{y}_q(k) = \tilde{\mathbf{d}}_q^H\tilde{\mathbf{d}}_q s_q(k) + \tilde{\mathbf{d}}_q^H\tilde{\mathbf{H}}_I\mathbf{e}(k) + \tilde{\mathbf{d}}_q^H\mathbf{v}(k). \tag{2.61}$$

Assume that the effective noise term in (2.61) can be approximately modeled as a Gaussian random variable. Following the same manner as (2.58), the approximate BER upper bound of the qth subcarrier, averaged over $\tilde{\mathbf{d}}_q(q)$ and $\tilde{\mathbf{d}}_q(Q+q)$, is thus obtained as

$$P_{AP,q} \le \frac{1}{2\sigma_d^2}\left(\frac{1}{\sigma_{c,q}^2}\right)^{-1}, \tag{2.62}$$

where $\sigma_{c,q}^2 = \sigma_{isi,q}^2 + \sigma_v^2$, in which $\sigma_{isi,q}^2 = E\left\{2\tilde{\mathbf{d}}_q^H\tilde{\mathbf{H}}_I\tilde{\mathbf{H}}_I^H\tilde{\mathbf{d}}_q\right\}$, averaged over $\overline{h}^{(m)}(l)$, $1 \le m \le 2$, $0 \le l \le L$, is the ISI/ICI power; in deriving (2.62), we have neglected the correlation between the signal and ICI, as suggested in [5]. As σ_v^2 approaches zero, the upper bound in (2.62) converges to a constant, so that

$$P_{AP,q} \le \frac{1}{2\sigma_d^2}\left(\frac{1}{\sigma_{isi,q}^2}\right)^{-1}. \tag{2.63}$$

Hence, as full channel correlation occurs, the proposed system loses the diversity gain and its BER performance is dominated by the un-cancelled ISI/ICI: significant error floor would occur even in the high SNR regime. The performance tendency of the downlink case is the same as in the uplink case, and is not repeated for brevity.

2.9 Simulation Results

A system platform similar to the IEEE 802.11a standard is considered: $Q = 64$ subcarriers, and the length of CP is set to be 16 for the CP-based transmission. The OFDM symbol lengths are $4 \, \mu s$ and $3.2 \, \mu s$ for the CP-based and CP-free systems, respectively. The source symbols are drawn from the QPSK constellation. The background channel characteristics is the standard wireless exponential decay model [50], in which the lth tap of the channel impulse response is modeled as a complex Gaussian random variable with variance $\sigma_l^2 = \left(1 - e^{-T/T_{rms}} \right) e^{-lT/T_{rms}}$, where T and T_{rms} are the source symbol duration and root mean square delay spread of the channel, respectively. The channel impulse response is normalized such that $\sum_{l=0}^{L} \left| h^{(m)}(l) \right|^2 = 1$. The input SNR at the mth receive antenna is defined as

$$SNR = \frac{\sum_{l=0}^{L} \left| h^{(m)}(l) \right|^2}{\sigma_v^2} = \frac{1}{\sigma_v^2} \, .$$

Unless otherwise mentioned, we set $T_{rms} = 40 \, ns$, which corresponds to a channel order $L = 8$ (it can be seen from [50]). The following two assumptions are made: perfect channel information is available at the AP, and the tap gains are constant over one OFDM symbol (quasi-static fading). For fair comparison, in the CP-based system, the average transmit power of each OFDM symbol (of length $Q+16$) is normalized to Q (this accounts for the power loss factor due to the CP redundancy [80]). The case of CP-based transmission without power normalization, in which each information symbol shares the same amount of power as in the CP-free case, is also included as the performance bound.

A. System BER Performance

This subsection shows the BER performance of the GSC-based equalizers and pre-equalizers as a function of the input SNR. Figure 2.3 compares the BER performances at the AP end in a CP-free OFDM system obtained by the following schemes: linear ZF and MMSE equalizers based on the ISI-free model (2.11), and the proposed FA and PA GSC equalizers with $M = 2$ and $M = 4$. The BER curves resulting from the CP-based transmissions, for $M = 2$ and $M = 4$ (using the MRC scheme to maximize SNR at the receiver), are included as the benchmark performances. As

can be seen from the figure, the performance of the FA GSC-based equalizer is best; the PA version of the GSC method reduces complexity at the cost of a slight SNR degradation. The ZF equalizer performs poorly due to noise enhancement, as stated in Section 2.4.2. On other hand, the MMSE equalizer performs worse than the proposed GSC-based equalizer, confirming that the two-stage method consumes extra degrees-of-freedom for suppressing the nulling induced ICI, which in turm degrades the diversity gain (see Section 2.4.2). Also, with either $M = 2$ or $M = 4$, the performance of the proposed solution (both FA and PA versions) is almost identical to that of the CP-based OFDM system. Therefore, it can be concluded that, for ISI/ICI suppression, the choice of $M = 2$ typically satisfies the performance requirement. Figure 2.4 compares the BER performances of the proposed GSC pre-equalizers and CP-based pre-MRC scheme [16] in the downlink stage. The pre-MRC scheme is known to provide the maximum SNR at the receiver, and can serve as a benchmark for the proposed pre-equalizer. The simulation results show that both the FA and PA GSC pre-equalizers, with either $M = 2$ or $M = 4$, have almost the same performance as the CP-based pre-MRC scheme. From the results, both in the uplink and downlink cases, the performances of the PA and FA implementations are very close. For clearer presentation, we will only show the BER curves of the PA versions of the proposed methods in the following figures.

B. Effect of Channel Estimation Error

In this subsection, the effect of channel estimation error on the proposed systems is demostrated. We consider the case of $M = 2$ and the estimated channel impulse responses $\hat{h}^{(m)}(l) = h^{(m)}(l) + \sigma_h \Delta h^{(m)}(l)$, where $\Delta h^{(m)}(l)$ is a complex Gaussian random variable with the same variance of $h^{(m)}(l)$, $1 \le m \le 2$, $0 \le l \le 8$, and σ_h is a measure of the relative channel estimation error. Figures 2.5 and 2.6 show that the BER performances of the proposed GSC-based equalizer and pre-equalizer for $\sigma_h = 0$, 0.1, 0.3 and 0.5. Compared with the error-free case, the performances of all methods are seen to degrade. For a fixed σ_h (excluding the case of $\sigma_h = 0$), the BER curves of the proposed GSC methods still closely match that of the CP-based method in the low-to-medium SNR region (SNR < 15 dB), but deteriorates for high SNR. This is because,

when channel estimation error occurs, the computed blocking matrix deviates from the actual solution. As a result, the desired signal cannot be completely blocked and thus leaks to the second branch. This leads to mutual cancellation of the desired signal at the receiver output and decreases the effective SINR.

C. Impacts of Channel Correlation on BER Performance

In this subsection, we address the impacts of channel correlation on the performance of the proposed schemes. For simplicity, we consider the case of $M = 2$; similar performance tendencies are observed for $M > 2$. The following three channel conditions are considered:

1. Medium correlation: $\rho = 0.35$.

2. High correlation: $\rho = 0.9$.

3. Full correlation: $\rho = 1$.

Figure 2.7 shows that, as the correlation of the two channel branches raises, the performances of all considered methods become worse. Moreover, the BER curve of the proposed GSC-based equalizer still well tracks that of the CP-based system even when ρ is as large as 0.9. However, the performance of the GSC solution degrades when full channel correlation occurs: the proposed interference cancellation mechanism breaks down. This confirms the results in Section 2.8. Figure 2.8 shows the results of the downlink case, whose tendencies are the same as in the uplink case. From the results, it can be concluded that the proposed ISI/ICI cancellation method is quite insensitive to channel correlation.

D. Resistance to Channel Delay Spread

This subsection shows the resistance of the GSC-based equalizers and pre-equalizers against channel delay spread. We fix $M = 2$ and $\text{SNR} = 15$ dB, and increase T_{rms} from 10 to 80 ns (this amounts to increasing the discrete-time channel order L from 2 to 16, which is the maximal tolerable channel order for the length of the inserted CP (=16). Figure 2.9 compares robustness of the various considered ISI/ICI suppression schemes against the increase in the channel delay spread. The simulation results show that the proposed GSC-based equalizer is quite resistant to a large channel delay spread: its BER curve maintains a small deviation from that of the CP-based system.

Figure 2.10 compares the proposed GSC-based pre-equalizer with the CP-based pre-MRC downlink transmission. As in the uplink case, the results show that the proposed GSC-based pre-equalizer exhibits a negligible degradation in performance compared to the CP-based pre-MRC scheme. This shows that the proposed pre-equalizer is also very resistant to a large channel delay spread.

2.10 Summary

CP-free OFDM transmission can dramatically increase the spectral efficiency, provided that the ISI/ICI induced by the channel delay spread can be effectively suppressed. In this chapter, we exploit the multi-antenna resource for ISI/ICI suppression based on the block OFDM system model. This naturally leads to a GSC-based framework for equalizer design in the uplink transmission of multi-antenna CP-free OFDM systems, as well as the pre-equalizer design in the corresponding downlink stage. Low-complexity PA implementations for both links are also derived, and the resultant simulated performances are seen to be comparable to the original FA realizations. To further access the performance of the proposed system, the output SINR at the AP and MU ends are also provided. The analysis shows that, with the proposed method for ISI/ICI suppression, M-fold receive and transmit diversity gains in the uplink and downlink stages are restored. Simulation results show that the proposed GSC-based transceiver compares favorably with the conventional CP-based ISI-free schemes and are highly resistant to a large channel delay spread.
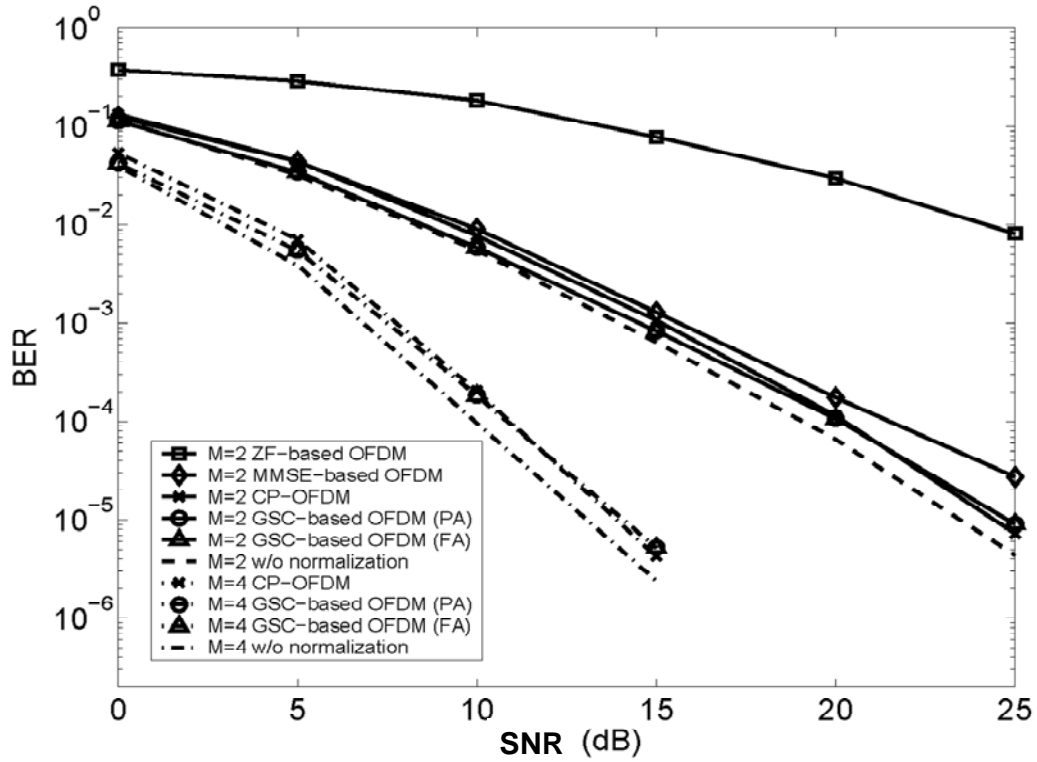
Figure 2.3. System BER performances at the AP end (uplink) versus receiver input SNR.
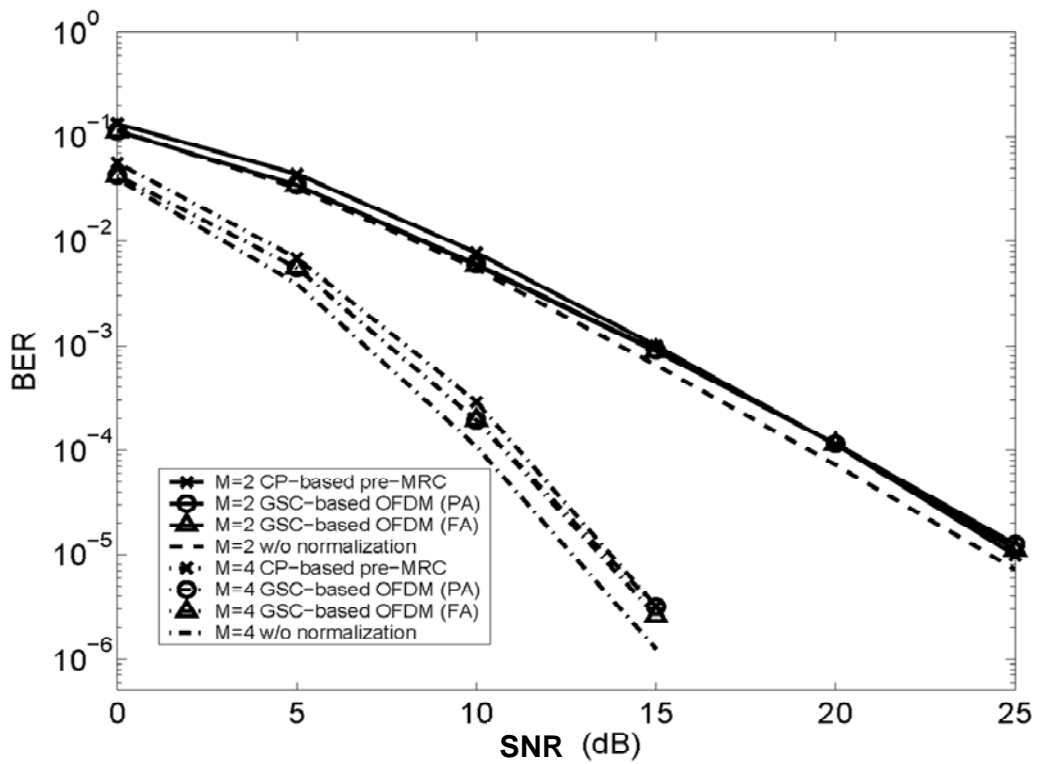
Figure 2.4. System BER performances at the MU end (downlink) versus receiver input SNR.
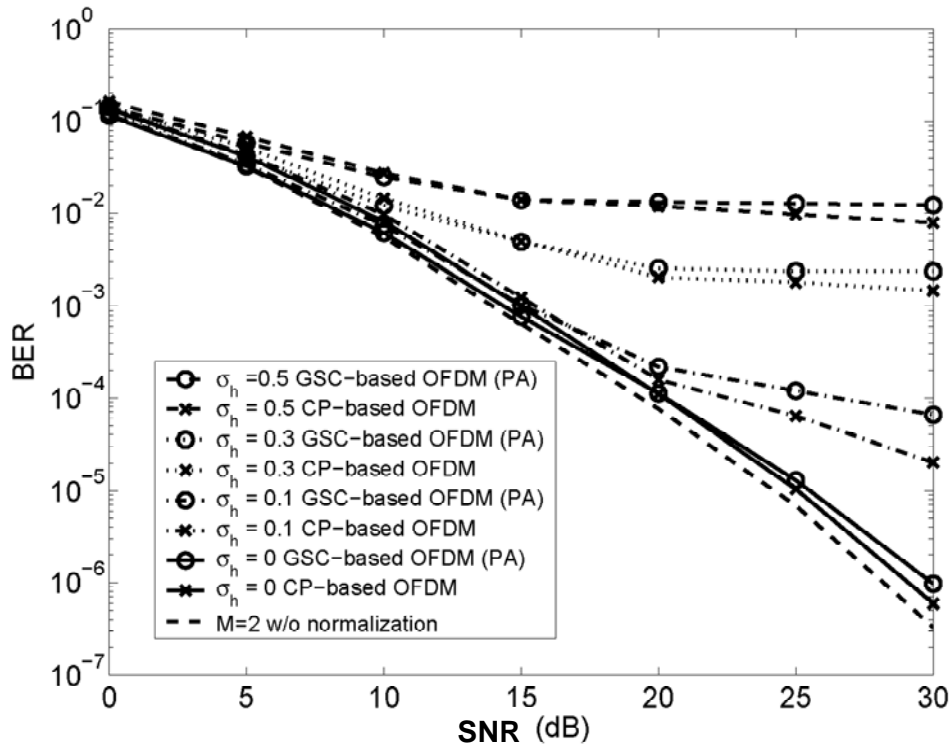
Figure 2.5. System BER performances under various channel estimation errors at the AP end (uplink) versus receiver input SNR.
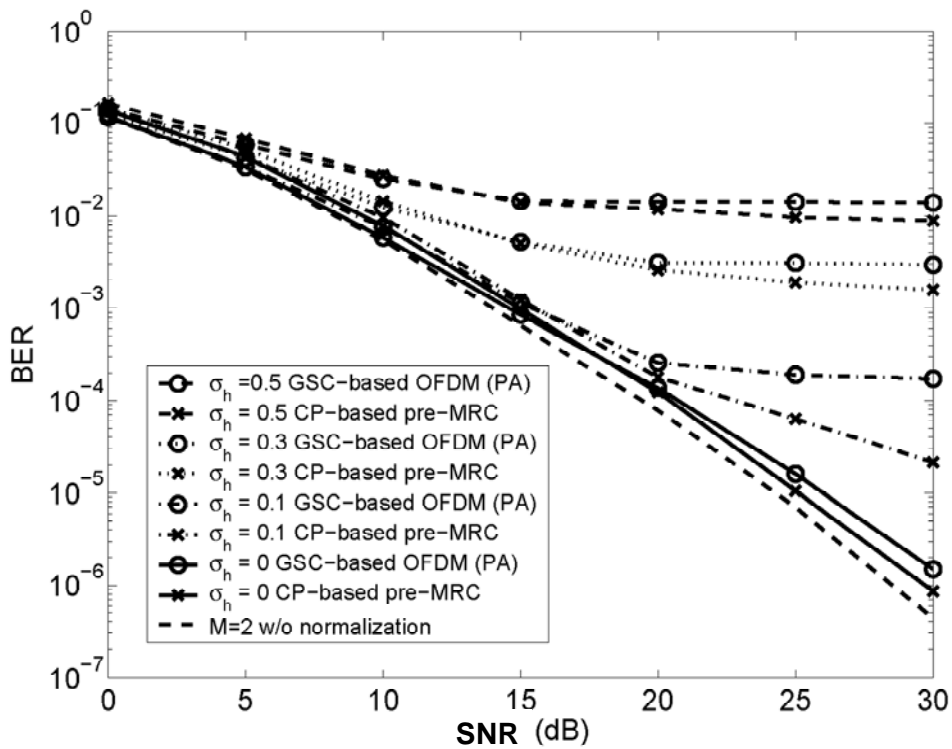
Figure 2.6. System BER performances under various channel estimation errors at the MU end (downlink) versus receiver input SNR.
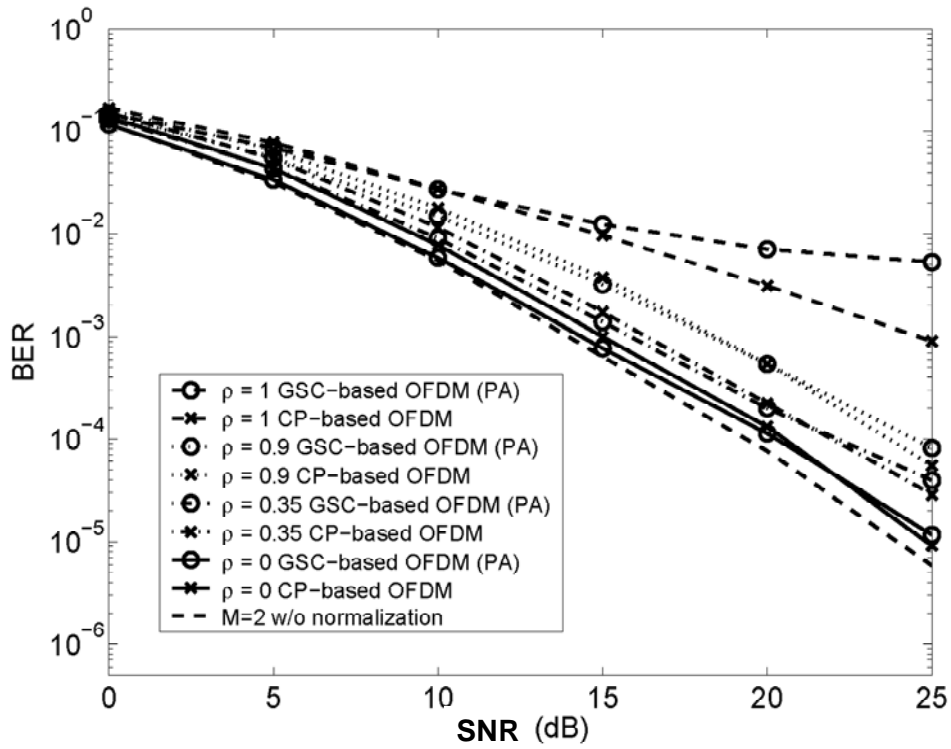
Figure 2.7. System BER performances under various channel correlation conditions at the AP end (uplink) versus receiver input SNR.
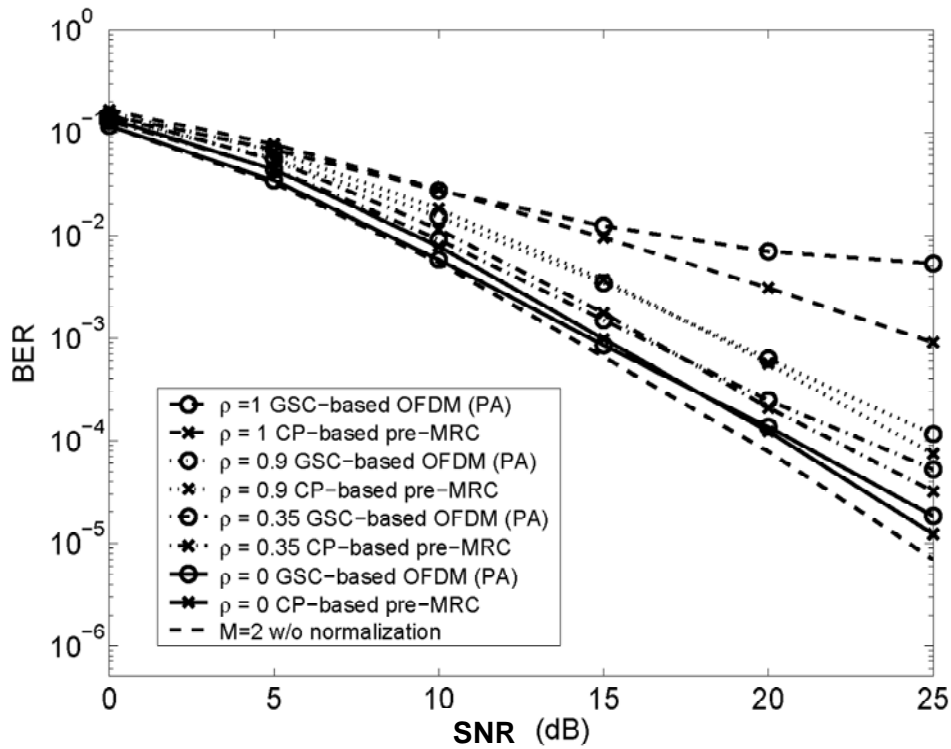
Figure 2.8. System BER performances under various channel correlation conditions at the MU end (downlink) versus receiver input SNR.
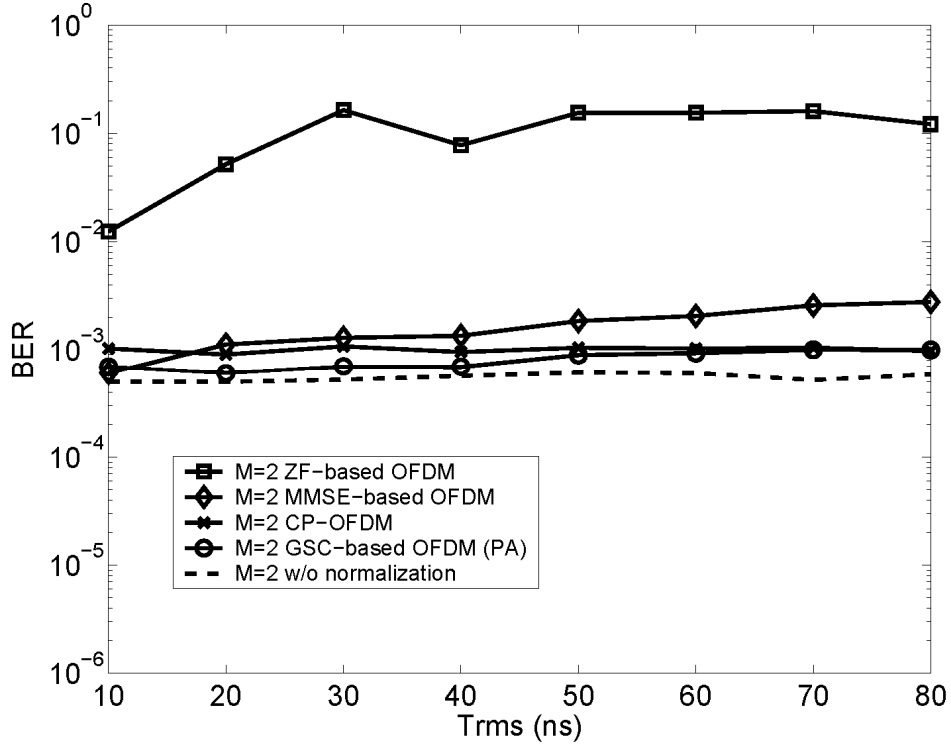
Figure 2.9. System BER performances at the AP end (uplink) versus T_{rms}, with $M = 2$ and SNR $= 15$ dB.
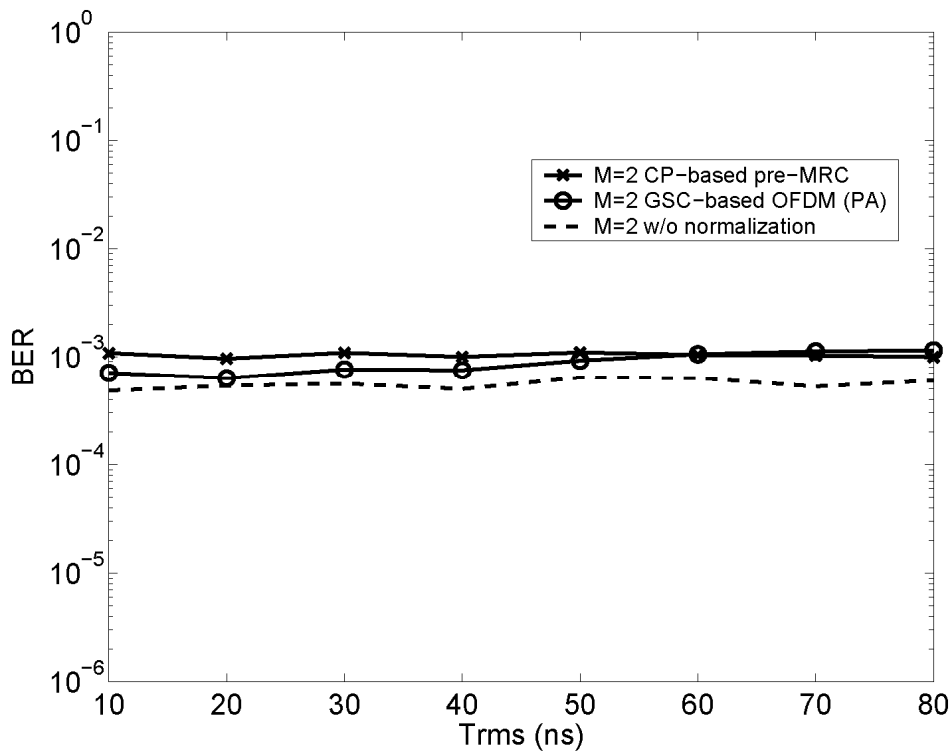
Figure 2.10. System BER performances at the MU end (downlink) versus T_{rms}, with $M = 2$ and SNR $= 15$ dB.

Chapter 3

Robust Receiver Design for MIMO-OFDM Under Channel Estimation Errors

3.1 Overview

This chapter generalizes the CP-free SIMO-OFDM framework proposed in Chapter 2 for MIMO-OFDM systems with variable CP length (shorter than the channel order), assuming that the channel parameters are not exactly known but are estimated via the commonly used LS training technique [6], [43]. To tackle the impact of channel estimation errors, one natural strategy in the considered scenario is to model the channel mismatch as a random variable and exploit the presumed LS channel error characteristic to derive a solution. The constraint-free GSC setup allows us to explicitly track each corrupted signal component resulting from imperfect channel estimation, and in turn leads to a very natural cost function for joint interference and channel error mitigation. We further leverage the perturbation technique [36], [83] to incorporate the channel parameter deviation into the solution equation; the optimal weighting matrix is then obtained by invoking the error characteristics of the LS channel estimate. The proposed optimal solution can mitigate the aggregate impacts due to channel estimation errors like signal leakage and other background parameter perturbation effects. Further analysis reveals that signal leakage turns out to be a dominant factor and a suboptimal solution, in the form of diagonal loading (DL), can attain a near-optimal performance. Also relying on the perturbation technique, we then derive a closed-form SINR expression associated with the suboptimal DL scheme; due to the near-optimal nature of the DL solution, the established result can well predict the actual SINR tendency attained by the optimal one. Our analytic formula can further quantify the achievable average SINR improvement over the non-robust GSC weight (i.e., the one derived under exact channel knowledge). In particular, we provide a closed-form expression of the SINR increment, and based on which several key features regarding the proposed robust solution can be inferred.

3.2 Preliminary

3.2.1 System Model and Basic Assumptions

We consider the discrete-time baseband model of a MIMO-OFDM system with Q subcarriers, N transmit antennas, and M receive antennas. At time k, the time-domain symbol to be sent from the nth transmit antenna is expressed as [79]

$$\tilde{\mathbf{s}}_n(k) = \mathbf{GF}^{-1}\mathbf{s}_n(k), \ \ 1 \le n \le N, \tag{3.1}$$

where $\mathbf{s}_n(k) \in \mathbb{C}^Q$ is the frequency-domain OFDM symbol of the nth transmit antenna, $\mathbf{F} \in \mathbb{C}^{Q \times Q}$ is the DFT matrix, $\mathbf{G} := [\overline{\mathbf{I}}_G^T \ \ \mathbf{I}_Q]^T \in \mathbb{C}^{(G+Q) \times Q}$, with $\overline{\mathbf{I}}_G \in \mathbb{C}^{G \times Q}$ being the last G rows of $\mathbf{I}_Q$, accounts for the insertion of a CP with length G. Let $h^{(m,n)}(\cdot)$ be the impulse response of the channel between the nth transmit antenna and the mth receive antenna; we assume without loss of generality that all the MN channels are of the same order L. Then the received $(G+Q) \times 1$ time-domain data vector from the mth antenna branch is [79]

$$\tilde{\mathbf{r}}_m(k) = \sum_{n=1}^N \tilde{\mathbf{H}}_0^{(m,n)} \tilde{\mathbf{s}}_n(k) + \sum_{n=1}^N \tilde{\mathbf{H}}_1^{(m,n)} \tilde{\mathbf{s}}_n(k-1) + \tilde{\mathbf{v}}_m(k), \tag{3.2}$$

where $\tilde{\mathbf{H}}_0^{(m,n)} \in \mathbb{C}^{(G+Q) \times (G+Q)}$ is lower triangular Toeplitz with $[h^{(m,n)}(0) \cdots h^{(m,n)}(L) \ 0 \cdots 0]^T$ as the first column, $\tilde{\mathbf{H}}_1^{(m,n)} \in \mathbb{C}^{(G+Q) \times (G+Q)}$ is upper triangular Toeplitz with $[0 \cdots 0 \ h^{(m,n)}(L) \cdots h^{(m,n)}(1)]$ as the first row, and $\tilde{\mathbf{v}}_m(k) \in \mathbb{C}^{(G+Q)}$ is the noise vector. To process the received data $\tilde{\mathbf{r}}_m(k)$, the G leading guard samples is first discarded; this corresponds to post-multiplying $\tilde{\mathbf{r}}_m(k)$ by the CP-removal matrix $\mathbf{M} := [\mathbf{O}_{Q \times G} \ \ \mathbf{I}_Q]^T \in \mathbb{C}^{(Q+G) \times Q}$ to get

$$\mathbf{r}_m(k) := \mathbf{M}^T \tilde{\mathbf{r}}_m(k) = \sum_{n=1}^N \underbrace{\mathbf{M}^T \tilde{\mathbf{H}}_0^{(m,n)} \mathbf{G}}_{:= \mathbf{H}_0^{(m,n)}} \mathbf{F}^{-1}\mathbf{s}_n(k) + \sum_{n=1}^N \underbrace{\mathbf{M}^T \tilde{\mathbf{H}}_1^{(m,n)} \mathbf{G}}_{:= \mathbf{H}_1^{(m,n)}} \mathbf{F}^{-1}\mathbf{s}_n(k-1) + \mathbf{M}^T \tilde{\mathbf{v}}_m(k).$$

$$\tag{3.3}$$

When the length of CP is no less than the channel order, i.e., $G \ge L$, the received signal $\mathbf{r}_m(k)$ is free from ISI and ICI, so that $\mathbf{H}_1^{(m,n)} = \mathbf{0}_Q$ and $\mathbf{H}_0^{(m,n)} := \mathbf{H}^{(m,n)}$, which is a circulant matrix with

$$\left[h^{(m,n)}(0) \ 0 \cdots 0 \ h^{(m,n)}(L) \cdots h^{(m,n)}(1) \right] \tag{3.4}$$

as the first row. We will focus on the case $G < L$; as such direct manipulation shows $\mathbf{H}_1^{(m,n)}$ becomes upper triangular Toeplitz with $[0 \cdots 0 \ h^{(m,n)}(L) \cdots h^{(m,n)}(G+1)]$ as the first row, and

$\mathbf{H}_0^{(m,n)}$ is instead obtained from $\mathbf{H}^{(m,n)}$ by setting $\left[\mathbf{H}^{(m,n)}\right]_{i,j} = 0$ for $1 \leq i \leq L-G$ and $Q - L + i \leq j \leq Q - G$. With (3.3), the received signal in the frequency domain reads

$$\mathbf{z}_m(k) := \mathbf{F}\mathbf{r}_m(k) = \sum_{n=1}^{N} \mathbf{F}\mathbf{H}_0^{(m,n)}\mathbf{F}^{-1}\mathbf{s}_n(k) + \sum_{n=1}^{N} \mathbf{F}\mathbf{H}_1^{(m,n)}\mathbf{F}^{-1}\mathbf{s}_n(k-1) + \underbrace{\mathbf{F}\mathbf{M}^T\tilde{\mathbf{v}}_m(k)}_{:=\mathbf{v}_m(k)}. \qquad (3.5)$$

We note that, on the right-hand-side (RHS) of (3.5), the second term is the ISI, whereas the first term is a mixture of the desired tone-by-tone signals and ICI. To realize low-complexity per-tone based signal recovery, which is one main advantage of MIMO-OFDM [51], a natural approach is to treat ISI and ICI as an overall composite interference and devise efficient schemes for *joint* ISI-ICI suppression [13]. For this it requires to further split ICI from the signal-ICI mixture in (3.5); since ICI is characterized via the deviation of $\mathbf{H}_0^{(m,n)}$ from the circulant matrix $\mathbf{H}^{(m,n)}$ [13], the splitting can be done according to the following decomposition

$$\mathbf{H}_0^{(m,n)} = \mathbf{H}^{(m,n)} - \underbrace{\mathbf{H}_1^{(m,n)}\mathbf{J}^G}_{:=\mathbf{H}_2^{(m,n)}}, \qquad (3.6)$$

where

$$\mathbf{H}_2^{(m,n)} := \mathbf{H}_1^{(m,n)}\mathbf{J}^G, \qquad (3.7)$$

with $\mathbf{J} \in \mathbb{C}^{Q \times Q}$ denoting the circulant permutation matrix with $[0 \cdots 0 \ 1] \in \mathbb{C}^{1 \times Q}$ as the first row (see Figure 3.1 for schematic descriptions of (3.6) and (3.7)).
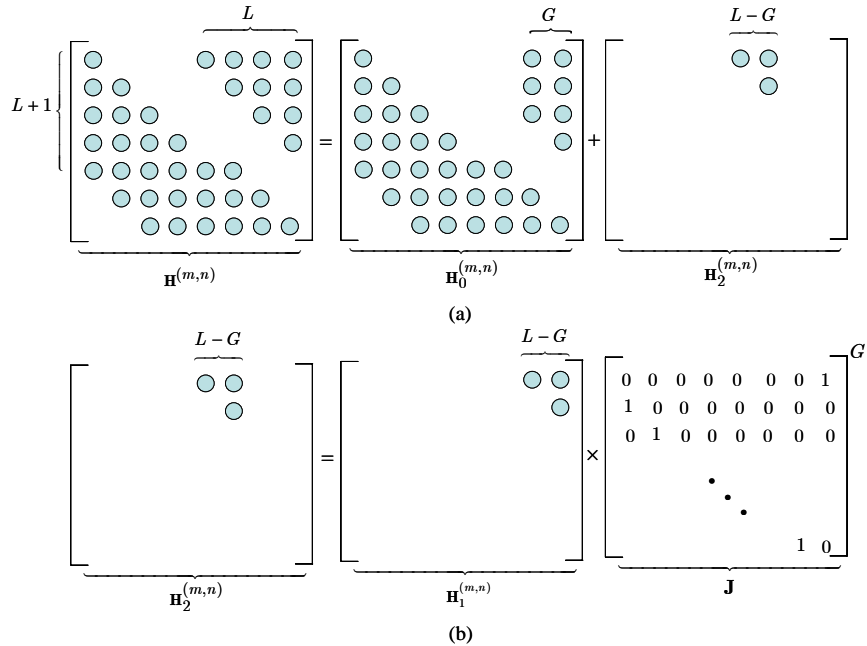
Figure 3.1. Schematic descriptions of (a) the decomposition (3.6) and (b) the relation (3.7).

From (3.6), we can rewrite $\mathbf{z}_m(k)$ in (3.5) as

$$\mathbf{z}_m(k) = \sum_{n=1}^{N} \underbrace{\mathbf{F}\mathbf{H}^{(m,n)}\mathbf{F}^{-1}}_{:=\mathbf{D}^{(m,n)}} \mathbf{s}_n(k) + \sum_{n=1}^{N} \mathbf{F}\mathbf{H}_1^{(m,n)}\mathbf{F}^{-1}\mathbf{s}_n(k-1) - \sum_{n=1}^{N} \mathbf{F}\mathbf{H}_2^{(m,n)}\mathbf{F}^{-1}\mathbf{s}_n(k) + \mathbf{v}_m(k),$$

(3.8)

where $\mathbf{D}^{(m,n)} \in \mathbb{C}^{Q \times Q}$ is diagonal with $[\mathbf{D}^{(m,n)}]_{ii} = \sum_{l=0}^{L} h^{(m,n)}(l)e^{-j2\pi l(i-1)/Q}$. The first term on the RHS of (2.8), which is composed of parallel tone-by-tone symbol streams from all transmit antennas, serves as the signal of interest. Since $G < L$, the symbols in each tone are contaminated by both ISI from the previously transmitted block (the second term in (3.8)) as well as ICI due to the loss of channel cyclicity (the third term in (3.8)). Based on (3.8), we propose a method for jointly suppressing ISI and ICI in the presence of channel mismatch. The following assumptions are made in the sequel.

1. The number of receive antennas is greater than the number of transmit antennas, i.e., $M > N$.

2. The source symbols of each transmit antenna $s_n(k)$ is zero mean, unit-variance, and $E\{s_{n_1}(k_1)s_{n_2}(k_2)^*\} = \delta(n_1 - n_2)\delta(k_1 - k_2)$, where $\delta(\cdot)$ is the Kronecker delta.

3. The elements of $\tilde{\mathbf{v}}_m(k)$'s are i.i.d. complex circular Gaussian with zero mean and variance σ_v^2.

3.2.2 Least-Squares Channel Estimation

We assume that $G \geq L$ during the training phase and the channels are estimated based on the LS training technique; see [6], [43] for detailed treatments. Let $\mathbf{h}^{(m,n)} := \left[h^{(m,n)}(0) \cdots h^{(m,n)}(L)\right]^T \in \mathbb{C}^{L+1}$, and $\hat{\mathbf{h}}^{(m,n)} := \mathbf{h}^{(m,n)} + \Delta\mathbf{h}^{(m,n)}$ be the corresponding optimal LS estimate, with $\Delta\mathbf{h}^{(m,n)}$ modeling the estimation error. With $\Delta\mathbf{h}^{(m)} := \left[\Delta\mathbf{h}^{(m,1)T} \cdots \Delta\mathbf{h}^{(m,N)T}\right]^T \in \mathbb{C}^{N(L+1)}$, $1 \leq m \leq M$, it is known that [6]

$$\mathbf{R}_{\mathbf{h}}^{(m)} := E\left\{\Delta\mathbf{h}^{(m)}\Delta\mathbf{h}^{(m)H}\right\} = \frac{\sigma_v^2}{P}\mathbf{I}_{N(L+1)},$$

(3.9)

where P is the transmit power dedicated to channel estimation; also, since the noises between different receive antennas are independent, we have

$$\mathbf{R}_{\mathbf{h}}^{(m_1,m_2)} := E\left\{\Delta\mathbf{h}^{(m_1)}\Delta\mathbf{h}^{(m_2)H}\right\} = \mathbf{0}_{N(L+1)}, \text{ for } m_1 \neq m_2.$$

(3.10)

The channel error properties (3.9) and (3.10) will be used in our robust equalizer design.

3.3 ISI-ICI Mitigation

This section shows how the receive diversity can be exploited for ISI-ICI suppression in the considered scenario. We will first collect all the M received signal branches to form a multi-channel system model, with which the interference subspace can be characterized; it will be seen that under quite mild conditions there will be sufficient inherent spatial and frequency degrees-of-freedom for combating ISI and ICI. Then, based on Section 2.4.3, we will show a GSC-based ISI-ICI suppression scheme under perfect channel knowledge assumption.

3.3.1 Multi-Channel System Representation

Let $\mathbf{s}(k) := [\mathbf{s}_1^T(k) \cdots \mathbf{s}_N^T(k)]^T \in \mathbb{C}^{NQ}$ be the vector containing all the N transmitted OFDM symbol blocks. By stacking $\mathbf{z}_m(k)$, $1 \le m \le M$, in (3.8) into a vector, we can form the MQ-dimensional multi-channel model:

$$\mathbf{z}(k) = \begin{bmatrix} \mathbf{z}_1^T(k) & \cdots & \mathbf{z}_M^T(k) \end{bmatrix}^T = \overline{\mathbf{D}}\mathbf{s}(k) + \underbrace{\mathbf{F}_M \mathbf{H}_1 \mathbf{F}_N^{-1}}_{:=\tilde{\mathbf{H}}_{ISI}} \mathbf{s}(k-1) - \underbrace{\mathbf{F}_M \mathbf{H}_2 \mathbf{F}_N^{-1}}_{:=\tilde{\mathbf{H}}_{ICI}} \mathbf{s}(k) + \mathbf{v}(k), \qquad (3.11)$$

where

$$\mathbf{H}_i := [\mathbf{H}_i^{(1)} \cdots \mathbf{H}_i^{(N)}] \in \mathbb{C}^{MQ \times NQ}, \ 1 \le i \le 2, \qquad (3.12)$$

with

$$\mathbf{H}_i^{(n)} := [\mathbf{H}_i^{(1,n)T} \cdots \mathbf{H}_i^{(M,n)T}]^T \in \mathbb{C}^{MQ \times Q}, \ 1 \le n \le N, \qquad (3.13)$$

$$\overline{\mathbf{D}} := [\mathbf{D}^{(1)} \cdots \mathbf{D}^{(N)}] \in \mathbb{C}^{MQ \times NQ}, \text{ with } \mathbf{D}^{(n)} := [\mathbf{D}^{(1,n)T} \cdots \mathbf{D}^{(M,n)T}]^T \in \mathbb{C}^{MQ \times Q}, \ 1 \le n \le N, \qquad (3.14)$$

$\mathbf{F}_p := \mathbf{I}_p \otimes \mathbf{F} \in \mathbb{C}^{pQ \times pQ}$, for $p \in \{M, N\}$, and $\mathbf{v}(k) := [\mathbf{v}_1^T(k) \cdots \mathbf{v}_M^T(k)]^T \in \mathbb{C}^{MQ}$. Since $M > N$, we may assume without loss of generality that the channel tone matrix $\overline{\mathbf{D}}$ is of full column rank NQ; this assumption is valid whenever the MN subchannels between all transmit and receive antennas are uncorrelated. Since $\mathbf{H}_1^{(m,n)}$ is upper triangular with all the nonzero entries clustering in the last $L-G$ columns, the matrix $\mathbf{H}_1^{(n)}$ in (3.13) will have the last $L-G$ columns nonzero; by definition, the "big" $\mathbf{H}_1$ in (3.12) then has only $N(L-G)$ nonzero columns, and is of rank at most $N(L-G)$. From (3.7) and (3.13), it is easy to check $\mathbf{H}_2^{(n)} = \mathbf{H}_1^{(n)} \mathbf{J}^G$, and hence we have from (3.12)

41

that

$$\mathbf{H}_2 = \left[\mathbf{H}_2^{(1)} \ \mathbf{H}_2^{(2)} \cdots \mathbf{H}_2^{(N)} \right] = \left[\mathbf{H}_1^{(1)}\mathbf{J}^G \ \mathbf{H}_1^{(2)}\mathbf{J}^G \cdots \mathbf{H}_1^{(N)}\mathbf{J}^G \right] = \mathbf{H}_1(\mathbf{I}_N \otimes \mathbf{J}^G), \qquad (3.15)$$

where the last equality follows by the definition of Kronecker product. This asserts that the rank of $\mathbf{H}_2$ does not exceed $N(L-G)$ either since it is obtained by permuting the columns of $\mathbf{H}_1$. The relation (3.15) also implies the interference subspace, spanned by the columns of $\mathbf{H}_{ISI}$ and $\mathbf{H}_{ICI}$, is of dimension no larger than $N(L-G)$. Indeed, from (3.15) and (3.11) it follows

$$\mathbf{H}_{ICI} = \mathbf{F}_M\mathbf{H}_2\mathbf{F}_N^{-1} = \mathbf{F}_M\mathbf{H}_1(\mathbf{I}_N \otimes \mathbf{J}^G)\mathbf{F}_N^{-1}. \qquad (3.16)$$

With (3.16) and since $\mathbf{H}_{ISI} = \mathbf{F}_M\mathbf{H}_1\mathbf{F}_N^{-1}$ (see (3.11)), the column spaces of both $\mathbf{H}_{ISI}$ and $\mathbf{H}_{ICI}$ coincide with that of $\mathbf{F}_M\mathbf{H}_1$, whose rank is at most $N(L-G)$ (as $\mathbf{F}_M$ is orthonormal). Hence, if *i)* the NQ-dimensional range space of the channel tone matrix $\mathbf{D}$ does not overlap with the ISI-ICI subspace and *ii)* $(M-N)Q \geq N(L-G)$, it is plausible to exploit the extra degrees-of-freedom provided by the multi-channel space-frequency model (3.11) for ISI-ICI suppression. We note that condition *i)* can be verified to hold unless all the MN subchannels lapse into the same direction, viz., $\mathbf{h}^{(m,n)} = \beta^{(m,n)}\mathbf{h}$ for some $\mathbf{h} \in \mathbb{C}^{L+1}$ and $\beta^{(m,n)} \in \mathbb{C}$; condition *ii)* is typically true since the number of subcarriers Q is often substantially larger than the channel order L.

3.3.2 GSC Based Interference Suppression: Perfect Channel Knowledge Case

It is assumed for the moment that the channel is perfectly known at the receiver. In what follows, we aim to seek for a linear weighting matrix $\mathbf{W} \in \mathbb{C}^{MQ \times NQ}$ which satisfies

$$\mathbf{W}^H\overline{\mathbf{D}} = \overline{\mathbf{D}}^H\overline{\mathbf{D}}, \qquad (3.17)$$

and minimizes the mean power of the filtered ISI-ICI, i.e., $E\left\{ \left\| \mathbf{W}^H \left(\mathbf{H}_{ISI}\mathbf{s}(k-1) - \mathbf{H}_{ICI}\mathbf{s}(k) + \mathbf{v}(k) \right) \right\|^2 \right\}$. With constraint (3.17), the optimal weight $\mathbf{W}$ will linearly combine the desired signal in the maximal-ratio sense (channel matched filtering), and suppress ISI-ICI via minimizes the total output interference-plus-noise power; the equalized signal then approximates

$$\mathbf{W}^H\mathbf{z}(k) \approx \overline{\mathbf{D}}^H\overline{\mathbf{D}}\mathbf{s}(k), \tag{3.18}$$

which can facilitate low-complexity tone-by-tone signal separation. Such a constrained-optimization can be transformed into an unconstrained one via the GSC principle, i.e., the weight matrix $\mathbf{W}$ is decomposed into

$$\mathbf{W} = \overline{\mathbf{D}} - \overline{\mathbf{B}}\mathbf{U}, \tag{3.19}$$

where the signal signature $\overline{\mathbf{D}}$ denotes the non-adaptive component for verifying constraint (3.17), $\mathbf{B} \in \mathbb{C}^{MQ\times(M-N)Q}$ is the signal blocking matrix determined by $\overline{\mathbf{B}}^H\overline{\mathbf{D}} = \mathbf{0}$, and $\mathbf{U} \in \mathbb{C}^{(M-N)Q\times NQ}$ forms the remaining free parameters to be determined. Following the same procedures presented in Section 2.4.3, the optimal solution of the adaptive portion $\mathbf{U}$ can be derived as

$$\mathbf{U}_g = \left(\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\right)^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{D}}, \tag{3.20}$$

where $\mathbf{R}_I := \mathbf{H}_{ISI}\mathbf{H}_{ISI}^H + \mathbf{H}_{ICI}\mathbf{H}_{ICI}^H + \sigma_v^2\mathbf{I}_{MQ} \in \mathbb{C}^{MQ\times MQ}$, and the optimal GSC weight is thus

$$\mathbf{W}_g = \overline{\mathbf{D}} - \overline{\mathbf{B}}\mathbf{U}_g = \overline{\mathbf{D}} - \overline{\mathbf{B}}\left(\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\right)^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{D}}. \tag{3.21}$$

We note that the matrix $\mathbf{W}_g$ in (3.21) is obtained based on the crucial perfect channel knowledge assumption; when channel parameter mismatch occurs due to imperfect estimation, the performance of solution (3.21) will degrade since it does not take into account channel error mitigation.

3.4 Proposed Robust Solution Against Channel Estimation Error

This section studies the problem of robust equalizer design against channel estimation error. We will first introduce the design formulation, and point out the challenge toward a solution. Then we will characterize the estimated blocking matrix via a perturbation analysis; this is crucial for solution derivation and subsequent performance analysis. Based on the established results and assuming the channel is estimated via the LS criterion, a closed-form optimal weighting matrix is obtained and the related discussions are given.

3.4.1 Problem Formulation

Our exposure will directly rely on the GSC setup. We first observe that, when only an estimated

$\widehat{\overline{\mathbf{D}}} \neq \overline{\mathbf{D}}$ is available, exact maximal-ratio combining of the desired signal $\overline{\mathbf{D}}\mathbf{s}(k)$ is impossible; the best we can do, however, is to linearly combine $\overline{\mathbf{D}}\mathbf{s}(k)$ through $\widehat{\overline{\mathbf{D}}}$ to get the approximation $\widehat{\overline{\mathbf{D}}}^{H}\overline{\mathbf{D}}\mathbf{s}(k)$. This suggests us to fix $\widehat{\overline{\mathbf{D}}}$ as the non-adaptive portion of the GSC solution, and decompose the weighting matrix into

$$\widehat{\mathbf{W}} = \widehat{\overline{\mathbf{D}}} - \widehat{\overline{\mathbf{B}}}\mathbf{U}, \tag{3.22}$$

where $\widehat{\overline{\mathbf{B}}}$ is the blocking matrix associated with $\widehat{\overline{\mathbf{D}}}$, that is, $\widehat{\overline{\mathbf{B}}}^{H}\widehat{\overline{\mathbf{D}}} = \mathbf{0}$, and $\mathbf{U} \in \mathbb{C}^{(M-N)Q \times NQ}$ is to be determined. With (3.22), the equalized output is instead

$$\widehat{\mathbf{W}}^{H}\mathbf{z}(k) = \overline{\mathbf{z}}_{d}(k) - \mathbf{U}^{H}\overline{\mathbf{z}}_{b}(k), \tag{3.23}$$

where

$$\overline{\mathbf{z}}_{d}(k) := \widehat{\overline{\mathbf{D}}}^{H}\mathbf{z}(k) = \widehat{\overline{\mathbf{D}}}^{H}\overline{\mathbf{D}}\mathbf{s}(k) + \widehat{\overline{\mathbf{D}}}^{H}\mathbf{H}_{ISI}\mathbf{s}(k-1) - \widehat{\overline{\mathbf{D}}}^{H}\mathbf{H}_{ICI}\mathbf{s}(k) + \widehat{\overline{\mathbf{D}}}^{H}\mathbf{v}(k). \tag{3.24}$$

and

$$\overline{\mathbf{z}}_{b}(k) := \widehat{\overline{\mathbf{B}}}^{H}\mathbf{z}(k) = \widehat{\overline{\mathbf{B}}}^{H}\overline{\mathbf{D}}\mathbf{s}(k) + \widehat{\overline{\mathbf{B}}}^{H}\mathbf{H}_{ISI}\mathbf{s}(k-1) - \widehat{\overline{\mathbf{B}}}^{H}\mathbf{H}_{ICI}\mathbf{s}(k) + \widehat{\overline{\mathbf{B}}}^{H}\mathbf{v}(k). \tag{3.25}$$

Due to inexact channel knowledge, the desired signal in the matched filtered branch $\overline{\mathbf{z}}_{d}(k)$ is non-coherently combined and the contaminating interference becomes

$$\overline{\mathbf{i}}(k) := \widehat{\overline{\mathbf{D}}}^{H}\mathbf{H}_{ISI}\mathbf{s}(k-1) - \widehat{\overline{\mathbf{D}}}^{H}\mathbf{H}_{ICI}\mathbf{s}(k) + \widehat{\overline{\mathbf{D}}}^{H}\mathbf{v}(k). \tag{3.26}$$

The channel estimation error also modifies the blocked signal characteristics in (3.25). In particular, since the estimated $\widehat{\overline{\mathbf{B}}}$ is otherwise determined via $\widehat{\overline{\mathbf{B}}}^{H}\widehat{\overline{\mathbf{D}}} = \mathbf{0}$ (and hence $\widehat{\overline{\mathbf{B}}}^{H}\overline{\mathbf{D}} \neq \mathbf{0}$ in general), there is a signal leakage $\widehat{\overline{\mathbf{B}}}^{H}\overline{\mathbf{D}}\mathbf{s}(k)$ into the blocking branch $\overline{\mathbf{z}}_{b}(k)$. To mitigate the aggregate impacts due to channel errors, a natural approach is to treat $\overline{\mathbf{z}}_{b}(k)$ as a composite interference, and to design $\mathbf{U}$ by minimizing

$$\overline{J} := E\left\{\left\|\overline{\mathbf{i}}(k) - \mathbf{U}^{H}\overline{\mathbf{z}}_{b}(k)\right\|^{2}\right\}, \tag{3.27}$$

where the expectation is taken with respect to the source signal, channel estimation error, and background noise (assuming all are mutually independent). Based on (3.25), (3.26) and (3.27), and by averaging the cost function $\overline{J}$ over the source signal and noise, we have

$$\bar{J} = Tr\left(\mathbf{U}^H E\{\hat{\overline{\mathbf{B}}}^H(\mathbf{R}_I - \overline{\mathbf{D}}\mathbf{H}_{ICI}^H - \mathbf{H}_{ICI}\overline{\mathbf{D}}^H + \overline{\mathbf{D}}\overline{\mathbf{D}}^H)\hat{\overline{\mathbf{B}}}\}\mathbf{U}\right) - Tr\left(\mathbf{U}^H E\{\hat{\overline{\mathbf{B}}}^H(\mathbf{R}_I - \overline{\mathbf{D}}\mathbf{H}_{ICI}^H)\hat{\overline{\mathbf{D}}}\}\right)$$

$$-Tr\left(E\{(\hat{\overline{\mathbf{D}}}^H(\mathbf{R}_I - \mathbf{H}_{ICI}\overline{\mathbf{D}}^H)\hat{\overline{\mathbf{B}}}\}\mathbf{U}\right) + Tr\left(E\{(\hat{\overline{\mathbf{D}}}^H\mathbf{R}_I\hat{\overline{\mathbf{D}}}\}\right). \tag{3.28}$$

Since, for a given $\mathbf{A}$ we have[1] $\partial Tr(\mathbf{U}^H\mathbf{A}\mathbf{U})/\partial\mathbf{U} = (\mathbf{U}^H\mathbf{A})^T$, $\partial Tr(\mathbf{A}\mathbf{U})/\partial\mathbf{U} = \mathbf{A}^T$, and $\partial Tr(\mathbf{U}^H\mathbf{A})/\partial\mathbf{U} = \mathbf{0}_{(M-N)Q\times NQ}$, the derivative of $\bar{J}$ with respect to $\mathbf{U}$ is thus

$$\frac{\partial}{\partial\mathbf{U}}\bar{J} = \left(\mathbf{U}^H E\{\hat{\overline{\mathbf{B}}}^H(\mathbf{R}_I - \overline{\mathbf{D}}\mathbf{H}_{ICI}^H - \mathbf{H}_{ICI}\overline{\mathbf{D}}^H + \overline{\mathbf{D}}\overline{\mathbf{D}}^H)\hat{\overline{\mathbf{B}}}\}\right)^T - \left(E\{(\hat{\overline{\mathbf{D}}}^H(\mathbf{R}_I - \mathbf{H}_{ICI}\overline{\mathbf{D}}^H)\hat{\overline{\mathbf{B}}}\}\right)^T. \tag{3.29}$$

With (3.29) and by setting $\partial\bar{J}/\partial\mathbf{U} = \mathbf{0}_{(M-N)Q\times NQ}$, the first-order necessary condition reads

$$\left[E\{\hat{\overline{\mathbf{B}}}^H\overline{\mathbf{D}}\overline{\mathbf{D}}^H\hat{\overline{\mathbf{B}}})\} + E\{\hat{\overline{\mathbf{B}}}^H(\mathbf{R}_I - \overline{\mathbf{D}}\mathbf{H}_{ICI}^H - \mathbf{H}_{ICI}\overline{\mathbf{D}}^H)\hat{\overline{\mathbf{B}}}\}\right]\mathbf{U} = E\{\hat{\overline{\mathbf{B}}}^H(\mathbf{R}_I - \overline{\mathbf{D}}\mathbf{H}_{ICI}^H)\hat{\overline{\mathbf{D}}}\}. \tag{3.30}$$

To determine the optimal $\mathbf{U}$ from (3.30), it is necessary to explicitly evaluate all the involved expectation terms. This can be done if we can establish an expression linking the estimated blocking matrix $\hat{\overline{\mathbf{B}}}$ with the channel matrix perturbation $\Delta\overline{\mathbf{D}}$. Since $\hat{\overline{\mathbf{B}}}$ is constructed as a basis of the left null space associated with $\hat{\overline{\mathbf{D}}}$ through SVD (recalling that $\hat{\overline{\mathbf{B}}}^H\hat{\overline{\mathbf{D}}} = \mathbf{0}$), an exact relation between $\hat{\overline{\mathbf{B}}}$ and $\Delta\overline{\mathbf{D}}$ appears highly intractable. In the next subsection, we will resort to the perturbation technique for developing an analytic (but approximate) expression.

3.4.2 Estimated Blocking Matrix: A Perturbation Analysis

Let us express the estimated channel tone matrix as

$$\hat{\overline{\mathbf{D}}} = \overline{\mathbf{D}} + \Delta\overline{\mathbf{D}}, \tag{3.31}$$

where $\Delta\overline{\mathbf{D}} \in \mathbb{C}^{MQ\times NQ}$ models the estimation error and is defined similarly as $\overline{\mathbf{D}}$ in (3.14), except that the (m, n)th $Q \times Q$ submatrix is

$$\Delta\mathbf{D}^{(m,n)} := diag\left\{\sqrt{Q}\mathbf{F}_L\Delta\mathbf{h}^{(m,n)}\right\}, \tag{3.32}$$

where $\mathbf{F}_L \in \mathbb{C}^{Q\times(L+1)}$ contains the first $L+1$ columns of the DFT matrix. Write an SVD of the exact signal matrix $\overline{\mathbf{D}}$ as

---

[1] These equalities follow immediately from the definition of the derivative of a real valued function with respect to a matrix, e.g., [78, Appendix A], together with some straightforward manipulations.

$$\overline{\mathbf{D}} = \begin{bmatrix} \mathbf{U}_D & \mathbf{U}_B \end{bmatrix} \begin{bmatrix} \Sigma_D & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_D^H \\ \mathbf{V}_B^H \end{bmatrix}, \tag{3.33}$$

and likewise for $\widehat{\overline{\mathbf{D}}}$,

$$\widehat{\overline{\mathbf{D}}} = \begin{bmatrix} \widehat{\mathbf{U}}_D & \widehat{\mathbf{U}}_B \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{\Sigma}}_D & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{V}}_D^H \\ \widehat{\mathbf{V}}_B^H \end{bmatrix}. \tag{3.34}$$

We note that each component matrix in (3.34) is of the same dimension as the corresponding noise-free counterpart in (3.33); the blocking matrices with, and without, channel parameter error are respectively $\widehat{\overline{\mathbf{B}}} = \widehat{\mathbf{U}}_B$ and $\overline{\mathbf{B}} = \mathbf{U}_B$, both of dimension $MQ \times (M-N)Q$. Let us further express $\widehat{\overline{\mathbf{B}}}$ as

$$\widehat{\overline{\mathbf{B}}} = \overline{\mathbf{B}} + \Delta\overline{\mathbf{B}}, \tag{3.35}$$

with $\Delta\mathbf{B}$ modeling the deviation. When $\left\|\Delta\overline{\mathbf{D}}\right\|$ is small, we have the following linear first-order approximation of $\widehat{\overline{\mathbf{B}}}$.

Lemma 3.1 [62]: The perturbed blocking matrix $\widehat{\overline{\mathbf{B}}}$ can be approximated by

$$\widehat{\overline{\mathbf{B}}} \approx \overline{\mathbf{B}} + \mathbf{U}_D \mathbf{P}, \tag{3.36}$$

where $\mathbf{P} \in \mathbb{C}^{NQ \times (M-N)Q}$ is such that $\|\mathbf{P}\| = O(\left\|\Delta\overline{\mathbf{D}}\right\|)$. $\qquad\square$

To completely specify $\widehat{\overline{\mathbf{B}}}$ in the form (3.36), it remains to determine $\mathbf{P}$. This can be done by further taking into account the equality $\widehat{\overline{\mathbf{B}}}^H \widehat{\overline{\mathbf{D}}} = \mathbf{0}$, which together with (3.31) and (3.36) implies

$$\widehat{\overline{\mathbf{B}}}^H \widehat{\overline{\mathbf{D}}} = (\overline{\mathbf{B}} + \mathbf{U}_D \mathbf{P})^H (\overline{\mathbf{D}} + \Delta\overline{\mathbf{D}}) = \mathbf{0}. \tag{3.37}$$

Since $\overline{\mathbf{B}}^H \overline{\mathbf{D}} = \mathbf{0}$ and $\mathbf{U}_D^H \overline{\mathbf{D}} = \Sigma_D \mathbf{V}_D^H$ (see (3.33)), equation (3.37) can be rearranged into

$$\mathbf{V}_D \Sigma_D \mathbf{P} + \Delta\overline{\mathbf{D}}^H \mathbf{U}_D \mathbf{P} = -\Delta\overline{\mathbf{D}}^H \mathbf{B}. \tag{3.38}$$

To determine $\mathbf{P}$ from (3.38), we further observe that

$$\left\|\Delta\overline{\mathbf{D}}^H \mathbf{U}_D \mathbf{P}\right\| \le \left\|\Delta\overline{\mathbf{D}}\right\| \cdot \left\|\mathbf{U}_D \mathbf{P}\right\| = \left\|\Delta\overline{\mathbf{D}}\right\| \cdot \|\mathbf{P}\| \propto O(\left\|\Delta\overline{\mathbf{D}}\right\|^2), \tag{3.39}$$

where the equality follows from the orthonormality of $\mathbf{U}_D$. Since $\|\mathbf{P}\| = O(\left\|\Delta\overline{\mathbf{D}}\right\|)$, inequality (3.39) asserts that $\left\|\Delta\overline{\mathbf{D}}^H \mathbf{U}_D \mathbf{P}\right\|$ is bounded from above by some quantity quadratic in $\left\|\Delta\overline{\mathbf{D}}\right\|$, which is

small with small $\left\| \Delta \overline{\mathbf{D}} \right\|$. We may thus neglect $\Delta \overline{\mathbf{D}}^H \mathbf{U}_D \mathbf{P}$ in (3.38) so that

$$\mathbf{V}_D \Sigma_D \mathbf{P} \approx -\Delta \overline{\mathbf{D}}^H \mathbf{B} ; \tag{3.40}$$

this technique is used in [36] for determining the first-order perturbation. With (3.40), the matrix $\mathbf{P}$ can thus be approximated by

$$\mathbf{P} = -\Sigma_D^{-1} \mathbf{V}_D^H \Delta \overline{\mathbf{D}}^H \overline{\mathbf{B}}. \tag{3.41}$$

This immediately implies $\Delta \overline{\mathbf{B}} = \mathbf{U}_D \mathbf{P} = -\mathbf{U}_D \Sigma_D^{-1} \mathbf{V}_D^H \Delta \overline{\mathbf{D}}^H \overline{\mathbf{B}}$, and

$$\widehat{\overline{\mathbf{B}}} = \overline{\mathbf{B}} \underbrace{-\mathbf{U}_D \Sigma_D^{-1} \mathbf{V}_D^H \Delta \overline{\mathbf{D}}^H \overline{\mathbf{B}}}_{:=\Delta \overline{\mathbf{B}}}. \tag{3.42}$$

Equation (3.42) provides a closed-form expression of $\widehat{\overline{\mathbf{B}}}$ *linear* in the estimation error $\Delta \overline{\mathbf{D}}$. The linearity nature can considerably simplify the derivation of the optimal solution, and will also lead to tractable procedures of performance analysis. We note that, instead of (3.36), more accurate approximation of $\widehat{\overline{\mathbf{B}}}$ can be obtained by incorporating higher order components [83]. Although this can improve the solution accuracy, the resultant analysis would however become intractable.

3.4.3 Optimal Solution

Based on (3.42) and assuming that the channel is estimated in the optimal LS sense, we can explicitly determine the expectation quantities involved in (3.30); these are summarized in the next lemma (see Appendix A for a proof), and will be used for deriving an optimal weighting matrix.

Lemma 3.2: The following results hold.

(1) $E\left\{ \widehat{\overline{\mathbf{B}}}^H (\mathbf{R}_I - \overline{\mathbf{D}} \mathbf{H}_{ICI}^H) \widehat{\overline{\mathbf{D}}} \right\} = \overline{\mathbf{B}}^H \mathbf{R}_I \overline{\mathbf{D}}$.

(2) $E\left\{ \widehat{\overline{\mathbf{B}}}^H \overline{\mathbf{D}} \overline{\mathbf{D}}^H \widehat{\overline{\mathbf{B}}} \right\} = \dfrac{N(L+1)\sigma_v^2}{P} \mathbf{I}_{(M-N)Q}$.

(3) Define the matrix

$$\mathbf{K} := \left(\overline{\mathbf{D}}^H \overline{\mathbf{D}} \right)^{-1} \overline{\mathbf{D}}^H (\mathbf{R}_I - \overline{\mathbf{D}} \mathbf{H}_{ICI}^H - \mathbf{H}_{ICI} \overline{\mathbf{D}}^H) \overline{\mathbf{D}} \left(\overline{\mathbf{D}}^H \overline{\mathbf{D}} \right)^{-1}. \tag{3.43}$$

Then we have

$$E\left\{\widehat{\overline{\mathbf{B}}}^H(\mathbf{R}_I - \overline{\mathbf{D}}\mathbf{H}_{ICI}^H - \mathbf{H}_{ICI}\overline{\mathbf{D}}^H)\widehat{\overline{\mathbf{B}}}\right\} = \overline{\mathbf{B}}^H(\mathbf{R}_I + \mathbf{R}_c)\overline{\mathbf{B}}, \tag{3.44}$$

in which $\mathbf{R}_c \in \mathbb{C}^{MQ \times MQ}$ is block diagonal, with the mth $Q \times Q$ block diagonal submatrix $\mathbf{R}_c^{(m)}$ given by

$$\mathbf{R}_c^{(m)} = \frac{Q\sigma_v^2}{P}\mathbf{F}_L\mathbf{F}_L^H \odot \sum_{n=1}^{N}\mathbf{K}^{(n)}, \tag{3.45}$$

and $\mathbf{K}^{(n)}$ denotes the nth $Q \times Q$ diagonal block of $\mathbf{K}$. $\qquad\square$

Based on (3.30) and Lemma 3.2, the optimal $\mathbf{U}_{opt}$ can be obtained as

$$\mathbf{U}_{opt} = \left(\frac{N(L+1)\sigma_v^2}{P}\mathbf{I}_{(M-N)Q} + \overline{\mathbf{B}}^H(\mathbf{R}_I + \mathbf{R}_c)\overline{\mathbf{B}}\right)^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{D}}, \tag{3.46}$$

where $\mathbf{R}_I$ and $\mathbf{R}_c$ are respectively defined in (3.15) and (3.45). Note that solution (3.46) is *on average* the optimal choice for ISI-ICI suppression under (white) LS channel estimation error assumption. In practical implementation when only an estimated channel is available, the sampled-version of the robust GSC weight is thus

$$\widehat{\mathbf{W}}_{opt} = \widehat{\overline{\mathbf{D}}} - \widehat{\overline{\mathbf{B}}}\widehat{\mathbf{U}}_{opt} = \widehat{\overline{\mathbf{D}}} - \widehat{\overline{\mathbf{B}}}\left(\frac{N(L+1)\sigma_v^2}{P}\mathbf{I}_{(M-N)Q} + \widehat{\overline{\mathbf{B}}}^H(\widehat{\mathbf{R}}_I + \widehat{\mathbf{R}}_c)\widehat{\overline{\mathbf{B}}}\right)^{-1}\widehat{\overline{\mathbf{B}}}^H\widehat{\mathbf{R}}_I\widehat{\overline{\mathbf{D}}}, \tag{3.47}$$

in which $\widehat{\mathbf{R}}_I$ and $\widehat{\mathbf{R}}_c$ are respectively the estimates of $\mathbf{R}_I$ and $\mathbf{R}_c$.

Discussions:

1. The proposed approach explicitly incorporates the channel mismatch effect into the GSC formulation; it aims for joint mitigation of ISI-ICI and the net parameter mismatch effects induced by channel estimation errors. The optimal $\mathbf{U}_{opt}$ differs from $\mathbf{U}_g$ in (3.20) in additional two terms, namely, $\frac{N(L+1)\sigma_v^2}{P}\mathbf{I}_{(M-N)Q}$ and $\overline{\mathbf{B}}^H\mathbf{R}_c\overline{\mathbf{B}}$, involved in matrix inversion; the former accounts for the signal leakage effect driven by the white-noise like LS channel estimation error (cf. (3.9) and (3.10)), whereas the latter is due to the parameter perturbation in the ISI-ICI signature matrices.

2. We should note that our design formulation is not exclusive to the case with LS channel

estimation (which produces a white channel estimation error); it does provide a unified framework for robust GSC filter design regardless of the adopted channel acquisition techniques. Indeed, as long as the channel error (with known covariance matrix) is independent of the source signal and noise, equation (3.30) remains true and the proposed approach will yield a solution of the form (3.46), except that all the involved matrices are accordingly modified based on the actual channel error characteristics.

3. Since $\overline{\mathbf{D}}^H\overline{\mathbf{D}}$ is the coherently combined signal signature, the magnitude of its non-zero entries would in general be substantially larger than those of $\mathbf{R}_I$ and $\overline{\mathbf{D}}\mathbf{H}_{ICI}^H$. This implies the entries of the matrix $\mathbf{K}$ in (3.43), and $\overline{\mathbf{B}}^H\mathbf{R}_c\overline{\mathbf{B}}$ as well, could be relatively small as compared with $P^{-1}N(L+1)\sigma_v^2$ (through simulations it is found that the entries of $\overline{\mathbf{B}}^H\mathbf{R}_c\overline{\mathbf{B}}$ are two-order less in magnitude in the medium-to-high SNR region). As a result, the achievable performance of $\mathbf{U}_{opt}$ in (3.46) can remain largely intact if we ignore the term $\overline{\mathbf{B}}^H\mathbf{R}_c\overline{\mathbf{B}}$, which reflects the ISI-ICI signature perturbation, and consider the sampled DL solution

$$\widehat{\mathbf{W}}_{dl} := \widehat{\overline{\mathbf{D}}} - \widehat{\overline{\mathbf{B}}}\widehat{\mathbf{U}}_{dl}, \tag{3.48}$$

with

$$\widehat{\mathbf{U}}_{dl} := \left(\frac{N(L+1)\sigma_v^2}{P}\mathbf{I}_{(M-N)Q} + \widehat{\overline{\mathbf{B}}}^H\widehat{\mathbf{R}}_I\widehat{\overline{\mathbf{B}}} \right)^{-1}\widehat{\overline{\mathbf{B}}}^H\widehat{\mathbf{R}}_I\widehat{\overline{\mathbf{D}}}. \tag{3.49}$$

This would indicate that the signal leakage, on the other hand, is the dominant effect incurred by channel estimation errors. An intuitive reason for this is that, the leaking signal component into the blocking branch (3.25) will cause undesirable signal cancellation via the two-branch GSC interference-rejection mechanism, leading to a loss of the effective SINR.

4. Robust constrained-optimization based beamformer design with background parameter error modeled as a white noise is addressed in [12]. The solution approach reported therein is to impose certain quadratic constraint on the beamforming weight so as to potentially keep down the white-noise amplification gain [12, pp. 1366-1367]. In light of this point, another plausible approach to robust GSC filter design in our context is thus

$$\min_{\mathbf{U}} E\left\{\left\|\mathbf{i}(k) - \mathbf{U}^H \mathbf{z}_b(k)\right\|^2\right\}, \text{ subject to } \|\mathbf{U}\| = \delta \text{ for some } \delta > 0 \,; \qquad (3.50)$$

the expectation in (3.50) is taken with respect to the source signal and measurement noise. The solution to (3.50) is known to be

$$\mathbf{U}_\gamma = \left(\gamma \mathbf{I}_{(M-N)Q} + \overline{\mathbf{B}}^H \mathbf{R}_I \overline{\mathbf{B}}\right)^{-1} \overline{\mathbf{B}}^H \mathbf{R}_I \overline{\mathbf{D}}, \text{ for some } \gamma \geq 0 \,. \qquad (3.51)$$

The performance of solution (3.51) depends crucially on the selection of γ [12], [42], [72], [75]; however, there are in general no tractable rules for explicitly determining an optimal γ, even when the uncertainty level δ is known [42]. We note that the suboptimal alternative (4.28) is exactly the sampled-version of the DL solution (3.51), with γ set to be

$$\overline{\gamma} := N(L+1)\sigma_v^2 / P \,. \qquad (3.52)$$

It is thus conjectured that, under white channel error assumption, $\overline{\gamma}$ derived based on the presented perturbation analysis is the best choice with respect to the design criterion (3.50). Our simulation results (see Simulation D) tend to confirm this postulation.

3.5 Performance Analysis

This section investigates the SINR performance of the proposed robust filter (3.47). We will first derive an approximate average SINR expression in closed-form. Then we will show the proposed solution can yield an improved SINR gain over $\mathbf{W}_g$ in (3.21); in particular, the achievable SINR increment will be quantified, and based on which several key features regarding the optimal solution can be inferred.

3.5.1 SINR Evaluation

To evaluate the average SINR attained by (3.47), we will resort to the perturbation based technique. Specifically, we will explicitly link the estimated solution with channel mismatch $\Delta \overline{\mathbf{D}}$, and then invoke the LS channel error property for mean SINR evaluation. To facilitate the underlying analysis, we will neglect the term $\widehat{\overline{\mathbf{B}}}^H \widehat{\mathbf{R}}_c \widehat{\overline{\mathbf{B}}}$ in (3.47) and consider the diagonal loading solution (3.48); through simulation tests the derived results based on the simplified solution form

are seen to well predict the actual SINR tendency attained by the optimal one (3.47).

To proceed, let

$$\widehat{\mathbf{W}}_{dl} = \mathbf{W}_{dl} + \Delta \mathbf{W}_{dl}, \tag{3.53}$$

where $\mathbf{W}_{dl}$ is the exact solution of $\widehat{\mathbf{W}}_{dl}$ (by substituting the true parameters $\overline{\mathbf{D}}$, $\overline{\mathbf{B}}$ and $\mathbf{R}_I$ in (3.48)) and $\Delta \mathbf{W}_{dl}$ models the deviation. With (3.53), the output from the robust GSC filter is

$$\mathbf{z}_{dl}(k) := \widehat{\mathbf{W}}_{dl}^H \mathbf{z}(k) = \widehat{\mathbf{W}}_{dl}^H \overline{\mathbf{D}}\mathbf{s}(k) + \widehat{\mathbf{W}}_{dl}^H \mathbf{H}_{ISI}\mathbf{s}(k-1) - \widehat{\mathbf{W}}_{dl}^H \mathbf{H}_{ICI}\mathbf{s}(k) + \widehat{\mathbf{W}}_{dl}^H \mathbf{v}(k)$$

$$= \underbrace{\mathbf{W}_{dl}^H \overline{\mathbf{D}}\mathbf{s}(k)}_{\mathbf{s}_{dl}(k)} + \underbrace{(\Delta \mathbf{W}_{dl}^H \overline{\mathbf{D}}\mathbf{s}(k) + \widehat{\mathbf{W}}_{dl}^H \mathbf{H}_{ISI}\mathbf{s}(k-1) - \widehat{\mathbf{W}}_{dl}^H \mathbf{H}_{ICI}\mathbf{s}(k) + \widehat{\mathbf{W}}_{dl}^H \mathbf{v}(k))}_{\mathbf{i}_{dl}(k)},$$

$$\tag{3.54}$$

were $\mathbf{s}_{dl}(k)$ is the desired signal component and $\mathbf{i}_{dl}(k)$ is the overall interference and noise. With (3.54), the average SINR is thus [34]

$$\text{SINR}_{dl} := \frac{E\left\{\|\mathbf{s}_{dl}(k)\|^2\right\}}{E\left\{\|\mathbf{i}_{dl}(k)\|^2\right\}}, \tag{3.55}$$

where the expectation is taken with respect to the source signal, noise, and channel estimation errors. The signal power $E\left\{\|\mathbf{s}_{dl}(k)\|^2\right\}$ can be directly computed by

$$E\left\{\|\mathbf{s}_{dl}(k)\|^2\right\} = E\left\{\left\|\mathbf{W}_{dl}^H \overline{\mathbf{D}}\mathbf{s}(k)\right\|^2\right\} = Tr\left(\mathbf{W}_{dl}^H \overline{\mathbf{D}}\overline{\mathbf{D}}^H \mathbf{W}_{dl}\right) = \left\|\overline{\mathbf{D}}^H \overline{\mathbf{D}}\right\|^2, \tag{3.56}$$

where the last equality in (3.56) follows from the definition of $\mathbf{W}_{dl}$ in (3.53) and $\overline{\mathbf{B}}^H \overline{\mathbf{D}} = \mathbf{0}$. The crucial step is to determine the composite interference power $E\left\{\|\mathbf{i}_{dl}(k)\|^2\right\}$. Based on further perturbation analysis, we have (see Appendix B for detailed derivations)

$$P_{I,dl} := E\left\{\|\mathbf{i}_{dl}(k)\|^2\right\} = Tr(\mathbf{W}_{dl}^H \mathbf{R}_I \mathbf{W}_{dl}) + Tr(E\{\Delta \mathbf{W}_{dl}^H \overline{\mathbf{D}}\overline{\mathbf{D}}^H \Delta \mathbf{W}_{dl}\}) + Tr(E\{\Delta \mathbf{W}_{dl}^H \mathbf{R}_{I,c} \Delta \mathbf{W}_{dl}\}),$$

$$\tag{3.57}$$

where

$$\mathbf{R}_{I,c} := \mathbf{R}_I - \overline{\mathbf{D}}\mathbf{H}_{ICI}^H - \mathbf{H}_{ICI}\overline{\mathbf{D}}^H; \tag{3.58}$$

the expectation involved on the RHS on (3.57) is with respect to the channel errors. The quantity $Tr(\mathbf{W}_{dl}^H \mathbf{R}_I \mathbf{W}_{dl})$ in $P_{I,dl}$ is the filtered interference power under perfect channel knowledge; the

remaining two arise due to channel mismatch and can be further computed as (see Appendix C for detailed derivations)

$$Tr(E\{\Delta \mathbf{W}_{dl}^H \overline{\mathbf{D}} \overline{\mathbf{D}}^H \Delta \mathbf{W}_{dl}\}) + Tr(E\{\Delta \mathbf{W}_{dl}^H \mathbf{R}_{I,c} \Delta \mathbf{W}_{dl}\})$$
$$= \bar{\gamma} \left\| \mathbf{U}_{dl} \right\|^2 + \sum_{i=1}^{6} Tr\left(\mathbf{X}_i \right) - 2 \operatorname{Re}\left\{ Tr\left(\mathbf{X}_7 \right) \right\} - 2 \operatorname{Re}\left\{ Tr\left(\mathbf{X}_8 \right) \right\} - 2 \operatorname{Re}\left\{ Tr\left(\mathbf{X}_9 \right) \right\}, \quad (3.59)$$

where the matrices $\mathbf{X}_i$, $1 \leq i \leq 9$, are provided in Table 3.1. It is noted that (3.59) keeps only the dominant terms. The average SINR attained by $\mathbf{W}_{dl}$ can then be evaluated based on (5.56), (5.57), and (5.59).

3.5.2 Achievable Performance Advantage

Based on (3.57), the proposed scheme can be shown to yield an SINR advantage over $\mathbf{W}_g$ in (3.21). To see this, let us write $\widehat{\mathbf{W}}_g = \mathbf{W}_g + \Delta \mathbf{W}_g$ as an estimate of $\mathbf{W}_g$. By splitting the filtered output $\mathbf{z}_g(k) := \widehat{\mathbf{W}}_g^H \mathbf{z}(k)$ in the form (3.54), it is straightforward to check that the signal power is

$$E\left\{ \left\| \mathbf{W}_g^H \overline{\mathbf{D}} \mathbf{s}(k) \right\|^2 \right\} = Tr\left(\mathbf{W}_g^H \overline{\mathbf{D}} \overline{\mathbf{D}}^H \mathbf{W}_g \right) = \left\| \overline{\mathbf{D}}^H \overline{\mathbf{D}} \right\|^2, \quad (3.60)$$

where the last equality follows from (3.21). Also, by going through essentially the same perturbation analysis as in Appendix B, the filtered interference power is verified to be

$$P_{I,g} := E\left\{ \left\| \mathbf{i}_g(k) \right\|^2 \right\} = Tr\left(\mathbf{W}_g^H \mathbf{R}_I \mathbf{W}_g \right) + Tr(E\{\Delta \mathbf{W}_g^H \overline{\mathbf{D}} \overline{\mathbf{D}}^H \Delta \mathbf{W}_g\}) + Tr(E\{\Delta \mathbf{W}_g^H \mathbf{R}_{I,c} \Delta \mathbf{W}_g\}).$$

$$(3.61)$$

From (3.56) and (3.60), we can see that the average signal levels sustained by $\mathbf{W}_{dl}$ and $\mathbf{W}_g$ are identical; the SINR are thus completely determined by the respective interference powers $P_{I,dl}$ and $P_{I,g}$. When σ_v^2 is small, it can be shown that $P_{I,dl} < P_{I,g}$. More precisely, we have the following result (see Appendix D for a proof).

Theorem 3.3: For small σ_v^2,

$$P_{I,g} - P_{I,dl} \approx \bar{\gamma}^2 Tr\left(\mathbf{U}_g^H \left(\bar{\gamma} \mathbf{I} + \overline{\mathbf{B}}^H \mathbf{R}_I \overline{\mathbf{B}} \right)^{-1} \mathbf{U}_g \right). \quad (3.62)$$

□

Equation (3.62) shows that, in the high SNR regime, the proposed robust solution can provide an SINR increment:

$$\Delta \text{SINR} := \text{SINR}_{dl} - \text{SINR}_g \approx \frac{\left\| \overline{\mathbf{D}}^H \overline{\mathbf{D}} \right\|^2}{P_{I,dl}} - \frac{\left\| \overline{\mathbf{D}}^H \overline{\mathbf{D}} \right\|^2}{P_{I,g}} = \frac{\overline{\gamma}^2 \left\| \overline{\mathbf{D}}^H \overline{\mathbf{D}} \right\|^2 Tr\left(\mathbf{U}_g^H \left(\overline{\gamma} \mathbf{I} + \overline{\mathbf{B}}^H \mathbf{R}_I \overline{\mathbf{B}} \right)^{-1} \mathbf{U}_g \right)}{P_{I,dl} P_{I,g}};$$

(3.63)

our simulation results show that (3.63) does remain valid for a wide range of SNR. The analytic SINR increment (3.63) not only quantifies the performance advantage of the robust scheme (3.47), but can also reveal several associated intrinsic features. To see this, we note that

$$Tr\left(\mathbf{U}_g^H \left(\overline{\gamma} \mathbf{I} + \overline{\mathbf{B}}^H \mathbf{R}_I \overline{\mathbf{B}} \right)^{-1} \mathbf{U}_g \right) \geq \frac{Tr\left(\mathbf{U}_g^H \mathbf{U}_g \right)}{\left(\overline{\gamma} + \lambda_{\max}(\overline{\mathbf{B}}^H \mathbf{R}_I \overline{\mathbf{B}}) \right)} = \frac{\left\| \mathbf{U}_g \right\|^2}{\left(\overline{\gamma} + \lambda_{\max}(\overline{\mathbf{B}}^H \mathbf{R}_I \overline{\mathbf{B}}) \right)}. \quad (3.64)$$

With (3.64), we can infer from (3.63) the inequality relation:

$$\Delta \text{SINR} \geq \frac{\overline{\gamma}^2 \left\| \mathbf{U}_g \right\|^2 \cdot \left\| \overline{\mathbf{D}}^H \overline{\mathbf{D}} \right\|^2}{\left[\lambda_{\max}(\overline{\mathbf{B}}^H \mathbf{R}_I \overline{\mathbf{B}}) + \overline{\gamma} \right] P_{I,dl} P_{I,g}}, \quad (3.65)$$

where $\lambda_{\max}(\overline{\mathbf{B}}^H \mathbf{R}_I \overline{\mathbf{B}})$ denotes the maximal eigenvalue associated with $\mathbf{B}^H \mathbf{R}_I \mathbf{B}$. The lower bound (3.65) leads to the following observations.

1. Let σ_v^2 be small and fixed. Through manipulation it can be shown that the incremental SINR lower bound (3.65) will increase as P decreases. Hence, when P is small and incurs severe channel mismatch, the proposed robust equalizer (3.47) would yield significant performance gain over solution (3.21). As P increases, and hence the estimation accuracy improves, the performance gain would however become negligible (this is also seen in our simulation).

2. We can also see from (3.65) that, for fixed P and σ_v^2, the performance improvement would be limited when $\lambda_{\max}(\overline{\mathbf{B}}^H \mathbf{R}_I \overline{\mathbf{B}})$, which reflects the maximal power of ISI and ICI with perfect channel knowledge, is large. This is intuitively reasonable since, under severe ISI and ICI, the equalizer will largely aim for interference suppression, rather than combating the channel mismatch effects.

3.6 Complexity Comparison

This section compares the algorithm complexity of the GSC solution (3.21) with that of the time-domain channel shortening/equalization (TEQ) approach [1] and the frequency-domain per-tone equalizer (PTEQ) [33]; we note that the complexity of the proposed robust solution (3.47) is essentially the same with that of (3.21).

The computational cost of solution (3.21) is in solving for the blocking matrix via $\overline{\mathbf{B}}^{H}\overline{\mathbf{D}} = \mathbf{0}$ and inverting the $(M-N)Q \times (M-N)Q$ matrix $\overline{\mathbf{B}}^{H}\mathbf{R}_{I}\overline{\mathbf{B}}$. A low-complexity scheme for obtaining $\overline{\mathbf{B}}$ which exploits the block diagonal structure of $\overline{\mathbf{D}}$ can be found in Section 2.5.1. As such, the total number of flop counts for computing the GSC solution (3.21) (in terms of the number of complex multiplications) is approximately

$$\mathrm{CM_{GSC}} = \frac{1}{3}(M-N)^3 Q^3 + \left[MN^2 + \left(2M + \frac{5}{2}\right)(M-N)^2 + M^2 \log_2^Q \right] Q^2$$
$$+ \left(\frac{1}{6}M^2 \left(2N(M-1) - N^2\right) - \frac{1}{6}\left(N^2 - 2MN\right)(M-N-1)^2 + N^3 \right) Q. \tag{3.66}$$

It is noted that the computational burden of the matrix inversion involved in (3.21) can be further alleviated by resorting to the partial adaptivity (PA) implementation (see Section 2.5.2); this instead calls for inverting an $N(L-G) \times N(L-G)$ matrix, and can limit the flop cost to

$$\mathrm{CM_{GSC,PA}} = \left[MN^2 + \left(2M + N(L-G)\right)(M-N)^2 + M^2 \log_2 Q \right] Q^2$$

$$+ \left[\frac{1}{6}M^2 \left(2N(M-1) - N^2\right) - \frac{1}{6}\left(N^2 - 2MN\right)(M-N-1)^2 + M(M-N) + N^3 \right] Q + MN^2(L-G)^2 \log_2^Q \tag{3.67}$$

The numbers of flop counts for TEQ [1] and PTEQ [33], respectively, are obtained as

$$\mathrm{CM_{TEQ}} = N^3 Q + NM(M+1)(L_t + L)L_t^2 + \frac{1}{3}N^3(G+1)^3 + \frac{7}{2}N^2(G+1)^2 + \frac{1}{6}N^2(G+1)^2, \tag{3.68}$$

and

$$\mathrm{CM_{PEQ}} = \left\{ \left[2(L_p + 1)(L + 2L_p + 1) + \frac{1}{3}(L + L_p + 1)^2 \right] N^3 + \frac{5}{2}(L + L_p + 1)N^2 + \frac{1}{6}N \right\}(L + L_p + 1)Q \tag{3.69}$$

where L_t and L_p are respectively the TEQ and PTEQ filter orders. From (3.66), (3.68) and (3.69), we have the following observations: *i)* TEQ method calls for the least algorithm complexity

among the three, *ii)* the complexities of the GSC and PTEQ methods are comparable for moderate numbers of subcarriers Q, *iii)* the GSC solution entails the highest computational cost when Q is large. As we will see in the simulation section, the proposed GSC approach, although incurring more algorithm complexity, does yield significant performance improvement over the other two comparative methods, even when perfect channel parameters are used for equalizer design.

3.7 Synchronization Issues

The synchronization issues of MIMO-OFDM systems are well discussed in the literature [63], [69]. This section will introduce a typical preamble-based method for an example. Assume that each transmit antenna transmits two identical preambles (all preambles are assumed the same) prior to a data burst, and the CP length in the synchronization phase, i.e., G_s, is larger than the channel order L. In general, a synchronization flow includes timing and frequency synchronizations, which contain the following four steps [63].

(1) Coarse timing synchronization: Coarse timing synchronization aims to roughly detect the start of an OFDM frame within only several samples away from the actual frame start point. Since the two transmitted preambles are identical, their CP portions are the same. When an OFDM frame arrives, the current G_s-sample block in the received sample sequence is highly correlated to the one at a distance of Q samples away. As a result, timing acquisition is usually performed via

$$\Phi_c = \arg\max_{\Phi} \left| \sum_{q=0}^{G_s-1} r^*_{\Phi+q} r_{\Phi+Q+q} \right|, \tag{3.70}$$

where r_q, $q \geq 0$, represents the time-domain received sample sequence. If Φ_c is larger than a prescribed threshold value, Φ_c is treated as the start instant of an OFDM frame.

(2) Time-domain frequency offset estimation: The frequency offset between the transmitter and receiver oscillator leads to a progressive phase shift $\theta = 2\pi\Delta$ imposed on the time-domain received sample sequence, where Δ denotes the frequency offset. Given Φ_c, the coarse frequency offset can be estimated via

$$\Delta_c = \frac{1}{2\pi} \angle \left\{ \sum_{q=0}^{G_s-1} r^*_{\Phi_c+q} r_{\Phi_c+Q+q} \right\}. \tag{3.71}$$

To compensate the frequency offset, each sample of the time-domain received sample sequence is multiplied by $\exp(-j2\pi q \Delta_c / Q)$.

(3) Residual frequency offset correction: Since the range of the time-domain frequency offset estimation in step 2 might be insufficient, residual frequency offset is left and has to be further compensated. Denote $\mathbf{p} = [p_0,...,p_{Q-1}]^T$ as the frequency-domain preamble sequence. The residual frequency offset can be estimated based on cyclically cross correlating the foreknown preamble sequence with the frequency-corrected OFDM symbol $\mathbf{z} = [z_0,...,z_{Q-1}]^T$ (the DFT of $[r_{\Phi_c} \exp(-j2\pi\Phi_c\Delta_c / Q),...,r_{\Phi_c+Q-1} \exp(-j2\pi(\Phi_c + Q - 1)\Delta_c / Q)]^T$), i.e.,

$$\Delta_f = \arg\max_{\Delta} \left| \sum_{q=0}^{Q-1} p^*_{(\Delta+q)_Q} z_q \right|, \ \ 0 \le \Delta \le Q-1. \tag{3.72}$$

With Δ_f, the residual phase rotation can be removed by further multiplying the time-domain sample sequence with $\exp(-j2\pi q \Delta_f / Q)$.

(4) Fine timing synchronization: After frequency offset is removed, the fine timing synchronization follows. The fine timing metric can be obtained by cross correlating the frequency-corrected time-domain sample sequence from step 3 ($r_{c,q} = r_q \exp(-j2\pi q(\Delta_c + \Delta_f)/Q)$, $q \ge 0$) with the time-domain preamble sequence $\mathbf{p}_t = [p_{t,0},...,p_{t,Q-1}]^T$ (the IDFT of $\mathbf{p}$), i.e.,

$$\Phi_f = \arg\max_{\Phi} \left| \sum_{u=1}^{2} \sum_{q=0}^{Q-1} p^*_{t,q} r_{c,\Phi+(u-1)Q+q} \right|. \tag{3.73}$$

In general, the coarse timing synchronization can estimate a timing metric close to the actual frame start instant. As a result, the fine timing synchronization is usually performed within a window centered around Φ_c to save computational burden.

3.8 Simulation Results

This section uses several numerical examples to illustrate the performance of the proposed method. We consider a MIMO-OFDM system with $N = 2$ transmit antennas, $M = 3$ receive

antennas, and $Q = 64$ subcarriers; the source symbols are drawn from the QPSK constellation. The background channel characteristics follow the standard wireless exponential decay model [50]: the channel impulse response is normalized such that $\sum_{n=1}^{N} \sum_{l=0}^{L} \left| h^{(m,n)}(l) \right|^2 = 1$. The input SNR at the mth receive antenna is defined as

$$SNR = \frac{\sum_{n=1}^{N} \sum_{l=0}^{L} \left| h^{(m,n)}(l) \right|^2}{\sigma_v^2} = \frac{1}{\sigma_v^2}.$$

We consider the quasi-static environment, in which the channels are assumed to remain constant during per coherent interval of 300 OFDM symbol periods, and can vary independently between different intervals. In each data burst the training pilots are placed in the entire heading OFDM symbol, and are designed according to [6]. The outputs of both TEQ and GSC filters are fed into an MMSE-VBLAST detector [29], [82] for further separating the multi-antenna transmitted signals on each tone. All the simulations results are averaged over 800 trials.

A. Comparison with Previous Works

We first compare the bit-error-rate (BER) performance of the proposed GSC based receiver with that of the TEQ [1] and PTEQ [33] approaches. For a given channel order L, the performances of TEQ and PTEQ depend crucially on the equalizer order and the allowable decision delay. The TEQ is implemented using an $(L+20)$-tap filter to shorten the composite channel order to the prescribed CP length G (through simulation it is found that further increase in the filter order does not seem to be able to improve performance); also, the resultant delay choice yielding the lowest BER is determined through exhaustive search and is then used in simulation. The order of PTEQ, and the associated decision delay, are both set to be $L - G$, as suggested in [33]. In the first simulation we consider the case when the channel is perfectly known at the receiver (the GSC equalizer (3.21) is adopted). For channel order $L = 13$ and CP length $G = 2$, Figure 3.2 shows the BER results of the three methods at different SNR levels. We can see that, among the three ISI-ICI mitigation schemes, the GSC equalizer (3.21) yields the best performance: it incurs no more than 1 dB penalty as compared with the ISI-free benchmark result, and even the associated low-complexity PA implementation can outperform TEQ and PTEQ. The performance advantage of

the GSC approach would lie in the exploitation of the joint space-frequency degrees-of-freedom for interference suppression. With SNR=20 dB, Figure 3.3 compares the respective BER as the CP length fixed at $G = 2$ and channel order L increases from 4 to 21; on the other hand, Figure 3.4 shows the BER results with L fixed at 13 and CP length G varying from 2 to 10. As we can see, the GSC filter in both cases yields the lowest BER. In the second simulation, we repeat the above three experiments but instead use the LS channel estimate for equalizer design; the results are shown in Figures 3.5, 3.6 and 3.7 (the power dedicated for training is $P = 14$, which is equally distributed to the pilot tones of the $N = 2$ transmit antennas). Compared with Figures 3.2, 3.3 and 3.4, we can see that TEQ, PTEQ, and the non-robust GSC filter (3.21) all suffer performance degradation due to imperfect channel estimation. The proposed robust solution (3.47) is seen to improve the performance over the one (3.21); still, it maintains less than 1 dB SNR gap with respect to the ISI-free case, and can also relieve the BER penalty against long delay spread channels and short CP lengths. Finally, we demonstrate the BER performances of the three comparative methods under distinct subchannel orders: $L_{1,1} = L_{1,2} = 13$, $L_{2,1} = L_{2,2} = 5$, and $L_{3,1} = L_{3,2} = 11$, where $L_{m,n}$ denotes the order of the subchannel between the nth transmit and mth receive antennas, $1 \leq m \leq 3$, $1 \leq n \leq 2$. In implementing the three methods we thus set $L = 13$ (for the subchannels with orders smaller than 13, $L - L_{m,n}$ zeros are padded in the respective impulse response tails). For $G = 2$, Figures 3.8 and 3.9 show the BER performances at different SNR under perfect and imperfect channel estimates, respectively; as we can see, the BER tendencies are essentially the same with those in the common subchannel order case (cf. Figures 3.2 and 3.5).

B. Performance of the Suboptimal Diagonal Loading Scheme (3.48)

This simulation illustrates the achievable performance of the suboptimal diagonal loading solution (3.48). For $L = 13$, $G = 2$, and two distinct transmit powers during the training phase $P = 14$ and 64, Figure 3.10 compares solution (3.48) with the optimal weighting matrix (3.47) in terms of SINR. The results show that the respective performances are almost identical. Since the diagonal loading weight aims exclusively for signal leakage reduction, this simulated results would imply that the leakage effect is the prime negative factor induced by channel mismatch.

C. Corroboration of the Analytic SINR Result

This simulation validates the predicted SINR results in Section 3.5.1. We consider two different transmit powers in the training phase $P = 14$ and $P = 64$. For $L = 13$ and $G = 2$, Figure 3.11 shows the theoretical SINR, computed using the formulas in Table 3.1, and the corresponding simulated outcomes. It can be seen that our analytic formula based on perturbation analysis well predicts the actual SINR tendency, even if the channel would be poorly estimated using a small transmit power. Figure 3.12 shows the SINR gain, in terms of difference in dB, attained by the proposed robust solution (3.47) over the one (3.21); the theoretical solution is computed based on (3.63). As we can see, although the theoretical solution is derived based on the high SNR assumption, it appears very close to the experimental results over the medium-to-high SNR region (>10 dB). It is also observed that, for a fixed SNR, the SINR gain is larger for smaller P (hence a less accurate channel estimate). This confirms the effectiveness of proposed robust GSC filter against severe parameter uncertainty; such a tendency has been deduced based on the lower bound relation (3.65) (see the first discussion following (3.65)).

D. On Selection of Regularization Factor

This simulation investigates the performance of the regularization based design (3.51) at different γ factors; the channel order, CP length and noise variance are respectively set to be $L = 13$, $G = 2$ and $\sigma_v^2 = 0.003162$ (this corresponds to SNR = 25 dB). We consider two different transmit powers $P = 2$ and $P = 14$ in the training phase; the respective conjectured optimal $\bar{\gamma}$ computed using (3.52) are 0.044268 and 0.006324. Figures 3.13 and 3.14 show the respective SINR performances of the regularized solution

$$\widehat{\mathbf{W}}_r := \widehat{\overline{\mathbf{D}}} - \widehat{\overline{\mathbf{B}}}\left(\gamma\mathbf{I}_{(M-N)Q} + \widehat{\overline{\mathbf{B}}}^H\widehat{\mathbf{R}}_I\widehat{\overline{\mathbf{B}}}\right)^{-1}\widehat{\overline{\mathbf{B}}}^H\widehat{\mathbf{R}}_I\widehat{\overline{\mathbf{D}}}. \tag{3.74}$$

at different values of γ. It can be seen that the SINR peaks attain at $\gamma = 0.044268$ for the $P = 2$ case, and at $\gamma = 0.006324$ when $P = 14$: this tends to indicate that $\bar{\gamma}$ in (3.52) is optimal with respect to the regularization based design under a (stochastic) white estimation error assumption (however, $\bar{\gamma}$ will no longer be optimal when different channel error models and design criteria are considered).

3.9 Summary

This chapter proposes a robust constrained-optimization based ISI-ICI mitigation scheme for supporting high-rate MIMO-OFDM transmission when the channels are not exactly known but are estimated via the LS training technique. The proposed constraint-free GSC design formulation yields a very natural cost function, and can facilitate the exploitation of the presumed LS channel error property toward a solution through simple first-order perturbation analysis. The proposed robust GSC filter can jointly mitigate ISI-ICI and the net detrimental factors caused by channel estimation errors. Numerical study reveals that the signal leakage is the dominant impairment and a suboptimal diagonal loading solution can attain almost all the performance gain. Based on perturbation techniques, we further derive a closed-form mean approximate SINR expression for the proposed robust scheme, and also an informative formula for quantifying the achievable SINR increment over the non-robust solution. The analytic SINR gain reveals that prominent performance advantage can be attained by the robust solution under severe channel mismatch. Simulation results confirm the effectiveness of the proposed GSC based equalizer: it outperforms existing methods under either perfect or imperfect channel assumption (at the cost of complexity) and, in both cases, yields a performance very close to the ISI-ICI free benchmark.

Table 3.1. Formulas of $\mathbf{X}_1 \sim \mathbf{X}_9$, with $\mathbf{X}_i = \mathbf{M}_i^H \mathbf{Y}_i \mathbf{M}_i$, $1 \le i \le 7$, $\mathbf{X}_8 = \mathbf{N}_8^H \mathbf{Y}_8 \mathbf{M}_8$, and $\mathbf{X}_9 = \mathbf{N}_9^H \mathbf{Y}_9 \mathbf{M}_9$.

$\mathbf{M}_1$	$\mathbf{I}_{NQ}$	$\mathbf{Y}_1^{(n)}$	$\dfrac{Q\sigma_v^2}{P} \mathbf{F}_L^* \mathbf{F}_L^T \odot \sum\limits_{m=1}^{M}\sum\limits_{n=1}^{N} \mathbf{D}^{(m,n)}\mathbf{D}^{(m,n)H}$
$\mathbf{M}_2$	$\mathbf{B}\mathbf{U}_{dl}$	$\mathbf{Y}_2^{(m)}$	$\dfrac{Q\sigma_v^2}{P} \mathbf{F}_L\mathbf{F}_L^H \odot \sum\limits_{n=1}^{N} \left[\mathbf{V}_D\Sigma_D^{-1}\mathbf{U}_D^H\mathbf{R}_{I,c}\mathbf{U}_D\Sigma_D^{-1}\mathbf{V}_D^H \right]^{(n)}$
$\mathbf{M}_3$	$\mathbf{I}_{NQ}$	$\mathbf{Y}_3^{(n)}$	$\dfrac{Q\sigma_v^2}{P} \mathbf{F}_L^* \mathbf{F}_L^T \odot \sum\limits_{m=1}^{M} \left[\mathbf{R}_{I,c} + \mathbf{R}_I\mathbf{B}(\mathbf{B}^H\mathbf{R}_{I,dl}\mathbf{B})^{-1}(\mathbf{B}^H\mathbf{R}_{I,c}\mathbf{B})(\mathbf{B}^H\mathbf{R}_{I,dl}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{R}_I \right]^{(m)}$
$\mathbf{M}_4$	$\mathbf{V}_D\Sigma_D^{-1}\mathbf{U}_D^H\mathbf{R}_I\mathbf{D}$	$\mathbf{Y}_4^{(n)}$	$\dfrac{Q\sigma_v^2}{P} \mathbf{F}_L^* \mathbf{F}_L^T \odot \sum\limits_{m=1}^{M} \left[\mathbf{B}(\mathbf{B}^H\mathbf{R}_{I,dl}\mathbf{B})^{-1}(\mathbf{B}^H\mathbf{R}_{I,c}\mathbf{B})(\mathbf{B}^H\mathbf{R}_{I,dl}\mathbf{B})^{-1}\mathbf{B}^H \right]^{(m)}$
$\mathbf{M}_5$	$\mathbf{V}_D\Sigma_D^{-1}\mathbf{U}_D^H\mathbf{R}_{I,dl}\mathbf{B}\mathbf{U}_{dl}$	$\mathbf{Y}_5^{(n)}$	$\dfrac{Q\sigma_v^2}{P} \mathbf{F}_L^* \mathbf{F}_L^T \odot \sum\limits_{m=1}^{M} \left[\mathbf{B}(\mathbf{B}^H\mathbf{R}_{I,dl}\mathbf{B})^{-1}(\mathbf{B}^H\mathbf{R}_{I,c}\mathbf{B})(\mathbf{B}^H\mathbf{R}_{I,dl}\mathbf{B})^{-1}\mathbf{B}^H \right]^{(m)}$
$\mathbf{M}_6$	$\mathbf{B}\mathbf{U}_{dl}$	$\mathbf{Y}_6^{(m)}$	$\dfrac{Q\sigma_v^2}{P} \mathbf{F}_L\mathbf{F}_L^H \odot \sum\limits_{n=1}^{N} \left[\mathbf{V}_D\Sigma_D^{-1}\mathbf{U}_D^H\mathbf{R}_{I,dl}\mathbf{B}(\mathbf{B}^H\mathbf{R}_{I,dl}\mathbf{B})^{-1}(\mathbf{B}^H\mathbf{R}_{I,c}\mathbf{B})(\mathbf{B}^H\mathbf{R}_{I,dl}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{R}_{I,dl}\mathbf{U}_D\Sigma_D^{-1}\mathbf{V}_D^H \right]^{(n)}$
$\mathbf{M}_7$	$\mathbf{B}\mathbf{U}_{dl}$	$\mathbf{Y}_7^{(m)}$	$\dfrac{Q\sigma_v^2}{P} \mathbf{F}_L\mathbf{F}_L^H \odot \sum\limits_{n=1}^{N} \left[\mathbf{V}_D\Sigma_D^{-1}\mathbf{U}_D^H\mathbf{R}_{I,c}\mathbf{B}(\mathbf{B}\mathbf{R}_{I,dl}\mathbf{B})^{-1}\mathbf{B}\mathbf{R}_{I,dl}\mathbf{U}_D\Sigma_D^{-1}\mathbf{V}_D^H \right]^{(n)}$
$\mathbf{M}_8$	$\mathbf{V}_D\Sigma_D^{-1}\mathbf{U}_D^H\mathbf{R}_I\mathbf{W}_{dl}$	$\mathbf{Y}_8^{(n)}$	$-\dfrac{Q\sigma_v^2}{P} \mathbf{F}_L^* \mathbf{F}_L^T \odot \sum\limits_{m=1}^{M} \left[\mathbf{R}_{I,c}\mathbf{B}(\mathbf{B}^H\mathbf{R}_{I,dl}\mathbf{B})^{-1}\mathbf{B}^H \right]^{(m)}$
$\mathbf{M}_9$	$\mathbf{V}_D\Sigma_D^{-1}\mathbf{U}_D^H\mathbf{R}_{I,dl}\mathbf{B}\mathbf{U}_{dl}$	$\mathbf{Y}_9^{(n)}$	$\dfrac{Q\sigma_v^2}{P} \mathbf{F}_L^* \mathbf{F}_L^T \odot \sum\limits_{m=1}^{M} \left[\mathbf{B}(\mathbf{B}^H\mathbf{R}_{I,dl}\mathbf{B})^{-1}(\mathbf{B}^H\mathbf{R}_{I,c}\mathbf{B})(\mathbf{B}^H\mathbf{R}_{I,dl}\mathbf{B})^{-1}\mathbf{B}^H \right]^{(m)}$
$\mathbf{N}_8$	$\mathbf{I}_{NQ}$	$\mathbf{N}_9$	$\mathbf{V}_D\Sigma_D^{-1}\mathbf{U}_D^H\mathbf{R}_I\mathbf{D}$
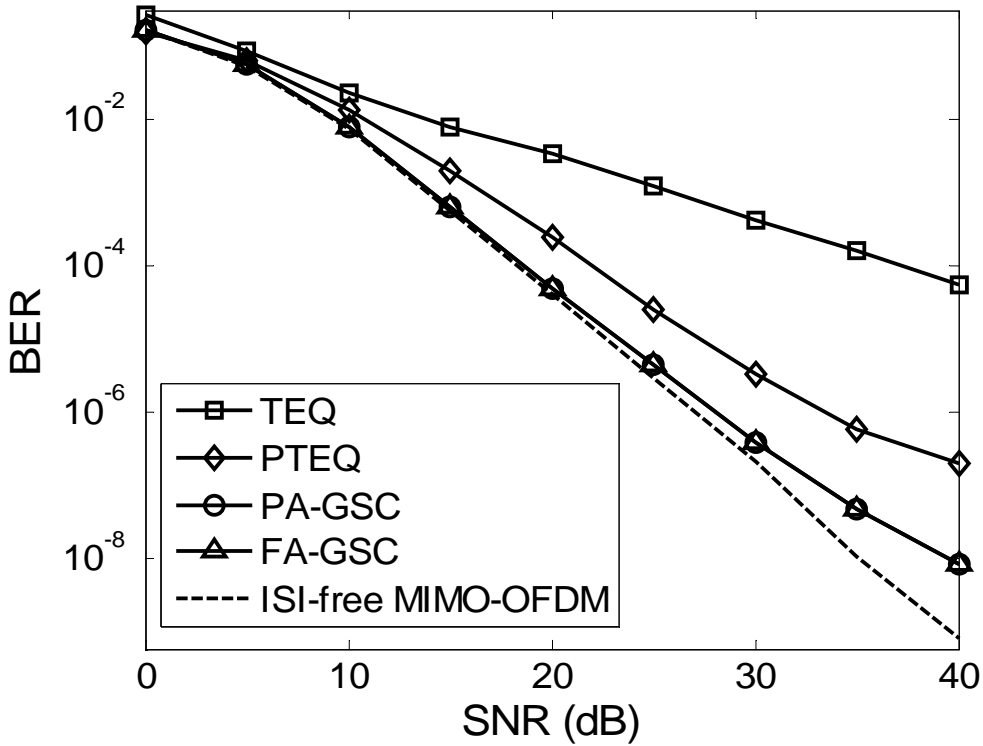
Figure 3.2. BER performances of the three methods (perfect channel knowledge).
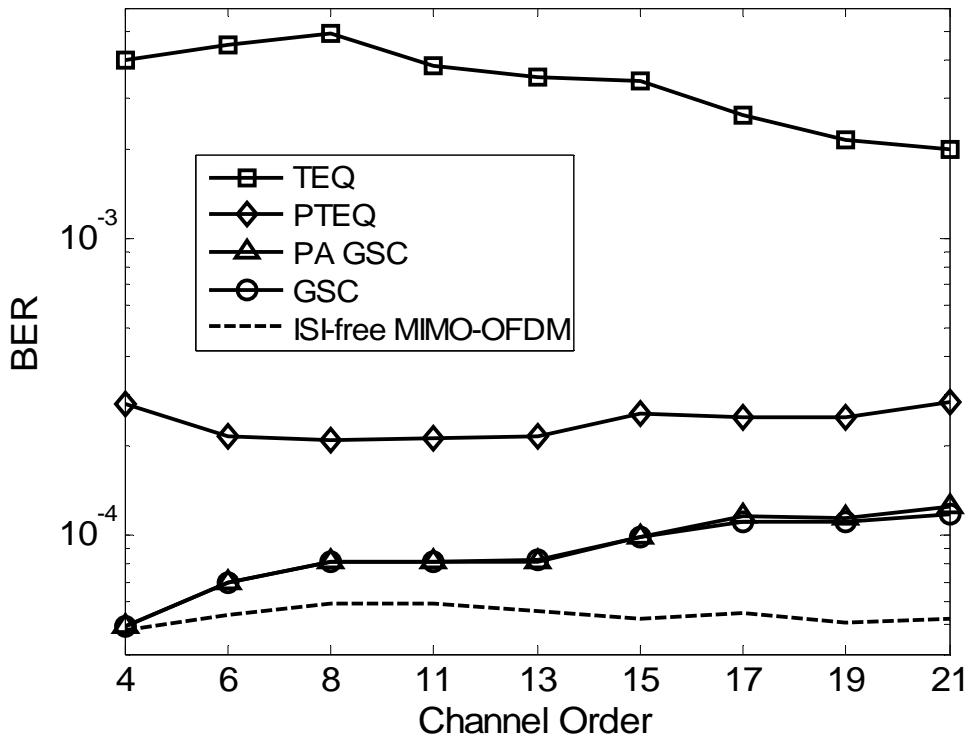
Figure 3.3. BER performances of the three methods at various channel orders with $\mathrm{SNR} = 20$ dB (perfect channel knowledge).
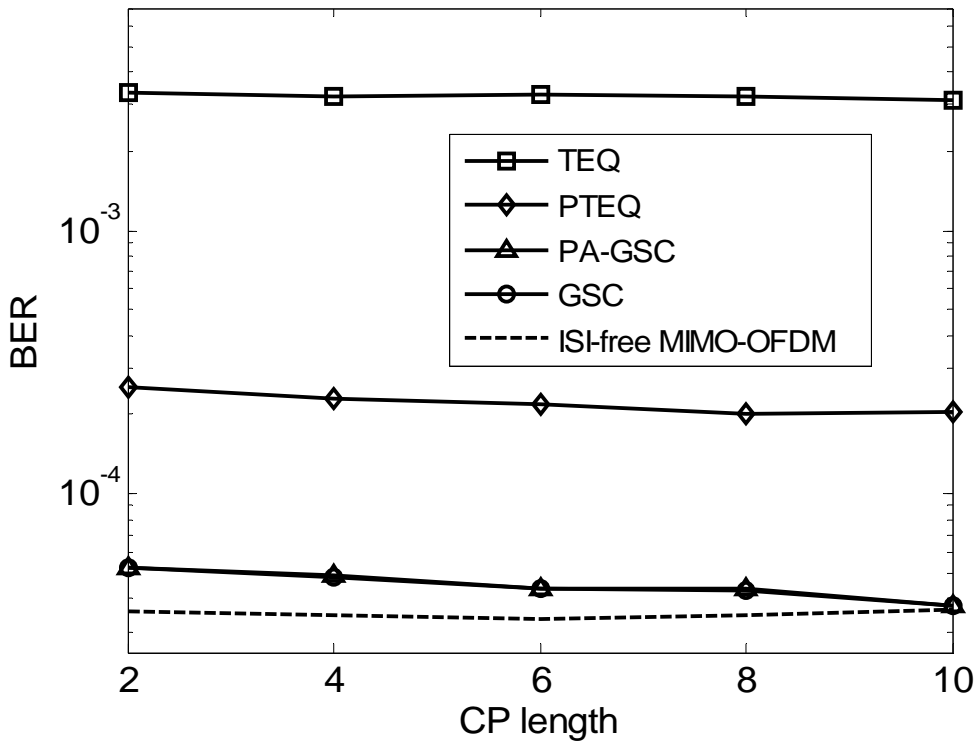
Figure 3.4. BER performances of the three methods at various CP lengths with $\mathrm{SNR} = 20$ dB (perfect channel knowledge).
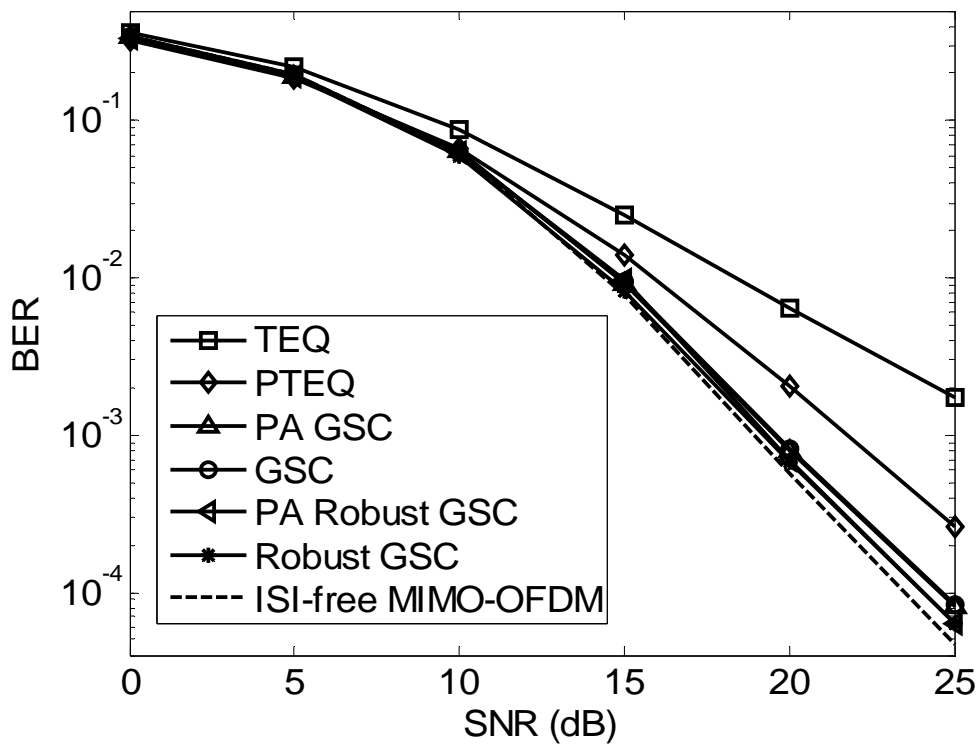
Figure 3.5. BER performances of the three methods (LS channel estimate).
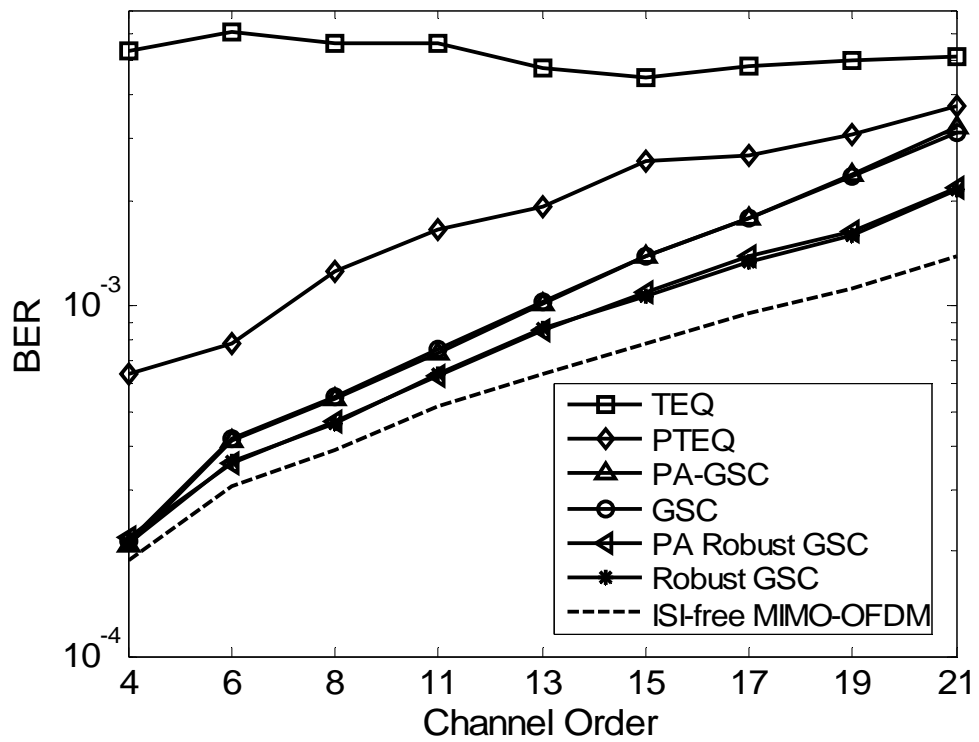
Figure 3.6. BER performances of the three methods at various channel orders with $\mathrm{SNR} = 20$ dB (LS channel estimate).
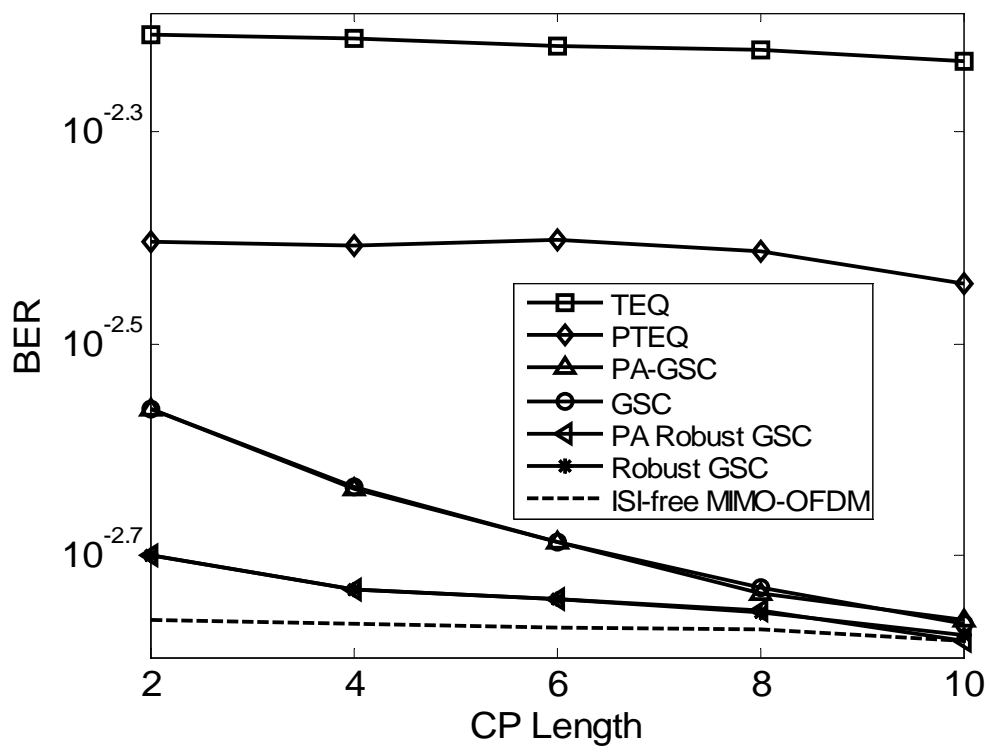
Figure 3.7. BER performances of the three methods at various CP lengths with $\mathrm{SNR} = 20$ dB (LS channel estimate).
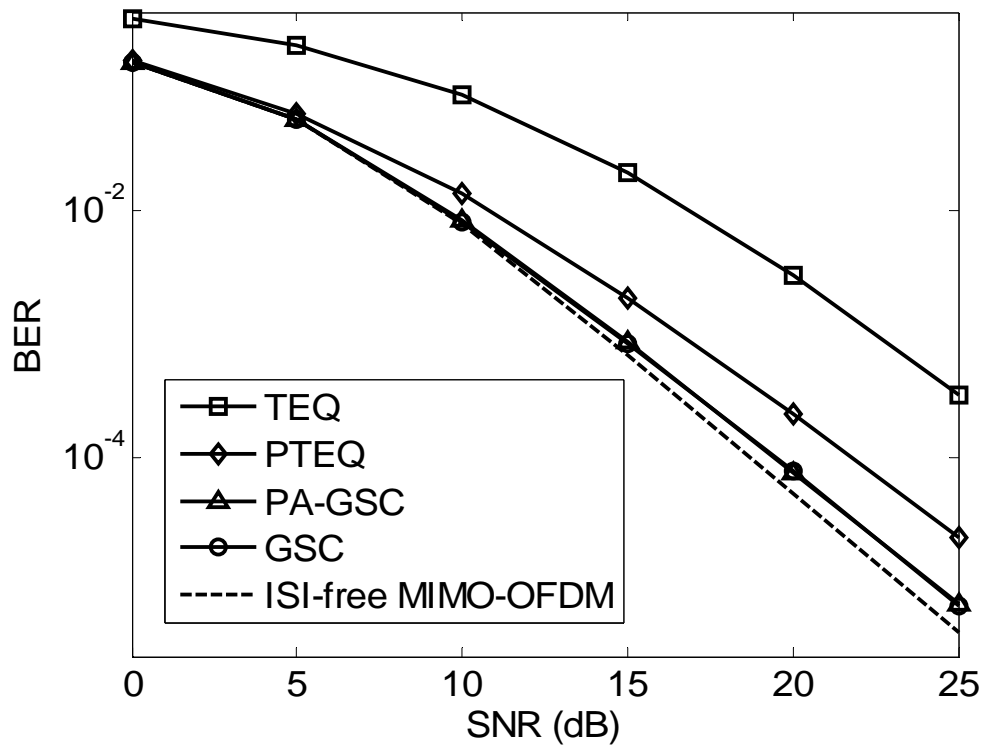
Figure 3.8. BER performances of the three methods with distinct subchannel orders (perfect channel knowledge).
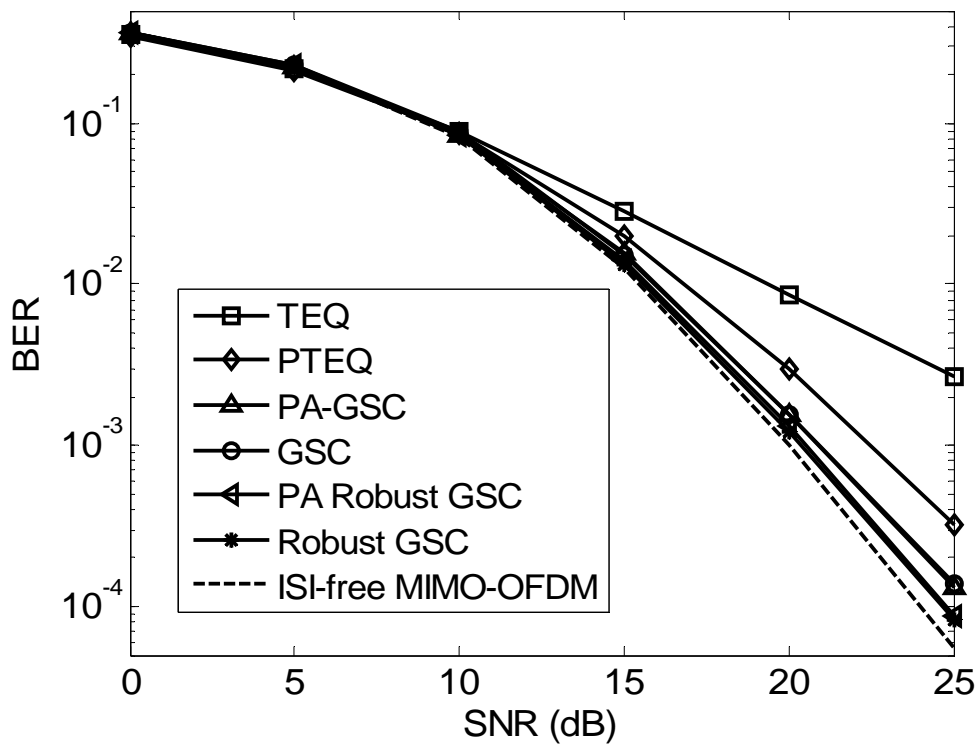
Figure 3.9. BER performances of the three methods with distinct subchannel orders (LS channel estimate).
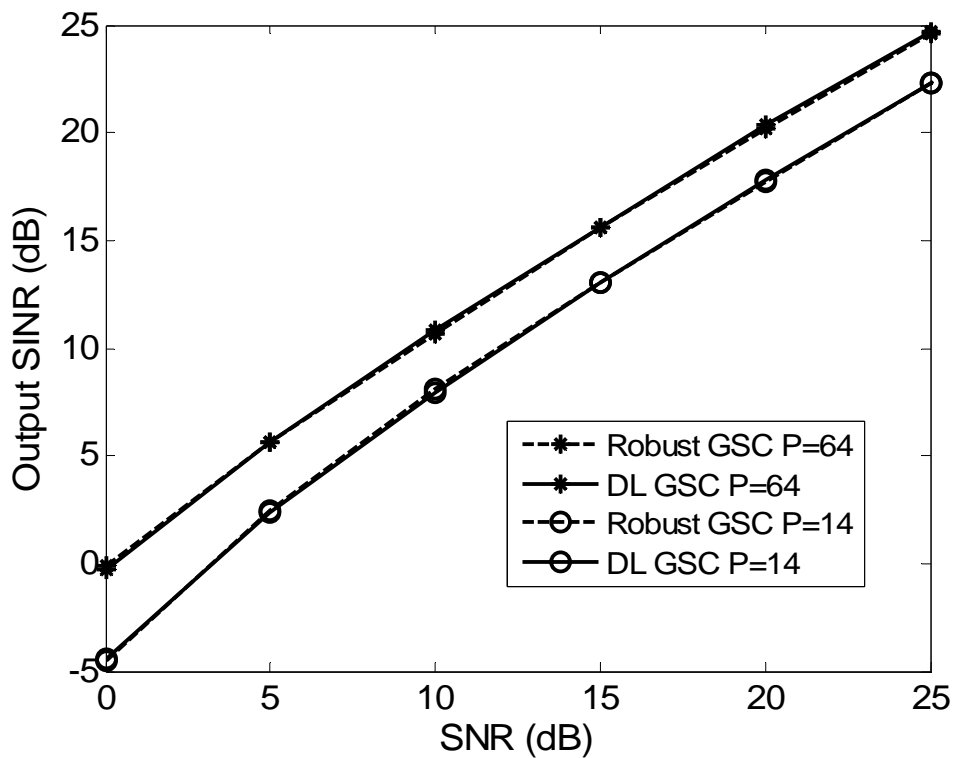
Figure 3.10. SINR performances of the optimal and suboptimal DL solutions.
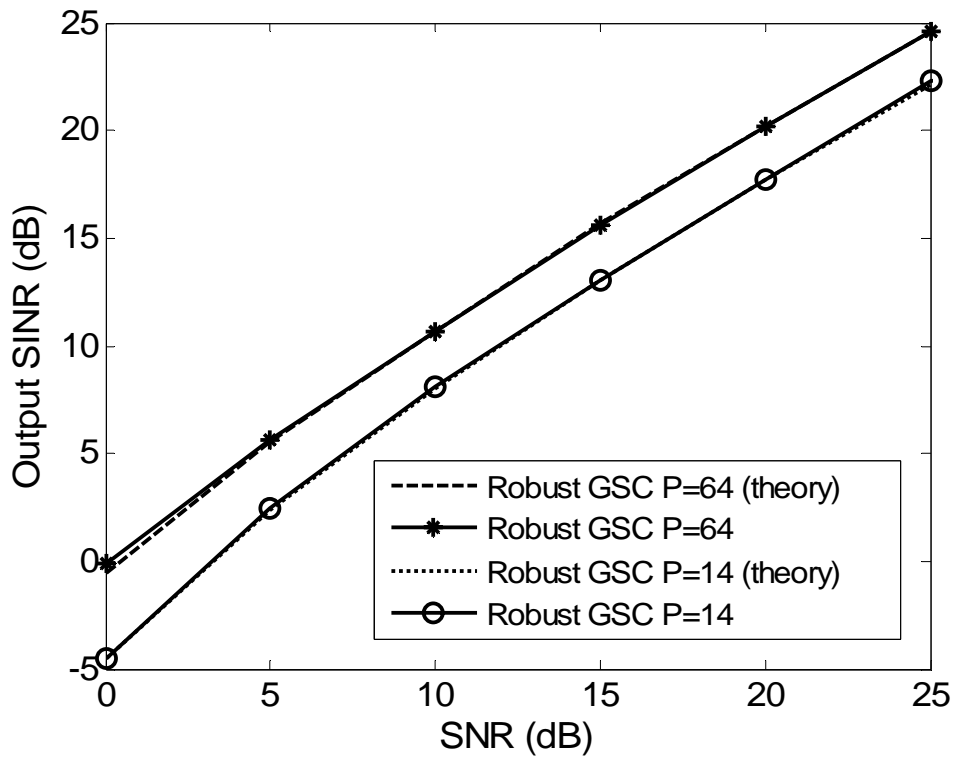
Figure 3.11. Output SINR at different transmit powers in the training phase.
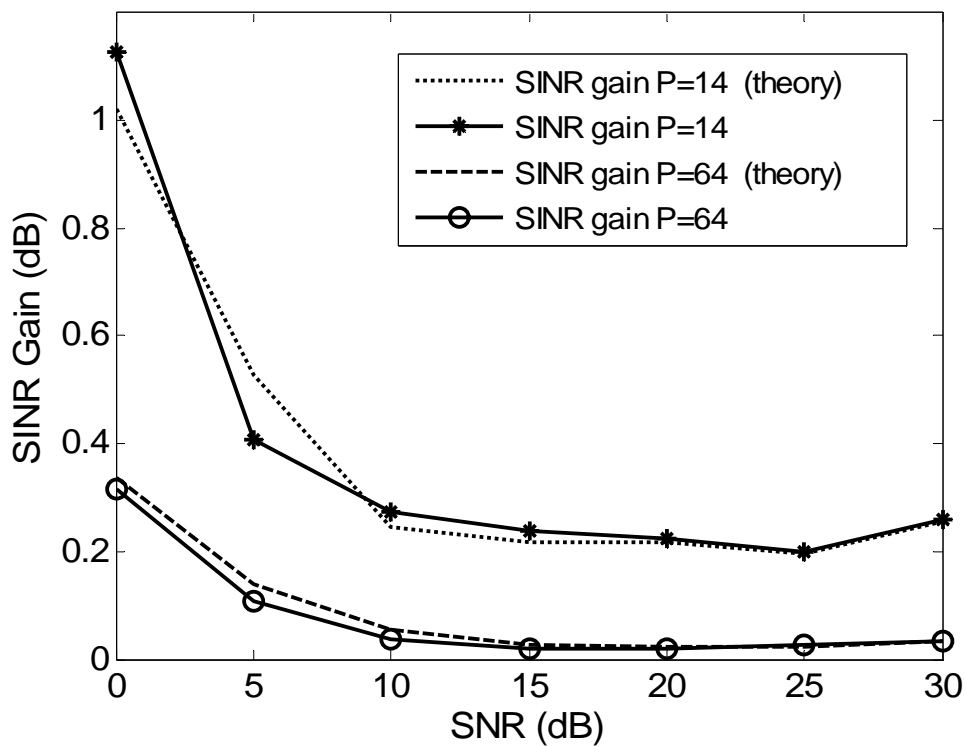
Figure 3.12. SINR gain at different transmit powers in the training phase.
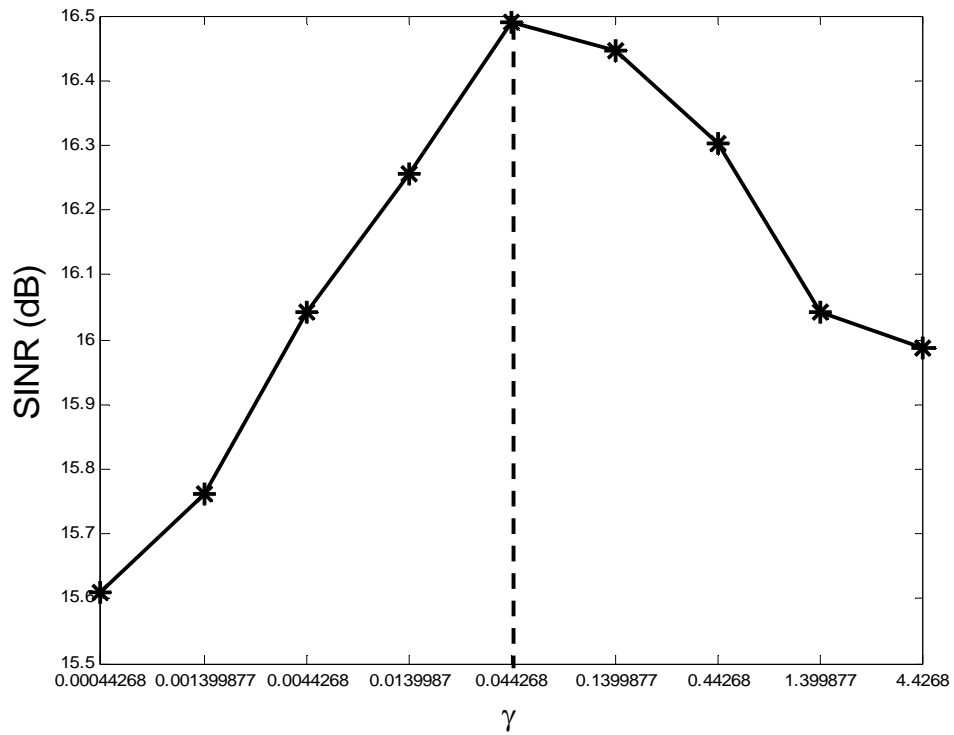
Figure 3.13. Output SINR at different regularization factors γ ($P = 2$).
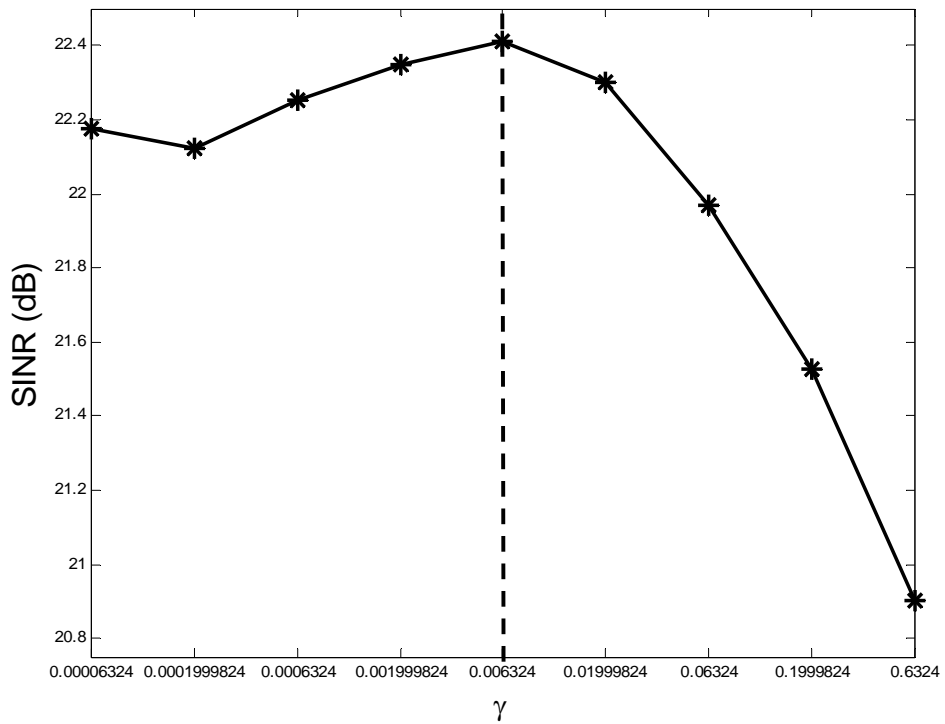
Figure 3.14. Output SINR at different regularization factors γ ($P = 14$).

Chapter 4

Robust Receiver Design for MIMO SC-CP over Time-Varying Dispersive Channels Under Imperfect Channel Knowledge

4.1 Overview

Channel temporal variation is also known to be a vital detrimental factor to the block transmission systems. Even with sufficient CP insertion to avoid ISI, additional ICI is also induced due to loss of orthogonality among signals in frequency domain. This chapter proposes a robust receiver design scheme for MIMO SC-CP systems when the multipath channels undergo time selectivity, and are not exactly known but estimated via the LS training technique [6], [14]. In lieu of relying on signal recovery in the frequency domain, we exploit the constrained optimization formulation [56], [72], based on the orthogonality structure of the time-domain channel matrix, to design an associated group-wise detection framework. To further take into account the mitigation of channel mismatch due to the channel temporal variation and the estimation errors, we leverage the GSC principle [23], [56], [74] to transform the constrained optimization problem into an equivalent unconstrained setup. It will be also shown that the perturbation analysis framework for the GSC filter proposed in Chapter 3 allows us *1)* to explicitly model the channel mismatch effect into the system equations and *2)* to exploit the underlying statistical assumptions on the channel mismatch to derive an *analytic* solution. The proposed scheme also can be combined with the SIC mechanism for further performance enhancement.

4.2 Preliminary

4.2.1 System Model and Basic Assumptions

Consider the discrete-time baseband model of a MIMO SC-CP system with N transmit antennas, M receive antennas, CP length G, and symbol block size Q. Let $h_{m,n}(k,l)$, $0 \le l \le L$, be the lth tap of the channel between the nth and mth transmit-receive antenna pair at time instant k, where L denotes the delay spread assumed common to all MN subchannels. Assuming $G \ge L$, after CP removal the received symbol block at the mth receive antenna can be expressed as

$$\mathbf{r}_m(t) = \sum_{n=1}^{N} \mathbf{H}_{m,n}(t)\mathbf{s}_n(t) + \mathbf{v}_m(t), \tag{4.1}$$

where $\mathbf{s}_n(t) := [s_{n,1}(t),...,s_{n,Q}(t)]^T \in \mathbb{C}^Q$ is the tth symbol block sent from the nth transmit antenna, $\mathbf{v}_m(t) \in \mathbb{C}^Q$ is the channel noise vector, and $\mathbf{H}_{m,n}(t) \in \mathbb{C}^{Q \times Q}$ is the channel matrix whose ith column, denoted by $\mathbf{c}_{m,n}^{(i)}(t)$, is

$$\mathbf{c}_{m,n}^{(i)}(t) = \mathbf{J}^{i-1}\left[\mathbf{h}_{m,n}^{(i)T}(t)\ 0\ \cdots\ 0\right]^T, 1 \le i \le Q, \tag{4.2}$$

in which $\mathbf{h}_{m,n}^{(i)}(t) := [h_{m,n}(\tilde{t} + (i-1)_Q, 0)\ h_{m,n}(\tilde{t} + (i)_Q, 1) \cdots h_{m,n}(\tilde{t} + (L + i - 1)_Q, L)]^T$ with $(\cdot)_Q$ denoting the modulo-Q operation and $\tilde{t} := (t-1)(Q+G) + G$, and

$$\mathbf{J} := \begin{bmatrix} \mathbf{0}_{1\times(Q-1)} & 1 \\ \mathbf{I}_{Q-1} & \mathbf{0}_{(Q-1)\times1} \end{bmatrix}. \tag{4.3}$$

In case that the channel is time-invariant, i.e., $h_{m,n}(k,l) = h_{m,n}(l)$ for some $h_{m,n}(l)$, $\mathbf{H}_{m,n}(t)$ is a circulant matrix and symbol recovery can be done via the tone-by-tone FDE technique [15], [85]. In the considered time-varying channel environment, $\mathbf{H}_{m,n}(t)$ is no longer circulant and the FDE facility is negated. Since there are no specific advantages of processing the data in the frequency domain, we will instead focus on the time-domain signal model (4.1) for receiver design; as will be shown next this can lead to an effective framework for addressing the robust signal recovery problem against imperfect channel knowledge. The following assumptions are made in the sequel.

(1) The number of receive antennas is equal to or greater than the number of transmit antennas, i.e., $M \ge N$.

(2) The source symbols of each transmit antenna $s_{n,q}(t)$ is zero mean, unit-variance, and

$E\{s_{n_1,q_1}(t_1)s_{n_2,q_2}(t_2)^*\} = \delta(n_1 - n_2)\delta(t_1 - t_2)\delta(q_1 - q_2)$, where $\delta(\cdot)$ is the Kronecker delta.

(3) The elements of $\mathbf{v}_m(t)$'s are i.i.d. complex circular Gaussian with zero mean and variance σ_v^2.

4.2.2 Time-Varying Channel Estimation & Equalization

The burst-by-burst transmission is considered such that *1)* each data burst consists of T symbol blocks, *2)* the leading block per burst serves as the training symbol for channel estimation, and *3)* the resultant channel estimate is used for receiver design to recover the subsequent $T-1$ source symbol blocks. Since channel estimation and equalization is done on a burst-wise basis, in the sequel we shall focus on the initial burst. The time-varying channel estimation scheme adopted in this chapter is briefly reviewed as below; the robust equalizer design which exploits the channel error characteristics will be discussed in Section 4.4.

We assume that the MIMO channel is estimated by using the LS training technique [6], [14], which, in a time-varying environment, is known to yield the optimal estimate of the "averaged" channel impulse response within one symbol duration [44], namely,

$$\mathbf{h}_{m,n}^{(av)} := [(1/Q)\sum_{q=0}^{Q-1} h_{m,n}(G+q,0) \cdots (1/Q)\sum_{q=0}^{Q-1} h_{m,n}(G+q,L)]^T. \tag{4.4}$$

Associated with each $1 \le m \le M$ let us define $\mathbf{h}_m^{(av)} = [\mathbf{h}_{m,1}^{(av)T} \cdots \mathbf{h}_{m,N}^{(av)T}]^T$, the resultant channel estimate is given by [44]

$$\hat{\mathbf{h}}_m^{(av)} = \mathbf{h}_m^{(av)} + \underbrace{\mathbf{A}^+(\mathbf{i}_m + \mathbf{v}_m)}_{:=\Delta\mathbf{h}_m} \in \mathbb{C}^{N(L+1)\times 1}, \ 1 \le m \le M, \tag{4.5}$$

where $\mathbf{A} = [diag\{\mathbf{t}_1\}\mathbf{F}_L,...,diag\{\mathbf{t}_N\}\mathbf{F}_L] \in \mathbb{C}^{Q\times N(L+1)}$ with $\mathbf{F}_L/\sqrt{Q}$ being the first $L+1$ columns of the $Q \times Q$ FFT matrix, $\mathbf{A}^+$ denotes the pseudoinverse of $\mathbf{A}$, $\mathbf{t}_n \in \mathbb{C}^Q$ is the frequency-domain training sequence for the nth transmit antenna, and $\mathbf{i}_m = \sum_{n=1}^{N} \mathbf{i}_{m,n} \in \mathbb{C}^Q$ with the qth entry [44]

$$[\mathbf{i}_{m,n}]_q = \sum_{q_1=1}^{Q-1}\sum_{l=0}^{L+1}\left[(1/Q)\sum_{q_2=0}^{Q-1} h_{m,n}(G+q_2,l)e^{(-j2\pi(q-q_1)l/Q)}\right][\mathbf{t}_n]_{(q-q_1)_Q} e^{-j2\pi q_1 q_2/Q}, 1 \le q \le Q. \tag{4.6}$$

Assuming that the channel variation is piecewise linear in time, $\hat{\mathbf{h}}_{m,n}^{(av)}$ can also be treated as the

optimal estimate of the channel parameters in the middle instant of the training period, i.e., $h_{m,n}(G + Q/2, l)$, $0 \le l \le L$ [44]. For fixed $1 \le m \le M$ and $1 \le q \le Q$ we assume that $h_{m,n}(q, l)$, $\forall n, l$, are independent circular complex Gaussian variables with zero-mean and variance σ_l^2. Then the channel estimation error $\Delta \mathbf{h}_m$ in (4.5) is zero mean with covariance matrix $\mathbf{R}_{\Delta\mathbf{h},m} := \mathbf{A}^+ \left(E\{\mathbf{i}_m \mathbf{i}_m^H\} + \sigma_v^2 \mathbf{I}_{N(L+1)} \right) \left(\mathbf{A}^+ \right)^H \in \mathbb{C}^{N(L+1) \times N(L+1)}$, in which the (q, q') th entry of $E\{\mathbf{i}_m \mathbf{i}_m^H\}$, $1 \le q, q' \le Q$, can be directly computed from (4.6) as

$$\left[E\{\mathbf{i}_m \mathbf{i}_m^H\} \right]_{q,q'} = \frac{1}{Q^2} \sum_{n=1}^{N} \sum_{\substack{q_1=0, q_2=0 \\ q_1 \ne q, q_2 \ne q'}}^{Q-1} \sum_{l=0}^{L} \sum_{q_3, q_4=0}^{Q-1} \mathbf{R}_h(q_3 - q_4, l) e^{\frac{-j2\pi\left[(q-q_1)-(q'-q_2)\right]l}{Q}} \left[\mathbf{t}_n \right]_{(q-q_1)_Q} \left[\mathbf{t}_n \right]_{(q'-q_2)_Q}^* e^{\frac{-j2\pi(q_1-q_2)l}{Q}},$$

(4.7)

where $\mathbf{R}_h(q, l) = E\{h_{m,n}(k, l) h_{m,n}^*(k + q, l)\} = \sigma_l^2 J_0(2\pi f_d q T_s)$, with J_0, f_d, and T_s respectively denoting the zeroth-order Bessel function, Doppler frequency, and sampling period. Moreover, assuming *i)* the taps among subchannels are mutually independent, and *ii)* the noise is spatially uncorrelated, we have

$$\mathbf{R}_{\Delta\mathbf{h},m_1,m_2} := E\{\Delta\mathbf{h}_{m_1} \Delta\mathbf{h}_{m_2}^H\} = \mathbf{0}_{N(L+1)}.$$

(4.8)

The results (4.7) and (4.8) will be used for robust equalizer design in Section 4.4.

4.3 Group-Wise Symbol Detection: Perfect Channel Knowledge

4.3.1 Motivation

This subsection highlights the motivation behind the proposed approach. Particularly, we will show that the time-domain channel matrix $\mathbf{H}_{m,n}(t)$ in (4.2) is imbedded with certain column-wise orthogonality structure; such an appealing feature will naturally lead to an inter-group interference cancellation framework followed by a low-complexity intra-group symbol recovery scheme. We shall first focus on the ideal case that the channel is perfectly known at the receiver, and will discuss more realistic situations in Section 4.4.

To proceed, we first observe from (4.2) that, for $2 \le i \le Q$, the ith column $\mathbf{c}_{m,n}^{(i)}(t)$ of the channel matrix $\mathbf{H}_{m,n}(t)$ is simply an $(i-1)$-step down-shifted version of the zero-padded channel impulse response vector $[\mathbf{h}_{m,n}^{(i)T}(t)\ 0 \cdots 0]^T$. As a result, if the symbol block size Q is chosen

to be an integer multiple of $L+1$, i.e., $Q = P(L+1)$ for some positive integer P, then for each fixed $1 \le i \le L+1$ we have

$$\mathbf{c}_{m,n}^{(i+p_1(L+1))}(t)^H \mathbf{c}_{m,n}^{(i+p_2(L+1))}(t) = 0 \quad \text{for} \quad 0 \le p_1 \ne p_2 \le P-1, \tag{4.9}$$

since the locations of the respective nonzero entries never overlap; a schematic description of such an orthogonality relation is depicted in Figure 4.1.
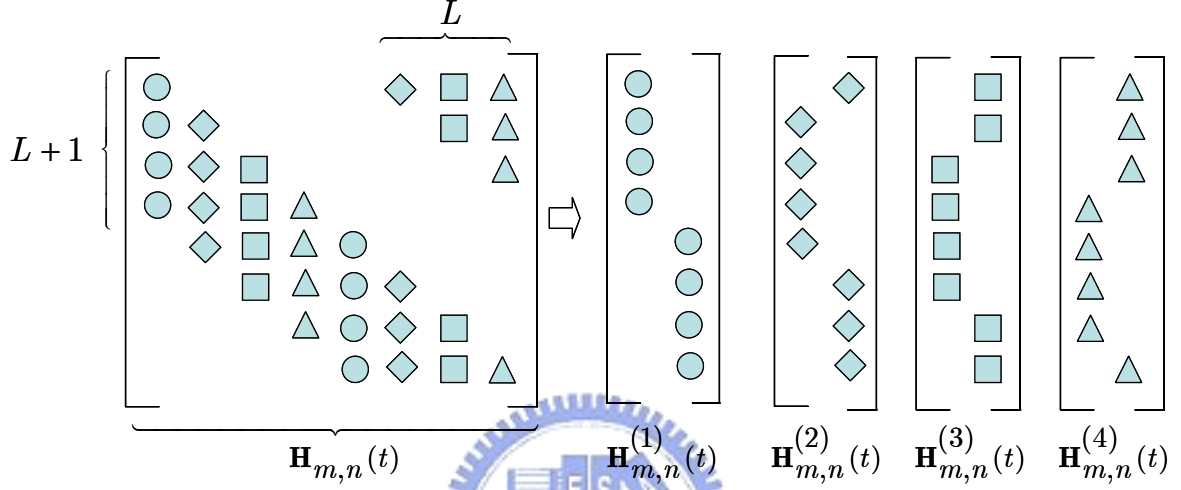
Figure 4.1. A schematic description of the orthogonality condition (4.9) with $L = 3$.

Equation (4.9) suggests that the Q columns of $\mathbf{H}_{m,n}(t)$ can be divided into $L+1$ groups of orthogonal vectors as follows

$$\mathbf{H}_{m,n}^{(i)}(t) := \left[\mathbf{c}_{m,n}^{(i)}(t) \quad \mathbf{c}_{m,n}^{(i+(L+1))}(t) \cdots \mathbf{c}_{m,n}^{(i+(P-1)(L+1))}(t) \right] \in \mathbb{C}^{Q \times P}, \quad 1 \le i \le L+1. \tag{4.10}$$

To further exploit the benefit from the orthogonality condition (4.9), for each fixed $1 \le i \le L+1$ let us stack $\mathbf{H}_{m,n}^{(i)}(t)$ for all $1 \le m \le M$ and $1 \le n \le N$ to form

$$\mathbf{H}_n^{(i)}(t) := \left[\mathbf{H}_{1,n}^{(i)}(t)^T \quad \mathbf{H}_{2,n}^{(i)}(t)^T \quad \cdots \quad \mathbf{H}_{M,n}^{(i)}(t)^T \right]^T \in \mathbb{C}^{MQ \times P},$$

and

$$\mathbf{H}^{(i)}(t) := \left[\mathbf{H}_1^{(i)}(t) \quad \mathbf{H}_2^{(i)}(t) \quad \cdots \quad \mathbf{H}_N^{(i)}(t) \right] \in \mathbb{C}^{MQ \times NP}, \quad 1 \le i \le L+1. \tag{4.11}$$

Collecting the M received signal blocks $\mathbf{r}_m$'s in (4.1) into a vector, the overall input-output relation can be rearranged as

$$\mathbf{r}(t) := \left[\mathbf{r}_1(t)^T \cdots \mathbf{r}_M(t)^T\right]^T = \sum_{i=1}^{L+1} \mathbf{H}^{(i)}(t)\mathbf{s}^{(i)}(t) + \underbrace{\left[\mathbf{v}_1(t)^T \cdots \mathbf{v}_M(t)^T\right]^T}_{:=\mathbf{v}(t)}, \tag{4.12}$$

where $\mathbf{s}^{(i)}(t) := \left[\mathbf{s}_1^{(i)}(t)^T \cdots \mathbf{s}_N^{(i)}(t)^T\right]^T \in \mathbb{C}^{NP}$,

$\mathbf{s}_n^{(i)}(t) := \left[s_{n,i}(t) \ s_{n,i+(L+1)}(t) \cdots s_{n,i+(L+1)(P-1)}(t)\right]^T \in \mathbb{C}^P$. Toward symbol extraction based on (4.12)
we propose to first design an *inter-group interference suppression* matrix $\mathbf{W}^{(j)}(t) \in \mathbb{C}^{MQ \times NQ}$ such
that

$$\mathbf{r}^{(j)}(t) := \mathbf{W}^{(j)}(t)^H \mathbf{r}(t) \approx \mathbf{H}^{(j)}(t)^H \mathbf{H}^{(j)}(t)\mathbf{s}^{(j)}(t) = \begin{bmatrix} \mathbf{H}_1^{(j)}(t)^H \mathbf{H}_1^{(j)}(t) & \cdots & \mathbf{H}_1^{(j)}(t)^H \mathbf{H}_N^{(j)}(t) \\ \vdots & \ddots & \vdots \\ \mathbf{H}_N^{(j)}(t)^H \mathbf{H}_1^{(j)}(t) & \cdots & \mathbf{H}_N^{(j)}(t)^H \mathbf{H}_N^{(j)}(t) \end{bmatrix} \mathbf{s}^{(j)}(t), \tag{4.13}$$

where the last equality in (4.13) follows directly from (4.11), and then recover $\mathbf{s}^{(j)}(t)$ via the
signal model (4.13). Benefiting from the orthogonality condition (4.9), the matched-filtered channel
matrix $\mathbf{H}^{(j)}(t)^H \mathbf{H}^{(j)}(t)$ in (4.13) exhibits an appealing structure. Indeed, since the nonzero entries
among the P columns of $\mathbf{H}_{m,n}^{(i)}(t)$ in (4.10) do not overlap and, for each $1 \le i \le L+1$, $\mathbf{H}_n^{(i)}(t)$
is obtained by stacking $\mathbf{H}_{m,n}^{(i)}(t)$ over all $1 \le m \le M$, it is easy to check that

$$\mathbf{H}_{n_1}^{(i)}(t)^H \mathbf{H}_{n_2}^{(i)}(t) = \mathbf{D}_{n_1 n_2}^{(i)}(t), \text{ where } \mathbf{D}_{n_1 n_2}^{(i)}(t) \in \mathbb{C}^{P \times P} \text{ is diagonal, } 1 \le n_1, n_2 \le N. \tag{4.14}$$

Equation (4.14) implies that, for subsequent *intra-group symbol recovery* through separating the
NP coupled streams in (4.13), the problem is reduced to solving a set of P independent linear
equations, each with dimension $N \times N$. Such a decoupled nature reduces computations, especially
when the block length Q (and hence $P = Q/(L+1)$) is large; it can also limit the error
propagation effect in the symbol recovery stage. The key challenge of the proposed two-stage
equalization scheme is the design of the interference suppression matrix $\mathbf{W}^{(j)}(t)$ for fulfilling
(4.13); the underlying mathematical formulation is discussed next.

4.3.2 Solution Based on Constrained Optimization Technique

Based on (4.13), the linear weighting matrix $\mathbf{W}^{(j)}(t)$ for recovering the *j*th symbol group
should be designed to minimize the inter-group interference power, and then extract the desired

signal component through space-time matched filtering. A typical technique toward fulfilling such a two-fold task is through constrained optimization [56], [72]; more precisely, we shall design $\mathbf{W}^{(j)}(t)$ by solving the following problem

$$\min_{\mathbf{W}^{(j)}(t)} E\left\{\left\|\mathbf{W}^{(j)}(t)^H\left(\sum_{i=1,i\neq j}^{L+1}\mathbf{H}^{(i)}(t)\mathbf{s}^{(i)}(t)+\mathbf{v}(t)\right)\right\|^2\right\}, \text{ s.t. } \mathbf{W}^{(j)}(t)^H\mathbf{H}^{(j)}(t)=\mathbf{H}^{(j)}(t)^H\mathbf{H}^{(j)}(t); \quad (4.15)$$

the recovery of the entire symbol vector $\mathbf{s}(t)$ is then done group-wise via repeatedly solving (4.15) for $1 \leq j \leq L+1$. By using the standard Lagrangian multiplier technique [8, p-215], the solution to (4.15) is given by

$$\mathbf{W}^{(j)}(t) = \mathbf{R}_I^{(j)}(t)^{-1}\mathbf{H}^{(j)}(t)\left(\mathbf{H}^{(j)}(t)^H\mathbf{R}_I^{(j)}(t)^{-1}\mathbf{H}^{(j)}(t)\right)^{-1}\mathbf{H}^{(j)}(t)^H\mathbf{H}^{(j)}(t), \quad (4.16)$$

where $\mathbf{R}_I^{(j)}(t) := \sum_{i=1,i\neq j}^{L+1}\mathbf{H}^{(i)}(t)\mathbf{H}^{(i)}(t)^H + \sigma_n^2\mathbf{I}_{MQ}$. An alternative approach to solving (4.15) lies in transforming the constrained optimization formulation into an equivalent *unconstrained* setup via the GSC principle [23]. This relies on the following decomposition of $\mathbf{W}^{(j)}(t)$:

$$\mathbf{W}^{(j)}(t) = \mathbf{H}^{(j)}(t) - \mathbf{B}^{(j)}(t)\mathbf{U}^{(j)}(t), \quad (4.17)$$

where $\mathbf{H}^{(j)}(t) \in \mathbb{C}^{MQ\times NP}$ represents the *non-adaptive* portion which verifies the desired space-time matched filtering constraint, $\mathbf{B}^{(j)}(t) \in \mathbb{C}^{MQ\times(MQ-NP)}$ is the signal blocking matrix with $\mathbf{B}^{(j)}(t)^H\mathbf{H}^{(j)}(t) = \mathbf{0}_{(MQ-NP)\times NP}$, and $\mathbf{U}^{(j)}(t) \in \mathbb{C}^{(MQ-NP)\times NP}$ is the *adaptive component* which forms the remaining free parameters to be determined. By following the standard procedures presented in Section 2.4.3, the solution of $\mathbf{U}^{(j)}(t)$ is given by

$$\mathbf{U}_{opt}^{(j)}(t) = \left(\mathbf{B}^{(j)}(t)^H\mathbf{R}_I^{(j)}(t)\mathbf{B}^{(j)}(t)\right)^{-1}\mathbf{B}^{(j)}(t)^H\mathbf{R}_I^{(j)}(t)\mathbf{H}^{(j)}(t); \quad (4.18)$$

the resultant optimal GSC weight is thus

$$\mathbf{W}_{opt}^{(j)}(t) = \mathbf{H}^{(j)}(t) - \mathbf{B}^{(j)}(t)\mathbf{U}_{opt}^{(j)}(t) := \mathbf{H}^{(j)}(t) - \mathbf{B}^{(j)}(t)\left(\mathbf{B}^{(j)}(t)^H\mathbf{R}_I^{(j)}(t)\mathbf{B}^{(j)}(t)\right)^{-1}\mathbf{B}^{(j)}(t)\mathbf{R}_I^{(j)}(t)\mathbf{H}^{(j)}(t).$$

$$(4.19)$$

We note that *i)* solutions (4.16) and (4.19) are obtained based on the crucial perfect channel knowledge assumption; when channel parameter mismatch occurs due to imperfect estimation and time variation, they are only suboptimal because the formulations do not take into account channel error mitigation, *ii)* as will be seen in the next section, the GSC framework is advantageous in that it

allows us to directly model the channel mismatch effect into the system equations for facilitating robust equalizer design.

4.4 Proposed Robust GSC Equalizer Against Channel Mismatch

Given only the channel estimate $\widehat{\mathbf{H}}_{av}^{(i)}$, $1 \leq i \leq L+1$, which is acquired through training at $t = 1$, one may simply choose to modify solution (4.16) via replacing $\mathbf{H}^{(i)}(t)$ by $\widehat{\mathbf{H}}_{av}^{(i)}$ to get the time-invariant equalizer

$$\widehat{\mathbf{W}}^{(j)} = (\widehat{\mathbf{R}}_I^{(j)})^{-1} \widehat{\mathbf{H}}_{av}^{(j)} \left(\widehat{\mathbf{H}}_{av}^{(j)H} (\widehat{\mathbf{R}}_I^{(j)})^{-1} \widehat{\mathbf{H}}_{av}^{(j)} \right)^{-1} \widehat{\mathbf{H}}_{av}^{(j)H} \widehat{\mathbf{H}}_{av}^{(j)}, \tag{4.20}$$

where $\widehat{\mathbf{R}}_I^{(j)} = \sum_{i=1,i \neq j}^{L+1} \widehat{\mathbf{H}}_{av}^{(i)} \widehat{\mathbf{H}}_{av}^{(i)H} + \sigma_n^2 \mathbf{I}_{MQ}$, and use (4.20) once and for all toward subsequent symbol recovery. Such a strategy, even though quite simple, fails to combat channel mismatch and could incur very poor equalization performance. The GSC principle, on the contrary, provides a very natural and effective framework for robust time-varying equalizer design, as is shown in this section. We will first introduce the problem formulation in Section 4.4.1, and then derive the solution in Section 4.4.2. Some discussions regarding the proposed robust scheme are given in Section 4.4.3.

4.4.1 Problem Formulation

Recall that the mechanism of GSC filter (4.17) basically involves space-time signal matched filtering and signal blocking followed by interference suppression. While the signal combining and nulling components ($\mathbf{H}^{(j)}(t)$ and $\mathbf{B}^{(j)}(t)$, respectively) are immediately fixed upon the (possibly imperfect) knowledge of $\mathbf{H}^{(j)}(t)$, the adaptive portion $\mathbf{U}^{(j)}(t)$ can nonetheless be allowed to be time-varying, and is then designed to tackle the channel parameter mismatch effects caused by time variation and estimation error. It is such inherent channel tracking capability that makes GSC principle a promising approach in the considered scenario.

Specifically, when only a channel estimate $\widehat{\mathbf{H}}_{av}^{(j)}$ is available, exact signal matched filtering over $2 \leq t \leq T$ is impossible; the best one can do, however, is to linearly combine $\mathbf{H}^{(j)}(t)\mathbf{s}^{(j)}(t)$ just with $\widehat{\mathbf{H}}_{av}^{(j)}$ to get the approximation $\widehat{\mathbf{H}}_{av}^{(j)H} \mathbf{H}^{(j)}(t)\mathbf{s}^{(j)}(t)$. This implies that the non-adaptive portion of the GSC weight should be set as $\widehat{\mathbf{H}}_{av}^{(j)}$ throughout $2 \leq t \leq T$ and, in turn, the blocking

matrix is likewise fixed according to the relation $\widehat{\mathbf{B}}_{av}^{(j)H}\widehat{\mathbf{H}}_{av}^{(j)} = \mathbf{0}_{(MQ-NP)\times NP}$. Hence, given the channel knowledge $\widehat{\mathbf{H}}_{av}^{(j)}$ only, the signal matching and blocking matrices are restricted to be time-invariant. Nevertheless, to reliably recover the desired signal against background time-varying interference, the adaptive component must account for the temporal variation. This thus suggests the following modified GSC decomposition

$$\widehat{\mathbf{W}}^{(j)}(t) = \widehat{\mathbf{H}}_{av}^{(j)} - \widehat{\mathbf{B}}_{av}^{(j)}\overline{\mathbf{U}}^{(j)}(t). \tag{4.21}$$

With (4.21), the equalized output instead reads

$$\overline{\mathbf{z}}^{(j)}(t) = \widehat{\mathbf{W}}^{(j)}(t)^{H}\mathbf{r}(t) = \overline{\mathbf{z}}_{d}^{(j)}(t) - \overline{\mathbf{U}}^{(j)}(t)^{H}\overline{\mathbf{z}}_{b}^{(j)}(t), \tag{4.22}$$

where

$$\overline{\mathbf{z}}_{d}^{(j)}(t) := \widehat{\mathbf{H}}_{av}^{(j)H}\mathbf{H}^{(j)}(t)\mathbf{s}^{(j)}(t) + \underbrace{\widehat{\mathbf{H}}_{av}^{(j)H}\sum_{i=1,i\neq j}^{L+1}\mathbf{H}^{(i)}(t)\mathbf{s}^{(i)}(t) + \widehat{\mathbf{H}}_{av}^{(j)H}\mathbf{v}(t)}_{:=\overline{\mathbf{i}}_{d}^{(j)}(t)}, \tag{4.23}$$

is the (approximate) space-time matched filtered signal, and

$$\overline{\mathbf{z}}_{b}^{(j)}(t) := \widehat{\mathbf{B}}_{av}^{(j)H}\mathbf{H}^{(j)}(t)\mathbf{s}^{(j)}(t) + \widehat{\mathbf{B}}_{av}^{(j)H}\sum_{i=1,i\neq j}^{L+1}\mathbf{H}^{(i)}(t)\mathbf{s}^{(i)}(t) + \widehat{\mathbf{B}}_{av}^{(j)H}\mathbf{v}(t) \tag{4.24}$$

is the corresponding blocking component. Due to the out-of-date channel knowledge, the signal of interest in the $\overline{\mathbf{z}}_{d}^{(j)}(t)$ is non-coherently combined and is corrupted by the interference $\overline{\mathbf{i}}_{d}^{(j)}(t)$; also, since the blocking matrix is determined via $\widehat{\mathbf{B}}_{av}^{(j)H}\widehat{\mathbf{H}}_{av}^{(j)} = \mathbf{0}_{(MQ-NP)\times NP}$, and hence $\widehat{\mathbf{B}}_{av}^{(j)H}\mathbf{H}^{(j)}(t) \neq \mathbf{0}_{(MQ-NP)\times NP}$ in general, there is a signal leakage term $\widehat{\mathbf{B}}_{av}^{(j)H}\mathbf{H}^{(j)}(t)\mathbf{s}^{(j)}(t)$ into the blocking branch $\overline{\mathbf{z}}_{b}^{(j)}(t)$. Toward signal recovery against interference, a natural strategy as suggested by the GSC principle is to treat $\overline{\mathbf{z}}_{b}^{(j)}(t)$ in (4.24) as an aggregate interference and to design $\overline{\mathbf{U}}^{(j)}(t)$ such that $\overline{\mathbf{U}}^{(j)}(t)^{H}\overline{\mathbf{z}}_{b}^{(j)}(t)$ is best close to $\overline{\mathbf{i}}_{d}^{(j)}(t)$.

For this we shall first note that in (4.23) and (4.24) only the channel estimate $\widehat{\mathbf{H}}_{av}^{(i)}$'s (acquired through training at $t=1$) are available but the true channel matrices $\mathbf{H}^{(i)}(t)$'s are unknown: the mismatches between $\widehat{\mathbf{H}}_{av}^{(i)}$'s and $\mathbf{H}^{(i)}(t)$'s are due to time variation as well as channel estimation errors. To facilitate subsequent analysis we must seek for explicit rules linking the unknown $\mathbf{H}^{(i)}(t)$ to the channel estimate $\widehat{\mathbf{H}}_{av}^{(i)}$. A commonly used model which specifies such channel

parameter deviation can be found in [31] and [47], and in terms of the current matrix formulation it reads

$$\mathbf{H}^{(i)}(t) = \mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)} + \delta\mathbf{H}^{(i)}(t), \quad 1 \leq i \leq L+1, \tag{4.25}$$

where $\mathbf{D}_c^{(i)}(t)$ is an $MQ \times MQ$ diagonal matrix depending on the time-varying channel characteristics, and $\delta\mathbf{H}^{(i)}(t)$ models the temporal channel variation whose entries are assumed to be zero-mean Gaussian random variables (explicit formulae for $\mathbf{D}_c^{(i)}(t)$ and the covariance of $\delta\mathbf{H}^{(i)}(t)$ are given in Appendix E). By using (4.25) we can rewrite (4.23) and (4.24) as

$$\overline{\mathbf{z}}_d^{(j)}(t) = \widehat{\mathbf{H}}_{av}^{(j)H}\mathbf{D}_c^{(j)}(t)\widehat{\mathbf{H}}_{av}^{(j)}\mathbf{s}^{(j)}(t) + \underbrace{\widehat{\mathbf{H}}_{av}^{(j)H}\sum_{i=1,i\neq j}^{L+1}\mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\mathbf{s}^{(i)}(t) + \widehat{\mathbf{H}}_{av}^{(j)H}\sum_{i=1}^{L+1}\delta\mathbf{H}^{(i)}(t)\mathbf{s}^{(i)}(t) + \widehat{\mathbf{H}}_{av}^{(j)H}\mathbf{v}(t)}_{:=\overline{\mathbf{i}}^{(j)}(t)} \quad ,$$

$$\tag{4.26}$$

and

$$\overline{\mathbf{z}}_b^{(j)}(t) = \widehat{\mathbf{B}}_{av}^{(j)H}\mathbf{D}_c^{(j)}(t)\widehat{\mathbf{H}}_{av}^{(j)}\mathbf{s}^{(j)}(t) + \widehat{\mathbf{B}}_{av}^{(j)H}\sum_{i=1,i\neq j}^{L+1}\mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\mathbf{s}^{(i)}(t) + \widehat{\mathbf{B}}_{av}^{(j)H}\sum_{i=1}^{L+1}\delta\mathbf{H}^{(i)}(t)\mathbf{s}^{(i)}(t) + \widehat{\mathbf{B}}_{av}^{(j)H}\mathbf{v}(t) \quad .$$

$$\tag{4.27}$$

Based on (4.26) and (4.27), we specifically propose to design $\overline{\mathbf{U}}^{(j)}(t)$ by minimizing the following cost function

$$\overline{J}(t) = E\left\{\left\|\overline{\mathbf{i}}^{(j)}(t) - \overline{\mathbf{U}}^{(j)}(t)^H\overline{\mathbf{z}}_b^{(j)}(t)\right\|^2\right\}, \tag{4.28}$$

where the expectation is taken with respect to the source symbol, background noise, channel estimation error, and channel temporal variation (assuming all are mutually independent).

4.4.2 Optimal Solution

Let us expand (4.28) into

$$\overline{J}(t) = \overline{\mathbf{U}}^{(j)}(t)^H E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{z}}_b^{(j)}(t)^H\right\}\overline{\mathbf{U}}^{(j)}(t) - \overline{\mathbf{U}}^{(j)}(t)^H E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{i}}^{(j)}(t)^H\right\}$$

$$-E\left\{\overline{\mathbf{i}}^{(j)}(t)\overline{\mathbf{z}}_b^{(j)}(t)^H\right\}\overline{\mathbf{U}}^{(j)}(t) + E\left\{\overline{\mathbf{i}}^{(j)}(t)\overline{\mathbf{i}}^{(j)}(t)^H\right\}. \tag{4.29}$$

As in Section 3.4.1, taking the first-order partial derivative of $\overline{J}(t)$ with respect to $\overline{\mathbf{U}}^{(j)}(t)$ yields

$$E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{z}}_b^{(j)}(t)^H\right\}\overline{\mathbf{U}}^{(j)}(t) = E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{i}}^{(j)}(t)^H\right\}. \tag{4.30}$$

The optimal $\overline{\mathbf{U}}^{(j)}(t)$ can then be obtained by solving the matrix equation (4.30), provided that the

covariance matrices $E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{z}}_b^{(j)}(t)^H\right\}$ and $E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{i}}^{(j)}(t)^H\right\}$ are available. With (4.26), (4.27), and by taking expectation with respect to source symbols, channel noise, and channel temporal variation characterized via (4.25), we can reach the following intermediate expressions (see Appendix E for detailed proof):

$$E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{z}}_b^{(j)}(t)^H\right\} = E_e\left\{\widehat{\mathbf{B}}_{av}^{(j)H}\left(\sum_{i=1}^{L+1}\mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\widehat{\mathbf{H}}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)^H + N\mathbf{I}_M\otimes\mathbf{D}_Q(t) + \sigma_v^2\mathbf{I}_{MQ}\right)\widehat{\mathbf{B}}_{av}^{(j)}\right\}, \quad (4.31)$$

and

$$E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{i}}^{(j)}(t)^H\right\} = E_e\left\{\widehat{\mathbf{B}}_{av}^{(j)H}\left(\sum_{i=1,i\neq j}^{L+1}\mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\widehat{\mathbf{H}}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)^H + N\mathbf{I}_M\otimes\mathbf{D}_Q(t) + \sigma_v^2\mathbf{I}_{MQ}\right)\widehat{\mathbf{H}}_{av}^{(j)}\right\},$$

$$(4.32)$$

where $\mathbf{D}_Q(t) := diag\left\{\sum_{l=0}^{L}\sigma_l^2(1-\left|\rho(\overline{t},l)\right|^2),...,\sum_{l=0}^{L}\sigma_l^2(1-\left|\rho(\overline{t}+Q-1,l)\right|^2)\right\}$, with $\rho(k,l) := \mathbf{R}_h(k,l)(\sigma_l\hat{\sigma}_l)^{-1}$, $\hat{\sigma}_l^2 = \sigma_l^2 + [\mathbf{R}_{\Delta\mathbf{h},m}]_{(n-1)(L+1)+l+1,(n-1)(L+1)+l+1}$, $\overline{t} := (t-1)(G+Q)-Q/2$, and $E_e\{\cdot\}$ denotes the expectation involving channel estimation error yet to be carried out. To explicitly determine the expectations in (4.31) and (4.32) we must further seek for tractable relations linking the channel estimates ($\widehat{\mathbf{H}}_{av}^{(i)}$ for $1\leq i\leq L+1$ and $\widehat{\mathbf{B}}_{av}^{(j)}$) and the background estimation errors. While $\widehat{\mathbf{H}}_{av}^{(i)}$ can be directly modeled as the actual channel parameter corrupted by the errors, namely,

$$\widehat{\mathbf{H}}_{av}^{(i)} = \mathbf{H}_{av}^{(i)} + \Delta\mathbf{H}_{av}^{(i)}, \ 1\leq i\leq L+1, \quad (4.33)$$

an exact expression of the blocking component $\widehat{\mathbf{B}}_{av}^{(j)}$ in terms of $\Delta\mathbf{H}_{av}^{(j)}$ remains formidable to characterize, since it is determined through $\widehat{\mathbf{B}}_{av}^{(j)H}\widehat{\mathbf{H}}_{av}^{(i)} = \mathbf{0}_{(MQ-NP)\times NP}$ and, thus, is obtained as an orthonormal basis of the left null space of $\widehat{\mathbf{H}}_{av}^{(i)}$. To resolve this difficulty, we will leverage the perturbation analysis technique [36] to get an approximate, but analytic, relation among $\widehat{\mathbf{B}}_{av}^{(j)}$ and the channel estimation error $\Delta\mathbf{H}_{av}^{(j)}$. Let $\mathbf{H}_{av}^{(j)} = \mathbf{U}_h^{(j)}\Sigma_h^{(j)}\mathbf{V}_h^{(j)H}$ be a singular value decomposition of the actual channel matrix $\mathbf{H}_{av}^{(j)}$. From (3.42), the blocking matrix $\widehat{\mathbf{B}}_{av}^{(j)}$ can be approximated by

$$\widehat{\mathbf{B}}_{av}^{(j)} \approx \mathbf{B}_{av}^{(j)} - \underbrace{\mathbf{U}_h^{(j)}\Sigma_h^{(j)-1}\mathbf{V}_h^{(j)H}\Delta\mathbf{H}_{av}^{(j)H}\mathbf{B}_{av}^{(j)}}_{:=\Delta\mathbf{B}_{av}^{(j)}}. \quad (4.34)$$

By substituting $\widehat{\mathbf{B}}_{av}^{(j)}$ in (4.34) into (4.31) and (4.32), we have the following main results (see Appendix F for detailed derivations).

Proposition 4.1: The covariance matrices involved in (4.30) can be expressed as

$$E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{z}}_b^{(j)}(t)^H\right\} = \mathbf{B}_{av}^{(j)H}\left(\mathbf{R}_L^{(j)}(t) + \overline{\mathbf{R}}_I^{(j)}(t) + \mathbf{R}_{e,1}^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)\right)\mathbf{B}_{av}^{(j)}, \qquad (4.35)$$

$$E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{i}}^{(j)}(t)^H\right\} = \mathbf{B}_{av}^{(j)H}\left(\overline{\mathbf{R}}_I^{(j)}(t) + \mathbf{R}_{e,2}^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)\right)\mathbf{H}_{av}^{(j)}, \qquad (4.36)$$

where

$$\mathbf{R}_L^{(j)}(t) := \mathbf{D}_c^{(j)}(t)\mathbf{H}_{av}^{(j)}\mathbf{H}_{av}^{(j)H}\mathbf{D}_c^{(j)}(t)^H, \qquad (4.37)$$

$$\overline{\mathbf{R}}_I^{(j)}(t) := \sum_{i=1,i\neq j}^{L+1}\mathbf{D}_c^{(i)}(t)\mathbf{H}_{av}^{(i)}\mathbf{H}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)^H + \sigma_v^2\mathbf{I}_{MQ}, \qquad (4.38)$$

and $\mathbf{R}_{e,1}^{(j)}(t) := \sum_{i=1}^4 \mathbf{X}_i(t)$ and $\mathbf{R}_{e,2}^{(j)}(t) := \sum_{k=1}^3 \mathbf{Y}_k(t)$, in which the component matrices $\mathbf{X}_i(t)$'s and $\mathbf{Y}_k(t)$'s are defined in Table 4.1. $\qquad\square$

Based on (4.30), (4.35), and (4.36), the robust solution of $\overline{\mathbf{U}}^{(j)}(t)$, which is *on average* optimal for mitigating channel temporal variation and estimation error, is thus

$$\overline{\mathbf{U}}_{opt}^{(j)}(t) = \left[\mathbf{B}_{av}^{(j)H}\left(\mathbf{R}_L^{(j)}(t) + \overline{\mathbf{R}}_I^{(j)}(t) + \mathbf{R}_{e,1}^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)\right)\mathbf{B}_{av}^{(j)}\right]^{-1} \times$$
$$\mathbf{B}_{av}^{(j)H}\left(\overline{\mathbf{R}}_I^{(j)}(t) + \mathbf{R}_{e,2}^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)\right)\mathbf{H}_{av}^{(j)}. \qquad (4.39)$$

In the practical situation when only a channel estimate is available, the sampled-version of the overall robust GSC weighting matrix is accordingly given by

$$\widehat{\mathbf{W}}_r^{(j)}(t) = \widehat{\mathbf{H}}_{av}^{(j)} - \widehat{\mathbf{B}}_{av}^{(j)}\widehat{\overline{\mathbf{U}}}_{opt}^{(j)}(t). \qquad (4.40)$$

4.4.3 Associated Discussions

1. The proposed design formulation aims for joint mitigation of channel temporal variation and estimation error. Compared with (4.18) obtained under exact channel knowledge, the main distinctive feature of the proposed robust scheme (4.39) lies in replacing the two $\mathbf{R}_I^{(j)}(t)$'s in (4.18) respectively by the "composite" interference covariance matrices $\mathbf{R}_L^{(j)}(t) + \overline{\mathbf{R}}_I^{(j)}(t) + \mathbf{R}_{e,1}^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)$ and $\overline{\mathbf{R}}_I^{(j)}(t) + \mathbf{R}_{e,2}^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)$. We should

note that (i) the common term $\overline{\mathbf{R}}_I^{(j)}(t)$ can be regarded as the analogue of $\mathbf{R}_I^{(j)}(t)$ with the availability of $\mathbf{H}_{av}^{(i)}$'s only (rather than the exact $\mathbf{H}^{(j)}(t)$), (ii) the term $\mathbf{R}_L^{(j)}(t)$ arises from signal leakage into the blocking branch due to channel mismatch, (iii) the matrices $\mathbf{R}_{e,1}^{(j)}(t)$ and $\mathbf{R}_{e,2}^{(j)}(t)$ are due to the aggregate impacts caused by channel estimation errors, and (iv) the quantity $N\mathbf{I}_M \otimes \mathbf{D}_Q(t)$ accounts for the background channel temporal variation.

2. Another scheme for estimating time-varying channels is to treat the channel estimates for two consecutive data bursts as the end points, and further leverage linear interpolation for acquiring the channel information within the entire time frame [44]; this approach applies whenever the channel varies linearly with respect to time. In such a scenario, the impacts due to channel temporal variation would be largely reduced and channel estimation error, instead, becomes the dominant adverse factor to be combated. The proposed design strategy can be used for constructing an associated error-resistant GSC equalizer. Indeed, the cost function (4.28) instead involves only the averages with respect to source, noise, and channel estimation errors; by following the procedures shown in Section 4.4.2, the sampled-version of the resultant robust GSC weight can be obtained by (the derivations are highlighted in Appendix G)

$$\widehat{\mathbf{W}}_a^{(j)}(t) = \widehat{\mathbf{H}}^{(j)}(t) - \left[\widehat{\mathbf{B}}^{(j)}(t)^H \left(\widehat{\mathbf{R}}_I^{(j)}(t) + \widehat{\mathbf{R}}_e(t)\right)\widehat{\mathbf{B}}^{(j)}(t)\right]^{-1} \widehat{\mathbf{B}}^{(j)}(t)^H \widehat{\mathbf{R}}_I^{(j)}(t)\widehat{\mathbf{H}}^{(j)}(t), \quad (4.41)$$

where $\widehat{\mathbf{H}}^{(j)}(t)$ and $\widehat{\mathbf{R}}_I^{(j)}(t)$ respectively denote the estimates of $\mathbf{H}^{(j)}(t)$ and $\mathbf{R}_I^{(j)}(t)$ (through linear interpolation), and $\widehat{\mathbf{R}}_e(t)$ is the estimated version of $\mathbf{R}_e(t)$ defined in Table 4.1.

3. The proposed group-wise symbol recovery scheme can be directly combined with the SIC mechanism for further performance improvement; the resultant solutions in each processing layer is essentially of the form (4.40), except that in $\overline{\mathbf{R}}_I^{(j)}(t)$, $\mathbf{R}_{e,1}^{(j)}(t)$, and $\mathbf{R}_{e,2}^{(j)}(t)$ the signature matrices corresponding to the previously detected signal components are removed. Through simulation, the SIC-based implementation is observed to yield up to a 5 dB SNR gain.

4.5 Algorithm Complexity

The computation of the proposed GSC weighting matrix basically involves solving

$\widehat{\mathbf{B}}_{av}^{(j)H}\widehat{\mathbf{H}}_{av}^{(j)} = \mathbf{0}_{(MQ-NP)\times NP}$ for the blocking matrix $\widehat{\mathbf{B}}_{av}^{(j)}$ followed by inverting an $(MQ-NP)\times(MQ-NP)$ matrix. The complexity for matrix inversion can be further reduced by leveraging the conjugate gradient [24] based reduced-rank (RR) implementation [27]. Table 4.2 summarizes the flop counts (measured in terms of the number of complex-valued additions and multiplications) of the GSC method, GSC with RR implementation as well as two comparative receivers, namely, the layered space-frequency equalization (LSFE) scheme [85] and the group-wise V-BLAST detector [73]. From the table we observe that the complexity of the three methods is roughly comparable. For the system parameters $(N, M, Q, L, L_b, N_b, I) = (2, 2, 64, 7, 1, 2, 10)$ adopted in the simulation section the flop counts are, respectively, $FC_{GSC} \approx 4.8\times 10^6$, $FC_{GSC}^{(RR)} \approx 10^6$, $FC_{LSFE} \approx 2\times 10^6$, and $FC_{GB} \approx 6.8\times 10^6$ (L_b and N_b respectively denote the tap order of feedback filter and the number of decision stages for the LSFE detector, and I the number of iterations involved in the RR implementation).

4.6 Channel Order Determination

To implement the proposed group-wise detection scheme, the information of channel order L is required. Since the conventional information theoretic criteria based channel order determination method [38] is based on the static assumption of the channel, they are no longer able to determine L under time-varying channels. In such a case, one feasible manner is to rely on the typical correlation-based method, which will be introduced in what follows. Since the orders of all subchannels are assumed common (see Section 4.2.1), we only require to estimate the order of a subchannel between one of the transmit-receive antenna pairs. Suppose that the nth transmit branch transmits a pseudo-random training sequence $\mathbf{p} = [p_0,...,p_{Q-1}]^T \in \mathbb{C}^{Q\times 1}$ prior to each SC burst. After passing through the channel, the received version of $\mathbf{p}$ at the mth receive branch is written as

$$\overline{\mathbf{r}}_m = \overline{\overline{\mathbf{H}}}_{m,n}\mathbf{p} + \overline{\mathbf{v}}_m,\tag{4.42}$$

where $\overline{\mathbf{v}}_m \in \mathbb{C}^{(Q+L)\times 1}$ denotes the noise vector, and $\overline{\overline{\mathbf{H}}}_{m,n} \in \mathbb{C}^{(Q+L)\times Q}$ the convolution matrix with the ith column defined as

$$\overline{\mathbf{c}}_{m,n}^{(i)} = \overline{\mathbf{J}}^{i-1}\Big[\mathbf{h}_{m,n}^{(i)}\; 0\cdots 0\Big]^T,\; 1\le i\le Q,\tag{4.43}$$

81

in which $\mathbf{h}_{m,n}^{(i)} = [h_{m,n}(i-1,0),...,h_{m,n}(i+L-1,L)]^T$ and

$$\bar{\mathbf{J}} = \begin{bmatrix} \mathbf{0}_{1 \times (Q+L-1)} & 1 \\ \mathbf{I}_{Q+L-1} & \mathbf{0}_{(Q+L-1) \times 1} \end{bmatrix}.$$

To determine L, $\bar{\mathbf{r}}_m$ is fed into a sliding correlator, which is matched to the foreknown training sequence $\mathbf{p}$. The sliding correlator then outputs $\bar{L}$ peak correlation values

$$\tau_l = \sum_{i=0}^{Q-1} |p_i|^2 h_{m,n}(l, i+l) + I_{1,l}(1 - \delta(l)) + I_{2,l}(1 - \delta(l-\bar{L})) + n_l, \ \ 0 \le l \le \bar{L}, \tag{4.44}$$

where n_l is the noise component,

$$I_{1,l} = \sum_{u_1=0}^{l-1} \sum_{i_1=0}^{Q-l+u_1-1} p_{i_1}^* p_{i_1+(l-u_1)} h_{m,n}\left(u_1, i_1 + l\right), \tag{4.45}$$

$$I_{2,l} = \sum_{u_2=0}^{L-l} \sum_{i_2=0}^{Q-u_2-1} p_{i_2+(L-l)}^* p_{i_2} h_{m,n}\left(l + u_2, i_2 + l + u_2\right). \tag{4.46}$$

Note that $h_{m,n}(i,l) = 0$, $\forall i$, if $l > L$. If the noise level is moderately low, the sliding correlator is able to resolve all of the $L+1$ channel taps, i.e., $\bar{L} = L+1$, and the channel order is immediately $L = \bar{L} - 1$.

4.7 Simulation Results

This section illustrates the simulated performance of the proposed scheme. We consider a MIMO SC-CP system with carrier frequency 5 GHz, transmission bandwidth 20 MHz, $N = 2$ transmit antennas, $M = 2$ receive antennas, symbol block size $Q = 64$, and CP length $G = 7$. The velocity of the moving transmitter is set to be 120 Km/h. The source symbols are drawn from the QPSK constellation. The channels are characterized by the Jakes's model [30] with order $L = 7$ and the impulse response is normalized such that $\sum_{n=1}^{2} \sum_{l=0}^{7} E\left\{\left|h^{(m,n)}(k,l)\right|^2\right\} = 1$ for each $1 \le m \le 2$ and $k \ge 0$. The input SNR at the mth receive antenna is defined as

$$SNR = \frac{\sum_{n=1}^{2} \sum_{l=0}^{7} E\left\{\left|h^{(m,n)}(k,l)\right|^2\right\}}{\sigma_v^2} = \frac{1}{\sigma_v^2},$$

and the data burst length is set to be $T = 15$. The number of iterations involved in the RR

implementation is set to be $I = 10$. After inter-group interference suppression the MMSE V-BLAST detector [82] is used for symbol recovery.

A. Impacts of SIC Based Implementation

We first compare the performances of the proposed GSC equalizers with and without the SIC implementation. Figures 4.2 and 4.3, respectively, show the simulated bit-error-rate (BER) with perfect and imperfect channel knowledge (for the latter the channels are estimated via the LS training technique [6] with training power $P = 32$). We note that, when channel is perfectly known, the robust GSC solution (4.40) reduces to (4.19). The results show that the proposed method combined with the SIC mechanism yields about a 2~3 dB gain with exact channel knowledge, and a 4~5 dB gain with LS channel estimate. This would benefit from the increased receive diversity attained by the SIC mechanism.

B. Comparison with Existing Works

We compare the BER performance of the proposed method (combined with SIC) with the two alternative solutions group-wise V-BLAST [73] and LSFE [85]. The group size of the group-wise V-BLAST is set to be 16 (this is the same as the group size in our scheme $NP = 2 \times 8 = 16$). For the LSFE detector, the number of decision stages and the tap order of the feedback filter are respectively $N_b = 2$ and $L_b = 1$ ($N_b = 2$ is the suggested optimal choice [85] for two transmit antennas, and through simulation further increasing L_b does not seem to improve performance). Figures 4.4 and 4.5, respectively, show the results with perfect and imperfect channel knowledges. As we can see, even in the ideal case the proposed approach can outperform the two comparative choices. When only a channel estimate is available, the performances of all equalizers degrade, but the robust solution (4.40) remains to yield the lowest BER. The reason behind is that in the step of intra-group symbol recovery the proposed approach can obtain further diversity gain through the V-BLAST processing (with moderately low additional complexity due to the symbol decoupling nature (see Section 4.3.1)). With a fixed SNR level (25 dB) Figure 4.6 illustrates the BER of all methods when the burst duration T increases from 13 to 25. The results show that the proposed

robust equalizer (4.40) can significantly limit the BER penalty and is thus quite resistant to the increase of T.

C. Channel Estimation with Linear Interpolation

Figure 4.7 further depicts the simulated BER when the channel information at each time instant is acquired through the linear interpolation technique [44]; in this case the proposed robust scheme is given by (4.41). Compared with Figure 4.5, for each receiver the incurred performance loss with respect to the exact channel knowledge case is less severe due to the availability of timely channel information; the proposed robust solution (4.41), as expected, achieves the best performance since it is capable of combating channel estimation errors.

4.8 Summary

This chapter studies the robust receiver design problem for MIMO SC-CP systems when the multipath channels are time-varying and are estimated through the LS training technique. By exploiting certain group-wise orthogonality structure imbedded in the time-domain channel matrix, the proposed receiver aims for inter-group interference suppression followed by a low-complexity intra-group symbol recovery scheme. The design of the interference rejection matrix is mathematically formulated via constrained optimization technique. To further tackle the adverse effects due to imperfect channel knowledge we leverage the GSC principle to reformulate the problem into an equivalent unconstrained setup. The constraint-free GSC formulation resorts to proper weighting matrix decomposition and has several unique advantages: 1) it provides a simple yet efficient channel tracking mechanism by setting the adaptive component to be time-varying, 2) it allows for a very natural cost function for weighting matrix design against channel uncertainty due to time variation and estimation errors, 3) it enables us to directly model the channel mismatch effect into the filtered signal model via the perturbation analysis and, accordingly, we can then exploit the channel error statistics to derive a closed-form solution. The proposed approach can be further implemented in an SIC fashion for performance enhancement; it can also be used for estimation-error resistant receiver design when the channel estimate at each time instant is acquired

via linear interpolation. Compared with the existing works the proposed scheme yields significantly improved simulated BER, with comparable algorithm complexity.

Table 4.1. Formulae of $\mathbf{X}^{(j)}(t)$, $\mathbf{X}^{(i,j)}(t)$, $\mathbf{X}_1(t) \sim \mathbf{X}_4(t)$, $\mathbf{Y}_1(t) \sim \mathbf{Y}_3(t)$, and $\mathbf{R}_e(t)$.

$\mathbf{X}^{(j)}(t)$	$\mathbf{V}_h^{(j)}\Sigma_h^{(j)-1}\mathbf{U}_h^{(j)H}(\mathbf{R}_L^{(j)}(t) + \overline{\mathbf{R}}_I^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)) \times \mathbf{U}_h^{(j)}\Sigma_h^{(j)-1}\mathbf{V}_h^{(j)H}$	$\mathbf{Y}_3(t)$	$\sum_{i=1,i\neq j}^{L+1}\mathbf{D}_{i,j}\left(\mathbf{X}^{(i,j)}(t)^H\right)\mathbf{D}_c^{(i)}(t)$
$\mathbf{X}^{(i,j)}(t)$	$\mathbf{H}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)\mathbf{U}_h^{(j)}\Sigma_h^{(j)-1}\mathbf{V}_h^{(j)H}$	$\mathbf{R}_e(t)$	$\mathbf{D}_{j,j}(\mathbf{A}^{(j)}(t)) + \left(\overline{\mathbf{D}}_{j,j}(\mathbf{A}^{(j)}(t)) - \mathbf{D}_{j,j}(\mathbf{A}^{(j)}(t))\right)\mathbf{D}_T(t)$ $+\mathbf{D}_T(t)\left(\overline{\mathbf{D}}_{j,j}^H(\mathbf{A}^{(j)}(t)) - \mathbf{D}_{j,j}(\mathbf{A}^{(j)}(t))\right)$ $+\mathbf{D}_T(t)\left(2\mathbf{D}_{j,j}(\mathbf{A}^{(j)}(t)) - \overline{\mathbf{D}}_{j,j}^H(\mathbf{A}^{(j)}(t)) - \overline{\mathbf{D}}_{j,j}(\mathbf{A}^{(j)}(t))\right)\mathbf{D}_T(t)$
$\mathbf{X}_1(t)$	$\sum_{i=1}^{L+1}\mathbf{D}_c^{(i)}(t)\mathbf{D}_H^{(i)}\mathbf{D}_c^{(i)}(t)^H$		
$\mathbf{X}_2(t)$	$\mathbf{D}_{j,j}\left(\mathbf{X}^{(j)}(t)\right)$	$\mathbf{A}^{(j)}(t)$	$\mathbf{V}_h^{(j)}\Sigma_h^{(j)-1}\mathbf{U}_h^{(j)H}\mathbf{R}_I^{(j)}(t)\mathbf{U}_h^{(j)}\Sigma_h^{(j)-1}\mathbf{V}_h^{(j)H}$
$\mathbf{X}_3(t)$	$\sum_{i=1}^{L+1}\mathbf{D}_c^{(i)}\mathbf{D}_{i,j}\left(\mathbf{X}^{(i,j)}(t)\right)$	$\left[\overline{\mathbf{D}}_{j,j}(\mathbf{Z}_1)\right]^{(m,m)}$	$\sum_{n_1=1}^{N}\sum_{n_2=1}^{N}\overline{\mathbf{J}}_i\left(\mathbf{Z}_1^{(n_1,n_2)} \otimes Diag\left\{\overline{\mathbf{R}}_{\Delta h,m}^{(n_1,n_2)}, \mathbf{0}_{L+1}\cdots\mathbf{0}_{L+1}\right\}\right)\overline{\mathbf{J}}_j^T$
$\mathbf{X}_4(t)$	$\mathbf{X}_3(t)^H$	$\overline{\mathbf{R}}_{\Delta h,m}$	$\mathbf{A}^+\mathbf{K}(\mathbf{A}^+)^H$
$\mathbf{Y}_1(t)$	$\sum_{i=1,i\neq j}^{L+1}\mathbf{D}_c^{(i)}(t)\mathbf{D}_H^{(i)}\mathbf{D}_c^{(i)}(t)^H$	$[\mathbf{K}]_{q,q'}$	$\frac{1}{Q^2}\sum_{n=1}^{N}\sum_{\substack{q_1=0,q_2=0\\q_1\neq q,q_2\neq q'}}^{Q-1}\sum_{l=0}^{L}\sum_{q_3,q_4=0}^{Q-1}\mathbf{R}_h(T(G+Q)+q_3-q_4,l)e^{\frac{-j2\pi(q_1-q_2)l}{Q}} \cdot$ $[\mathbf{t}_n]_{(q-q_1)_Q}[\mathbf{t}_n]^*_{(q'-q_2)_Q}e^{\frac{-j2\pi[(q-q_1)-(q'-q_2)]l}{Q}}$
$\mathbf{Y}_2(t)$	$\sum_{i=1,i\neq j}^{L+1}\mathbf{D}_c^{(i)}(t)\mathbf{H}_{av}^{(i)}\mathbf{K}_{i,j}\left(\mathbf{D}_c^{(i)}(t)^H\right)$		
$\left[\mathbf{D}_H^{(i)}\right]^{(m,m)}$	$\sum_{n=1}^{N}\sum_{p=1}^{P}\mathbf{J}^{i+(p-1)(L+1)-1}Diag\left\{\mathbf{R}_{\Delta h,m}^{(n,n)},\mathbf{0}_{L+1},...,\mathbf{0}_{L+1}\right\}\left(\mathbf{J}^{i+(p-1)(L+1)-1}\right)^T$		
$\left[\mathbf{D}_{i,j}(\mathbf{Z}_1)\right]^{(m,m)}$	$\sum_{n_1=1}^{N}\sum_{n_2=1}^{N}\overline{\mathbf{J}}_i\left(\mathbf{Z}_1^{(n_1,n_2)} \otimes Diag\left\{\mathbf{R}_{\Delta h,m}^{(n_1,n_2)},\mathbf{0}_{L+1}\cdots\mathbf{0}_{L+1}\right\}\right)\overline{\mathbf{J}}_j^T$		
$\left[\mathbf{K}_{i,j}(\mathbf{Z}_2)\right]^{(n_1,n_2)}_{p_1,p_2}$	$\sum_{m=1}^{M}vec\left(\left[\overline{\mathbf{J}}_i^T\mathbf{Z}_2^{(m,m)}\overline{\mathbf{J}}_j\right]_1^{(p_1,p_2)}\right)^T vec\left(\mathbf{R}_{\Delta h,m}^{(n_1,n_2)}\right)^*$	$vec(.)$ stacks a matrix into a vector columnwise and $\left[\overline{\mathbf{J}}_i^T\mathbf{Z}_2^{(m,m)}\overline{\mathbf{J}}_j\right]_1^{(p_1,p_2)}$ denotes the first $(L+1)\times(L+1)$ diagonal block submatrix of the (p_1,p_2)th block submatrix of $\overline{\mathbf{J}}_i^T\mathbf{Z}_2^{(m,m)}\overline{\mathbf{J}}_j \in \mathbb{C}^{PQ\times PQ}$	

Table 4.2. Flop count comparison.

FC_{GSC}	$(MQ-NP)\left(\frac{2}{3}M^2Q^2 + \frac{2}{3}N^2P^2 + 2\left(M^2L(L+1)+MNL\right)Q\right)$
$FC_{GSC}^{(RR)}$	$\left(M^2L + MN(MQ-NP)\right)Q + INP\left(7(MQ-NP) + 2M(L+1)(MQ-NP) + M^2LQ\right)$
FC_{LSFE}	$\left(\frac{2}{3}M^3 + N\right)Q^3 + \left(\frac{1}{2}\left(M^4+M^3+M^2\right) + N_bL_b\right)Q^2 + \left(\frac{1}{2}M^3 + M^2 + \frac{5}{6}M\right)Q$
FC_{GB}	$\left(M^3(L+1) - \frac{3}{2}M^2LN + \frac{1}{2}MN^2(2L+1) - \frac{1}{4}N^3L\right)Q^3$ $+ \left(\frac{1}{2}M^2(L+1) - \frac{1}{2}MNL + \frac{1}{6}N^2(2L+1)\right)Q^2 + \left(\frac{5}{6}M - \frac{5}{12}NL\right)Q$
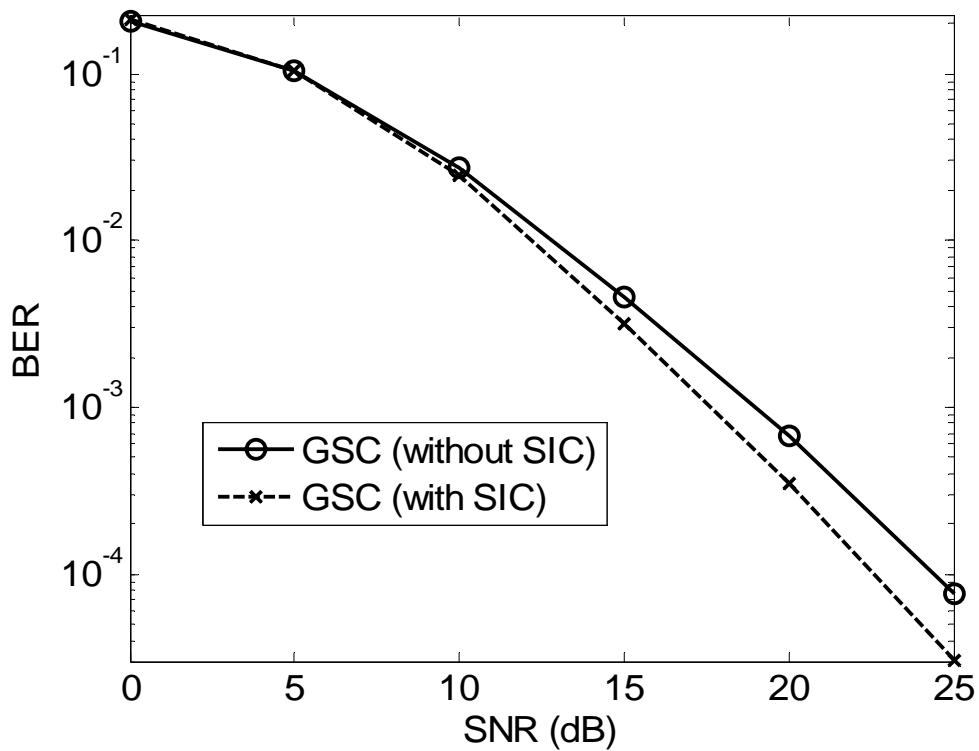
Figure 4.2. BER performances of the proposed method with and without the SIC mechanism (perfect channel knowledge).
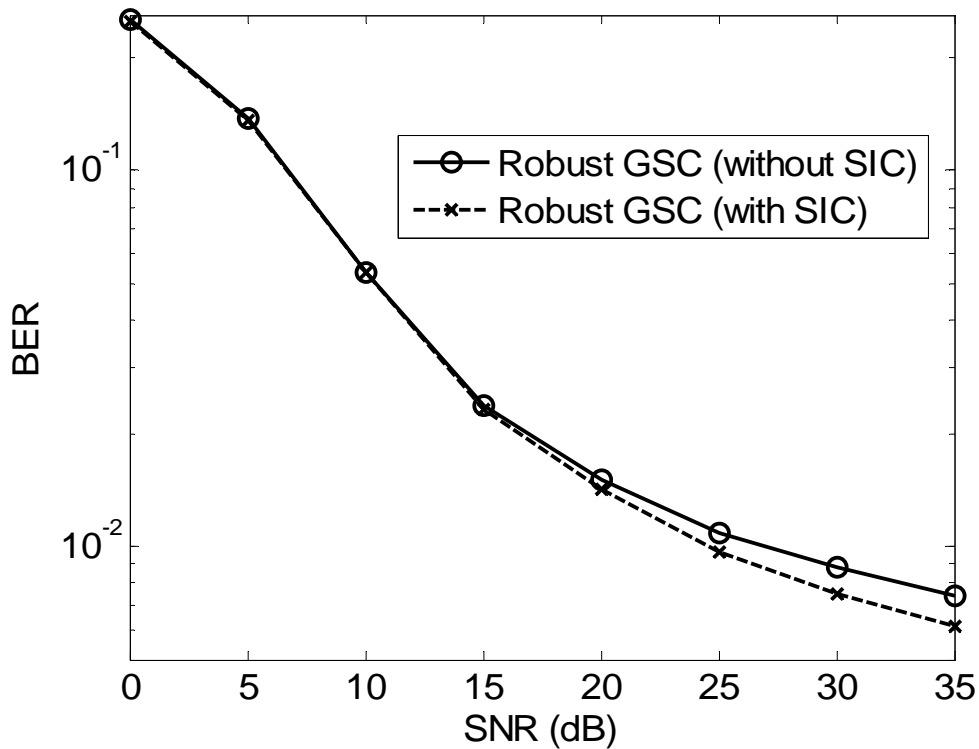
Figure 4.3. BER performances of the proposed method with and without the SIC mechanism (LS channel estimate).
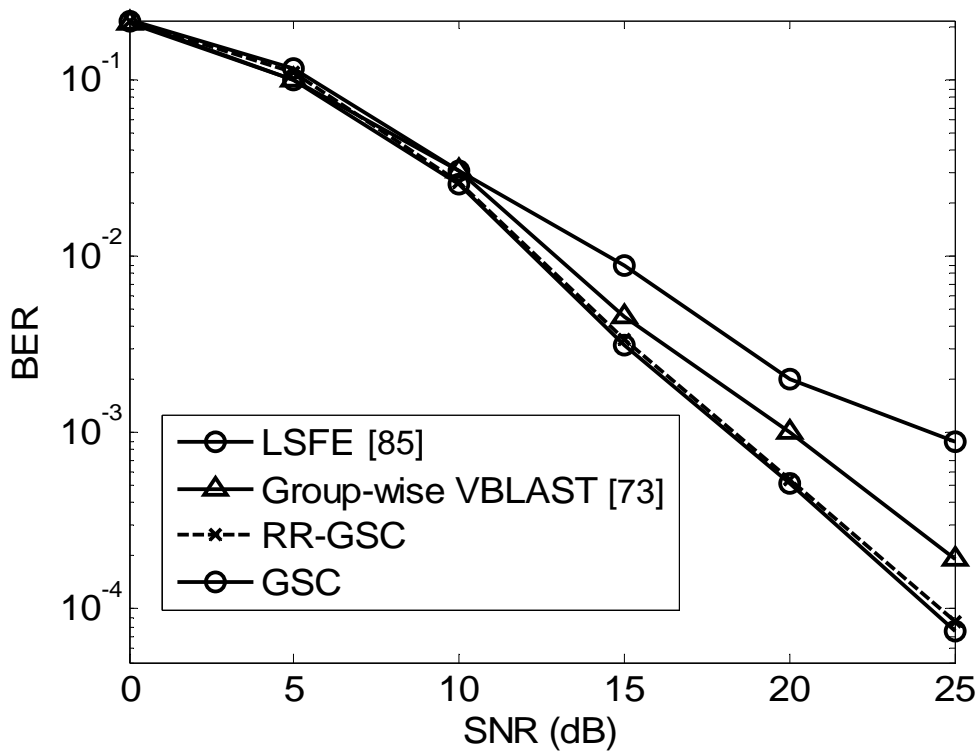
Figure 4.4. BER performances of the three methods (perfect channel knowledge).
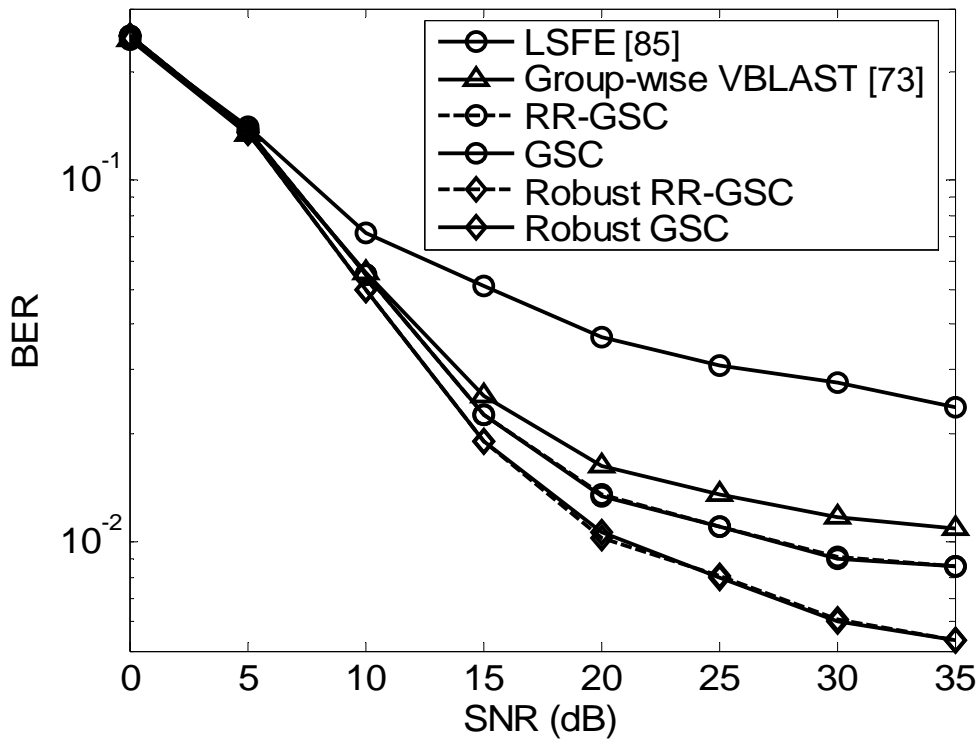
Figure 4.5. BER performances of the three methods (LS channel estimate).
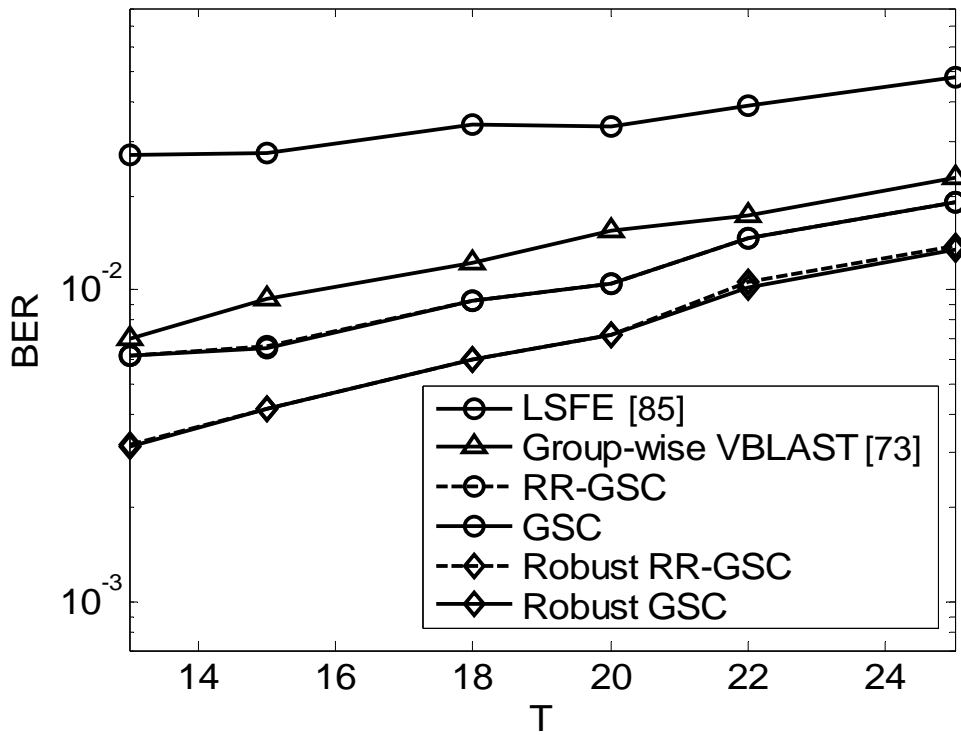
Figure 4.6. BER performances of the three method with respect to various burst length (LS channel estimate).
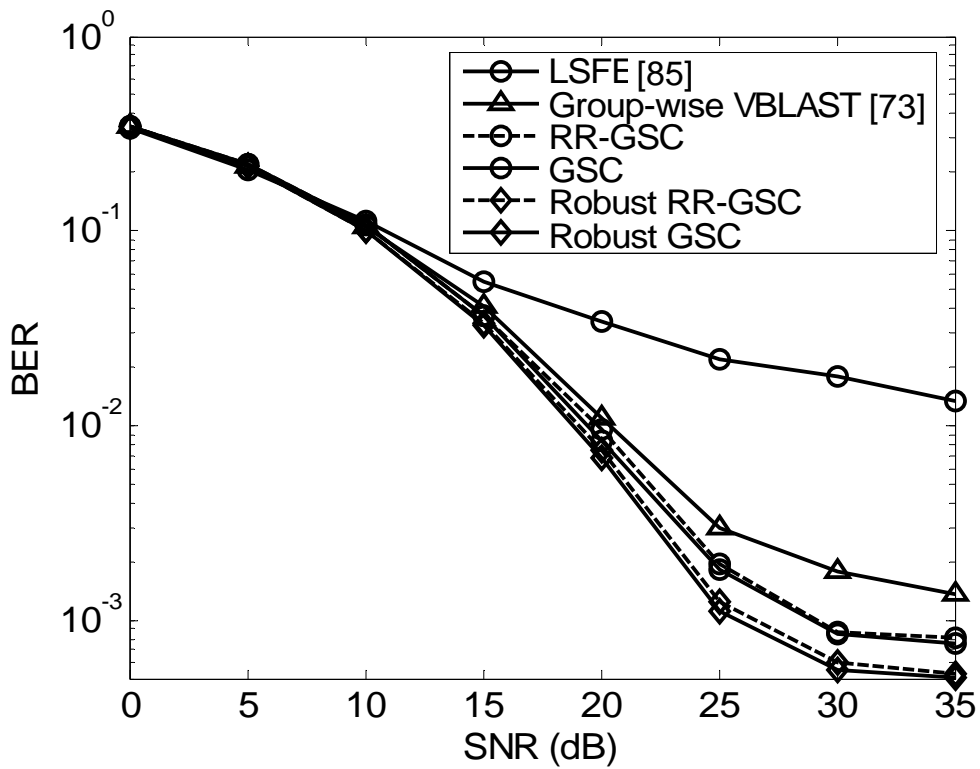
Figure 4.7. BER performances of the three methods (interpolation-based channel estimate).

Chapter 5

Conclusions and Future Works

5.1 Summary of Dissertation

This dissertation mainly addresses the robust receiver designs for MIMO-OFDM and MIMO SC-CP systems against channel mismatch. The main contribution lies in that we propose to exploit the perturbation analysis technique to incorporate the negative effects of channel mismatch into the considered problem formulations; to derive the tractable closed-form solutions only requires the first- and second-order statistical characteristics of the channel mismatch.

The introductory chapter includes the background overview, literature review, and contributions of this dissertation. In Chapter 2, a GSC-based equalization framework for the CP-free SIMO OFDM systems is proposed for ISI/ICI suppression. Low-complexity PA implementation is also derived based on the interference signature characteristics, and the resultant simulated performances are seen to be comparable to the original FA realizations.

In Chapter 3, we extend the equalizer proposed in Chapter 2 to develop a robust ISI/ICI mitigation scheme for supporting high-rate MIMO-OFDM transmission, under the assumption that the channels are not exactly known but are estimated via the LS training technique. Combined with the perturbation analysis technique, the negative effects caused by channel estimation error can be explicitly characterized and be considered in the GSC-based framework. Accordingly, we can then exploit the statistical characteristics of the estimation error to derive a closed-form robust solution. Another important contribution is to derive a closed-form mean SINR expression for the proposed robust scheme, and also an informative formula for quantifying the achievable SINR increment over the non-robust solution.

In Chapter 4, we study the robust receiver design problem for MIMO SC-CP systems when the

multipath channels are instead time-varying and are estimated through the LS training technique. By exploiting certain group-wise orthogonality structure imbedded in the time-domain channel matrix, the proposed receiver aims for inter-group interference suppression followed by a low-complexity intra-group symbol recovery scheme. The design of the interference suppression mechanism is also based on the GSC-based framework. It is shown that the robustness analysis method proposed in Chapter 3 enables us to take the channel mismatch, jointly caused by the channel temporal variation and the channel estimation error, in the cost function. By incorporating the channel error statistics, a closed-form robust solution against the channel variation and the channel estimation error is derived.

5.2 Future Works

Cooperative communication [41], [48] is a generalized MIMO system and has been widely recognized to be effective for improving system performance. Various research activities about cooperative communication are like wildfire proceeding; some related techniques are considered to be adopted in the next-generation wireless communication standards (e.g., 802.16j). To realize cooperative communication the destination end generally requires *1)* to synchronize all cooperative users and *2)* to estimate the channel impulse response of each cooperative link. However, it is impossible to perfectly perform the both tasks due to the existence of channel noise (or interference). As a result, the parameter estimation errors are expected to result in serious system performance degradation. In order to have more reliable cooperative links, the developments of the robustness techniques are currently believed to be one of the most important works for implementing cooperative communication. The following are some future research steps:

1. Analysis of the detrimental effects, caused by the system parameter estimation errors, on cooperative communication systems.

2. Developments of the robust receiver architectures based on the statistical (or deterministic) characteristics of the parameter errors.

Appendix

A Proof of Lemma 3.2

[Proof of (1)]: With $\widehat{\overline{\mathbf{D}}} = \overline{\mathbf{D}} + \Delta\overline{\mathbf{D}}$ and $\widehat{\overline{\mathbf{B}}}$ defined in (3.42), we have

$$E\{\widehat{\overline{\mathbf{B}}}^{H}(\mathbf{R}_I - \overline{\mathbf{D}}\mathbf{H}_{ICI}^{H})\widehat{\overline{\mathbf{D}}}\} = \overline{\mathbf{B}}^{H}(\mathbf{R}_I - \overline{\mathbf{D}}\mathbf{H}_{ICI}^{H})\overline{\mathbf{D}} + \overline{\mathbf{B}}^{H}(\mathbf{R}_I - \overline{\mathbf{D}}\mathbf{H}_{ICI}^{H})E\{\Delta\overline{\mathbf{D}}\}$$

$$-\mathbf{B}^{H}E\{\Delta\overline{\mathbf{D}}\}\mathbf{V}_D\Sigma_D^{-1}\mathbf{U}_D^{H}(\mathbf{R}_I - \overline{\mathbf{D}}\mathbf{H}_{ICI}^{H})\overline{\mathbf{D}} - \overline{\mathbf{B}}^{H}E\{\Delta\overline{\mathbf{D}}\mathbf{V}_D\Sigma_D^{-1}\mathbf{U}_D^{H}(\mathbf{R}_I - \overline{\mathbf{D}}\mathbf{H}_{ICI}^{H})\Delta\overline{\mathbf{D}}\}. \quad \text{(A.1)}$$

Since the elements of the noise $\mathbf{v}_m(\cdot)$ is circular Gaussian, the entries of the LS channel estimation error vector $\Delta\mathbf{h}^{(m)}$ are i.i.d. circular Gaussian (see (3.9)). By definitions of $\Delta\overline{\mathbf{D}}$ (see (3.32)) and $\Delta\overline{\mathbf{B}}$ (see (3.42)) and with the circularity condition of $\Delta\mathbf{h}^{(m)}$, the last three terms on the RHS of (A.1) is identically zero. Also, since $\mathbf{v}_m(\cdot)$ is zero-mean, so are $\Delta\mathbf{h}^{(m)}$ and $\Delta\overline{\mathbf{D}}$. Equation (A.1) thus reduces to

$$E\{\widehat{\overline{\mathbf{B}}}^{H}(\mathbf{R}_I - \overline{\mathbf{D}}\mathbf{H}_{ICI}^{H})\widehat{\overline{\mathbf{D}}}\} = \overline{\mathbf{B}}^{H}(\mathbf{R}_I - \overline{\mathbf{D}}\mathbf{H}_{ICI}^{H})\overline{\mathbf{D}} = \overline{\mathbf{B}}^{H}\mathbf{R}_I\overline{\mathbf{D}}, \quad \text{(A.2)}$$

where the last equality follows since $\overline{\mathbf{B}}^{H}\overline{\mathbf{D}} = \mathbf{0}$.

[Proof of (2)]: With $\widehat{\overline{\mathbf{B}}}$ given in (3.42), we have $E\{\widehat{\overline{\mathbf{B}}}^{H}\overline{\mathbf{D}}\overline{\mathbf{D}}^{H}\widehat{\overline{\mathbf{B}}}\} = \overline{\mathbf{B}}^{H}E\{\Delta\overline{\mathbf{D}}\Delta\overline{\mathbf{D}}^{H}\}\overline{\mathbf{B}}$, where the equality follows by definition of $\Delta\overline{\mathbf{B}}$ in (3.42) and since $\overline{\mathbf{D}} = \mathbf{U}_D\Sigma_D\mathbf{V}_D^{H}$. Denote by $\Delta\overline{\mathbf{D}}^{(m,n)} \in \mathbb{C}^{Q \times Q}$ the (m,n)th $Q \times Q$ block submatrix of $\Delta\overline{\mathbf{D}}$, for $1 \leq m \leq M$ and $1 \leq n \leq N$; then we have

$$\Delta\mathbf{D}^{(m,n)} = diag\{\hat{\mathbf{g}}^{(m,n)}\}, \text{ where } \hat{\mathbf{g}}^{(m,n)} := \sqrt{Q}\mathbf{F}_L\Delta\mathbf{h}^{(m,n)} \in \mathbb{C}^{Q}. \quad \text{(A.3)}$$

For $\mathbf{X} \in \mathbb{C}^{m \times m}$, let $Diag\{\mathbf{X}\}$ be the diagonal matrix obtained by keeping only the diagonal entries of $\mathbf{X}$. With (A.3), it follows

$$E\{\Delta\mathbf{D}^{(m_1,n)}\Delta\mathbf{D}^{(m_2,n)H}\} = E\{Diag\{\hat{\mathbf{g}}^{(m_1,n)}\hat{\mathbf{g}}^{(m_2,n)H}\}\} = Diag\{Q\mathbf{F}_L E\{\Delta\mathbf{h}^{(m_1,n)}\Delta\mathbf{h}^{(m_2,n)H}\}\mathbf{F}_L^{H}\}. \quad \text{(A.4)}$$

Equations (3.9) and (3.10) together imply $E\{\Delta\mathbf{h}^{(m_1,n)}\Delta\mathbf{h}^{(m_2,n)H}\} = \frac{\sigma_v^2}{P}\mathbf{I}_{L+1}\delta(m_1 - m_2)$, and (A.4) thus becomes

92

$$E\{\Delta\mathbf{D}^{(m_1,n)}\Delta\mathbf{D}^{(m_2,n)H}\} = \frac{Q\sigma_v^2}{P}Diag\{\mathbf{F}_L\mathbf{F}_L^H\}\delta(m_1 - m_2) = \frac{\sigma_v^2(L+1)}{P}\mathbf{I}_Q\delta(m_1 - m_2), \quad (A.5)$$

where the last equality follows due to $Diag\{\mathbf{F}_L\mathbf{F}_L^H\} = Q^{-1}(L+1)\mathbf{I}_Q$. Based on (A.5), the (p,q) th $Q \times Q$ block submatrix of $E\{\Delta\overline{\mathbf{D}}\Delta\overline{\mathbf{D}}^H\}$ is given by

$$\sum_{n=1}^{N} E\{\Delta\mathbf{D}^{(p,n)}\Delta\mathbf{D}^{(q,n)H}\} = \begin{cases} \dfrac{N(L+1)\sigma_v^2}{P}\mathbf{I}_Q, & p = q, \\ \mathbf{0}_Q, & \text{if } p \neq q. \end{cases} \quad (A.6)$$

The result follows directly from (A.6) and since $\overline{\mathbf{B}}^H\overline{\mathbf{B}} = \mathbf{I}_{(M-N)Q}$.

[Proof of (3)]: With $\widehat{\overline{\mathbf{B}}}$ given in (3.42) and since $\overline{\mathbf{B}}^H\overline{\mathbf{D}} = \mathbf{0}$, direct manipulation shows

$$E\{\widehat{\overline{\mathbf{B}}}^H(\mathbf{R}_I - \overline{\mathbf{D}}\mathbf{H}_{ICI}^H - \mathbf{H}_{ICI}\overline{\mathbf{D}}^H)\widehat{\overline{\mathbf{B}}}\} = \overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}} + \mathbf{B}^H\underbrace{E\{\Delta\overline{\mathbf{D}}\mathbf{K}\Delta\overline{\mathbf{D}}^H\}}_{=\mathbf{R}_c}\mathbf{B}, \quad (A.7)$$

where the matrix $\mathbf{K}$ is defined in (3.43). We note that the (m_1, m_2) th block submatrix of $\mathbf{R}_c$ equals $\mathbf{R}_c^{(m_1,m_2)} = \sum_{n_1=1}^{N}\sum_{n_2=1}^{N} E\{\Delta\mathbf{D}^{(m_1,n_1)}\mathbf{K}^{(n_1,n_2)}\Delta\mathbf{D}^{(m_2,n_2)H}\}$, $1 \leq m_1, m_2 \leq M$. Since the channel estimation errors between different receive antennas are independent (From (3.9) and (3.10)), we have $E\{\Delta\mathbf{D}^{(m_1,n_1)}\mathbf{K}^{(n_1,n_2)}\Delta\mathbf{D}^{(m_2,n_2)H}\} = \mathbf{0}_Q$, $m_1 \neq m_2$, $n_1 \neq n_2$, which implies

$$\mathbf{R}_c^{(m_1,m_2)} = \begin{cases} \sum_{n=1}^{N} E\{\Delta\mathbf{D}^{(m_1,n)}\mathbf{K}^{(n,n)}\Delta\mathbf{D}^{(m_1,n)H}\}, & m_1 = m_2 \\ \mathbf{0}_Q & , \quad \text{otherwise.} \end{cases} \quad (A.8)$$

Equation (A.8) shows $\mathbf{R}_c$ is a block diagonal matrix, and (3.44) follows by expanding the block diagonal terms in (A.8) and using $\Delta\mathbf{D}^{(m,n)} = diag\{\sqrt{Q}\mathbf{F}_L\Delta\mathbf{h}^{(m,n)}\}$. $\quad\square$

B Derivation of Equation (3.57)

To derive (3.57), we need the following lemma.

Lemma B.1: Let $\mathbf{R}_I$ be defined in (3.20) and $\mathbf{U}_{dl}$ be the exact solution of $\widehat{\mathbf{U}}_{dl}$ defined in (3.49). Also, let an SVD of the channel tone matrix $\overline{\mathbf{D}}$ be given in (3.33). Then the first-order approximation to $\Delta\mathbf{W}_{dl}$ is

$$\Delta\mathbf{W}_{dl} = \Delta\overline{\mathbf{D}} + \mathbf{U}_D\Sigma_D^{-1}\mathbf{V}_D^H\Delta\overline{\mathbf{D}}^H\overline{\mathbf{B}}\mathbf{U}_{dl} - \overline{\mathbf{B}}(\overline{\mathbf{B}}^H\mathbf{R}_{I,dl}\overline{\mathbf{B}})^{-1}(\overline{\mathbf{B}}^H\mathbf{R}_I\Delta\overline{\mathbf{D}} - \overline{\mathbf{B}}^H\Delta\overline{\mathbf{D}}\mathbf{V}_D\Sigma_D^{-1}\mathbf{U}_D^H\mathbf{R}_I\overline{\mathbf{D}} + \Delta\mathbf{R}\mathbf{U}_{dl}),$$

$$(B.1)$$

with

$$\mathbf{R}_{I,dl} := \mathbf{R}_I + \overline{\gamma}\mathbf{I}_{MQ}, \tag{B.2}$$

and

$$\Delta\mathbf{R} := \overline{\mathbf{B}}^H\mathbf{R}_{I,dl}\mathbf{U}_D\Sigma_D^{-1}\mathbf{V}_D^H\Delta\overline{\mathbf{D}}^H\overline{\mathbf{B}} + \overline{\mathbf{B}}^H\Delta\overline{\mathbf{D}}\mathbf{V}_D\Sigma_D^{-1}\mathbf{U}_D^H\mathbf{R}_{I,dl}\overline{\mathbf{B}}. \tag{B.3}$$

$\square$

[Proof]: By definition, the estimated $\widehat{\mathbf{W}}_{dl}$ can be expressed as

$$\widehat{\mathbf{W}}_{dl} = (\overline{\mathbf{D}} + \Delta\overline{\mathbf{D}}) - (\overline{\mathbf{B}} + \Delta\overline{\mathbf{B}})\left((\overline{\mathbf{B}} + \Delta\overline{\mathbf{B}})^H(\mathbf{R}_{I,dl} + \Delta\mathbf{R}_{I,dl})(\overline{\mathbf{B}} + \Delta\overline{\mathbf{B}})\right)^{-1}(\overline{\mathbf{B}} + \Delta\overline{\mathbf{B}})^H(\mathbf{R}_I + \Delta\mathbf{R}_{I,dl})(\overline{\mathbf{D}} + \Delta\overline{\mathbf{D}}),$$

$$\tag{B.4}$$

where

$$\Delta\mathbf{R}_{I,dl} := \Delta\mathbf{H}_{ISI}\mathbf{H}_{ISI}^H + \mathbf{H}_{ISI}\Delta\mathbf{H}_{ISI}^H + \Delta\mathbf{H}_{ISI}\Delta\mathbf{H}_{ISI}^H + \Delta\mathbf{H}_{ICI}\mathbf{H}_{ICI}^H + \mathbf{H}_{ICI}\Delta\mathbf{H}_{ICI}^H + \Delta\mathbf{H}_{ICI}\Delta\mathbf{H}_{ICI}^H. \tag{B.5}$$

We note that

$$(\overline{\mathbf{B}} + \Delta\overline{\mathbf{B}})^H(\mathbf{R}_{I,dl} + \Delta\mathbf{R}_{I,dl})(\overline{\mathbf{B}} + \Delta\overline{\mathbf{B}}) = \overline{\mathbf{B}}^H\mathbf{R}_{I,dl}\overline{\mathbf{B}} + \overline{\mathbf{B}}^H\Delta\mathbf{R}_{I,dl}\Delta\overline{\mathbf{B}} + \Delta\overline{\mathbf{B}}^H\mathbf{R}_{I,dl}\overline{\mathbf{B}} + \overline{\mathbf{B}}^H\mathbf{R}_{I,dl}\Delta\overline{\mathbf{B}} + \overline{\mathbf{B}}^H\Delta\mathbf{R}_{I,dl}\overline{\mathbf{B}}$$

$$+\Delta\overline{\mathbf{B}}^H\Delta\mathbf{R}_{I,dl}\overline{\mathbf{B}} + \Delta\overline{\mathbf{B}}^H\mathbf{R}_{I,dl}\Delta\overline{\mathbf{B}} + \Delta\overline{\mathbf{B}}^H\Delta\mathbf{R}_{I,dl}\Delta\overline{\mathbf{B}}. \tag{B.6}$$

For small channel estimation error, $\Delta\mathbf{H}_{ISI}$ and $\Delta\mathbf{H}_{ICI}$ are small, and $\Delta\mathbf{R}_{I,dl}$ will be close to a zero matrix. This implies equation (B.6) can be well approximated by

$$(\overline{\mathbf{B}} + \Delta\overline{\mathbf{B}})^H(\mathbf{R}_{I,dl} + \Delta\mathbf{R}_{I,dl})(\overline{\mathbf{B}} + \Delta\overline{\mathbf{B}}) \approx \overline{\mathbf{B}}^H\mathbf{R}_{I,dl}\overline{\mathbf{B}} + \Delta\overline{\mathbf{B}}^H\mathbf{R}_{I,dl}\overline{\mathbf{B}} + \overline{\mathbf{B}}^H\mathbf{R}_{I,dl}\Delta\overline{\mathbf{B}}. \tag{B.7}$$

From (B.7) and since $(\mathbf{A} + \Delta\mathbf{A})^{-1} \approx \mathbf{A}^{-1} - \mathbf{A}^{-1}\Delta\mathbf{A}\mathbf{A}^{-1}$ for small $\Delta\mathbf{A}$, we have

$$\left((\overline{\mathbf{B}} + \Delta\overline{\mathbf{B}})^H(\mathbf{R}_{I,dl} + \Delta\mathbf{R}_{I,dl})(\overline{\mathbf{B}} + \Delta\overline{\mathbf{B}})\right)^{-1} \approx (\overline{\mathbf{B}}^H\mathbf{R}_{I,dl}\overline{\mathbf{B}} + \Delta\overline{\mathbf{B}}^H\mathbf{R}_{I,dl}\overline{\mathbf{B}} + \overline{\mathbf{B}}^H\mathbf{R}_{I,dl}\Delta\overline{\mathbf{B}})^{-1}$$

$$= (\overline{\mathbf{B}}^H\mathbf{R}_{I,dl}\overline{\mathbf{B}})^{-1} - (\overline{\mathbf{B}}^H\mathbf{R}_{I,dl}\overline{\mathbf{B}})^{-1}(\Delta\overline{\mathbf{B}}^H\mathbf{R}_{I,dl}\overline{\mathbf{B}} + \overline{\mathbf{B}}^H\mathbf{R}_{I,dl}\Delta\overline{\mathbf{B}})(\overline{\mathbf{B}}^H\mathbf{R}_{I,dl}\overline{\mathbf{B}})^{-1}. \tag{B.8}$$

Substituting (B.8) into (B.4) and using (3.42), the result follows by keeping only the first order terms. $\square$

[Derivation of (3.57)]: By definitions of $\Delta\overline{\mathbf{D}}$ (see (3.32)), it is easy to verify from (B.1) that $E\{\Delta\mathbf{W}_{dl}\} = \mathbf{0}$; equation (3.57) can be obtained via substituting $\Delta\mathbf{W}_{dl}$ into $\mathbf{i}_{dl}(k)$ in (3.54) followed by some direct manipulations. $\square$

C   Derivation of Equation (3.59)

With $\Delta\mathbf{W}_{dl}$ given in (B.1), we can expand

$$Tr(E\{\Delta\mathbf{W}_{dl}^H \overline{\mathbf{D}}\overline{\mathbf{D}}^H \Delta\mathbf{W}_{dl}\}) + Tr(E\{\Delta\mathbf{W}_{dl}^H \mathbf{R}_{I,c} \Delta\mathbf{W}_{dl}\}) = \frac{N(L+1)\sigma_v^2}{P}\|\mathbf{U}_{dl}\|^2 + Tr\left(\underbrace{E\{\Delta\overline{\mathbf{D}}^H \overline{\mathbf{D}}\overline{\mathbf{D}}^H \Delta\overline{\mathbf{D}}\}}_{\mathbf{x}_1}\right)$$

$$+Tr(\underbrace{\mathbf{U}_{dl}^H \overline{\mathbf{B}}^H E\{\Delta\overline{\mathbf{D}}\mathbf{V}_D \Sigma_D^{-1} \mathbf{U}_D^H \mathbf{R}_{I,c} \mathbf{U}_D \Sigma_D^{-1} \mathbf{V}_D^H \Delta\overline{\mathbf{D}}^H\}\overline{\mathbf{B}}\mathbf{U}_{dl}}_{\mathbf{x}_2})$$

$$+Tr(\underbrace{E\{\Delta\overline{\mathbf{D}}^H (\mathbf{R}_{I,c} + \mathbf{R}_I \overline{\mathbf{B}}(\overline{\mathbf{B}}^H \mathbf{R}_{I,dl}\overline{\mathbf{B}})^{-1}(\overline{\mathbf{B}}^H \mathbf{R}_{I,c}\overline{\mathbf{B}})(\overline{\mathbf{B}}^H \mathbf{R}_{I,dl}\overline{\mathbf{B}})^{-1}\overline{\mathbf{B}}^H \mathbf{R}_I)\Delta\overline{\mathbf{D}}\}}_{\mathbf{x}_3})$$

$$+Tr(\underbrace{\overline{\mathbf{D}}^H \mathbf{R}_I \mathbf{U}_D \Sigma_D^{-1} \mathbf{V}_D^H E\{\Delta\overline{\mathbf{D}}^H \overline{\mathbf{B}}(\overline{\mathbf{B}}^H \mathbf{R}_{I,dl}\overline{\mathbf{B}})^{-1}(\overline{\mathbf{B}}^H \mathbf{R}_{I,c}\overline{\mathbf{B}})(\overline{\mathbf{B}}^H \mathbf{R}_{I,dl}\overline{\mathbf{B}})^{-1}\overline{\mathbf{B}}^H \Delta\overline{\mathbf{D}}\}\mathbf{V}_D \Sigma_D^{-1} \mathbf{U}_D^H \mathbf{R}_I \overline{\mathbf{D}}}_{\mathbf{x}_4})$$

$$+Tr(\underbrace{\mathbf{U}_{dl}^H \overline{\mathbf{B}}^H \mathbf{R}_{I,dl}\mathbf{U}_D \Sigma_D^{-1} \mathbf{V}_D^H E\{\Delta\overline{\mathbf{D}}^H \overline{\mathbf{B}}(\overline{\mathbf{B}}^H \mathbf{R}_{I,dl}\overline{\mathbf{B}})^{-1}(\overline{\mathbf{B}}^H \mathbf{R}_{I,c}\overline{\mathbf{B}})(\overline{\mathbf{B}}^H \mathbf{R}_{I,dl}\overline{\mathbf{B}})^{-1}\overline{\mathbf{B}}^H \Delta\overline{\mathbf{D}}\}\mathbf{V}_D \Sigma_D^{-1} \mathbf{U}_D^H \mathbf{R}_{I,dl}\overline{\mathbf{B}}\mathbf{U}_{dl}}_{\mathbf{x}_5})$$

$$+Tr(\underbrace{\mathbf{U}_{dl}^H \overline{\mathbf{B}}^H E\{\Delta\overline{\mathbf{D}}\mathbf{V}_D \Sigma_D^{-1} \mathbf{U}_D^H \mathbf{R}_{I,dl}\overline{\mathbf{B}}(\overline{\mathbf{B}}^H \mathbf{R}_{I,dl}\overline{\mathbf{B}})^{-1}(\overline{\mathbf{B}}^H \mathbf{R}_{I,c}\overline{\mathbf{B}})(\overline{\mathbf{B}}^H \mathbf{R}_{I,dl}\overline{\mathbf{B}})^{-1}\overline{\mathbf{B}}^H \mathbf{R}_{I,dl}\mathbf{U}_D \Sigma_D^{-1} \mathbf{V}_D^H \Delta\overline{\mathbf{D}}^H\}\overline{\mathbf{B}}\mathbf{U}_{dl}}_{\mathbf{x}_6})$$

$$-2\,\mathrm{Re}\left\{Tr(\underbrace{\mathbf{U}_{dl}^H \overline{\mathbf{B}}^H E\{\Delta\overline{\mathbf{D}}\mathbf{V}_D \Sigma_D^{-1} \mathbf{U}_D^H \mathbf{R}_{I,c}\overline{\mathbf{B}}(\overline{\mathbf{B}}^H \mathbf{R}_{I,dl}\overline{\mathbf{B}})^{-1}\overline{\mathbf{B}}^H \mathbf{R}_{I,dl}\mathbf{U}_D \Sigma_D^{-1} \mathbf{V}_D^H \Delta\overline{\mathbf{D}}^H\}\overline{\mathbf{B}}\mathbf{U}_{dl}}_{\mathbf{x}_7})\right\}$$

$$-2\,\mathrm{Re}\left\{Tr(\underbrace{-E\{\Delta\overline{\mathbf{D}}^H \mathbf{R}_{I,c}\overline{\mathbf{B}}(\overline{\mathbf{B}}^H \mathbf{R}_{I,dl}\overline{\mathbf{B}})^{-1}\overline{\mathbf{B}}^H \Delta\overline{\mathbf{D}}\}\mathbf{V}_D \Sigma_D^{-1} \mathbf{U}_D^H \mathbf{R}_{I,dl}\mathbf{W}_{dl}}_{\mathbf{x}_8})\right\}$$

$$-2\,\mathrm{Re}\left\{Tr(\underbrace{\overline{\mathbf{D}}^H \mathbf{R}_I \mathbf{U}_D \Sigma_D^{-1} \mathbf{V}_D^H E\{\Delta\overline{\mathbf{D}}^H \mathbf{U}_B(\overline{\mathbf{B}}^H \mathbf{R}_{I,dl}\overline{\mathbf{B}})^{-1}(\overline{\mathbf{B}}^H \mathbf{R}_{I,c}\overline{\mathbf{B}})(\overline{\mathbf{B}}^H \mathbf{R}_{I,dl}\overline{\mathbf{B}})^{-1}\mathbf{U}_B^H \Delta\overline{\mathbf{D}}\}\mathbf{V}_D \Sigma_D^{-1} \mathbf{U}_D^H \mathbf{R}_{I,dl}\overline{\mathbf{B}}\mathbf{U}_{dl}}_{\mathbf{x}_9})\right\}$$

$$(C.1)$$

We note that (C.1) only shows the dominant terms. To explicitly evaluate the RHS of (C.1), we observe that the involved expectations are either of the form $E\{\Delta\overline{\mathbf{D}}\mathbf{Z}_1\Delta\overline{\mathbf{D}}^H\}$ or $E\{\Delta\overline{\mathbf{D}}^H \mathbf{Z}_2\Delta\overline{\mathbf{D}}\}$, for some given matrix $\mathbf{Z}_1 \in \mathbb{C}^{NQ\times NQ}$ and $\mathbf{Z}_2 \in \mathbb{C}^{MQ\times MQ}$. By following the same procedures in the third part in Appendix A (see *[Proof of (3)]*), it can be verified that both $E\{\Delta\overline{\mathbf{D}}\mathbf{Z}_1\Delta\overline{\mathbf{D}}^H\} \in \mathbb{C}^{MQ\times MQ}$ and $E\{\Delta\overline{\mathbf{D}}^H \mathbf{Z}_2\Delta\overline{\mathbf{D}}\} \in \mathbb{C}^{NQ\times NQ}$ are block diagonal. More precisely, let $\mathbf{Z}_1^{(n)}$ and $\mathbf{Z}_2^{(n)}$ be the nth $Q\times Q$ block diagonal submatrix of $\mathbf{Z}_1$ and $\mathbf{Z}_2$. Then the jth $Q\times Q$ block diagonal submatrix of $E\{\Delta\overline{\mathbf{D}}\mathbf{Z}_1\Delta\overline{\mathbf{D}}^H\}$ and $E\{\Delta\overline{\mathbf{D}}^H \mathbf{Z}_2\Delta\overline{\mathbf{D}}\}$ can be respectively obtained as

$$\left[E\left\{\Delta\overline{\mathbf{D}}\mathbf{Z}_1\Delta\overline{\mathbf{D}}^H\right\}\right]^{(j)} = \frac{Q\sigma_v^2}{P}\mathbf{F}_L\mathbf{F}_L^H \odot \sum_{n=1}^N \mathbf{Z}_1^{(n)}, \tag{C.2}$$

and

$$\left[E\left\{ \Delta \overline{\mathbf{D}}^H \mathbf{Z}_2 \Delta \overline{\mathbf{D}} \right\} \right]^{(j)} = \frac{Q\sigma_v^2}{P} \mathbf{F}_L^* \mathbf{F}_L^T \odot \sum_{m=1}^{M} \mathbf{Z}_2^{(m)} . \tag{C.3}$$

Based on (C.2) and (C.3), the matrices $\mathbf{X}_1 \sim \mathbf{X}_9$ are then explicitly computed in Table 3.1. $\qquad\square$

D Proof of Theorem 3.3

We shall note that the matrix $\mathbf{R}_{I,c}$ will be sparse for small σ_v^2 (cf. (3.58) and (3.20)). This implies that, as $\Delta \mathbf{W}_{dl}$ and $\Delta \mathbf{W}_g$ are small, both $Tr(E\{\Delta \mathbf{W}_{dl}^H \mathbf{R}_{I,c} \Delta \mathbf{W}_{dl}\})$ and $Tr(E\{\Delta \mathbf{W}_g^H \mathbf{R}_{I,c} \Delta \mathbf{W}_g\})$ are close to zero in the high SNR region, and hence

$$P_{I,g} - P_{I,dl} \approx Tr\left(\mathbf{W}_g^H \mathbf{R}_I \mathbf{W}_g \right) + Tr(E\{\Delta \mathbf{W}_g^H \overline{\mathbf{D}}\overline{\mathbf{D}}^H \Delta \mathbf{W}_g\}) - Tr(\mathbf{W}_{dl}^H \mathbf{R}_I \mathbf{W}_{dl}) - Tr(E\{\Delta \mathbf{W}_{dl}^H \overline{\mathbf{D}}\overline{\mathbf{D}}^H \Delta \mathbf{W}_{dl}\}) . \tag{D.1}$$

With $\Delta \mathbf{W}_{dl}$ in (B.1) and since $\overline{\mathbf{B}}^H \overline{\mathbf{D}} = \mathbf{0}$, we have $\Delta \mathbf{W}_{dl}^H \overline{\mathbf{D}} = \Delta \overline{\mathbf{D}}^H \overline{\mathbf{D}} - \mathbf{U}_{dl}^H \overline{\mathbf{B}}^H \Delta \overline{\mathbf{D}}$, and similarly $\Delta \mathbf{W}_g^H \overline{\mathbf{D}} = \Delta \overline{\mathbf{D}}^H \overline{\mathbf{D}} - \mathbf{U}_g^H \overline{\mathbf{B}}^H \Delta \overline{\mathbf{D}}$. This implies

$$\begin{aligned} Tr\left(E\{\Delta \mathbf{W}_{dl}^H \overline{\mathbf{D}}\overline{\mathbf{D}}^H \Delta \mathbf{W}_{dl}\} \right) &= Tr\left(E\{\Delta \overline{\mathbf{D}}^H \overline{\mathbf{D}}\overline{\mathbf{D}}^H \Delta \overline{\mathbf{D}}\} \right) + Tr\left(\mathbf{U}_{dl}^H \overline{\mathbf{B}}^H E\{\Delta \overline{\mathbf{D}}\Delta \overline{\mathbf{D}}^H\} \overline{\mathbf{B}} \mathbf{U}_{dl} \right) \\ &\quad + Tr\left(E\{\Delta \overline{\mathbf{D}}^H \overline{\mathbf{D}}\Delta \overline{\mathbf{D}}^H\} \overline{\mathbf{B}} \mathbf{U}_{dl} \right) + Tr\left(\mathbf{U}_{dl}^H \overline{\mathbf{B}}^H E\{\Delta \overline{\mathbf{D}}\overline{\mathbf{D}}^H \Delta \overline{\mathbf{D}}\} \right) \end{aligned} \tag{D.2}$$

and

$$\begin{aligned} Tr\left(E\{\Delta \mathbf{W}_g^H \overline{\mathbf{D}}\overline{\mathbf{D}}^H \Delta \mathbf{W}_g\} \right) &= Tr\left(E\{\Delta \overline{\mathbf{D}}^H \overline{\mathbf{D}}\overline{\mathbf{D}}^H \Delta \overline{\mathbf{D}}\} \right) + Tr\left(\mathbf{U}_g^H \overline{\mathbf{B}}^H E\{\Delta \overline{\mathbf{D}}\Delta \overline{\mathbf{D}}^H\} \overline{\mathbf{B}} \mathbf{U}_g \right) \\ &\quad + Tr\left(E\{\Delta \overline{\mathbf{D}}^H \overline{\mathbf{D}}\Delta \overline{\mathbf{D}}^H\} \overline{\mathbf{B}} \mathbf{U}_g \right) + Tr\left(\mathbf{U}_g^H \overline{\mathbf{B}}^H E\{\Delta \overline{\mathbf{D}}\overline{\mathbf{D}}^H \Delta \overline{\mathbf{D}}\} \right) \end{aligned} . \tag{D.3}$$

The circularity condition of $\Delta \mathbf{h}^{(m)}$ implies the last two terms on the RHS of both (D.2) and (D.3) are identically zero. From (D.2), (D.3), and since $\overline{\mathbf{B}}^H E\{\Delta \overline{\mathbf{D}}\Delta \overline{\mathbf{D}}^H\} \overline{\mathbf{B}} = \overline{\gamma}$ (see (A.3) and (A.6)), we have

$$Tr\left(E\{\Delta \mathbf{W}_g^H \overline{\mathbf{D}}\overline{\mathbf{D}}^H \Delta \mathbf{W}_g\} \right) - Tr\left(E\{\Delta \mathbf{W}_{dl}^H \overline{\mathbf{D}}\overline{\mathbf{D}}^H \Delta \mathbf{W}_{dl}\} \right) = \overline{\gamma} \left\{ Tr\left(\mathbf{U}_g^H \mathbf{U}_g \right) - Tr\left(\mathbf{U}_{dl}^H \mathbf{U}_{dl} \right) \right\} . \tag{D.4}$$

With (D.1) and (D.4), it follows

$$P_{I,g} - P_{I,dl} \approx Tr\left(\mathbf{W}_g^H \mathbf{R}_I \mathbf{W}_g \right) - Tr\left(\mathbf{W}_{dl}^H \mathbf{R}_I \mathbf{W}_{dl} \right) + \overline{\gamma} \left\{ Tr\left(\mathbf{U}_g^H \mathbf{U}_g \right) - Tr\left(\mathbf{U}_{dl}^H \mathbf{U}_{dl} \right) \right\} . \tag{D.5}$$

According to the matrix inversion lemma [58], we have

$$\left(\overline{\mathbf{B}}^H \mathbf{R}_I \overline{\mathbf{B}} + \overline{\gamma} \mathbf{I}_{(M-N)Q} \right)^{-1} = \left(\overline{\mathbf{B}}^H \mathbf{R}_I \overline{\mathbf{B}} \right)^{-1} - \left(\overline{\mathbf{B}}^H \mathbf{R}_I \overline{\mathbf{B}} \right)^{-1} \left[\overline{\gamma}^{-1} \mathbf{I}_{(M-N)Q} + \left(\overline{\mathbf{B}}^H \mathbf{R}_I \overline{\mathbf{B}} \right)^{-1} \right]^{-1} \left(\mathbf{B}^H \mathbf{R}_I \mathbf{B} \right)^{-1} . \tag{D.6}$$

Using (D.6) and by definitions of $\mathbf{U}_{dl}$ and $\mathbf{U}_g$ (see Lemma B.1 and (3.20)), it can be verified that

$$\mathbf{U}_{dl} = \mathbf{U}_g - \left(\mathbf{I}_{(M-N)Q} + \overline{\gamma}^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\right)^{-1}\mathbf{U}_g, \tag{D.7}$$

and hence

$$\mathbf{W}_{dl} = \overline{\mathbf{D}} - \overline{\mathbf{B}}\mathbf{U}_{dl} = \underbrace{\overline{\mathbf{D}} - \overline{\mathbf{B}}\mathbf{U}_g}_{\mathbf{W}_g} + \mathbf{B}\left(\mathbf{I}_{(M-N)Q} + \overline{\gamma}^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\right)^{-1}\mathbf{U}_g. \tag{D.8}$$

Substituting (D.7) and (D.8) into (D.5), we have

$$P_{I,g} - P_{I,dl} = 2\overline{\gamma}Tr\left(\mathbf{U}_g^H\left(\mathbf{I}_{(M-N)Q} + \overline{\gamma}^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\right)^{-1}\mathbf{U}_g\right) - \overline{\gamma}Tr\left(\mathbf{U}_g^H\left(\mathbf{I}_{(M-N)Q} + \overline{\gamma}^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\right)^{-2}\mathbf{U}_g\right)$$

$$-Tr\left(\mathbf{U}_g^H\left(\mathbf{I}_{(M-N)Q} + \overline{\gamma}^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\right)^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\left(\mathbf{I}_{(M-N)Q} + \overline{\gamma}^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\right)^{-1}\mathbf{U}_g\right); \tag{D.9}$$

We observe that the second and the third terms on the RHS of (D.9) can be further combined into

$$-\overline{\gamma}Tr\left(\mathbf{U}_g^H\left(\mathbf{I}_{(M-N)Q} + \overline{\gamma}^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\right)^{-2}\mathbf{U}_g\right)$$

$$-Tr\left(\mathbf{U}_g^H\left(\mathbf{I}_{(M-N)Q} + \overline{\gamma}^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\right)^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\left(\mathbf{I}_{(M-N)Q} + \overline{\gamma}^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\right)^{-1}\mathbf{U}_g\right)$$

$$= -\overline{\gamma}Tr\left(\mathbf{U}_g^H\left(\mathbf{I}_{(M-N)Q} + \overline{\gamma}^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\right)^{-1}\mathbf{U}_g\right). \tag{D.10}$$

With (D.9) and (D.10), it follows that

$$P_{I,g} - P_{I,dl} \approx 2\overline{\gamma}Tr\left(\mathbf{U}_g^H\left(\mathbf{I}_{(M-N)Q} + \overline{\gamma}^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\right)^{-1}\mathbf{U}_g\right) - \overline{\gamma}Tr\left(\mathbf{U}_g^H\left(\mathbf{I}_{(M-N)Q} + \overline{\gamma}^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\right)^{-1}\mathbf{U}_g\right)$$

$$= \overline{\gamma}Tr\left(\mathbf{U}_g^H\left(\mathbf{I}_{(M-N)Q} + \overline{\gamma}^{-1}\overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\right)^{-1}\mathbf{U}_g\right) = \overline{\gamma}^2Tr\left(\mathbf{U}_g^H\left(\overline{\gamma}\mathbf{I}_{(M-N)Q} + \overline{\mathbf{B}}^H\mathbf{R}_I\overline{\mathbf{B}}\right)^{-1}\mathbf{U}_g\right).$$

$$\square$$

E   Derivations of (4.31) and (4.32)

With (4.26), (4.27) and by averaging over source symbols and channel noises, we have

$$E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{z}}_b^{(j)}(t)^H\right\} = E_e\left\{\widehat{\mathbf{B}}_{av}^{(j)H}\left(\sum_{i=1}^{L+1}\mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\widehat{\mathbf{H}}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)^H + \sigma_v^2\mathbf{I}_{MQ} + E_h\{\mathbf{R}_{\delta\mathbf{H}}(t)\}\right)\widehat{\mathbf{B}}_{av}^{(j)}\right\}, \tag{E.1}$$

and

$$E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{i}}^{(j)}(t)^H\right\} = E_e\left\{\widehat{\mathbf{B}}_{av}^{(j)H}\left(\sum_{i=1,i\neq j}^{L+1}\mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\widehat{\mathbf{H}}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)^H + \sigma_v^2\mathbf{I}_{MQ} + E_h\{\mathbf{R}_{\delta\mathbf{H}}^{(j)}(t)\}\right)\widehat{\mathbf{H}}_{av}^{(j)}\right\}, \tag{E.2}$$

where $E_e\{\cdot\}$ and $E_h\{\cdot\}$ denote the expectations taken with respect to the channel estimation error and channel temporal variation, respectively,

$$\mathbf{R}_{\delta\mathbf{H}}(t) = \sum_{i=1}^{L+1} \delta\mathbf{H}^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)^H + \sum_{i=1}^{L+1}\mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\delta\mathbf{H}^{(i)}(t)^H + \sum_{i=1}^{L+1}\delta\mathbf{H}^{(i)}(t)\delta\mathbf{H}^{(i)}(t)^H \, , \quad \text{(E.3)}$$

and

$$\mathbf{R}_{\delta\mathbf{H}}^{(j)}(t) = \sum_{i=1,i\neq j}^{L+1} \delta\mathbf{H}^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)^H + \sum_{i=1,i\neq j}^{L+1}\mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\delta\mathbf{H}^{(i)}(t)^H + \sum_{i=1}^{L+1}\delta\mathbf{H}^{(i)}(t)\delta\mathbf{H}^{(i)}(t)^H \, ,$$

$$\text{(E.4)}$$

in which

$$\mathbf{D}_c^{(i)}(t) := diag\left\{\mathbf{J}^{i-1}[\mathbf{a}_1(t)\cdots\mathbf{a}_P(t)]^T\right\} \otimes \mathbf{I}_M \in \mathbb{C}^{MQ\times MQ} \, , \quad \text{(E.5)}$$

with

$$\mathbf{a}_p(t) = \left[\rho\left(\bar{t} + ((p-1)(L+1)+i-1)_Q, 0\right)\frac{\sigma_0}{\widehat{\sigma}_0}, ..., \rho\left(\bar{t} + ((p-1)(L+1)+L+i-1)_Q, L\right)\frac{\sigma_L}{\widehat{\sigma}_L}\right] \in \mathbb{R}^{1\times(L+1)} \, ,$$

$\rho(t,l) = \mathbf{R}_h(t,l)(\sigma_l\widehat{\sigma}_l)^{-1}$, $\widehat{\sigma}_l^2 = \sigma_l^2 + [\mathbf{R}_{\Delta\mathbf{h},m}]_{(n-1)(L+1)+l+1,(n-1)(L+1)+l+1}$, and $\delta\mathbf{H}^{(i)}(t) \in \mathbb{C}^{MQ\times NP}$ has the (m,n)th $Q\times P$ block submatrix whose pth column $\delta\mathbf{c}_{m,n}^{(i,p)}(t) \in \mathbb{C}^Q$ is given by

$$\delta\mathbf{c}_{m,n}^{(i,p)}(t)$$
$$= \mathbf{J}^{i+(p-1)(L+1)-1}[\underbrace{\delta h_{m,n}(\bar{t} + ((p-1)(L+1)+i-1)_Q,0),...,\delta h_{m,n}(\bar{t} + ((p-1)(L+1)+i+L-1)_Q,L),0,...,0}_{:=\delta\mathbf{h}_{m,n}^{(i,p)T}}]^T$$

$$\text{(E.6)}$$

where, for a fixed t, $\delta h_{m,n}(t,l)$, $0 \le l \le L$, are zero-mean Gaussian random variables with variance $\sigma_l^2\left(1 - |\rho(t,l)|^2\right)$. We claim that

$$E_h\left\{\mathbf{R}_{\delta\mathbf{H}}(t)\right\} = E_h\left\{\mathbf{R}_{\delta\mathbf{H}}^{(j)}(t)\right\} = N\mathbf{I}_M \otimes \mathbf{D}_Q(t) \, , \quad \text{(E.7)}$$

the result then follows from (E.1), (E.2), and (E.7).

[Proof of (E.7)]: Since the nonzero entries of $\delta\mathbf{H}^{(i)}(t)$ are zero-mean Gaussian random variables, we have $E_h\{\delta\mathbf{H}^{(i)}(t)\} = \mathbf{0}_{MQ\times NP}$, for $1 \le i \le L+1$, and from (E.3) and (E.4) we have

$$E_h\left\{\mathbf{R}_{\delta\mathbf{H}}(t)\right\} = E_h\left\{\mathbf{R}_{\delta\mathbf{H}}^{(j)}(t)\right\} = E_h\left\{\sum_{i=1}^{L+1}\delta\mathbf{H}^{(i)}(t)\delta\mathbf{H}^{(i)}(t)^H\right\} . \quad \text{(E.8)}$$

Let $Diag\{\mathbf{C}_1,...,\mathbf{C}_m\} \in \mathbb{C}^{mn\times mn}$ be the block diagonal matrix with the diagonal block submatrices $\mathbf{C}_p \in \mathbb{C}^{n\times n}$, $1 \le p \le m$. By the definition of $\delta\mathbf{H}^{(i)}(t) \in \mathbb{C}^{MQ\times NP}$ in (E.6), the (m_1,m_2)th $Q\times Q$ block submatrix of $\sum_{i=1}^{L+1}\delta\mathbf{H}^{(i)}(t)\delta\mathbf{H}^{(i)}(t)^H \in \mathbb{C}^{MQ\times MQ}$, $1 \le m_1,m_2 \le M$, is given by

$$\left[\sum_{i=1}^{L+1}\delta\mathbf{H}^{(i)}(t)\delta\mathbf{H}^{(i)}(t)^H\right]^{(m_1,m_2)} = \sum_{i=1}^{L+1}\sum_{n=1}^{N}\left[\delta\mathbf{H}^{(i)}(t)\right]^{(m_1,n)}\left[\delta\mathbf{H}^{(i)}(t)\right]^{(m_2,n)H}$$

$$= \sum_{i=1}^{L+1}\sum_{n=1}^{N}\sum_{p=1}^{P}\mathbf{J}^{i+(p-1)(L+1)-1}Diag\left(\delta\mathbf{h}_{m_1,n}^{(i,p)}(t)\delta\mathbf{h}_{m_2,n}^{(i,p)}(t)^H, \mathbf{0}_{L+1},...,\mathbf{0}_{L+1}\right)\left(\mathbf{J}^{i+(p-1)(L+1)-1}\right)^T.$$

(E.9)

Since the elements of $\delta\mathbf{h}_{m,n}^{(i,p)}(t)$, $\forall m,n$, are independent (see (E.6)), we have

$$E_h\left\{\delta\mathbf{h}_{m_1,n}^{(i,p)}(t)\delta\mathbf{h}_{m_2,n}^{(i,p)}(t)^H\right\} = \begin{cases}\mathbf{D}_p^{(i)}(t), & m_1 = m_2 \\ \mathbf{0}_{L+1}, & m_1 \neq m_2\end{cases},$$

(E.10)

$\mathbf{D}_p^{(i)}(t) \in \mathbb{C}^{(L+1)\times(L+1)}$ is a diagonal matrix with the lth diagonal element $[\mathbf{D}_p^{(i)}(t)]_{l,l} = \sigma_{l-1}^2(1 - |\rho(\bar{t}+(i+(p-1)(L+1)+l-2)_Q, l-1)|^2)$ for $1 \leq l \leq L+1$. Equation (E.10) directly implies that $E_h\left\{\sum_{i=1}^{L+1}\delta\mathbf{H}^{(i)}(t)\delta\mathbf{H}^{(i)}(t)^H\right\}$ is a diagonal matrix with the mth $Q \times Q$ diagonal block given by

$$E_h\left\{\left[\sum_{i=1}^{L+1}\delta\mathbf{H}^{(i)}(t)\delta\mathbf{H}^{(i)}(t)^H\right]^{(m,m)}\right\} = N\sum_{i=1}^{L+1}\sum_{p=1}^{P}\mathbf{J}^{i+(p-1)(L+1)-1}Diag\{\mathbf{D}_p^{(i)}(t), \mathbf{0}_{L+1},...,\mathbf{0}_{L+1}\}\left(\mathbf{J}^{i+(p-1)(L+1)-1}\right)^T.$$

(E.11)

By some direct rearrangements and since $Q=P(L+1)$, (E.11) becomes

$$E_h\left\{\left[\sum_{i=1}^{L+1}\delta\mathbf{H}^{(i)}(t)\delta\mathbf{H}^{(i)}(t)^H\right]^{(m,m)}\right\} = N\mathbf{D}_Q(t),$$

(E.12)

where $\mathbf{D}_Q(t) := diag\left\{\sum_{l=0}^{L}\sigma_l^2\left(1 - |\rho(\bar{t},l))|^2\right),...,\sum_{l=0}^{L}\sigma_l^2\left(1 - |\rho(\bar{t}+Q-1,l))|^2\right)\right\} \in \mathbb{C}^{MQ\times MQ}$. $\square$

Equation (E.7) directly follows from (E.12).

F Derivations of (4.35) and (4.36)

To derive (4.35) and (4.36), we have to determine the expectation quantities associated with the channel estimation error in (4.31) and (4.32). Based on (4.2), the (m,n)th block submatrices of $\mathbf{H}_{av}^{(i)} \in \mathbb{C}^{MQ\times NP}$ and $\Delta\mathbf{H}_{av}^{(i)} \in \mathbb{C}^{MQ\times NP}$ are respectively defined as

$$\left[\mathbf{H}_{av}^{(i)}\right]^{(m,n)} = \bar{\mathbf{J}}_i\left(\left[\mathbf{h}_{m,n}^{(av)T}, 0,...,0\right]^T \otimes \mathbf{I}_P\right) \in \mathbb{C}^{Q\times P},$$

(F.1)

and

$$\left[\Delta\mathbf{H}_{av}^{(i)}\right]^{(m,n)} = \overline{\mathbf{J}}_i\left(\left[\Delta\mathbf{h}_{m,n}^T, 0,...,0\right]^T \otimes \mathbf{I}_P\right) \in \mathbb{C}^{Q\times P}, \ 1 \le n \le N, \ 1 \le m \le M, \qquad \text{(F.2)}$$

where $\Delta\mathbf{h}_{m,n} \in \mathbb{C}^{L+1}$ is the nth subvector of $\Delta\mathbf{h}_m$ (see (4.5)) and $\overline{\mathbf{J}}_i \in \mathbb{R}^{Q\times PQ}$ denotes a matrix with the pth $Q \times Q$ block submatrix given by $\left[\overline{\mathbf{J}}_i\right]^{(1,p)} = \mathbf{J}^{i+(p-1)(L+1)-1}$ ($\mathbf{J}$ is defined in (4.3)). Assuming that the channel is slowly varying and SNR is high, the channel estimation error is thus small and $\Delta\mathbf{H}_{av}^{(i)}$ and $\Delta\mathbf{B}_{av}^{(j)}$ (see (4.34)) are close to zero matrices. Substituting (4.34) into (4.31) and using the fact that $E_e\{\Delta\mathbf{B}_{av}^{(j)H}\} = \mathbf{0}_{MQ\times(MQ-NP)}$ (cf. (4.34)) and the circularity condition of $\Delta\mathbf{h}_{m,n}^{(av)}$, by keeping only the first- and second-order terms of estimation error (4.31) can be expanded as

$$E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{z}}_b^{(j)}(t)^H\right\} = \mathbf{B}_{av}^{(j)H}\left(\mathbf{R}_L^{(j)}(t) + \overline{\mathbf{R}}_I^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)\right)\mathbf{B}_{av}^{(j)} + \mathbf{B}_{av}^{(j)H}E_e\left\{\Delta\mathbf{H}_{av}^{(j)}\mathbf{X}^{(j)}(t)\Delta\mathbf{H}_{av}^{(j)H}\right\}\mathbf{B}_{av}^{(j)}$$

$$+\mathbf{B}_{av}^{(j)H}\left(\sum_{i=1}^{L+1}\mathbf{D}_c^{(i)}(t)E_e\left\{\Delta\mathbf{H}_{av}^{(i)}\Delta\mathbf{H}_{av}^{(i)H}\right\}\mathbf{D}_c^{(i)}(t)^H\right)\mathbf{B}_{av}^{(j)} + \mathbf{B}_{av}^{(j)H}\left(\sum_{i=1}^{L+1}\mathbf{D}_c^{(i)}(t)E_e\left\{\Delta\mathbf{H}_{av}^{(i)}\mathbf{X}^{(i,j)}(t)\Delta\mathbf{H}_{av}^{(j)H}\right\}\right)\mathbf{B}_{av}^{(j)}$$

$$+\mathbf{B}_{av}^{(j)H}\left(\sum_{i=1}^{L+1}E_e\left\{\Delta\mathbf{H}_{av}^{(j)}\mathbf{X}^{(i,j)}(t)^H\Delta\mathbf{H}_{av}^{(i)H}\right\}\mathbf{D}_c^{(i)}(t)^H\right)\mathbf{B}_{av}^{(j)}, \qquad \text{(F.3)}$$

where $\mathbf{R}_L^{(j)}(t)$ and $\overline{\mathbf{R}}_I^{(j)}(t)$ are defined in (4.37) and (4.38), respectively, and the matrices $\mathbf{X}^{(j)}(t)$ and $\mathbf{X}^{(i,j)}(t)$ are defined in Table 4.1. Similarly, (4.32) also can be expanded as

$$E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{i}}^{(j)}(t)^H\right\} = \mathbf{B}_{av}^{(j)H}\left(\overline{\mathbf{R}}_I^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)\right)\mathbf{H}_{av}^{(j)} + \mathbf{B}_{av}^{(j)H}\left(\sum_{i=1,i\ne j}^{L+1}\mathbf{D}_c^{(i)}(t)E_e\left\{\Delta\mathbf{H}_{av}^{(i)}\Delta\mathbf{H}_{av}^{(i)H}\right\}\mathbf{D}_c^{(i)}(t)\right)\mathbf{H}_{av}^{(j)}$$

$$+\mathbf{B}_{av}^{(j)H}\left(\sum_{i=1,i\ne j}^{L+1}\mathbf{D}_c^{(i)}(t)\mathbf{H}_{av}^{(i)}E_e\left\{\Delta\mathbf{H}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)^H\Delta\mathbf{H}_{av}^{(j)}\right\}\right) + \mathbf{B}_{av}^{(j)H}\left(\sum_{i=1,i\ne j}^{L+1}E_e\left\{\Delta\mathbf{H}_{av}^{(j)}\mathbf{X}^{(i,j)}(t)^H\Delta\mathbf{H}_{av}^{(i)H}\right\}\mathbf{D}_c^{(i)}(t)^H\right)\mathbf{H}_{av}^{(j)}$$

$$\text{(F.4)}$$

We observe that, in (F.3) and (F.4), each of the involved expectations admit one of the following forms: $E_e\{\Delta\mathbf{H}_{av}^{(i)}\Delta\mathbf{H}_{av}^{(i)H}\}$, $E_e\{\Delta\mathbf{H}_{av}^{(i)}\mathbf{Z}_1\Delta\mathbf{H}_{av}^{(j)H}\}$ and $E_e\{\Delta\mathbf{H}_{av}^{(i)H}\mathbf{Z}_2\Delta\mathbf{H}_{av}^{(j)}\}$, for some given matrices $\mathbf{Z}_1 \in \mathbb{C}^{NP\times NP}$ and $\mathbf{Z}_2 \in \mathbb{C}^{MQ\times MQ}$. To proceed, we require the following lemma.

Lemma F.1: The following results hold.

(F1) $\mathbf{D}_H^{(i)} := E_e\left\{\Delta\mathbf{H}_{av}^{(i)}\Delta\mathbf{H}_{av}^{(i)H}\right\} \in \mathbb{C}^{MQ\times MQ}$ is a block diagonal matrix with the mth $Q \times Q$ diagonal block defined as

$$\left[\mathbf{D}_H^{(i)}\right]^{(m,m)} := \sum_{n=1}^N\sum_{p=1}^P \mathbf{J}^{i+(p-1)(L+1)-1}Diag\left\{\mathbf{R}_{\Delta\mathbf{h},m}^{(n,n)}, \mathbf{0}_{L+1},...,\mathbf{0}_{L+1}\right\}\left[\mathbf{J}^{i+(p-1)(L+1)-1}\right]^T, \quad \text{(F.5)}$$

in which $\mathbf{R}_{\Delta\mathbf{h},m}^{(n,n)} \in \mathbb{C}^{(L+1)\times(L+1)}$ denotes the nth diagonal block submatrix of $\mathbf{R}_{\Delta\mathbf{h},m} \in \mathbb{C}^{N(L+1)\times N(L+1)}$ (see Section 4.2.2).

(F2) $\mathbf{D}_{i,j}(\mathbf{Z}_1) := E_e\left\{\Delta\mathbf{H}_{av}^{(i)}\mathbf{Z}_1\Delta\mathbf{H}_{av}^{(j)H}\right\} \in \mathbb{C}^{MQ\times MQ}$ is a block diagonal matrix with the mth $Q \times Q$ diagonal block given by

$$\left[\mathbf{D}_{i,j}(\mathbf{Z}_1)\right]^{(m,m)} = \sum_{n_1=1}^{N}\sum_{n_2=1}^{N} \overline{\mathbf{J}}_i\left(\mathbf{Z}_1^{(n_1,n_2)} \otimes Diag\left\{\mathbf{R}_{\Delta\mathbf{h},m}^{(n_1,n_2)}, \mathbf{0}_{L+1}\cdots\mathbf{0}_{L+1}\right\}\right)\overline{\mathbf{J}}_j^T . \tag{F.6}$$

(F3) $\mathbf{K}_{i,j}(\mathbf{Z}_2) := E_e\left\{\Delta\mathbf{H}_{av}^{(i)H}\mathbf{Z}_2\Delta\mathbf{H}_{av}^{(j)}\right\} \in \mathbb{C}^{NP\times NP}$ is a matrix with the (n_1,n_2)th $P \times P$ block submatrix whose (p_1,p_2)th entry given by

$$\left[\mathbf{K}_{i,j}(\mathbf{Z}_2)\right]_{p_1,p_2}^{(n_1,n_2)} = \sum_{m=1}^{M} vec\left(\left[\overline{\mathbf{J}}_i^T\mathbf{Z}_2^{(m,m)}\overline{\mathbf{J}}_j\right]_1^{(p_1,p_2)}\right)^T vec\left(\mathbf{R}_{\Delta\mathbf{h},m}^{(n_1,n_2)}\right)^* , \tag{F.7}$$

where $vec(.)$ stacks a matrix into a vector columnwise and $[\overline{\mathbf{J}}_i^T\mathbf{Z}_2^{(m,m)}\overline{\mathbf{J}}_j]_1^{(p_1,p_2)}$ denotes the first $(L+1)\times(L+1)$ diagonal block submatrix of the (p_1,p_2)th $Q \times Q$ block submatrix of $\overline{\mathbf{J}}_i^T\mathbf{Z}_2^{(m,m)}\overline{\mathbf{J}}_j \in \mathbb{C}^{PQ\times PQ}$. $\qquad\qquad\square$

[Proof of (F1)]: By the definition of $\Delta\mathbf{H}_{av}^{(i)}$ in (F.2), the (m_1,m_2)th $Q \times Q$ submatrix of $E_e\left\{\Delta\mathbf{H}_{av}^{(i)}\Delta\mathbf{H}_{av}^{(i)H}\right\}$ is given by

$$\left[E\left\{\Delta\mathbf{H}_{av}^{(i)}\Delta\mathbf{H}_{av}^{(i)H}\right\}\right]^{(m_1,m_2)} = \sum_{n=1}^{N}\sum_{p=1}^{P}\mathbf{J}^{i+(p-1)(L+1)-1}Diag\left\{E\left\{\Delta\mathbf{h}_{m_1,n}\Delta\mathbf{h}_{m_2,n}^H\right\}, \mathbf{0}_{L+1},...,\mathbf{0}_{L+1}\right\}\left[\mathbf{J}^{i+(p-1)(L+1)-1}\right]^T .$$
$$\tag{F.8}$$

Based on (4.7) and (4.8), we have

$$E\left\{\Delta\mathbf{h}_{m_1,n}\Delta\mathbf{h}_{m_2,n}^H\right\} = \begin{cases} \mathbf{R}_{\Delta\mathbf{h},m_1}^{(n,n)}, & if\ m_1 = m_2 \\ \mathbf{0}_{L+1}, & if\ m_1 \neq m_2 \end{cases}, \tag{F.9}$$

Equation (F.9) implies that $E_e\left\{\Delta\mathbf{H}_{av}^{(i)}\Delta\mathbf{H}_{av}^{(i)H}\right\}$ is a block diagonal and (F.5) can be directly obtained by back substitution of (F.9) into (F.8).

[Proof of (F2)]: The (m_1,m_2)th $Q \times Q$ block submatrix of $\mathbf{D}_{i,j}(\mathbf{Z}_1)$ is expressed as

$$\left[\mathbf{D}_{i,j}(\mathbf{Z}_1)\right]^{(m_1,m_2)} = \sum_{n_1=1}^{N}\sum_{n_2=1}^{N} E_e\left\{\left[\Delta\mathbf{H}_{av}^{(i)}\right]^{(m_1,n_1)}\mathbf{Z}_1^{(n_1,n_2)}\left[\Delta\mathbf{H}_{av}^{(j)}\right]^{(m_2,n_2)H}\right\}, \tag{F.10}$$

where $\mathbf{Z}_1^{(n_1,n_2)} \in \mathbb{C}^{P\times P}$ denotes the (n_1,n_2)th block submatrix of $\mathbf{Z}_1$. Since the channel estimation errors between different receive antennas are independent (see (4.8)), we have $E_e\left\{\left[\Delta\mathbf{H}_{av}^{(i)}\right]^{(m_1,n_1)}\mathbf{Z}_1^{(n_1,n_2)}\left[\Delta\mathbf{H}_{av}^{(j)}\right]^{(m_2,n_2)H}\right\} = \mathbf{0}_{MQ}$ for $m_1 \neq m_2$, which implies

101

$$\left[\mathbf{D}_{i,j}(\mathbf{Z}_1) \right]^{(m_1,m_2)} = \begin{cases} \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} E_e \left\{ \left[\Delta\mathbf{H}_{av}^{(i)} \right]^{(m_1,n_1)} \mathbf{Z}_1^{(n_1,n_2)} \left[\Delta\mathbf{H}_{av}^{(j)} \right]^{(m_1,n_2)H} \right\}, m_1 = m_2 \\ \mathbf{0}_{MQ}, m_1 \neq m_2 \end{cases}. \quad \text{(F.11)}$$

From (F.11), $\mathbf{D}_{i,j}(\mathbf{Z}_1)$ is seen to be a block diagonal matrix. Furthermore, with (F.2), we have

$$E_e \left\{ \left[\Delta\mathbf{H}_{av}^{(i)} \right]^{(m,n_1)} \mathbf{Z}_1^{(n_1,n_2)} \left[\Delta\mathbf{H}_{av}^{(j)} \right]^{(m,n_2)H} \right\} = \bar{\mathbf{J}}_i \mathbf{R}_m(n_1, n_2) \bar{\mathbf{J}}_j^T, \quad \text{(F.12)}$$

in which the (p_1, p_2)th $Q \times Q$ block submatrix of $\mathbf{R}_m(n_1, n_2) \in \mathbb{C}^{PQ \times PQ}$, $1 \leq p_1, p_2 \leq P$, is given by

$$\left[\mathbf{R}_m(n_1, n_2) \right]^{(p_1,p_2)} = Diag \left\{ \left[\mathbf{Z}_1^{(n_1,n_2)} \right]_{p_1,p_2} \mathbf{R}_{\Delta\mathbf{h},m}^{(n_1,n_2)}, \mathbf{0}_{L+1}, ..., \mathbf{0}_{L+1} \right\}. \quad \text{(F.13)}$$

By back substitution of (F.12) and (F.13) into (F.11), (F.6) can be obtained via some rearrangements.

[Proof of (F3)]: With $\Delta\mathbf{H}_{av}^{(i)}$ given in (F.2), the (n_1, n_2)th $P \times P$ block submatrix of $\mathbf{K}_{i,j}(\mathbf{Z}_2)$, $1 \leq p_1, p_2 \leq P$, is expressed as

$$\left[\mathbf{K}_{i,j}(\mathbf{Z}_2) \right]^{(n_1,n_2)} = \sum_{m_1=1}^{M} \sum_{m_2=1}^{M} E_e \left\{ \left[\Delta\mathbf{H}_{av}^{(i)} \right]^{(m_1,n_1)H} \mathbf{Z}_2^{(m_1,m_2)} \left[\Delta\mathbf{H}_{av}^{(j)} \right]^{(m_2,n_2)} \right\}. \quad \text{(F.14)}$$

Since the channel estimation errors between different receive antennas are independent (see (4.8)), we have $E_e \left\{ [\Delta\mathbf{H}_{av}^{(i)}]^{(m_1,n_1)H} \mathbf{Z}_2^{(m_1,m_2)} [\Delta\mathbf{H}_{av}^{(j)}]^{(m_2,n_2)} \right\} = \mathbf{0}_P$ for $m_1 \neq m_2$, which implies

$$\left[\mathbf{K}_{i,j}(\mathbf{Z}_2) \right]^{(n_1,n_2)} = \sum_{m=1}^{M} E_e \left\{ \left[\Delta\mathbf{H}_{av}^{(i)} \right]^{(m,n_1)H} \mathbf{Z}_2^{(m,m)} \left[\Delta\mathbf{H}_{av}^{(j)} \right]^{(m,n_2)} \right\}. \quad \text{(F.15)}$$

Furthermore, based on (F.2), direct expansion results in

$$E_e \left\{ [\Delta\mathbf{H}_{av}^{(i)}]^{(m,n_1)H} \mathbf{Z}_2^{(m,m)} [\Delta\mathbf{H}_{av}^{(j)}]^{(m,n_2)} \right\} = E_e \left\{ \left[[\Delta\mathbf{h}_{m,n_1}^{H}, 0, ..., 0] \otimes \mathbf{I}_P \right] \bar{\mathbf{J}}_i^T \mathbf{Z}_2^{(m,m)} \bar{\mathbf{J}}_j \left[[\Delta\mathbf{h}_{m,n_2}, 0, ..., 0]^T \otimes \mathbf{I}_P \right] \right\}$$

$$\text{(F.16)}$$

By some direct manipulations on the RHS of (F.16), the (p_1, p_2)th entry of $E_e \left\{ [\Delta\mathbf{H}_{av}^{(i)}]^{(m_1,n)H} \mathbf{Z}_2^{(m,m)} [\Delta\mathbf{H}_{av}^{(j)}]^{(m_2,n)} \right\}$, $1 \leq p_1, p_2 \leq P$, is given by

$$\left[E_e \left\{ [\Delta\mathbf{H}_{av}^{(i)}]^{(m,n_1)H} \mathbf{Z}_2^{(m,m)} [\Delta\mathbf{H}_{av}^{(j)}]^{(m,n_2)} \right\} \right]_{p_1,p_2} = vec \left(\left[\bar{\mathbf{J}}_i^T \mathbf{Z}_2^{(m,m)} \bar{\mathbf{J}}_j \right]_1^{(p_1,p_2)} \right)^T vec \left(\mathbf{R}_{\Delta\mathbf{h},m}^{(n_1,n_2)} \right)^*. \quad \text{(F.17)}$$

Back substitution of (F.17) into (F.15) leads to (F.7). $\square$

With (F3), (F.4), and lemma F.1, direct rearrangements lead to (4.35) and (4.36).

G   Derivation of (4.41)

Assuming that the channel temporal variation is piecewise linear in time, the channel estimate at each time instant k (within one burst) can be acquired via the following relation [44]

$$\hat{h}_{m.n}(k,l) = \hat{h}_{m,n}^{(av)}(l) + \hat{\alpha}_l(k - (G + Q/2)), \quad G + Q \le k \le T(G + Q) - 1, \tag{G.1}$$

where $\hat{h}_{m,n}^{(av)}(l)$ denotes the LS estimate of the lth channel tap, $\hat{\alpha}_l = (\hat{h}_{m,n}^{(av,next)}(l) - \hat{h}_{m,n}^{(av)}(l))/T(G + Q)$ is the estimated variation slope with $\hat{h}_{m,n}^{(av,next)}(l)$ being the channel estimate obtained in the next burst. Substituting $\hat{h}_{m,n}^{(av)}(l) = h_{m,n}^{(av)}(l) + \Delta h_{m,n}(l)$ and $\hat{h}_{m,n}^{(av,next)}(l) = h_{m,n}^{(av,next)}(l) + \Delta h_{m,n}^{(next)}(l)$, where $\Delta h_{m,n}(l)$ and $\Delta h_{m,n}^{(next)}(l)$ represent the estimation errors, into (G.1), we have

$$\hat{h}_{m.n}(k,l) = h_{m.n}(k,l) + \Delta h_{m.n}(k,l), \quad G + Q \le k \le T(G + Q) - 1, \tag{G.2}$$

in which $\Delta h_{m.n}(k,l) = \Delta h_{m,n}(l) + \Delta\alpha_l(k - (G + Q/2))$ is the composite channel estimation error with $\Delta\alpha_l = (\Delta h_{m,n}^{(next)}(l) - \Delta h_{m,n}(l))/T(G + Q)$. Based on (G.2) and (4.11), the corresponding estimated signature matrix is then expressed as

$$\hat{\mathbf{H}}^{(i)}(t) = \mathbf{H}^{(i)}(t) + \Delta\mathbf{H}^{(i)}(t), \quad 1 \le i \le L + 1, \tag{G.3}$$

where $\Delta\mathbf{H}^{(i)}(t) = \Delta\mathbf{H}_{av}^{(i)} + \mathbf{D}_T(t)\left(\Delta\mathbf{H}_{av,next}^{(i)} - \Delta\mathbf{H}_{av}^{(i)}\right)$,

$\mathbf{D}_T(t) = \left(T(G + Q)\right)^{-1}\mathbf{I}_M \otimes diag\{\bar{t},...,\bar{t} + Q - 1\}$ and $\Delta\mathbf{H}_{av,next}^{(i)}$ can be formed by simply replacing $\Delta\mathbf{h}_{m,n}$ with $\Delta\mathbf{h}_{m,n}^{(next)} = \left[\Delta h_{m,n}^{(next)}(l),...,\Delta h_{m,n}^{(next)}(L)\right]^T$ in the definition of $\Delta\mathbf{H}_{av}^{(i)}$ (see (F.2)). By following the same procedure in Appendix F, the optimal robust GSC weight for the jth group is given by $\mathbf{W}_a^{(j)}(t) = \mathbf{H}^{(j)}(t) - \mathbf{B}^{(j)}(t)\mathbf{U}_a^{(j)}(t)$ with

$$\mathbf{U}_a^{(j)}(t) = \left(\mathbf{B}^{(j)}(t)^H\left(\mathbf{R}_I^{(j)}(t) + \mathbf{R}_e(t)\right)\mathbf{B}^{(j)}(t)\right)^{-1}\mathbf{B}^{(j)}(t)^H\mathbf{R}_I^{(j)}(t)\mathbf{H}^{(j)}(t), \tag{G.4}$$

where $\mathbf{R}_e(t) \in \mathbb{C}^{MQ \times MQ}$ is defined in Table 4.1. Based on (G.4), the sampled version of the overall GSC weight is shown in (4.41).

Bibliography

[1] N. Al-Dhahir, "FIR channel-shortening equalizers for MIMO ISI channels," *IEEE Trans. Communications*, vol. 49, no. 2, pp. 213-218, Feb. 2001.

[2] N. Al-Dhahir and J. Cioffi, "Efficiently computed reduced-parameter input-aided MMSE equalizers for ML detection: A unified approach," *IEEE Trans. Inform. Theory*, vol. 42, pp. 903–915, May 1996.

[3] N. Al-Dhahir and J. M. Cioffi, "Optimum finite-length equalization for multicarrier transceivers," *IEEE Trans. Commun.*, vol. 44, pp. 56–64, Jan. 1996.

[4] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE JSAC*, vol. 16, no. 8, pp. 1451-1458, Oct. 1998.

[5] G. Arslan, B. L. Evans and S. Kiaei, "Equalization for discrete multitone transceiver to maximize bit rate," *IEEE Trans. Signal Processing*, vol. 49, no. 12, pp. 3123-3135, Dec. 2001.

[6] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO-OFDM systems in mobile wireless channels," *IEEE Trans. Signal Processing*, vol. 51, no. 6, pp. 1615-1624, June 2003.

[7] J. Benesty, Y. Huang and J. Chen, "A fast recursive algorithm for optimum sequential signal detection in a BLAST system," *IEEE Trans. Signal Processing*, vol. 51, no. 7, pp. 1722-1730, July 2003.

[8] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[9] B. R. Breed and J. Strauss, "A short proof of the equivalence of LCMV and GSC beamforming," *IEEE Signal Processing Letters*, vol. 9, no. 6, pp. 168-169, June 2002.

[10] H. Bolcskel, A. J. Paulraj, K. V. S. Hari, R. U. Nabar, and W. W. Lu, "Fixed broadband wireless access: state of the art, challenges, and future directions," *IEEE Communication Mag.*, vol. 39, no. 1, pp. 100-108, Jan. 2001.

[11] B. D. Carlson, "Covariance estimation errors and diagonal loading in adaptive arrays," *IEEE*

Trans. Aerospace and Electronic Systems, vol. 24, no. 4, pp. 397-401, July 1988.

[12] H. Cox, R. Zeskind, and M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 10, pp. 1365–1376, Oct. 1987.

[13] S. Chen and C. Zhu, "ICI and ISI analysis and mitigation for OFDM systems with insufficient cyclic prefix in time-varying channels," *IEEE Trans. Consumer Electronics*, vol. 50, no. 2, pp. 78-83, Feb. 2004.

[14] J. Coon, M. Beach, and J. McGeehan, "Optimal training sequences for channel estimation in cyclic-prefix-based single-carrier systems with transmit diversity," *IEEE Signal Processing Letters*, vol. 11, no. 9, pp. 729-732, Sept. 2004.

[15] J. Coon, S. Armour, M. Beach, and J. McGeehan, "Adaptive frequency-domain for single-carrier multiple-input multiple-output wireless transmissions," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 3247-3256, Aug. 2005.

[16] Ruly Lai-U Choi, K. B. Lataief and R. D. Murch, "Frequency domain pre-equalization with transmit diversity for MISO broadband wireless communications," *Proc. IEEE VTC 2002-Fall*, vol. 3, pp. 1787-1791, Sept. 2002.

[17] H. Ekstrom, A. Furuskar, J. Karlsson, M. Meyer, S. Parkvall, J. Torsner, and M. Wahlqvist, "Technical solutions for the 3G long-term evolution," *IEEE Communications Magazine*, vol. 44, no. 3, pp. 38-45, March 2006.

[18] O. L. Frost, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, no. 8, pp. 926-935, Aug. 1972.

[19] D. D. Falconer, S. L. Ariyavistitakul, A. Benyamin-Seeyar and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 40, no. 4, pp. 58-66, Apr. 2002.

[20] D. D. Falconer and S. L. Ariyavistitakul, "Broadband wireless using single carrier and frequency domain equalization," *IEEE Proc. WPMC 2002*, vol. 1, pp. 27-36, Oct. 2002.

[21] F. H. P. Fitzek and M. D. Katz, *Cooperation in Wireless Networks: Principles and Applications*, Springer, 2006.

[22] J. S. Goldstein and I. S. Reed, "Subspace selection for partially adaptive sensor array processing," *IEEE Trans. Aerospace Electronic Syst.*, vol. 33, no. 2, pp. 539-544, Apr. 1997.

[23] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antenna and Propagation*, vol. AP-30, no. 1, pp. 27-34, Jan. 1982.

[24] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Cambridge University Press, Third-Edition, 1996.

[25] S. Haykin and A. Steinhardt, *Adaptive Radar Detection and Estimation*, NY: John Wiley and Sons, 1992.

[26] R. A. Horn and C. R. Johnson, *Matrix Analysis*, NY: Cambridge University Press, 1990.

[27] C. L. Ho, G. J. Lin, and T. S. Lee, "An OSIC based reduced-rank MIMO equalizer using conjugate gradient algorithm," *IEICE Trans. Communications*, vol. E86-B, no. 9, pp. 2656-2664, Sept. 2003.

[28] W. S. Hou and B. S. Chen, "ICI cancellation for OFDM communication systems in time-varying multipath fading channels," *IEEE Trans. Wireless Communications*, vol. 4, no. 5, pp. 2100-2110, Sept. 2005.

[29] H. Jafarkhani, *Space-Time Coding: Theory and Practice*, Cambridge University Press, 2005.

[30] W. C. Jakes, *Microwave Mobile Communications*, IEEE Press, 1974.

[31] N. Khaled, G. Leus, C. Desset, and H. De Man, "A robust joint linear precoder and decoder MMSE design for slowly time-varying MIMO channels," *Proc. IEEE ICCASP*, vol. 4, pp. iv.485-iv.488, May 2004.

[32] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*, NY: Cambridge University Press, 2003.

[33] G. Leus and M. Moonen, "Per-tone equalization for MIMO OFDM systems," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2965-2975, Nov. 2003.

[34] A. Leke and J. M. Cioffi, "Impact of imperfect channel knowledge on the performance of multicarrier systems," *Proc. IEEE GLOBECOM 1998*, vol. 2, pp. 951-955, Nov. 1998.

[35] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE*

Trans. Signal Processing, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.

[36] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DOA estimation algorithms: unification, simplification, and observations," *IEEE Trans. Aerospace and Electronic Systems*, vol. 29, no. 4, pp. 1170-1184, Oct. 1993.

[37] Y. P. Lin, L. H. Liang, P. J. Chung, and S. M. Phoong, "An eigen-based TEQ design for VDSL systems," *IEEE Trans. Signal Processing*, vol. 55., no. 1, pp. 290-298, Jan. 2007.

[38] A. P. Liavas, P. A. Regalia, and J. P. Delmas, "Blind channel approximation: effective channel order determination," *IEEE Trans. Signal Processing*, vol. 47, no. 12, pp. 3336-3344, Dec. 1999.

[39] C. Y. Chen and S.-M. Phoong, "Bit rate optimized time-domain equalizers for DMT systems," *Proc. Int. Symp. Circuits Syst.*, 2003.

[40] J. T. Liu, "Performance of multiple space-time coded MIMO in spatially correlated channels," *Proc. IEEE WCNC'03*, vol. 1, pp. 349-353, March 2003.

[41] P. Liu, Z. Tao, Z. Lin, E. Erkip, and S. Panwar, "Cooperative wireless communications: a cross-layer approach," *IEEE Trans. Wireless Communications*, vol. 13,　no. 4, pp. 84-92, Aug. 2006.

[42] R. G. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Processing*, vol. 53, no. 5, pp. 1684-1696, May 2005.

[43] J. H. Manton, "Optimal training sequences and pilot tones for OFDM systems," *IEEE Communications Letter*, vol. 5, no. 4, pp. 151-153, April 2001.

[44] Y. Mostofi and D. C. Donald, "ICI mitigation for pilot-aided OFDM mobile systems," *IEEE Trans. Wireless Communications*, vol. 4, no. 2, pp. 765-774, March 2005.

[45] H. G. Myung, L. J. Lim, and D. J. Goodman, "Single carrier FDMA for uplink wireless transmission," *IEEE Vehicular Technology Magazine*, vol. 1, no. 3, pp. 30-38, Sept. 2006.

[46] P. Melsa, R. Younce, and C. Rohrs, "Impulse response shortening for discrete multitone transceivers," *IEEE Trans. Commun.*, vol. 44, pp. 1662–1672, Dec. 1996.

[47] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in

multiple-antenna data transmission over fading channels," *IEEE JSAC*, vol. 16, no. 18, pp. 1423-1436, Oct. 1998.

[48] A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Communications Mag*., vol. 42, no. 10, pp. 74-82, Oct. 2004.

[49] R. van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*, Boston/London: Artech House, 1999.

[50] B. O'Hara and A. Petrick, *The IEEE 802.11 Handbook: A Designer's Companion*, NY: IEEE Press, 1999.

[51] A. J. Paulraj, D. Gore, R. U. Nabar, and H. Boelcskei, "An overview of MIMO communications-A key to Gigabit wireless," *Proc. IEEE*, vol. 92, No. 2, pp. 198-218, Feb. 2004.

[52] Y. Rong, S. A. Vorobyov, and A. B. Gershman, "Robust linear receivers for multi-access space-time block coded MIMO systems: a probabilistically constrained approach," *IEEE JSAC*, vol. 24, no. 8, pp. 1560-1570, Aug. 2006.

[53] Y. Rong, S. A. Vorobyov, and A. B. Gershman, "A robust linear receiver for multi-access space-time block-coded MIMO systems based on probability-constrained optimization," *Proc. IEEE VTC*, vol. 1, pp. 118-122, May 2004.

[54] Y. Rong, S. Shahbazpanahi, and A. B. Gershman, "Robust linear receivers for space-time block coded multi-access MIMO systems with imperfect channel state information," *IEEE Trans. Signal Processing*, vol. 53, no. 8, part 2, pp. 3081-3090, Aug. 2005.

[55] L. Rugini, P. Banelli, and G. Leus, "Simple equalization of time-varying channels for OFDM," *IEEE Communications Letters*, vol. 9, no. 7, pp. 619-621, July 2005.

[56] J. B. Schodorf and D. B. Williams, "A constrained optimization approach to multiuser detection," *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp. 258-262, Jan. 1997.

[57] S. Shahbazpanahi and A. B. Gershman, Z. Luo, and K. M. Wong, "Robust adaptive beamforming for general-rank signal models," *IEEE Trans. Signal Processing*, vol. 51, no. 9, pp. 2257–2269, Nov. 2004.

[58] J. R. Schott, *Matrix Analysis for Statistics*, Second Edition, John Wiley & Sons, 2005.

[59] A. Seyedi and G. J. Saulnier, "General ICI self-cancellation scheme for OFDM systems," *IEEE Trans. Vehicular Technology*, vol. 54, no. 1, pp. 198-210, Jan. 2005.

[60] C. Z. W. H. Swetman, J. S. Thompson, B. Mulgrew and P. M. Grant, "A comparison of the MMSE detector and its BLAST versions for MIMO channels," *IEE Seminar on MIMO: Communications Systems from Concept to Implementations*, pp. 19/1-19/6, Dec. 2001.

[61] M. D. Srinath, P. K. Rajasekaran and R. Viswanathan, *Introduction to Statistical Signal Processing with Applications*, NJ: Pentice Hall, 1996.

[62] G. W. Stewart, "Error and perturbation bounds for subspaces associated with certain eigenvalue problems," *SIAM Review*, 15-33, pp. 1413-1416, 1973.

[63] G. L. Stuber, J. R. Barry, S. W. Mclaughlin, Y. Li, M. A. Ingram, and T. G. Pratt, "Broadband MIMO-OFDM wireless communications," *Proc. IEEE,* vol. 92, no.2, pp. 271-294, Feb. 2004.

[64] S. Shahbazpanahi and A. B. Gershman, Z. Luo, and K. M. Wong, "Robust adaptive beamforming for general-rank signal models," *IEEE Trans. Signal Processing*, vol. 51, no. 9, pp. 2257–2269, Nov. 2004.

[65] T. Starr, J. Cioffi, and P. Silvermann, *Understanding Digital Subscriber Line Technology*, Englewood Cliffs, NJ: Prentice-Hall, 1999.

[66] A. Salvekar, S. Sandhu, Q. Li, M. A. Vuong, and X. Qian, "Multi-antenna technology in WiMAX systems," *Intel Technology Journal*, vol. 8, no. 3, pp. 229-240, Aug. 2004.

[67] P. Schniter and H. Liu, "Iterative frequency-domain equalization for single-carrier systems in doubly-dispersive channels," *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 667 - 671, Nov. 2004.

[68] P. Schniter, "Low-complexity equalization of OFDM in doubly selective channels," *IEEE Trans. Signal Processing*, vol. 52, no. 4, pp. 1002-1011, April 2004.

[69] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Communications*, vol. 45, pp. 1613–1621, Dec. 1997.

[70] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal

designs," *IEEE Trans. Information Theory*, vol. 45, no. 5, pp. 1456-1467, July 1999.

[71] J. Terry and J. Heiskala, *OFDM Wireless LANs: A Theoretical and Practical Guide*, Indiana: SAMS, 2001.

[72] Z. Tian, K. L. Bell, and H. L. Van Trees, "Robust constrained linear receivers for CDMA wireless systems," *IEEE Trans. Signal Processing*, vol. 49, no. 7, pp. 1510–1522, July 2001.

[73] X. N. Tran and T. Fujino, "Groupwise successive ICI cancellation for OFDM systems in time-varying channels," *Proc. IEEE International Symposium on Signal Processing and Information Theory*, pp. 489-494, Dec. 2005.

[74] B. D. Van Veen and K. M. Buckley, "Beamforming: a versatile approach to spatial filtering", *IEEE ASSP Mag.*, vol. 5, pp. 4-24, Apr. 1988.

[75] S. A. Vorobyov, A. Gershman, and Z. Luo, "Robust adaptive beamforming using worst-case performance optimization: a solution to the signal mismatch problem," *IEEE Trans. Signal Processing*, vol. 51, no. 2, pp. 313–324, Feb. 2003.

[76] K. Vanbleu, G. Ysebaert, G. Cuypers, M. Moonen, and K. Van Acker, "Bitrate maximizing time-domain equalizer design for DMT-based systems," *IEEE Trans. Commun.*, vol. 52, pp. 871–876, Jun. 2004.

[77] K. Vanbleu, G. Ysebaert, G. Cuypers, and M. Moonen, "Adaptive bit rate maximizing time-domain equalizer design for DMT-based systems," *IEEE Trans. Signal Processing*, vol. 54, no. 2, pp. 483-498, Feb. 2006.

[78] H. L. Van Trees, *Optimum Array Processing*, NY: John Wiley & Sons, 2002.

[79] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications: where Fourier meets Shannon," *IEEE Signal Processing Magazine*, vol. 17, pp. 29-48, May 2000.

[80] Z. Wang, X. Ma, and G. B. Giannakis, "OFDM or single-carrier block transmissions?," *IEEE Trans. Communications*, vol. 52, no. 3, pp. 380-394, March 2004.

[81] F. Wang, A. Ghosh, R. Love, K. Stewart, R. Ratasuk, R. Bachu, and Y. Sun, and Q. Zhao, "IEEE 802.16e system performance: analysis and simulations," *Proc. IEEE PIMRC 2005*, vol. 2, pp. 900-904, Sept. 2005.

[82] P. W. Wolniansky, G. J. Foschini, G. D. Golden, R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over rich scattering wireless channels," *Proc. IEEE ISSSE-98*, Italy, pp. 295-300, Sept., 1998.

[83] Z. Xu, "Perturbation analysis for subspace decomposition with applications in subspace-based algorithms," *IEEE Trans. Signal Processing*, vol. 50, no. 11, pp. 2820-2830, Nov. 2002.

[84] K. Zarifi, S. Shahbazpanahi, A. B. Gershman, and Z. Luo, "Robust blind multiuser detection based on the worse-case performance optimization o the MMSE receiver," *IEEE Trans. Signal Processing*, vol. 53, no. 1, pp. 295-305, Jan. 2005.

[85] X. Zhu and R. D. Murch, "Layered space-frequency equalization in a single-carrier MIMO system for frequency-selective channels," *IEEE Trans. Wireless Communications*, vol. 3, no. 3, pp. 701-708, May 2004.

[86] Y. Zhao and S. G. Haggman, "Intercarrier interference self-cancellation scheme for OFDM mobile communication systems," *IEEE Trans. Communications*, vol. 49, no. 7, pp. 1185-1191, July 2001.

簡　　歷

姓　名　：　林志遠

性　別　：　男

出生年月日　：　民國 68 年 8 月 1 日

籍　貫　：　宜蘭縣五結鄉中興路三段 51 巷 1 號

學　歷　：　國立中興大學電機工程系畢業 (86/9～90/6)

　　　　　　國立交通大學電信研究所博士班（90/9～96/5)

論文題目：於通道偏差下單載波及多載波區塊式傳輸系統之強健

　　　　　式接收機設計

Publication List

Journal Paper

[J1] C. Y. Lin, J . Y. Wu, T. S. Lee, "Robust Constrained-Optimization Based Linear Receiver for High-Rate MIMO OFDM Against Channel Parameter Mismatch," *IEEE Trans. on Signal Processing*, vol. 55, no. 6, pp. 2628-2645, June 2007.

[J2] C. Y. Lin, J. Y. Wu and T. S. Lee, "A Near-Optimal Low-Complexity Transceiver for CP-Free Multi-Antenna OFDM Systems," *IEICE Trans. Communications*, vol. E89-B, no. 1, pp. 88-99, Jan. 2006.

[J3] C. Y. Lin, J. Y. Wu, and T. S. Lee, "Robust Receiver Design for MIMO Single-Carrier Block Transmission over Time-Varying Dispersive Channels Against Imperfect Channel Knowledge," submitted to *IEEE Trans. Wireless Communications*.

Conference Paper

[C1] C. Y. Lin, J. Y. Wu, and T. S. Lee, " Robust Receiver Design for MIMO Single-Carrier Block Transmission over Time-Varying Dispersive Channels Against Imperfect Channel Knowledge," submitted to *IEEE GLOBECOM 2007*.

[C2] C. Y. Lin, J. Y. Wu, and T. S. Lee, "Robust Linear Receiver for High-Rate MIMO OFDM under Channel Parameter Mismatch," *Proc. IEEE GLOBECOM 2006*, San Francisco, USA, Nov. 2006.

[C3] C. H. Hung, C. Y. Lin and T. S. Lee, "Intercarrier Interference Compensation in Multiband OFDM UWB Systems," *Proc. ISCOM 2005*, Kaohsiung, Taiwan, Dec. 2005.

[C4] C. Y. Lin and T. S. Lee, "An Efficient Interference Cancellation Scheme for CP-free MIMO-OFDM Systems," *Proc. IEEE VTC2005-fall*, Dallas, USA, vol. 1, pp. 439-443, Sept. 2005.

[C5] C. Y. Lin, J. Y. Wu and T. S. Lee, "GSC-Based Frequency-Domain Equalizer for CP-Free OFDM Systems," *Proc. IEEE ICC 2005*, Seoul, Korea, pp. 1132-1136, May 2005.

[C6] J. H. Deng, C. Y. Lin and T. S. Lee, "A Low Complexity Group-Wise PIC Based Mixed-Mode MIMO-OFDM System," *Proc. IEEE MWC 2004*, Shanghai, China, pp. 481-484, May 2004.