國 立 交 通 大 學

電信工程學系

博 士 論 文

適應性決策迴授廣義旁波帶消除器：
效能分析與應用

Adaptive Decision Feedback
Generalized Sidelobe Canceller:
Performance Analysis and Applications

研究生：李　彥　文

指導教授：吳　文　榕　博士

中　華　民　國　九十五　年　四　月

適應性決策迴授廣義旁波帶消除器：效能分析與應用

Adaptive Decision Feedback Generalized Sidelobe Canceller: Performance Analysis and Applications

研究生：李彥文　　Student: Yinman Lee

指導教授：吳文榕博士　　Advisor: Dr. Wen-Rong Wu

國 立 交 通 大 學

電信工程學系

博 士 論 文

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

April 2006
Hsinchu, Taiwan, Republic of China

中　華　民　國　九十五　年　四　月

適應性決策迴授廣義旁波帶消除器：效能分析與應用

研究生：李彥文　　指導教授：吳文榕博士

國　立　交　通　大　學
電　信　工　程　學　系

摘　要

適應性廣義旁波帶消除器（GSC）被廣泛應用在波束形成之抗干擾技術上，由於內在構造的特性，其收斂頗為緩慢，同時，亦對於模型錯配頗為敏感，當發生模型錯配時，會造成訊號抵消（signal cancellation）而令效能嚴重下降，這些問題影響了適應性廣義旁波帶消除器於時變系統之應用，並使其實現更為複雜。在本論文中，我們提出一新技術能有效地對付這些問題。主要的構想是於廣義旁波帶消除器中加入一單權重決策主導等化器（decision-directed equalizer）和一單權重迴授濾波器（feedback filter），而成為決策迴授廣義旁波帶消除器（DFGSC）。當中我們選用最小平方法（LMS）以適應性地實現此方案，並對其收斂特性作全面之分析，這樣證明了其收斂速度會大大地得到改善且訊號抵消現象能避免發生。

於無線通訊中，同頻干擾（CCI）和符元干擾（ISI）為影響系統效能之兩大因素。一般可以使用波束形成器（beamformer）以消除同頻干擾和等化器以消除符元干擾，兩者結合可被稱為時空等化器（STE）。這種裝置通常需要利用訓練序列（training sequence）對其加以訓練後才可以被使用，但在一些特定應用上，接收用戶的空間資訊是可以得知，反而訓練序列不一定存在。在此我們對決策迴授廣義旁波帶消除器進行推廣，提出一適應性決策迴授時空等化器以對付這種情況。我們的方案包含了一適應性決策迴授廣義旁波帶消除器、一盲目決策迴授等化器（blind DFE）和一通道估測器。由於其特殊之

構造，此盲目決策迴授等化器可藉由通道估測之輔助，從而有效地對抗決策迴授等化器所造成的錯誤蔓延（error propagation）現象。當中我們也是使用最小平方法適應性地實現此方案，同時對其收斂特性作出分析。實驗模擬證明了我們所提出之適應性決策迴授時空等化器即使處於苛刻的通道環境中亦能有效地對抗同頻干擾和符元干擾。另外，此通道輔助盲目決策迴授等化器之架構也可以應用於適應性之最小均方誤差（MMSE）決策迴授等化器中而成為適應性通道輔助決策迴授等化器（ACA-DFE）。我們亦會證明此通道輔助決策迴授等化器亦適用於多輸入多輸出（MIMO）系統而使其效能優於一般適應性多輸入多輸出之決策迴授等化器。

適應性之平行式干擾消除（PIC）技術近年亦被用於多輸入多輸出系統之訊號檢測。傳統之平行式干擾消除技術使用最小均方誤差之法則來進行優化，但最小均方誤差並不能達致最小位元錯誤率（MBER）之效果。另外，當運作於時變環境中，其效能會深受錯誤蔓延之影響。最後，於本論文中我們提出一基於最小變異（MV）法則之適應性二階段（two-stage）平行式干擾消除檢測法以解決這些問題。其中我們使用決策迴授廣義旁波帶消除器以適應性地實現此演算法。在第一階段之消除時，我們提出使用一雙重的決策迴授廣義旁波帶消除器（dual-DFGSC）架構使其更適用於時變之通道環境。由於第一階段之良好效能，第二階段只需使用一簡單之匹配濾波器便可達接近最優之效能。我們亦選用最小平方法以適應性地實現此方案並分析其收斂特性。實驗模擬證明此適應性二階段平行式干擾消除檢測法之效能大大地優於傳統基於最小均方誤差之平行式干擾消除檢測法。

Adaptive Decision Feedback Generalized Sidelobe Canceller: Performance Analysis and Applications

Student: Yinman Lee Advisor: Dr. Wen-Rong Wu

Department of Communication Engineering
National Chiao Tung University

Abstract

The adaptive generalized sidelobe canceller (GSC) is a commonly used device for interference cancellation in array beamforming. However, due to its inherent structure, its convergence is slow. Also, it is sensitive to model mismatch. When model mismatch exists, a phenomenon called signal cancellation will occur, and the performance of the adaptive GSC can be seriously affected. These problems limit the use of adaptive GSC in time-variant systems and also complicate real-world implementations. In this dissertation, we propose a new scheme that can effectively solve these problems. The main idea is to introduce a single-tap decision-directed equalizer and a single-tap feedback filter in the GSC structure, resulting in a decision feedback GSC (DFGSC). The least-mean-square (LMS) algorithm is used for adaptation and the convergence behavior of the adaptive DFGSC is fully analyzed. It is shown that the convergence rate can be greatly enhanced and signal cancellation can be completely avoided.

In wireless communications, co-channel interference (CCI) and inter-symbol interference (ISI) are two main factors limiting the system performance. Conventionally, a beamformer is used to reduce CCI while an equalizer is used to compensate ISI. These two devices can be combined into one named space-time equalizer (STE). A training sequence is usually required to train the STE prior to its use. In some applications, however, spatial information corresponding to a desired user is available, but the training sequence is not. Extending the DFGSC approach, we

then propose an adaptive decision feedback STE to cope with this problem. Our scheme consists of an adaptive DFGSC, a blind decision feedback equalizer (DFE), and a channel estimator. With a specially designed structure, the proposed blind DFE, aided by the estimated channel, can better resist error propagation effect inherent in a DFE. As previously, adaptation operations are implemented with the LMS algorithm and convergence analysis is given as well. Simulations show that the proposed adaptive decision feedback STE is effective in mitigating both CCI and ISI even in severe channel environments. The channel-aided blind DFE structure can be further applied to the general adaptive minimum mean-squared-error DFE (MMSE-DFE), called the adaptive channel-aided DFE (ACA-DFE). We also demonstrate that the ACA-DFE can be extended to multiple-input multiple-out (MIMO) systems and it can outperform the conventional adaptive MIMO MMSE-DFE.

Adaptive parallel interference cancellation (PIC) has been recently introduced in the signal detection of MIMO systems. Conventional PIC uses the MMSE criterion for parameter adaptation. However, it is known that the MMSE criterion cannot achieve the minimum bit-error-rate (MBER). Also, it suffers from the error propagation problem when operated in time-variant channels. In the last part of the dissertation, an adaptive two-stage PIC detection scheme with the minimum variance (MV) criterion is proposed to solve the problems. Adaptation with the MV criterion is then realized with the DFGSC. In the first-stage cancellation, a dual-DFGSC configuration being effective in time-variant channel environments is developed. Due to the good performance of the first-stage processing, only matched filtering is required in the second stage to achieve near optimum results. The LMS algorithm is employed and its convergence behavior is also examined. Simulation results show that the proposed two-stage DFGSC-PIC detection significantly outperforms the conventional MMSE-PIC detection.

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisor, Prof. Wen-Rong Wu, for his help and support throughout my research. In the period of my study, he gave me encouragement and guided me to the right research direction. I wish there were enough words for me to thank him.

I have been blessed with friendships that have been precious sources of support, enjoyment and encouragement especially in times of difficulty. I would like to express my deepest gratitude to all my teachers, friends, lab-mates and office-mates I met these years.

Finally, I am in-debted to my family and my girl friend, Miss Ines Sou, for their mental support. Their unlimited love is the best encouragement in my life.

Contents

Chinese Abstract i

English Abstract iii

Acknowledgements v

Contents vi

List of Tables ix

List of Figures x

Abbreviations xiii

Mathematical Notations xv

1 Introduction 1

1.1 Basics of Beamforming 1

1.2 Basics of Space-Time Equalization 2

1.3 Basics of Spatial Multiplexing in Multiple-Input Multiple-Output Systems 3

1.4 Outline of Dissertation 4

2 Adaptive Decision Feedback Generalized Sidelobe Cancellation 9

2.1 Introduction 9

2.2 Background 11

2.2.1 Signal Model 11

2.2.2 Conventional Generalized Sidelobe Canceller 12

2.3 Decision Feedback Generalized Sidelobe Canceller 15
2.4 DOA Mismatch Analysis 19
2.4.1 Mismatch Signal Model 19
2.4.2 Mismatch Analysis with GSC and DFGSC 20
2.4.3 A Start-Up Approach for DFGSC with Mismatch 22
2.5 Convergence Analysis 24
2.5.1 MSE in Steady State 24
2.5.2 SINR in Steady State 25
2.6 Simulations 26
2.6.1 Exactly Known Desired Signal's DOA 27
2.6.2 Desired Signal's DOA Mismatch 28
2.7 Summary 29

3 Adaptive Decision Feedback Space-Time Equalization with Generalized Sidelobe Cancellation 36

3.1 Introduction 36
3.2 Space-Time Signal Model 38
3.3 Hybrid of GSC and LBDFE 40
3.3.1 GSC 40
3.3.2 LBDFE 41
3.4 Proposed Hybrid STE 43
3.4.1 DFGSC 43
3.4.2 CBDFE 47
3.5 Spatial Multipath and Spatial Signature Mismatch 50
3.6 Simulations 52
3.6.1 Channels with Temporal ISI Only 52
3.6.2 Channels with Temporal and Spatial ISI 54
3.7 Summary 54

4 Adaptive Channel-Aided Decision Feedback Equalization for SISO and MIMO Systems 62

4.1 Introduction . . . 62
4.2 Background . . . 64
4.2.1 Conventional DFE for SISO Systems . . . 64
4.2.2 Conventional DFE for MIMO Systems . . . 66
4.3 Proposed Adaptive Channel-Aided DFE . . . 68
4.3.1 ACA-DFE for SISO Systems . . . 68
4.3.2 ACA-DFE for MIMO Systems . . . 70
4.4 Simulations . . . 71
4.4.1 SISO channel environments . . . 71
4.4.2 MIMO channel environments . . . 73
4.5 Summary . . . 73

5 Adaptive Two-Stage GSC-Based PIC Detection for Time-Varying MIMO Channels 78

5.1 Introduction . . . 78
5.2 MIMO Signal Model and Parallel Interference Cancellation . . . 79
5.3 DFGSC-PIC and Its Adaptive Realization . . . 82
5.4 Convergence Analysis . . . 88
5.5 Simulations . . . 92
5.6 Summary . . . 94

6 Conclusions and Future Work 101

6.1 Conclusions . . . 101
6.2 Future Work . . . 102

References 104

List of Tables

3.1 Parameters used for simulations in Section 3.6.1: (a) Filter length, and (b) Step sizes 61

3.2 Parameters used for simulations in Section 3.6.2: (a) Step sizes, and (b) Channel settings 61

List of Figures

1.1 Block diagram of narrowband beamformer. 7
1.2 Block diagram of STE with tap delay line for each antenna element. 7
1.3 Block diagram of STE with one spatial filter and one temporal filter. . . . 8
1.4 Block diagram of MIMO system. 8

2.1 Block diagram of adaptive DFGSC. 31
2.2 Learning curves for GSC and DFGSC with same step size. 31
2.3 Beam patterns of GSC and DFGSC in Fig. 2.2 after 200 snapshots. 32
2.4 Learning curves for GSC and DFGSC with same SINR target. 32
2.5 Steady-state SINR performance in different SNR environments. 33
2.6 Steady-state SINR performance with different step sizes. 33
2.7 Learning curves for GSC with DOA mismatch utilizing point (pt) constraint only and point and first-order derivative (pt & 1d) constraints. 34
2.8 Learning curves for DFGSC with DOA mismatch utilizing point (pt) constraint only and point and first-order derivative (pt & 1d) constraints. . . . 34
2.9 Learning curves for DFGSC with transition from point and first-order derivative (pt & 1d) constraints to point (pt) constraint only. 35
2.10 Steady-state SINR performance for DFGSC and other robust adaptive beamforming approaches in different SNR environments. 35

3.1 Hybrid of GSC and LBDFE (with LBDFE in start-up period). 56
3.2 Proposed hybrid STE. 56
3.3 Learning curves of GSC output SINR for different schemes in suppressing CCI. 57

3.4 Beam patterns (enlarged region around CCI's DOA) of different schemes in Fig. 3.3 after 5000 iterations. 57
3.5 Learning curves of equalizer output MSE for different schemes. 58
3.6 Learning curves of GSC output SINR for different schemes with 5° DOA mismatch. 58
3.7 Beam patterns of different schemes in Fig. 3.6 after 5000 iterations with 5° DOA mismatch. 59
3.8 Learning curves of equalizer output MSE for different schemes with 5° DOA mismatch. 59
3.9 Learning curves of equalizer output MSE for different schemes with temporal and spatial ISI. 60
3.10 Learning curves of equalizer output MSE for different schemes with temporal and spatial ISI (artificial errors added). 60

4.1 ACA-DFE in decision-directed mode for SISO systems. 74
4.2 ACA-DFE in decision-directed mode for MIMO systems. 74
4.3 MSE learning curves for static Proakis C channel. 75
4.4 SER for static Proakis C channel with different step sizes. 75
4.5 SER for time-varying Proakis C channel ($f_{\mathrm{d}}T_{\mathrm{s}} = 5 \times 10^{-4}$) with different step sizes. 76
4.6 MSE learning curves for static MIMO channel. 76
4.7 SER for static MIMO channel with different step sizes. 77
4.8 SER for time-varying MIMO channel ($f_{\mathrm{d}}T_{\mathrm{s}} = 2 \times 10^{-4}$) with different step sizes. 77

5.1 General two-stage PIC detection scheme for MIMO systems. 96
5.2 Adaptive DFGSC for realization of weight vector in the first stage of PIC. 96
5.3 Adaptive DFGSC with steering matrix $\mathbf{P}(k)$ and channel estimator $\mathbf{q}(k)$. Dual-DFGSC works complementarily for each branch in the first stage of DFGSC-PIC. 97
5.4 Update and decision time relation for dual-DFGSC. 97

5.5 Interference cancellation MSE with different μ_a. 98
5.6 Learning curves of PIC output SINR with $f_d T_s = 1 \times 10^{-4}$ and SNR = 30 dB. 98
5.7 Learning curves of PIC output SINR with $f_d T_s = 5 \times 10^{-4}$ and SNR = 30 dB. 99
5.8 SER for various SNR values with $f_d T_s = 1 \times 10^{-4}$. 99
5.9 SER for various SNR values with $f_d T_s = 5 \times 10^{-4}$. 100
5.10 SER for various $f_d T_s$ values with SNR = 30 dB. 100

Abbreviations

- ACA-DFE: adaptive channel-aided decision feedback equalizer/equalization
- AWGN: additive white Gaussian noise
- BER: bit-error-rate
- CBDFE: channel-aided blind decision feedback equalizer/equalization
- CCI: co-channel interference
- CCINR: co-channel interference-to-noise ratio
- CDMA: code division multiple access
- CMA: constant modulus algorithm
- CSI: channel state information
- DFE: decision feedback equalizer/equalization
- DFGSC: decision feedback generalized sidelobe canceller/cancellation
- DOA: direction of arrival
- DOF: degree of freedom
- FIR: finite impulse response
- GSC: generalized sidelobe canceller/cancellation
- IIR: infinite impulse response
- INR: interference-to-noise ratio
- ISI: inter-symbol interference
- LBDFE: blind decision feedback equalizer/equalization proposed by Labat *et al.*
- LCMP: linearly constrained minimum power
- LCMV: linearly constrained minimum variance
- LMS: least-mean-square
- LSMI: loaded sample matrix inversion
- MAI: multiple access interference
- MBER: minimum bit-error-rate
- MIMO: multiple-input multiple-output
- MLSE: maximum likelihood sequence estimator
- MMA: multimodulus algorithm

- MMSE: minimum mean-squared-error
- MSE: mean-squared-error
- MSER: minimum symbol-error-rate
- MV: minimum variance
- NLMS: normalized least-mean-square
- OFDM: orthogonal frequency division multiplexing
- PIC: parallel interference cancellation
- QAM: quadrature amplitude modulation
- QPSK: quadrature phase-shift keying
- RLS: recursive least-squares
- SER: symbol-error-rate
- SDMA: space division multiple access
- SGA: stochastic gradient algorithm
- SIC: successive interference cancellation
- SINR: signal-to-interference plus noise ratio
- SISO: single-input single-output
- SNR: signal-to-noise ratio
- SOCP: second-order cone programming
- STE: space-time equalizer/equalization
- ULA: uniform linear array
- V-BLAST: vertical Bell Laboratories layered space-time
- WCPO: worst-case performance optimization
- ZF: zero-forcing

Mathematical Notations

- $(\cdot)^*$: complex conjugate
- $(\cdot)^T$: transpose
- $(\cdot)^H$: Hermitian transposition
- $(\cdot)_{\mathrm{r}}$: real part of the complex number
- $(\cdot)_{\mathrm{i}}$: imaginary part of the complex number
- $\min\{\cdot\}$: minimization
- $\mathrm{Dec}\{\cdot\}$: decision
- $\mathrm{E}\{\cdot\}$: expectation
- $\|\cdot\|$: two-norm
- $\mathrm{Diag}\{\cdot\}$: diagonal matrix whose main diagonal is the included vector
- $\mathrm{Tr}\{\cdot\}$: matrix trace
- $\otimes$: convolution
- $\mathrm{post}\{\cdot\}$: postcursor-taking operation

Chapter 1

Introduction

1.1 Basics of Beamforming

Recently, beamforming techniques have received great interests in wireless communications and they can achieve performance and capacity enhancement without the need for additional power or spectrum allocation [1]-[3]. A beamformer is used with an array of antennas to realize the idea of spatial filtering. This spatial processing leads to more degrees of freedom (DOFs) in the system design, which can help improving the overall performance of the system. The objective of beamforming is to estimate the signal arriving from a desired direction in the presence of interference and noise. A block diagram of a typical narrowband beamformer is shown in Fig. 1.1. The term adaptive beamformer is used when the weights are adjusted in a dynamic fashion, as required by a performance index of the system [4]. The adaptation process is normally under the control of the system. An adaptive beamformer may be used in a variety of ways to improve the performance of a wireless communication system. Perhaps most important is its capability to cancel interference. It works on the premise that the desired signal and unwanted interference arrive from different directions. The beam pattern is adjusted to suit the desired signal and to reject the interference from different antennas with appropriate spatial filtering. The scheme needs to differentiate the desired signal from the interference and this normally requires either the knowledge of a training signal, or the direction of arrival (DOA) of the desired signal and/or the interference to achieve this aim.

The linearly constrained minimum variance (LCMV) considered by Frost [5] is one of the commonly used criteria for interference and noise suppression provided that the DOA

information of the desired signal is known *a priori*. The generalized sidelobe canceller (GSC) proposed in [6] is often used for the realization of the LCMV beamformer. However, the conventional GSC is known to be quite sensitive even to a slight mismatch of the desired signal's DOA, which can easily occur in practice as a consequence of signal pointing errors. When a mismatch exists, the GSC tends to misinterpret the desired signal component in input as interference and to suppress this component instead of maintaining distortionless response towards it. This phenomenon is called signal cancellation and may cause severe degradation of the beamforming performance [7]. Also, due to the special structure of the GSC, when an adaptive algorithm like the least-mean-square (LMS) is applied for weight adaptation, the error signal, used in the LMS algorithm, always contains the desired signal (even in steady state). This *large* error signal magnifies the stochastic gradient in the LMS algorithm resulting a large mean-squared-error (MSE). In order to reduce the MSE, the step size, a parameter controlling the LMS convergence, must be small and it essentially makes the LMS converge slowly. Seeking robust and efficient beamforming structure is still an on-going research topic even nowadays.

1.2 Basics of Space-Time Equalization

In wireless communications, the co-channel interference (CCI) due to multiple access and the inter-symbol interference (ISI) due to multipath channel often cause severe signal distortion and limit the system performance [8], [9]. Lately, there has been a growing interest in applying adaptive antenna arrays and space-time signal processing techniques to solve these problems [3]. The common approach is to use a beamformer for CCI reduction and an equalizer for ISI compensation. These two devices can be further combined into one named space-time equalizer (STE) [10], [11]. The application of STE is beneficial to communication quality and system capacity, subsequently improving the overall performance even in severe channel environments.

The optimum STE is known to be a maximum likelihood sequence estimator (MLSE) operated in space-time domain [11]. However, the MLSE is notorious for high computational complexity. A suboptimum approach with a hybrid of a linear filter and a Viterbi equalizer was proposed in [12]. Even so, the implementation complexity is still high and

it limits the MLSE-like structure in real-world applications. The STE performing both beamforming and equalization as mentioned previously is what researchers consider most. The general structure of this kind of STE consists of an antenna array and a temporal filter bank [13], [14], as depicted in Fig. 1.2. Either a linear equalizer or a decision feedback equalizer (DFE) can be applied to the structure. Although the performance of this structure is satisfactory, its computational complexity is quite high. The other problem is that when an adaptive algorithm is applied, the convergence is slow especially operating under a large number of antenna elements and a severe fading channel. To ease these problems, another structure being a hybrid (or cascade) of a spatial filter and a temporal filter was proposed for the STE [15], [16], as illustrated in Fig. 1.3. It requires lower computational complexity and the convergence is faster. However, the space-time information of the received signal will not be fully exploited and there will be some performance loss.

A training sequence is usually required to train the STE prior to its use. In some applications, however, spatial information corresponding to a desired user is available, but the training sequence is not. A typical example is the beamforming approach in space division multiple access (SDMA) systems. This *a priori* spatial information may be utilized to avoid the need of the training sequence and enhance the performance of the STE.

1.3 Basics of Spatial Multiplexing in Multiple-Input Multiple-Output Systems

In recent years, much attention is paid in the development of multiple-input multiple-out (MIMO) systems. With the use of multiple antennas at both transmitter and receiver, the spectral efficiency of a communication system can be increased dramatically [17]. Fig. 1.4 shows a typical configuration of a MIMO system. MIMO signaling can improve wireless communications in two different ways: diversity methods and spatial multiplexing. Diversity methods improve the robustness of the communication system by suitably combining the multiple paths between transmit and receive antennas. Spatial multiplexing, on the other hand, emerged from the fact that in a rich scattering environment it is possible for the receiver to detect different signals that are transmitted simultaneously from multiple

antennas. Our focus here is on the later case. A successive interference cancellation (SIC) approach, known as the vertical Bell Laboratories layered space-time (V-BLAST) system, is commonly used for spatial multiplexing to achieve a substantial portion of the Shannon capacity for MIMO channels [18]. However, it requires high computational complexity, and the ordering operation inherent in the SIC structure often increases the processing delay and restricts the use of adaptive realization. In general, the V-BLAST system can only be blockwisely updated in time-varying channel environments [19]. Even with the blockwise update, re-calculation of tap weights and re-ordering often make real-world implementations difficult. Recently, an alternative adaptive SIC-based detection scheme was proposed in [20]. Nevertheless, in order to maintain good detection performance for time-varying MIMO channels, it requires many efforts and this makes the whole process rather complex. All these make the V-BLAST like approach a bit difficult to realize in rapid changing environments.

Lately, parallel interference cancellation (PIC) detection schemes were proposed for signal detection in MIMO systems [21]-[25]. The principle of PIC has its origin in code division multiple access (CDMA) multiuser detection [26], [27]. It has the advantages of low computational complexity and low processing delay. Contrast to SIC, PIC detects different data symbols from different transmit antennas in parallel and it is generally implemented with a multistage structure. Since PIC does not require the ordering operation, it is more adequate for adaptive implementations. The conventional PIC can provide satisfactory performance in time-invariant channel environments. However, its performance can be significantly degraded in time-variant channel environments. This is due to the error propagation effect inherent in the multistage PIC scheme.

1.4 Outline of Dissertation

As mentioned, the convergence of the adaptive GSC is slow. Also, it is sensitive to model mismatch. Due to these problems, application of the adaptive GSC in time-variant systems and real-world implementations may be difficult. In this dissertation, we propose a new approach that can effectively solve these problems. The main idea is to introduce a decision feedback operation such that the desired signal of the GSC is cancelled from

the error signal. With the operation, the adaptation of the GSC can be accelerated and signal cancellation due to model mismatch can be avoided. We call the new scheme the adaptive decision feedback GSC (DFGSC). The DFGSC is then extended to the STE and the MIMO PIC detection problems. This dissertation contains five chapters in addition to this introductory chapter.

In Chapter 2, we develop the DFGSC overcoming the aforementioned problems of insufficient robustness and slow convergence. We introduce a single-tap decision-directed equalizer and a single-tap feedback filter in the GSC. We will show that robustness against DOA mismatch is improved. The LMS algorithm, well-known for its simplicity and effectiveness, is applied for weight adaptation. The convergence behavior of the LMS-based DFGSC is fully analyzed and the analytic signal-to-interference plus noise ratio (SINR) is derived. Simulation results demonstrate that while the proposed structure can considerably enhance the overall performance, it has greatly improved robustness as compared to other existing robust adaptive beamformers.

In Chapter 3, we consider the space-time communication scenario in which the spatial information corresponding to a desired user is available, but the training sequence is not. The desired signal is corrupted by both CCI and ISI from different space and time. We propose an adaptive decision feedback STE to cope with this problem. Our scheme consists of an adaptive DFGSC, a blind DFE, and a channel estimator. The proposed decision feedback STE is not only superior to the conventional GSC, but also robust to multipath channel propagation and spatial signature error. Theoretical results are also derived for optimum solutions, convergence behavior, and robustness properties. With the special channel-aided architecture, the proposed blind DFE can reduce the error propagation effect and be more stable than the conventional blind DFE. Simulation results show that the proposed STE is effective in mitigating both CCI and ISI even in severe channel environments.

Chapter 4 focuses on the extension of the channel-aided structure developed in Chapter 3 to the general LMS-based minimum mean-squared-error DFE (MMSE-DFE) problem. The main objective is to improve the robustness against error propagation. To do that, a specifically designed channel estimator is introduced in the decision-directed

mode. Unlike the conventional DFE, the proposed adaptive channel-aided DFE (ACA-DFE) only adapts the feedforward filter with the LMS algorithm. The feedback filter, however, is obtained from the postcursors of the estimated channel convolved with the feedforward filter. As a result, the proposed ACA-DFE can reduce the error propagation effect and perform better than the conventional adaptive DFE, especially in time-variant channels. We also demonstrate that the ACA-DFE can be extended to MIMO systems improving the performance of the conventional adaptive MIMO DFE.

In Chapter 5, an adaptive two-stage PIC detection scheme with the minimum variance (MV) criterion is proposed to enhance the performance of signal detection in time-variant MIMO systems. Adaptation with the MV criterion is realized with the DFGSC. In the first stage cancellation, a dual-DFGSC configuration being effective in time-variant channel environments is developed. Due to the good performance of the first-stage processing, only matched filtering is required in the second stage to achieve near optimum results. All adaptation operations are implemented with the LMS algorithm. Optimum solutions are provided and convergence behavior is also analyzed. With the proposed structure, the estimated channel information can be utilized to improve the detection process and the probability of lost track in changing environments is substantially reduced.

Finally, conclusions and future work are given in Chapter 6. Throughout the dissertation, we utilize the superscripts/subscripts $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, $(\cdot)_{\mathrm{r}}$ and $(\cdot)_{\mathrm{i}}$ to denote conjugation, transposition, Hermitian transposition, real part and imaginary part of the complex number, respectively. Also, the operators $\min\{\cdot\}$, $\mathrm{Dec}\{\cdot\}$, $\mathrm{E}\{\cdot\}$, $\|\cdot\|$, $\mathrm{Diag}\{\cdot\}$, $\mathrm{Tr}\{\cdot\}$, $\otimes$, and $\mathrm{post}\{\cdot\}$ represent minimization, decision, expectation, two-norm, diagonal matrix constructed by the included vector, matrix trace, convolution, and postcursor-taking operation, respectively.

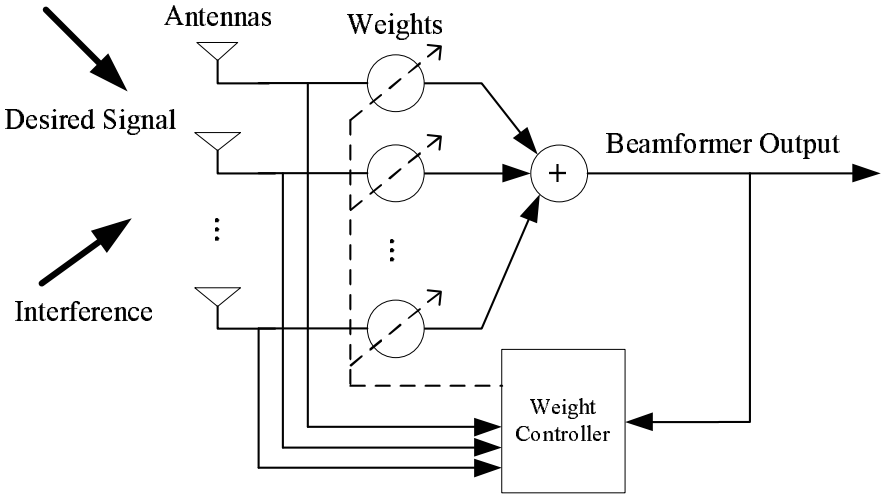

Figure 1.1: Block diagram of narrowband beamformer.

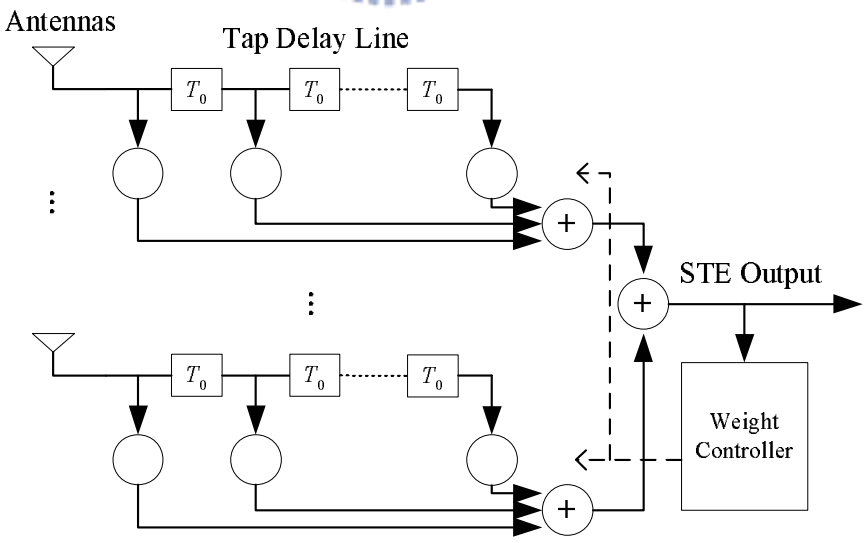

Figure 1.2: Block diagram of STE with tap delay line for each antenna element.

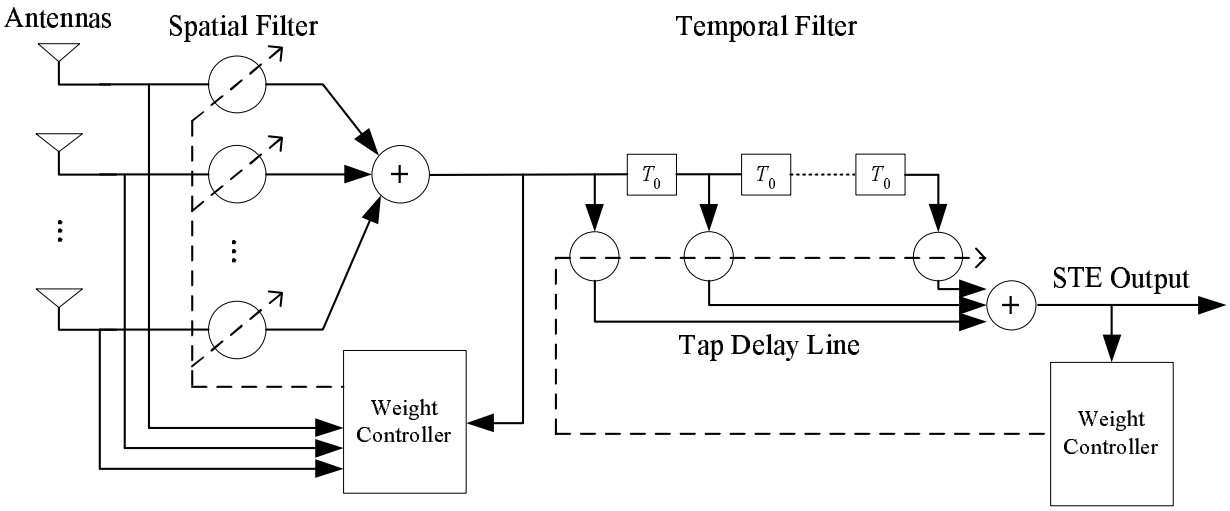

Figure 1.3: Block diagram of STE with one spatial filter and one temporal filter.

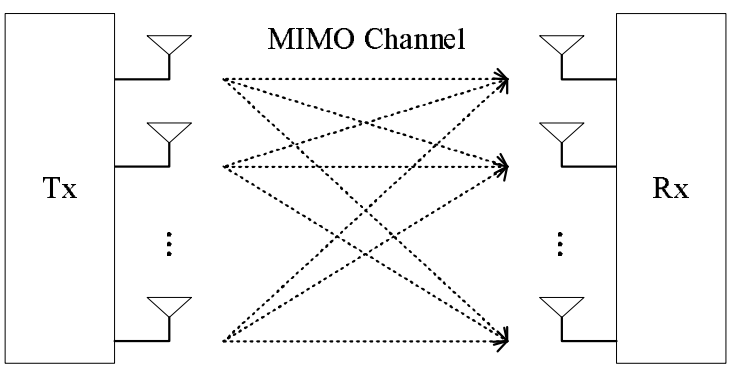

Figure 1.4: Block diagram of MIMO system.

Chapter 2

Adaptive Decision Feedback Generalized Sidelobe Cancellation

2.1 Introduction

Beamforming technology [4], [28], [29] plays an important role in radar, sonar, microphone array speech processing, and, more recently, wireless communications [1]-[3]. The LCMV is a commonly used criterion for a beamformer to suppress interference and noise. The GSC makes the implementation of the LCMV much more efficient. It can effectively reduce the computational cost, especially implemented with adaptive algorithms. However, whenever the actual and presumed DOA of the desired signal is different, the performance of the adaptive GSC degrades significantly. The conventional approach to the design of a beamformer assumes that the desired signal is absent in the training period. In such a case, the beamformer is known to be sufficiently robust against mismatch errors in the array response [30]. Unfortunately, in typical applications, including wireless communications, the signal-free training snapshots are difficult to obtain or even not available. It makes the instinctive robustness absent from these applications.

In [31], a spatial domain notch filter has been incorporated into the conventional GSC to exclude the desired signal component in the beamformer input. Robustness is then guaranteed because of the signal-free operation. It also lessens the deviation between the adaptive and optimum weights, and improves the convergence rate. The main disadvantage of this approach is the need of a sharp notch filter and a slave array for recovering the desired signal. There are several *ad hoc* approaches to the design of robust adaptive

beamformer, e.g., exploiting the eigenspace-based structure [32], [33], main beam constraints [34]-[36], diagonal loading [37] and matrix tapers [38] (see also [30] and [39] for a good summary). Other well-defined methods for designing robust adaptive beamformer include modifying the original optimization problems [40], [41]. They employ additional constraints in optimization and hence the beamforming capabilities lessen. In [42], the signal cyclostationarity was exploited for providing robustness for array processing; however, the implementation of the algorithm needs *a priori* information on some auxiliary parameters, and suffers from high computational complexity. A mathematically tractable method called the worst-case performance optimization (WCPO) was proposed recently, and formulated by the second-order cone programming (SOCP) [43]. Later, a more general approach of the algorithm was analyzed and the online implementation complexity was reduced to the order of $O(N^2)$ [44], where N is the dimension of the beamformer. Unfortunately, the WCPO approach cannot be applied to the GSC scheme directly. Also, the computational complexity is still rather high.

To reduce the implementation complexity, the weights in the GSC can be estimated using adaptive methods and these result in adaptive beamforming structures. The LMS algorithm is widely used in adaptive processing. It is well-known for its simplicity and robustness [45]. However, due to the special structure inherent in the GSC, the convergence rate of the LMS algorithm may be quite slow. Although we can use the recursive least-squares (RLS) algorithm to improve the convergence rate, the computational complexity will be substantially increased. In [46]-[48], a sophisticated direct data-domain least-squares algorithm solved by the conjugate-gradient method was proposed for adaptive beamforming. With reduced computational complexity, this method can still provide fast response to dynamic environments.

In this chapter, we present an adaptive DFGSC to overcome the insufficient robustness and slow convergence problems mentioned previously. In wireless communications, the transmitted symbols have discrete values. We can then take advantage of this characteristic and employ a decision feedback scheme. We introduce a decision-directed equalizer and a feedback filter in the GSC structure. This structure can eliminate the desired signal component from the error signal. With this modified error signal, the proposed DFGSC

can avoid the signal cancellation problem and provide extra robustness against a DOA mismatch. The computational complexity can be kept on the order of $O(N)$. Meanwhile, the modified error signal allows the use of a large step size and the convergence rate of the LMS algorithm can be greatly accelerated.

This chapter is organized as follows. In Section 2.2, the signal model of a narrowband GSC beamformer and its classic solution are described. In Section 2.3, we propose the DFGSC structure and derive its optimum solution. In Section 2.4, we give the mismatch model and analyze the behavior of both the GSC and DFGSC with a DOA mismatch. Section 2.5 analyzes the convergence behavior of both the adaptive GSC and adaptive DFGSC (with the LMS algorithm). The weakness of the GSC structure in this adaptive implementation and the amendment by the decision feedback technique is examined. Finally, simulation results and summary are given in Section 2.6 and 2.7, respectively.

2.2 Background

2.2.1 Signal Model

Consider a uniform linear array (ULA) of N antenna elements. Let a desired signal from far field impinge on the array from a *known* DOA θ_0 along with M uncorrelated interfering signals from unknown DOAs $\{\theta_1, \theta_2, \cdots, \theta_M\}$, respectively. With the first element as the reference point, the $N\times 1$ desired signal's steering vector is given by

$$\mathbf{a}(\theta_0) = [1, \mathrm{e}^{i\tau_{\theta_0}}, \mathrm{e}^{i2\tau_{\theta_0}}, \cdots, \mathrm{e}^{i(N-1)\tau_{\theta_0}}]^T \tag{2.1}$$

where $i = \sqrt{-1}$ and $\tau_{\theta_0} = (2\pi d/L_\lambda)\sin\theta_0$, in which d is the element spacing and L_λ is the signal wavelength. Similarly, $\mathbf{a}(\theta_m) = [1, \mathrm{e}^{i\tau_{\theta_m}}, \mathrm{e}^{i2\tau_{\theta_m}}, \cdots, \mathrm{e}^{i(N-1)\tau_{\theta_m}}]^T$ $(1 \le m \le M)$ corresponds to the interference arriving from direction θ_m, with $\tau_{\theta_m} = (2\pi d/L_\lambda)\sin\theta_m$. Then, the kth snapshot of the $N\times 1$ received equivalent baseband signal vector at the ULA can be written as

$$\begin{aligned}\mathbf{x}(k) &= \mathbf{a}(\theta_0)s_0(k) + \sum_{m=1}^{M}\mathbf{a}(\theta_m)s_m(k) + \mathbf{n}(k) \\ &\triangleq \mathbf{s}(k) + \mathbf{i}(k) + \mathbf{n}(k)\end{aligned} \tag{2.2}$$

where $s_0(k)$ denotes the desired signal, $s_m(k)$ $(1 \leq m \leq M)$ the mth interfering signal, $\mathbf{s}(k) = \mathbf{a}(\theta_0)s_0(k)$, $\mathbf{i}(k) = \sum_{m=1}^{M} \mathbf{a}(\theta_m)s_m(k)$, and $\mathbf{n}(k)$ the additive noise vector in the array. Each noise component is assumed here to be spatially white and Gaussian with a variance of σ_n^2. As shown in (2.2), we have decomposed the received signal into three components: desired signal, interference, and noise. For communication applications, $s_0(k)$ corresponds to the channel output which is a convolution of the channel response and the original transmitted symbol sequence. For simplicity, we only consider quadrature phase-shift keying (QPSK) modulation and flat fading channel environments in this chapter. Thus, we can have $s_0(k) = h(k)b_0(k)$, where $h(k)$ is the channel response and $b_0(k)$ is the desired transmitted symbol. The channel response can be further expressed as $h(k) = \beta(k)\mathrm{e}^{i\phi(k)}$, where $\beta(k)$ is the amplitude response and $\phi(k)$ is the phase response.

2.2.2 Conventional Generalized Sidelobe Canceller

The narrowband beamformer output at time instant k, i.e., $y(k)$, can be expressed as $y(k) = \mathbf{w}^H\mathbf{x}(k)$, where $\mathbf{w}$ is the weight vector. The LCMV beamformer determines $\mathbf{w}$ by minimizing the interference and noise output power under appropriate linear weight constraints, which is given as

$$\min_{\mathbf{w}} \quad \mathbf{w}^H\mathbf{R_{in}}\mathbf{w}, \qquad \text{subject to} \quad \mathbf{C}^H\mathbf{w} = \mathbf{f} \tag{2.3}$$

where $\mathbf{R_{in}} = \mathrm{E}\{(\mathbf{i}(k)+\mathbf{n}(k))(\mathbf{i}(k)+\mathbf{n}(k))^H\}$ is the input correlation matrix of interference-plus-noise, $\mathbf{C}$ is an $N \times P$ constraint matrix, and $\mathbf{f}$ is a $P \times 1$ response vector, with P being the number of constraints. Several different philosophies can be employed for choosing the constraint matrix and the response vector [39]. In many communication applications, the correlation matrix $\mathbf{R_{in}}$ is usually not available. Consequently, the input correlation matrix $\mathbf{R_x} = \mathrm{E}\{\mathbf{x}(k)\mathbf{x}^H(k)\}$ at the receiver is used instead. The optimization problem then becomes

$$\min_{\mathbf{w}} \quad \mathbf{w}^H\mathbf{R_x}\mathbf{w}, \qquad \text{subject to} \quad \mathbf{C}^H\mathbf{w} = \mathbf{f} \tag{2.4}$$

which is called the linearly constrained minimum power (LCMP) criterion [39]. The following development will be based on this workable criterion. It is not difficult to see that the optimum weight vectors solved by (2.3) and (2.4) yield the same gain and

the same output spectrum when the signal is perfectly matched and the distortionless constraint is imposed [39].

The GSC is an alternative formulation of the LCMV/LCMP beamformer. It has been shown that the GSC can convert the constrained optimization in (2.3) or (2.4) into an unconstrained one [6]. The structure of the conventional GSC is illustrated in the left part of Fig. 2.1. As shown in the figure, the upper path includes the quiescent signal matched filter $\mathbf{w}_{\mathrm{q}}$. The lower path includes the blocking matrix $\mathbf{B}$ and the interference cancelling filter $\mathbf{w}_{\mathrm{a}}$. Ideally, the span of $\mathbf{B}$ is in the null space of $\mathbf{C}^H$. From the left part of Fig. 2.1, we can obtain the output of the GSC expressed as

$$y(k) = (\mathbf{w}_{\mathrm{q}} - \mathbf{B}\mathbf{w}_{\mathrm{a}})^H \mathbf{x}(k) \tag{2.5}$$

where $\mathbf{w}_{\mathrm{q}}$ is of dimension $N \times 1$, $\mathbf{B}$ of dimension $N \times (N-P)$, and $\mathbf{w}_{\mathrm{a}}$ of dimension $(N-P) \times 1$. Using the constraint in (2.4), we can solve $\mathbf{w}_{\mathrm{q}}$ as

$$\mathbf{w}_{\mathrm{q}} = \mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{f}. \tag{2.6}$$

Let $J = \mathrm{E}\{|y(k)|^2\}$ denote a cost function of MSE. The constrained optimization problem in (2.4) can then be rewritten as the following unconstrained optimization problem.

$$\min_{\mathbf{w}_{\mathrm{a}}} \quad J = \min_{\mathbf{w}_{\mathrm{a}}} \quad (\mathbf{w}_{\mathrm{q}} - \mathbf{B}\mathbf{w}_{\mathrm{a}})^H \mathbf{R}_{\mathbf{x}} (\mathbf{w}_{\mathrm{q}} - \mathbf{B}\mathbf{w}_{\mathrm{a}}). \tag{2.7}$$

Then, we can find the optimum $\mathbf{w}_{\mathrm{a}}$ as [6]

$$\mathbf{w}_{\mathrm{a,opt}} = (\mathbf{B}^H \mathbf{R}_{\mathbf{x}} \mathbf{B})^{-1} \mathbf{B}^H \mathbf{R}_{\mathbf{x}} \mathbf{w}_{\mathrm{q}}. \tag{2.8}$$

Let $\mathbf{w}_{\mathrm{opt}} = \mathbf{w}_{\mathrm{q}} - \mathbf{B}\mathbf{w}_{\mathrm{a,opt}}$. The MMSE for (2.7), denoted as $J_{\min}$, is calculated as

$$\begin{aligned} J_{\min} &= \mathbf{w}_{\mathrm{opt}}^H \mathbf{R}_{\mathbf{x}} \mathbf{w}_{\mathrm{opt}} \\ &= \mathbf{w}_{\mathrm{q}}^H \mathbf{R}_{\mathbf{x}} \mathbf{w}_{\mathrm{opt}} \\ &= \mathbf{f}^H (\mathbf{C}^H \mathbf{R}_{\mathbf{x}}^{-1} \mathbf{C})^{-1} \mathbf{f}. \end{aligned} \tag{2.9}$$

It is simple to see that $J_{\min}$ is just the minimum GSC output power, denoted as $P_{\mathrm{o,min}}$. When the number of antennas is larger than the number of interfering sources, i.e., the DOFs are high enough, and the distortionless constraint is set to the desired signal's

DOA, all interference tends to be cancelled in the GSC output [5]. The MMSE will then be dominated by the output desired signal power.

$$\begin{aligned} J_{\min} &\simeq \sigma_{s_0}^2 |\mathbf{w}_{\mathrm{q}}^H \mathbf{a}(\theta_0)|^2 \\ &\triangleq P_{\mathbf{s}} \end{aligned} \tag{2.10}$$

where $\sigma_{s_0}^2$ and $P_{\mathbf{s}}$ denote the desired signal power (variance) in the beamformer input and output, respectively.

The output SINR is a widely accepted performance measure for beamforming. The optimum SINR can be written as

$$\mathrm{SINR}_{\mathrm{opt}} = \frac{\mathrm{E}\{|\mathbf{w}_{\mathrm{opt}}^H \mathbf{s}(k)|^2\}}{\mathrm{E}\{|\mathbf{w}_{\mathrm{opt}}^H \mathbf{x}(k) - \mathbf{w}_{\mathrm{opt}}^H \mathbf{s}(k)|^2\}}. \tag{2.11}$$

For the conventional GSC, the numerator of the SINR expression is just the power of the desired signal in the output, as shown in (2.10), and the denominator of the SINR is the power of the interference-plus-noise in the output, which can be found to be

$$\begin{aligned} \mathrm{E}\{|\mathbf{w}_{\mathrm{opt}}^H \mathbf{x}(k) - \mathbf{w}_{\mathrm{opt}}^H \mathbf{s}(k)|^2\} &= P_{\mathrm{o,min}} - P_{\mathbf{s}} \\ &= \mathbf{w}_{\mathrm{q}}^H \mathbf{R}_{\mathbf{x}} \mathbf{w}_{\mathrm{opt}} - \sigma_{s_0}^2 |\mathbf{w}_{\mathrm{q}}^H \mathbf{a}(\theta_0)|^2 \\ &= \mathbf{w}_{\mathrm{q}}^H \mathbf{R}_{\mathbf{in}} \mathbf{w}_{\mathrm{opt}}. \end{aligned} \tag{2.12}$$

The general expression for the optimum SINR can then be summarized as

$$\mathrm{SINR}_{\mathrm{opt}} = \frac{P_{\mathbf{s}}}{P_{\mathrm{o,min}} - P_{\mathbf{s}}}. \tag{2.13}$$

Since $\mathbf{C}$ and $\mathbf{f}$ are known *a priori*, $\mathbf{w}_{\mathrm{q}}$ can be calculated offline using (2.6). The optimum $\mathbf{w}_{\mathrm{a}}$, however, depends on the input correlation matrix which cannot be known in advance. A simple alternative to find the optimum $\mathbf{w}_{\mathrm{a}}$ is to use an adaptive training method. The LMS algorithm, being one of the stochastic gradient methods, is known to be a simple yet effective adaptive algorithm [45]. Using the cost function in (2.7), we can calculate the stochastic gradient with respect to $\mathbf{w}_{\mathrm{a}}^*$. The LMS update equation for $\mathbf{w}_{\mathrm{a}}$ is written as

$$\mathbf{w}_{\mathrm{a}}(k+1) = \mathbf{w}_{\mathrm{a}}(k) + \mu_{\mathrm{a}} \mathbf{v}(k) e^*(k) \tag{2.14}$$

where $\mathbf{w}_{\mathrm{a}}(k)$ is the estimate of $\mathbf{w}_{\mathrm{a,opt}}$ at the kth snapshot, $\mathbf{v}(k) = \mathbf{B}^H\mathbf{x}(k)$ is the filter input vector (the output vector from the blocking matrix), μ_{a} is the step size controlling the convergence rate, and $e(k)$ is an error signal between the desired and actual outputs. For GSC applications, we have $e(k) = y(k)$.

As mentioned, the conventional GSC is sensitive to a DOA mismatch which can easily occur in practical circumstances. Also, as seen from (2.14), the update term of the LMS algorithm involves $y(k)$, which will approach the output desired signal $\mathbf{w}_{\mathrm{q}}^H\mathbf{s}(k)$ ideally. This indicates that the stochastic gradient in (2.14), i.e., $\mathbf{v}(k)y^*(k)$, will not be close to zero even for optimum weights. As a result, the excess MSE induced by the LMS algorithm will be large. It has been shown that the excess MSE is roughly proportional to $\mu_{\mathrm{a}}J_{\min}$ [45], where $J_{\min}$ is the MMSE value in (2.9). To reduce the excess MSE, we then have to use a small step size and this slows the convergence of the conventional LMS-based adaptive GSC.

2.3 Decision Feedback Generalized Sidelobe Canceller

In this section, we propose a new decision feedback scheme to improve the robustness and performance of the conventional GSC. Fig. 2.1 shows the whole structure of the proposed DFGSC. The idea is to introduce a decision-directed equalizer and a feedback filter to modify the GSC output $y(k)$, which is then used as the error signal in the LMS adaptation. The purpose of the decision-directed equalizer is to equalize the channel and the DOA mismatch effect, and the feedback filter is to cancel any desired signal component in the LMS error signal. Due to the scenario we considered, we only need one weight for the equalizer and one weight for the feedback filter. Note that the decision-directed equalizer and the feedback filter are trained by different error signals. The equalizer is trained by the error signal between the output and input signals of the decision device while the feedback filter is trained by the error signal between the GSC and the feedback filter outputs. The advantage of this structure is that there will be no coupling effect between these two filters. The decision-directed equalizer can recover $b_0(k)$ with a phase ambiguity and we will show that this is sufficient for our use. Although any other blind equalization algorithms can be applied, we found that they may not be more effective for the scenario

considered here. The optimum weight for the decision-directed equalizer is obtained by minimizing an MSE criterion shown below.

$$\min_{w_{\rm m}} \quad J = \min_{w_{\rm m}} \quad {\rm E}\{|e_{\rm m}(k)|^2\} \tag{2.15}$$

where $e_{\rm m}(k) = \hat{b}_0(k) - w_{\rm m}^* y(k)$, with $\hat{b}_0(k)$ being the detected symbol and $w_{\rm m}$ being the equalizer tap weight. Taking the stochastic gradient of J in (2.15) and applying the LMS algorithm, we can obtain the update equation for $w_{\rm m}$ as

$$w_{\rm m}(k+1) = w_{\rm m}(k) + \mu_{\rm m} y(k) e_{\rm m}^*(k) \tag{2.16}$$

where $\mu_{\rm m}$ is the step size controlling the convergence behavior of $w_{\rm m}$. It can be easily shown that the tap weight in (2.16) can equalize the channel effect up to a phase ambiguity of $L\pi/2$, where $L = 1, 2, 3$ for QPSK modulation. Since our purpose is just to cancel the desired signal component in the GSC output, knowledge of the exact transmitted symbol $b_0(k)$ is not required. We now state the reason: Let $w_{\rm b}$ denote the feedback tap weight and $w_{\rm b,opt}$ denote the optimum choice for it. We assume that $w_{\rm b}^* = w_{\rm b,opt}^*$ and the decision is correct, i.e., $\hat{b}_0(k) = b_0(k)$. Then, let the decision have a phase ambiguity of ρ, i.e., $\hat{b}_0(k) = b_0(k)e^{j\rho}$. It is simple to see that if we make $w_{\rm b}^* = e^{-j\rho} w_{\rm b,opt}^*$, the output of the feedback filter will remain exactly the same.

With the proposed feedback structure, the cost function for the DFGSC is changed to

$$\begin{aligned} J &= {\rm E}\{|e(k)|^2\} \\ &= {\rm E}\{|\mathbf{w}_{\rm q}^H \mathbf{x}(k) - [\mathbf{w}_{\rm a}^H \quad w_{\rm b}^*] \begin{bmatrix} \mathbf{B}^H \mathbf{x}(k) \\ \hat{b}_0(k) \end{bmatrix} |^2\}. \end{aligned} \tag{2.17}$$

As mentioned, $\hat{b}_0(k)$ may have a phase ambiguity with respect to $b_0(k)$, but this will not affect the final result. For convenience, we simply assume that the decision is correct, i.e., $\hat{b}_0(k) = b_0(k)$ in the following analysis. Similarly, the channel effect $h(k)$ does not have any impact in our analysis and is assumed to be $h(k) = 1$. We may write the detected desired signal as $\hat{s}_0(k) = s_0(k) = \hat{b}_0(k) = b_0(k)$. With the point distortionless constraint, the minimization of the cost function in (2.17) can be written as

$$\begin{aligned} \min_{\mathbf{w}_{\rm a},\, w_{\rm b}} \quad J = \min_{\mathbf{w_c}} \quad & \mathbf{w}_{\rm q}^H \mathbf{R_x} \mathbf{w}_{\rm q} - \mathbf{w}_{\rm q}^H \begin{bmatrix} \mathbf{R_x B} & \mathbf{p} \end{bmatrix} \mathbf{w_c} \\ & - \mathbf{w_c}^H \begin{bmatrix} \mathbf{B}^H \mathbf{R_x} \\ \mathbf{p}^H \end{bmatrix} \mathbf{w}_{\rm q} + \mathbf{w_c}^H \mathbf{R_c} \mathbf{w_c} \end{aligned} \tag{2.18}$$

in which we let

$$\mathbf{w_c} = \begin{bmatrix} \mathbf{w}_\mathrm{a} \\ w_\mathrm{b} \end{bmatrix} \tag{2.19}$$

$$\mathbf{R_c} = \mathrm{E}\{\begin{bmatrix} \mathbf{B}^H\mathbf{x}(k) \\ \hat{s}_0(k) \end{bmatrix} \begin{bmatrix} \mathbf{x}^H(k)\mathbf{B} & \hat{s}_0^*(k) \end{bmatrix}\} = \begin{bmatrix} \mathbf{B}^H\mathbf{R_x}\mathbf{B} & \mathbf{0} \\ \mathbf{0}^H & \sigma_{s_0}^2 \end{bmatrix} \tag{2.20}$$

and

$$\mathbf{p} = \mathrm{E}\{\mathbf{x}(k)\hat{s}_0^*(k)\} = \sigma_{s_0}^2\mathbf{a}(\theta_0) \tag{2.21}$$

where $\mathbf{0}$ denotes a zero vector with dimension $(N-1)\times 1$. We can set $\mathbf{w}_\mathrm{q}$ the same as that in the conventional GSC. Taking the derivative of the cost function with respect to $\mathbf{w_c^*}$ and setting the result to zero, we can obtain the optimum $\mathbf{w_c}$.

$$\frac{\partial J}{\partial \mathbf{w_c^*}} = -2\begin{bmatrix} \mathbf{B}^H\mathbf{R_x} \\ \mathbf{p}^H \end{bmatrix}\mathbf{w}_\mathrm{q} + 2\mathbf{R_c}\mathbf{w_c} = 0. \tag{2.22}$$

Thus,

$$\mathbf{w}_{\mathbf{c},\mathrm{opt}} = \mathbf{R_c}^{-1}\begin{bmatrix} \mathbf{B}^H\mathbf{R_x} \\ \mathbf{p}^H \end{bmatrix}\mathbf{w}_\mathrm{q}. \tag{2.23}$$

Utilizing the special structure of $\mathbf{R_c}$, we can decompose $\mathbf{w}_{\mathbf{c},\mathrm{opt}}$ back into the two weights

$$\mathbf{w}_{\mathrm{a,opt}} = (\mathbf{B}^H\mathbf{R_x}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{R_x}\mathbf{w}_\mathrm{q} \tag{2.24}$$

$$w_{\mathrm{b,opt}} = \frac{\mathbf{p}^H}{\sigma_{s_0}^2}\mathbf{w}_\mathrm{q} = \mathbf{a}^H(\theta_0)\mathbf{w}_\mathrm{q}. \tag{2.25}$$

With $\mathbf{w}_\mathrm{q}$, $\mathbf{w}_{\mathrm{a,opt}}$, $w_{\mathrm{b,opt}}$, and the result in (2.9), the minimum J of (2.17) (which is not the minimum output power $P_{\mathrm{o,min}}$ in this case) for the DFGSC becomes

$$\begin{aligned} J_{\min} &= \mathbf{w}_\mathrm{q}^H\mathbf{R_x}\mathbf{w}_\mathrm{opt} - \sigma_{s_0}^2|\mathbf{w}_\mathrm{q}^H\mathbf{a}(\theta_0)|^2 \\ &= \mathbf{w}_\mathrm{q}^H\mathbf{R_{in}}\mathbf{w}_\mathrm{opt}. \end{aligned} \tag{2.26}$$

We can see that the desired signal component is totally excluded from the MMSE expression for the DFGSC and the resultant MMSE can thus be small. From the equations given above, we find three notable features of the DFGSC.

- The expression of $\mathbf{w}_{\mathrm{a,opt}}$ in (2.24) is the same as that in the conventional GSC.
- The effect of the additional feedback tap weight is only to reduce the minimum value of the cost function. Since $\mathbf{w}_\mathrm{q}$ and $\mathbf{w}_{\mathrm{a,opt}}$ remain the same, the minimum output power is not affected.

- From (2.25), the output of the optimum feedback filter $w_{\mathrm{b,opt}}^{*}\hat{s}_0(k)$ can be shown to equal $\mathbf{w}_{\mathrm{q}}^{H}\mathbf{s}(k)$, which is exactly the desired signal component in the output upon ideal error-free conditions.

Equation (2.13) is still valid for expressing the optimum output SINR for the DFGSC, where $P_{\mathrm{o,min}}$ used is the minimum output power as given in (2.9). Therefore, the optimum GSC performance is not enhanced by the decision feedback operation. However, when there is a DOA mismatch or the LMS is used to estimate those optimum weights, the performance can be greatly improved by the decision feedback structure. This will be elaborated in the next two sections. Similarly, the LMS update equations for the tap weights of the DFGSC can be written as

$$\mathbf{w}_{\mathrm{a}}(k+1) = \mathbf{w}_{\mathrm{a}}(k) + \mu_{\mathrm{a}}\mathbf{v}(k)e^{*}(k) \tag{2.27}$$

$$w_{\mathrm{b}}(k+1) = w_{\mathrm{b}}(k) + \mu_{\mathrm{b}}\hat{s}_0(k)e^{*}(k) \tag{2.28}$$

where μ_{a} is the step size for $\mathbf{w}_{\mathrm{a}}$, μ_{b} is the step size for w_{b}, $\mathbf{v}(k)$ is the filter input vector, and $e(k) = y(k) - w_{\mathrm{b}}^{*}(k)\hat{s}_0(k)$. Unlike the conventional GSC, the steady-state $e(k)$ will exclude the desired signal component and hence can be quite small. It is where the improvement of the DFGSC stems from. The derivations given above are based on the assumption that the decision is correct. Actually, decision errors occur sometimes. In general, they can be seen as some *sparse* noise added to the error signal $e(k)$. If the error rate is low, this will only increase the MSE slightly. We will show by simulation that even when the decision error rate is high (signal-to-noise ratio (SNR) is low), the overall performance of the DFGSC is still better than (or at least the same as) that of the conventional GSC.

In implementations, the GSC is usually come with the steering delays. They ensure that the desired signal effectively appears at 0^{o} and the computation of $\mathbf{w}_{\mathrm{q}}$ and $\mathbf{B}$ can be much simplified. For example, $\mathbf{w}_{\mathrm{q}}$ becomes an all-one vector. For the case of $N = 2^{l}$, where l is any nonnegative integer, a simple choice for $\mathbf{B}$ is the Hadamard matrix (excluding the first column). In this case, with the point distortionless constraint only, the complex multiplications required for the conventional adaptive GSC are $2(N-1)$ per snapshot and those for the adaptive DFGSC are $2(N+1)$ per snapshot. Four additional complex multiplications are used for the decision-directed equalizer and the feedback filter. For

$N = 8$, 16 and 32, the increases of complex multiplications in the DFGSC are 28.6%, 13.3% and 6.5%, respectively. We can see that the growth in complexity becomes negligible as the number of antennas is large.

2.4 DOA Mismatch Analysis

2.4.1 Mismatch Signal Model

From Section 2.2, we have defined the expression of the signal steering vector under perfectly matched conditions. If there is a mismatch between the actual and presumed desired signal's DOA, i.e., $\theta_0 = \hat{\theta}_0 + \Delta$, where $\hat{\theta}_0$ is the estimated DOA and Δ is the mismatch value, we may rewrite the steering vector as

$$\begin{aligned}\mathbf{a}(\theta_0) &= \mathbf{a}(\hat{\theta}_0 + \Delta) \\ &= [1, \mathrm{e}^{i\alpha \sin(\hat{\theta}_0+\Delta)}, \mathrm{e}^{i2\alpha \sin(\hat{\theta}_0+\Delta)}, \cdots, \mathrm{e}^{i(N-1)\alpha \sin(\hat{\theta}_0+\Delta)}]^T \end{aligned} \tag{2.29}$$

with $\alpha = 2\pi d/L_\lambda$. With modern DOA estimation methods [2], [39], Δ is generally small (if existed). Thus, we can have

$$\sin(\hat{\theta}_0 + \Delta) \simeq \sin(\hat{\theta}_0) + \Delta \cos(\hat{\theta}_0). \tag{2.30}$$

We may approximate the mismatch steering vector as

$$\begin{aligned}\mathbf{a}(\theta_0) \simeq [1, &\mathrm{e}^{i\alpha \sin\hat{\theta}_0} \cdot \mathrm{e}^{i\alpha\Delta \cos\hat{\theta}_0}, \ \mathrm{e}^{i2\alpha \sin\hat{\theta}_0} \cdot \mathrm{e}^{i2\alpha\Delta \cos\hat{\theta}_0}, \cdots, \\ &\mathrm{e}^{i(N-1)\alpha \sin\hat{\theta}_0} \cdot \mathrm{e}^{i(N-1)\alpha\Delta \cos\hat{\theta}_0}]^T \end{aligned} \tag{2.31}$$

in which each element in the mismatch steering vector is composed of an ordinary steering term multiplied by another mismatch steering term. Without loss of generality, we let the system be pre-steered, i.e., $\hat{\theta}_0$ is adjusted to 0°. The mismatch steering vector can then be simplified to

$$\begin{aligned}\mathbf{a}(\theta_0) &= \mathbf{a}(\Delta) \\ &\simeq [1, \mathrm{e}^{i\alpha\Delta}, \mathrm{e}^{i2\alpha\Delta}, \cdots, \mathrm{e}^{i(N-1)\alpha\Delta}]^T. \end{aligned} \tag{2.32}$$

We use (2.32) as the desired signal's steering vector for the following analysis. Then, $\mathbf{s}(k)$, being equal to $\mathbf{a}(\theta_0)s_0(k)$, becomes $\mathbf{a}(\Delta)s_0(k)$.

2.4.2 Mismatch Analysis with GSC and DFGSC

With an estimation error of the desired signal's DOA, $\mathbf{w}_{\mathrm{q}}$ is not matched to the desired signal's steering vector, and $\mathbf{B}$ cannot obstruct the desired signal entering the input of the interference cancelling filter. If there are enough DOFs, the filter will cancel the desired signal from the quiescent filter output. The so-called signal cancellation occurs. In the case of the conventional GSC, even with the optimum weight vector (under distortionless constraint), it does not provide the ability to maximize the SINR anymore [7]. The expression of the optimum output SINR for the conventional GSC is again the same as given in (2.13), but now $P_{\mathbf{s}}$ changes. The output desired signal power in the conventional GSC becomes

$$P_{\mathbf{s}} = \sigma_{s_0}^2 |(\mathbf{w}_{\mathrm{q}} - \mathbf{B}\mathbf{w}_{\mathrm{a,opt}})^H \mathbf{a}(\Delta)|^2 \tag{2.33}$$

which is different from (2.10) because $\mathbf{w}_{\mathrm{a,opt}}^H \mathbf{B}^H \mathbf{a}(\Delta)$ is not zero now. The actual amount of signal attenuation depends upon the power of the signal and the amount of error [49]. The minimum output power $P_{\mathrm{o,min}}$ for calculating the optimum output SINR of the conventional GSC in mismatch is the same as (2.9).

On the other hand, we can show that the signal cancellation phenomenon is avoided in the DFGSC. Since a correlation exists between the two signal paths in the GSC structure whenever a mismatch exists, the optimum solutions for $\mathbf{w}_{\mathrm{a}}$ and w_{b} in the DFGSC are coupled together. From (2.23), we have

$$\mathbf{w}_{\mathbf{c},\mathrm{opt}} = \begin{bmatrix} \mathbf{w}_{\mathrm{a,opt}} \\ w_{\mathrm{b,opt}} \end{bmatrix} = \mathbf{R}_{\mathbf{c}}^{-1} \begin{bmatrix} \mathbf{B}^H \mathbf{R}_{\mathbf{x}} \\ \mathbf{p}^H \end{bmatrix} \mathbf{w}_{\mathrm{q}} \tag{2.34}$$

with

$$\mathbf{p} = \sigma_{s_0}^2 \mathbf{a}(\Delta) \tag{2.35}$$

and

$$\mathbf{R}_{\mathbf{c}} = \begin{bmatrix} \mathbf{B}^H \mathbf{R}_{\mathbf{x}} \mathbf{B} & \mathbf{m} \\ \mathbf{m}^H & \sigma_{s_0}^2 \end{bmatrix} \tag{2.36}$$

where $\mathbf{m} = \sigma_{s_0}^2 \mathbf{B}^H \mathbf{a}(\Delta)$, which is the correlation between the blocking matrix output $\mathbf{B}^H \mathbf{x}(k)$ and the decision $\hat{s}_0(k)$. By using the inversion identity for subblock matrices [50], we can find the inverse of $\mathbf{R}_{\mathbf{c}}$ and so $\mathbf{w}_{\mathbf{c},\mathrm{opt}}$ in (2.34). For reference, we give the

inversion form used here as

$$\begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{S}_{\mathrm{D}}^{-1} & -\mathbf{S}_{\mathrm{D}}^{-1}\mathbf{A}_2\mathbf{A}_4^{-1} \\ -\mathbf{A}_4^{-1}\mathbf{A}_3\mathbf{S}_{\mathrm{D}}^{-1} & \mathbf{A}_4^{-1} + \mathbf{A}_4^{-1}\mathbf{A}_3\mathbf{S}_{\mathrm{D}}^{-1}\mathbf{A}_2\mathbf{A}_4^{-1} \end{bmatrix} \tag{2.37}$$

where $\mathbf{S}_{\mathrm{D}} = \mathbf{A}_1 - \mathbf{A}_2\mathbf{A}_4^{-1}\mathbf{A}_3$ is the Schur complement of $\mathbf{A}_4$. After some manipulation, we have

$$\mathbf{w}_{\mathbf{c},\mathrm{opt}} = \begin{bmatrix} (\mathbf{B}^H\mathbf{R}_{\mathbf{in}}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{R}_{\mathbf{in}} \\ \mathbf{a}^H(\Delta)[\mathbf{I} - \mathbf{B}(\mathbf{B}^H\mathbf{R}_{\mathbf{in}}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{R}_{\mathbf{in}}] \end{bmatrix} \mathbf{w}_{\mathrm{q}} \tag{2.38}$$

and thus the coupled $\mathbf{w}_{\mathrm{a,opt}}$ and $w_{\mathrm{b,opt}}$ can be written as

$$\mathbf{w}_{\mathrm{a,opt}} = (\mathbf{B}^H\mathbf{R}_{\mathbf{in}}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{R}_{\mathbf{in}}\mathbf{w}_{\mathrm{q}} \tag{2.39}$$

$$w_{\mathrm{b,opt}} = \mathbf{a}^H(\Delta)(\mathbf{w}_{\mathrm{q}} - \mathbf{B}\mathbf{w}_{\mathrm{a,opt}}). \tag{2.40}$$

The MMSE of the DFGSC with mismatch can then be solved to be

$$\begin{aligned} J_{\min} &= \mathbf{w}_{\mathrm{q}}^H\mathbf{R}_{\mathbf{x}}\mathbf{w}_{\mathrm{q}} - \mathbf{w}_{\mathrm{q}}^H \begin{bmatrix} \mathbf{R}_{\mathbf{x}}\mathbf{B} & \mathbf{p} \end{bmatrix} \mathbf{R}_{\mathbf{c}}^{-1} \begin{bmatrix} \mathbf{B}^H\mathbf{R}_{\mathbf{x}} \\ \mathbf{p}^H \end{bmatrix} \mathbf{w}_{\mathrm{q}} \\ &= \mathbf{w}_{\mathrm{q}}^H\mathbf{R}_{\mathbf{in}}\mathbf{w}_{\mathrm{q}} - \mathbf{w}_{\mathrm{q}}^H\mathbf{R}_{\mathbf{in}}\mathbf{B}(\mathbf{B}^H\mathbf{R}_{\mathbf{in}}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{R}_{\mathbf{in}}\mathbf{w}_{\mathrm{q}} \\ &= \mathbf{w}_{\mathrm{q}}^H\mathbf{R}_{\mathbf{in}}\mathbf{w}_{\mathrm{opt}}. \end{aligned} \tag{2.41}$$

It is equivalent to say that the error signal contains no desired signal. This result can also be seen as the fact that from (2.40), the output of the optimum feedback filter $w_{\mathrm{b,opt}}^*\hat{s}_0(k)$ is equal to $(\mathbf{w}_{\mathrm{q}} - \mathbf{B}\mathbf{w}_{\mathrm{a,opt}})^H\mathbf{s}(k)$, which is again exactly the output desired signal. It makes the interference cancelling filter have no means to cancel the desired signal in the GSC output. Here, we notice that the optimum weights are modified. From (2.24) and (2.39), we observe that the only difference between the two solutions of $\mathbf{w}_{\mathrm{a,opt}}$ is that the correlation matrix involved is changed from $\mathbf{R}_{\mathbf{x}}$ to $\mathbf{R}_{\mathbf{in}}$. The expression in (2.39) can be explained as the solution of the originally unreachable LCMV optimization problem given in (2.3), or its GSC implementation as

$$\min_{\mathbf{w}_{\mathrm{a}}} \quad (\mathbf{w}_{\mathrm{q}} - \mathbf{B}\mathbf{w}_{\mathrm{a}})^H\mathbf{R}_{\mathbf{in}}(\mathbf{w}_{\mathrm{q}} - \mathbf{B}\mathbf{w}_{\mathrm{a}}). \tag{2.42}$$

As said previously, this criterion is generally not workable in practical situations for wireless communications because $\mathbf{R}_{\mathbf{in}}$ is not available at the receiver. However, the proposed method can equivalently minimize this criterion. This is where the robustness comes from and why signal cancellation can be avoided.

Although the optimum interference cancelling filter is changed, the expression in (2.13) is still valid for calculating the optimum output SINR for the DFGSC. The power of the desired signal in the output is the same as (2.33), but for the DFGSC in mismatch, the minimum output power should be calculated as

$$P_{\text{o,min}} = \mathbf{w}_{\text{opt}}^H \mathbf{R}_{\mathbf{x}} \mathbf{w}_{\text{opt}}. \tag{2.43}$$

In Section 2.6, we will use simulation results to depict the output SINR difference between the conventional GSC and DFGSC.

2.4.3 A Start-Up Approach for DFGSC with Mismatch

In the LMS-based adaptive implementation, a remaining problem is how to acquire correct decisions initially whenever a DOA mismatch occurs. In this situation, the desired signal may be too weak to initiate the decision-directed equalizer. A straightforward method is to use training. From experiments, we find that the number of training snapshots needed is usually small for both the feedback filter and equalizer. An alternative method is to add derivative constraints, which is a classic approach for providing robustness against mismatch. With suitably chosen derivative constraints, the DFGSC can exclude the use of training snapshots and acquire correct decisions chiefly. The disadvantage of this approach is that the DOFs are reduced and so the interference and noise suppression ability degrades.

We now develop a scheme which can let the DFGSC skip the use of training symbols in the initial phase and reach the good SINR performance through adaptation. The idea is to use enough number of derivative constraints for the DFGSC initially and then release the constraints gradually. The initially broadened main beam can tolerate a certain amount of DOA mismatch value and ensure that there is enough desired signal strength in the quiescent filter output. Then, the equalizer can effectively converge without training. After convergence, the DFGSC may release the constraints one by one. This increment in the DOFs makes the beamformer have the opportunity to strengthen the interference and noise suppression capability. The detailed operation of this start-up approach is stated as follows. For illustration, we assume that the number of point-plus-derivative constraints

is changed from p to $p-1$, and so the DOFs increase from $N-p$ to $N-(p-1)$. We denote the blocking matrix, the quiescent signal matched filter, and the interference cancelling filter when the DOFs are $N-p$ as $\mathbf{B}_{N-p}$, $\mathbf{w}_{\mathrm{q},N-p}$ and $\mathbf{w}_{\mathrm{a},N-p}$. With 0^{o} pre-steering, $\mathbf{w}_{\mathrm{q}}$ can be calculated beforehand with any DOF and $\mathbf{B}$ can be arbitrarily chosen according to the subspace constraints. To add the blocking matrix only one column and increase the DOFs one at a time, we need the following preliminaries.

Initialization:

1) Prepare the projection matrix for $\mathbf{B}_{N-(p-1)}$ as

$$\mathbf{P}_{N-(p-1)} = \mathbf{B}_{N-(p-1)}\mathbf{B}^{H}_{N-(p-1)}. \tag{2.44}$$

2) Calculate the projection of the ready to throw away column of the derivative constraint in $\mathbf{C}$ onto the subspace spanned by columns of $\mathbf{B}^{H}_{N-(p-1)}$ as

$$\mathbf{b}_{N-(p-1)} = \mathbf{P}_{N-(p-1)}\mathbf{c}_{p}. \tag{2.45}$$

3) Prepare a vector holding the difference between the new and old quiescent weight vector

$$\mathbf{d}_{N-(p-1)} = \mathbf{w}_{\mathrm{q},N-(p-1)} - \mathbf{w}_{\mathrm{q},N-p} \tag{2.46}$$

add a zero tap weight to form the new interference cancelling filter

$$\mathbf{w}_{\mathrm{a},N-(p-1)} = [\mathbf{w}_{\mathrm{a},N-p} \quad 0] \tag{2.47}$$

and define a sequence $\mathbf{\Gamma}$ with elements monotonically increasing from 0 to 1, e.g., $\mathbf{\Gamma} = \{\gamma_0, \gamma_1, \cdots, \gamma_{T-1}\}$, where T is the sequence length. We may use a simple sequence $\gamma_k = 1 - ((T-1)-k)/(T-1)$, with $0 \le k \le T-1$.

Note that the length T indicates the number of snapshots during the transition. The longer the length is, the smoother the transition becomes. As a rule of thumb, it may be proportional to the average time constant of the LMS algorithm, which is approximated as [45]

$$\tau_{\mathrm{av}} \approx \frac{1}{2\mu_{\mathrm{a}}\lambda_{\mathrm{av}}} \tag{2.48}$$

where λ_{av} is the average eigenvalue for the underlying correlation matrix. Since the transient response settles in about four times of this average time constant, the length of

$\mathbf{\Gamma}$ can be chosen to be somewhat larger than $4\tau_{\text{av}}$, e.g., $5\tau_{\text{av}}$. It guarantees that there is enough smoothness and no waste of snapshots during transition. After the initialization, the DFGSC can adopt the new adaptation settings as below.

Transition:

1) Estimate the new weight vector $\mathbf{w}_{\text{a},N-(p-1)}$ iteratively according to the new blocking matrix and quiescent weight vector as

$$\mathbf{B}_{N-(p-1)} = [\mathbf{B}_{N-p} \quad \gamma_k \mathbf{b}_{N-(p-1)}] \tag{2.49}$$

$$\mathbf{w}_{\text{q},N-(p-1)} = \mathbf{w}_{\text{q},N-p} + \gamma_k \mathbf{d}_{N-(p-1)} \tag{2.50}$$

with γ_k acting as the kth element in the sequence $\mathbf{\Gamma}$ used for the kth snapshot.

2) Repeat step 1) until $\gamma_k = 1$.

Using the above procedure, the DFGSC can skip the use of training without degrading the SINR performance eventually.

2.5 Convergence Analysis

In this section, we give the convergence analysis for both the adaptive GSC and adaptive DFGSC.

2.5.1 MSE in Steady State

Let the MSE in steady state of the LMS algorithm be denoted as $J(\infty)$. Then

$$J(\infty) = J_{\text{min}} + J_{\text{ex}}(\infty) \tag{2.51}$$

where J_{min} is the MMSE solved by Wiener equations and $J_{\text{ex}}(\infty)$ is the excess MSE caused by the LMS adaptation. Also define the weight-error vector as

$$\boldsymbol{\epsilon}(k) = \mathbf{w}_{\text{a}}(k) - \mathbf{w}_{\text{a,opt}}. \tag{2.52}$$

Using the the direct averaging method [45], we have

$$\boldsymbol{\epsilon}(k+1) = (\mathbf{I} - \mu_{\text{a}} \mathbf{B}^H \mathbf{R}_{\mathbf{x}} \mathbf{B}) \boldsymbol{\epsilon}(k) + \mu_{\text{a}} \mathbf{B}^H \mathbf{x}(k) e^*_{\text{opt}}(k) \tag{2.53}$$

where $e_{\text{opt}}(k)$ denotes the error signal produced with the optimum weights. Define the correlation matrix of the weight-error vector as

$$\mathbf{K}(k) = \mathrm{E}\{\boldsymbol{\epsilon}(k)\boldsymbol{\epsilon}^H(k)\}. \tag{2.54}$$

Invoking the independence assumption [45], we can obtain the recursive relation of $\mathbf{K}(k)$ as

$$\mathbf{K}(k+1) = (\mathbf{I} - \mu_{\text{a}}\mathbf{B}^H\mathbf{R}_{\mathbf{x}}\mathbf{B})\mathbf{K}(k)(\mathbf{I} - \mu_{\text{a}}\mathbf{B}^H\mathbf{R}_{\mathbf{x}}\mathbf{B}) + \mu_{\text{a}}^2 J_{\min}\mathbf{B}^H\mathbf{R}_{\mathbf{x}}\mathbf{B}. \tag{2.55}$$

Under this premise, the excess MSE is written as

$$J_{\text{ex}}(k) = \mathrm{Tr}\{\mathbf{B}^H\mathbf{R}_{\mathbf{x}}\mathbf{B}\mathbf{K}(k)\}. \tag{2.56}$$

As $k \to \infty$, the excess MSE is given by

$$J_{\text{ex}}(\infty) = J_{\min}\sum_{l=1}^{N-P}\frac{\mu_{\text{a}}\lambda_l(\mathbf{B}^H\mathbf{R}_{\mathbf{x}}\mathbf{B})}{2 - \mu_{\text{a}}\lambda_l(\mathbf{B}^H\mathbf{R}_{\mathbf{x}}\mathbf{B})} \tag{2.57}$$

where $\lambda_l(\mathbf{B}^H\mathbf{R}_{\mathbf{x}}\mathbf{B})$ indicates the lth eigenvalue of $\mathbf{B}^H\mathbf{R}_{\mathbf{x}}\mathbf{B}$. Hence, the excess MSE is roughly proportional to the resultant MMSE and the step size used.

2.5.2 SINR in Steady State

The output SINR in steady state is used as the performance measure for the adaptive GSC and adaptive DFGSC. The transient SINR of both schemes can be written as

$$\mathrm{SINR}(k) = \frac{\mathrm{E}\{|\mathbf{w}^H(k)\mathbf{s}(k)|^2\}}{\mathrm{E}\{|\mathbf{w}^H(k)\mathbf{x}(k) - \mathbf{w}^H(k)\mathbf{s}(k)|^2\}}. \tag{2.58}$$

Without mismatch, as $k \to \infty$, the numerator of the SINR expression is the same as $P_{\mathbf{s}}$ given previously in (2.10); however, the denominator of the SINR is changed to

$$\begin{aligned}
&\mathrm{E}\{|\mathbf{w}^H(\infty)\mathbf{x}(\infty) - \mathbf{w}^H(\infty)\mathbf{s}(\infty)|^2\} \\
&= \mathbf{w}_{\text{q}}^H\mathbf{R}_{\mathbf{x}}\mathbf{w}_{\text{opt}} + \mathrm{Tr}\{\mathbf{B}^H\mathbf{R}_{\mathbf{x}}\mathbf{B}\mathbf{K}(\infty)\} - P_{\mathbf{s}} \\
&= P_{\text{o,min}} + J_{\text{ex}}(\infty) - P_{\mathbf{s}}
\end{aligned} \tag{2.59}$$

where $P_{\text{o,min}}$ specifically denotes the minimum output power as given in (2.9). Thus, the steady-state SINR with the LMS algorithm can be summarized as

$$\mathrm{SINR}_{\text{LMS}} = \frac{P_{\mathbf{s}}}{P_{\text{o,min}} + J_{\text{ex}}(\infty) - P_{\mathbf{s}}}. \tag{2.60}$$

Equation (2.60) shows how $J_{\text{ex}}(\infty)$ affects the steady-state SINR. The smaller the excess MSE value is, the larger the steady-state SINR becomes. From (2.57), we can see that $J_{\text{ex}}(\infty)$ is proportional to $J_{\min}$. For the conventional GSC and DFGSC, their MMSEs are shown in (2.9) and (2.26), respectively. It is apparent that the MMSE of the DFGSC is much smaller. As a result, with the same step size, the output SINR of the adaptive DFGSC is higher than that of the conventional adaptive GSC.

In the DOA mismatch case, we define $J_{\text{ex}}^{\mathbf{s}}(\infty)$ as the excess MSE of the leaky desired signal component present in the lower path. The corresponding correlation matrix of this component is $\mathbf{B}^H\mathbf{R}_{\mathbf{s}}\mathbf{B}$ where $\mathbf{R}_{\mathbf{s}} = \mathrm{E}\{\mathbf{s}(k)\mathbf{s}^H(k)\}$. Using (2.57), we can have

$$J_{\text{ex}}^{\mathbf{s}}(\infty) = J_{\min}\frac{\mu_{\text{a}}\lambda_1(\mathbf{B}^H\mathbf{R}_{\mathbf{s}}\mathbf{B})}{2-\mu_{\text{a}}\lambda_1(\mathbf{B}^H\mathbf{R}_{\mathbf{s}}\mathbf{B})} \tag{2.61}$$

since only one eigenvalue of $\mathbf{B}^H\mathbf{R}_{\mathbf{s}}\mathbf{B}$ is nonzero. The expression of the steady-state output SINR with mismatch is then slightly modified as

$$\text{SINR}_{\text{LMS}} = \frac{P_{\mathbf{s}} + J_{\text{ex}}^{\mathbf{s}}(\infty)}{P_{\text{o,min}} + J_{\text{ex}}(\infty) - (P_{\mathbf{s}} + J_{\text{ex}}^{\mathbf{s}}(\infty))}. \tag{2.62}$$

In (2.62), $P_{\mathbf{s}}$ for both the conventional GSC and DFGSC and $P_{\text{o,min}}$ for the DFGSC should be changed to (2.33) and (2.43), respectively. Another well-known adaptive algorithm is the normalized LMS (NLMS) algorithm. It has the advantages of better step size control and faster convergence. However, its computational complexity is also higher. Though the convergence behavior of the NLMS algorithm is similar to that of the LMS algorithm, the theoretical analysis is more involved.

2.6 Simulations

Computer simulations are conducted to verify our analytic results and demonstrate the effectiveness of the proposed algorithm. In all cases, we assume a ULA with N=16 omnidirectional antennas spaced half a wavelength apart. The transmitted symbols are randomly generated from $(\pm 1 \pm i)$. We consider one desired source and three uncorrelated interfering sources coming from 0^{o}, 20^{o}, 50^{o} and -35^{o}, respectively. The total interference-to-noise ratio (INR) is 60 dB, with 20 dB per interference. The SNR is 0 dB and the step size for $\mathbf{w}_{\text{a}}$ is 1×10^{-5}, unless specified otherwise. The step size for w_{b} is fixed at 0.01. In all figures, at least 500 simulation runs are averaged to obtain each simulated result.

2.6.1 Exactly Known Desired Signal's DOA

In this set of simulations, only the point distortionless constraint is considered. The flat fading channel coefficients for those signal vectors are independently and randomly generated, and remain unchanged in each simulation run. For comparison purpose, all these coefficients are scaled to let the average INR and SNR stick to the requirements. Since the desired signal's DOA is exactly known (no mismatch), no training is needed for the DFGSC. First, we use the same step size for both the conventional adaptive GSC and adaptive DFGSC and observe their convergence and steady-state SINR. Fig. 2.2 shows the learning curves of both algorithms. Also shown is the optimum SINR value calculated with the Wiener solution. From Fig. 2.2, we see that the adaptive DFGSC can achieve higher SINR than the conventional adaptive GSC, and both algorithms are comparable in convergence rate. As expected, the DFGSC can approach the optimum SINR much closely, which means that the effect of the excess MSE induced by the LMS algorithm is small. Fig. 2.3 reveals the beam patterns of both adaptive schemes after 200 snapshots. It is clear that the adaptive DFGSC performs better than the adaptive GSC in nulling the interference. The difference between the two schemes is almost 10 dB for each interfering source.

We then fix a target SINR, i.e., 11 dB, and choose suitable step sizes for both schemes ($\mu_a = 3 \times 10^{-6}$ and $\mu_a = 4 \times 10^{-5}$ for the conventional adaptive GSC and the adaptive DFGSC, respectively). It is to compare the convergence rate of both algorithms. Fig. 2.4 demonstrates the results. As we can see, the DFGSC converges around 150 snapshots while the GSC converges around 350 snapshots. The DFGSC converges much faster.

Next, we present the steady-state SINR achievable by the conventional adaptive GSC and the adaptive DFGSC under different SNR environments. Fig. 2.5 shows the results. We see that the achievable SINR for the DFGSC is proportional to the SNR and that for the GSC is saturated when the SNR is high. The performance gap becomes significant in high SNR regions.

Afterwards, we show the SINR performance with different step sizes used for $\mathbf{w}_a$ in the LMS algorithm. Fig. 2.6 gives the simulation results. It is clear that the larger the step size, the lower the SINR performance. However, we notice that the SINR degradation due

to large step sizes in the adaptive DFGSC is much smaller than that in the conventional adaptive GSC. For this reason, the DFGSC can work with a large step size to achieve fast convergence, which is very useful in time-varying environments. From all these figures described above, we can conclude that when the LMS algorithm is used for adaptation, the DFGSC can achieve higher SINR for the same convergence rate or faster convergence rate for the same SINR. We also see that our theoretical SINR analysis is quite accurate.

2.6.2 Desired Signal's DOA Mismatch

In this part, a scenario with a DOA mismatch is considered. We assume that there is a 2^o difference between the estimated and actual desired signal's DOA. Here, the step size for $\mathbf{w}_a$ is chosen to be 5×10^{-5} for fast convergence. Fig. 2.7 shows the learning curves for the conventional adaptive GSC with different constraint settings. From the figure, with the point constraint only, the SINR will eventually degrade to around -8 dB. It means that the desired signal is almost entirely cancelled out by the interference cancelling filter. Again, in the same figure, we see that with the additional first-order derivative constraint, the conventional GSC exhibits some robustness against the mismatch, but the signal cancellation still occurs. The steady-state SINR for this case is about 2 dB.

We repeat the same experiment described previously with the adaptive DFGSC and show the results in Fig. 2.8. Here, the first 50 snapshots are used as training for the scenario with the point constraint only and no training is required for the scenario with the point and first-order derivative constraints. We see that these schemes with the two different constraint settings achieve SINR about 9 dB and 7.5 dB, respectively. No signal cancellation seems to occur. Adding the additional constraint lowers the SINR for the DFGSC because DOFs are reduced. We conclude that the proposed algorithm can provide notable robustness for the GSC structure. From Fig. 2.7 and Fig. 2.8, we also notice that the simulated curves match the analytic ones well.

From the experiment in Fig. 2.8, we observe that for the 2^o DOA mismatch, the point and first-order derivative constraints are enough for the adaptive DFGSC to bypass the training period. However, the output SINR lessens. To keep the blind nature for the DFGSC structure and achieve higher SINR, we may adopt the approach described in Sec-

tion 2.4.3. Fig. 2.9 shows the learning curve with the start-up approach for our mismatch example. The DFGSC works with the point and first-order derivative constraints at the beginning. After 100 snapshots, it smoothly switches to use the point constraint only. The duration of transition is set to 80 snapshots (about $5\tau_{\text{av}}$). We see that the DFGSC converges after about 200 snapshots without using any training snapshots.

Finally, we compare the performance for the DFGSC with some well-known robust adaptive beamforming schemes against the DOA mismatch. The step size for $\mathbf{w}_{\text{a}}$ is chosen to be 1×10^{-5} here. The loaded sample matrix inversion (LSMI) algorithm, WCPO algorithm, and eigenspace-based beamformer are chosen for comparison. The number of training snapshots required by these robust methods is set to 100, which is large enough for providing good performance [43]. The diagonal loading factor for the LSMI is taken to be $10\sigma_n^2$, where σ_n^2 is the noise power in a single antenna element. The parameter ε in the WCPO algorithm is selected to provide nearly the optimum performance [43]. Fig. 2.10 shows the simulation results. From the figure, we observe that the performance of the DFGSC is almost the same as the WCPO and LSMI algorithms in low SNR regions, but the DFGSC outperforms all other algorithms in middle to high SNR regions. Based on the results, we conclude that the DFGSC is the only robust algorithm whose performance is consistently close to the optimum from low to high SNR regions.

2.7 Summary

In this chapter, a new LMS-based adaptive DFGSC has been proposed. The DFGSC introduces a decision-directed equalizer and a feedback filter in the GSC structure. We theoretically show that the optimum interference cancelling filters for both the DFGSC and conventional GSC are the same with perfectly known DOA. Robustness analysis for the two schemes with a DOA mismatch is also given. We derive the Wiener solution for the interference cancelling filter when the DFGSC is in mismatch and show that the signal cancellation phenomenon can be avoided. On the other hand, when the optimum weights are estimated by the LMS algorithm, the DFGSC gives significantly better results than that of the conventional GSC. We have examined the convergence behavior of the conventional adaptive GSC and the proposed adaptive DFGSC under perfectly matched

and mismatched DOA scenarios. Simulation results verify that the DFGSC can achieve higher SINR value for the same convergence rate or faster convergence for the same SINR, and the DFGSC can keep the high SINR performance even in mismatch.

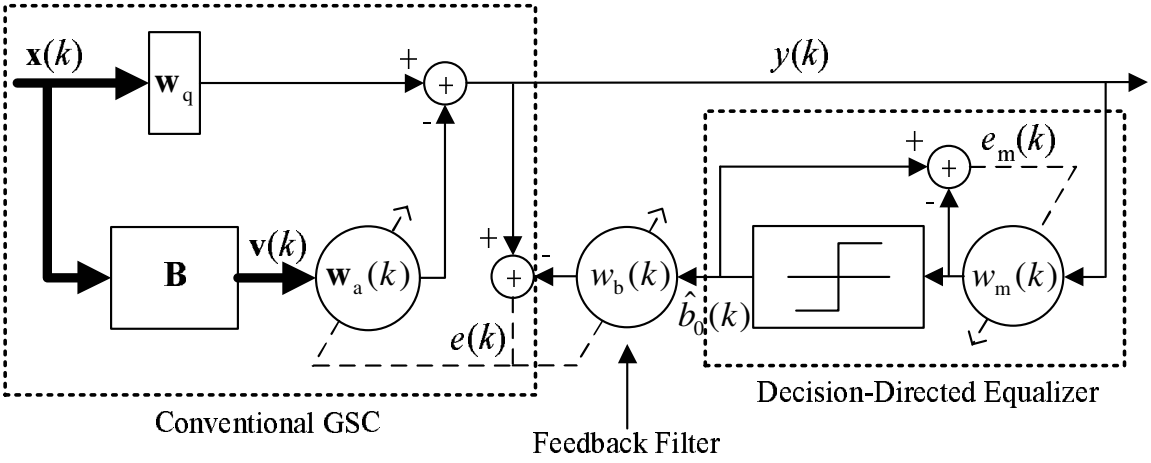

Figure 2.1: Block diagram of adaptive DFGSC.

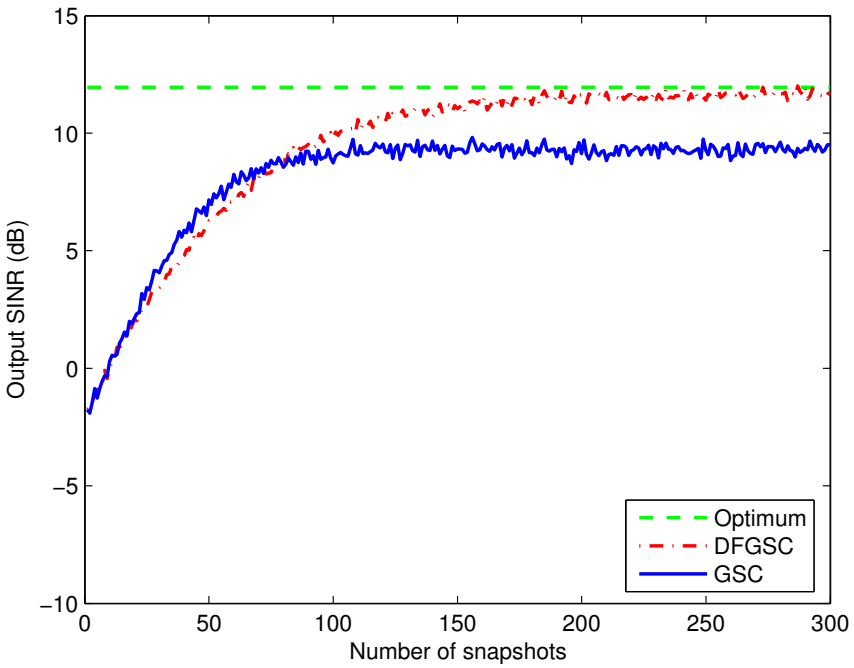

Figure 2.2: Learning curves for GSC and DFGSC with same step size.

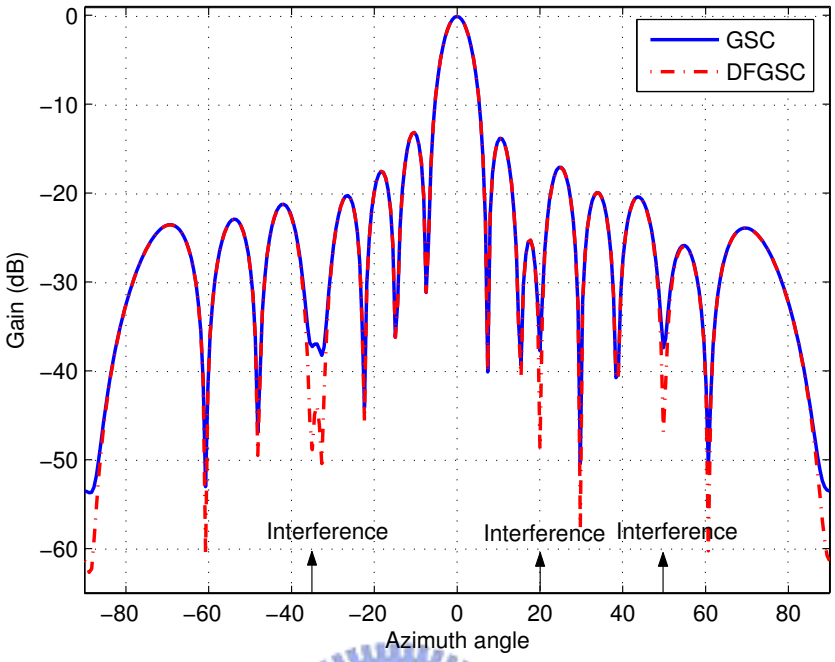

Figure 2.3: Beam patterns of GSC and DFGSC in Fig. 2.2 after 200 snapshots.

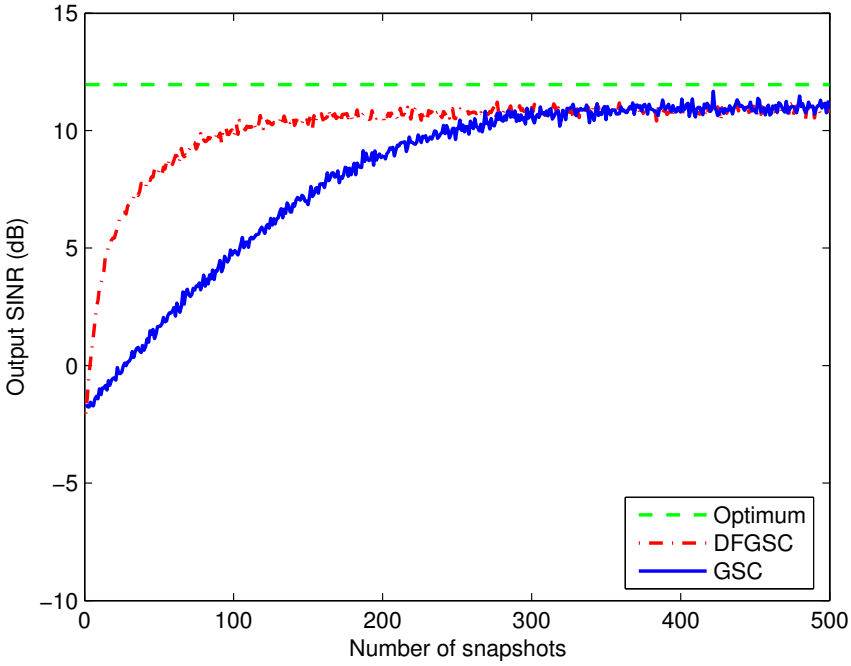

Figure 2.4: Learning curves for GSC and DFGSC with same SINR target.

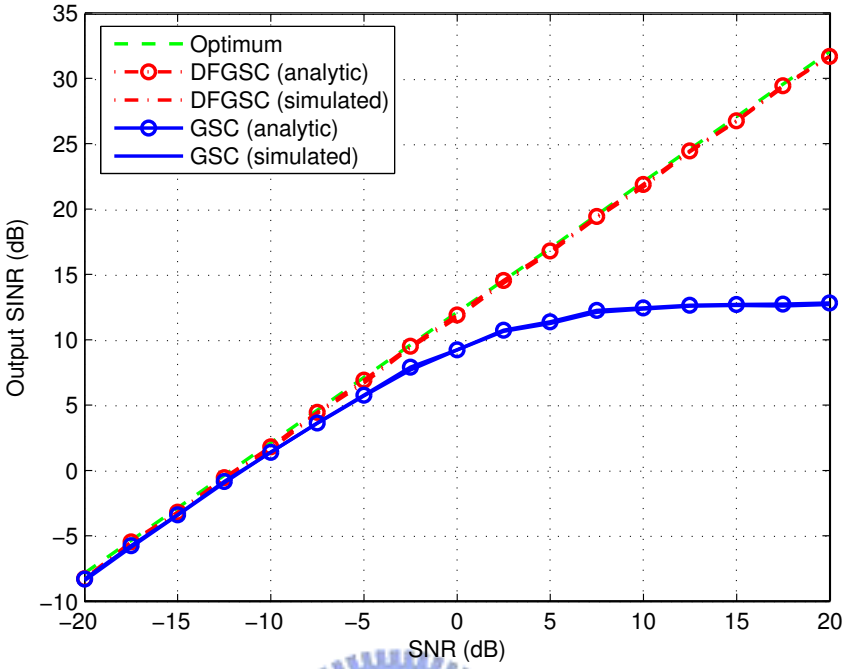

Figure 2.5: Steady-state SINR performance in different SNR environments.

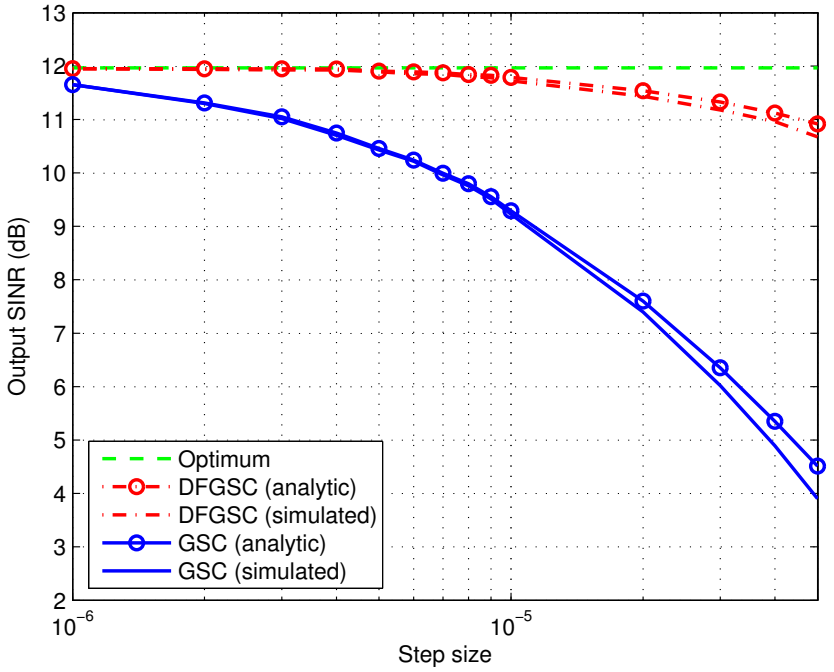

Figure 2.6: Steady-state SINR performance with different step sizes.

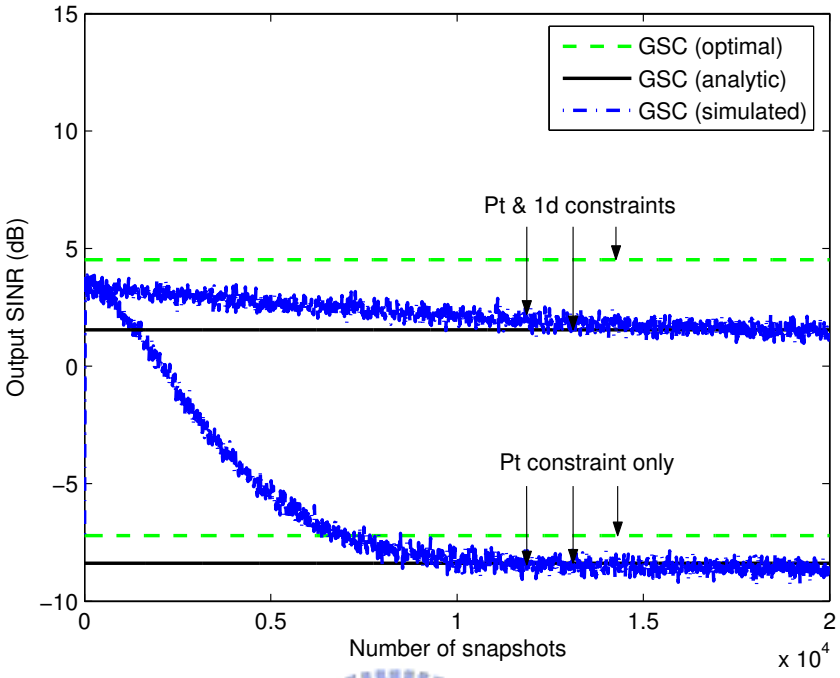

Figure 2.7: Learning curves for GSC with DOA mismatch utilizing point (pt) constraint only and point and first-order derivative (pt & 1d) constraints.

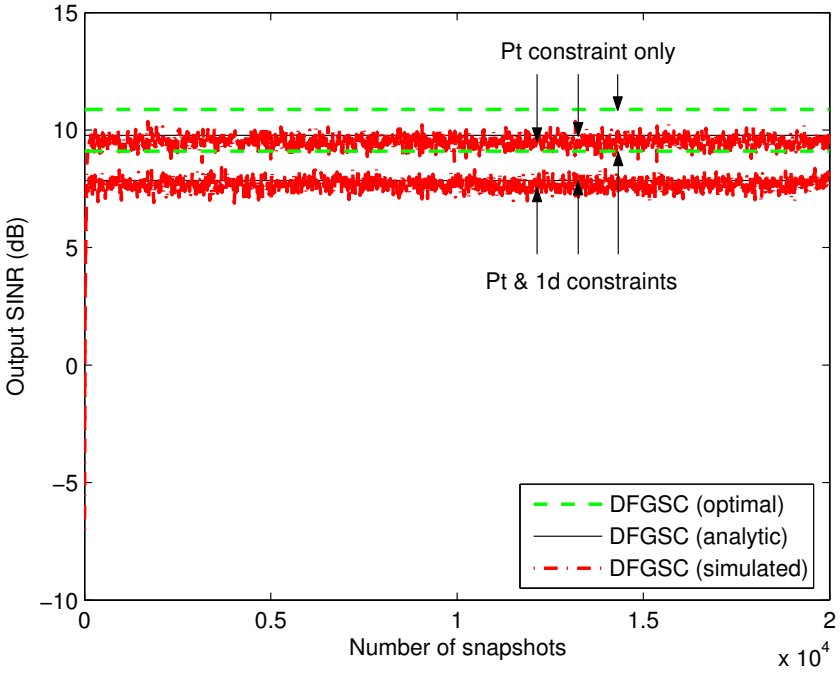

Figure 2.8: Learning curves for DFGSC with DOA mismatch utilizing point (pt) constraint only and point and first-order derivative (pt & 1d) constraints.

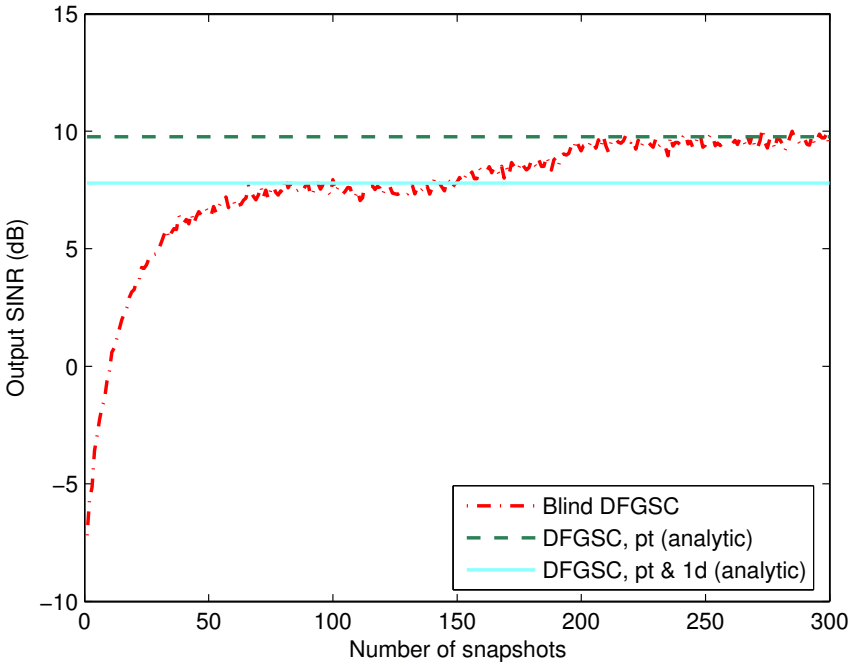

Figure 2.9: Learning curves for DFGSC with transition from point and first-order derivative (pt & 1d) constraints to point (pt) constraint only.

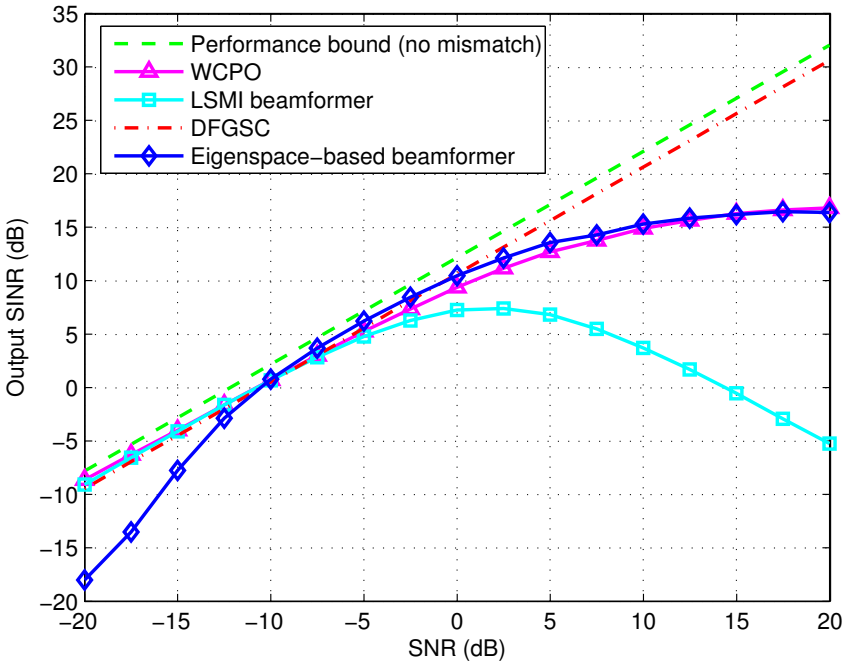

Figure 2.10: Steady-state SINR performance for DFGSC and other robust adaptive beamforming approaches in different SNR environments.

Chapter 3

Adaptive Decision Feedback Space-Time Equalization with Generalized Sidelobe Cancellation

3.1 Introduction

In wireless communications, CCI and ISI often cause severe signal distortion. STE is commonly used to tackle this problem. STE consists of an antenna array, combined with signal processing for both space and time domains. As stated in Chapter 1, there exists various kind of STEs. Nevertheless, all STEs mentioned previously are training-based. In other words, we have to transmit training sequences before actually using them. It is well-known that transmission of training sequences will reduce the bandwidth utilization efficiency and may not be always possible. In some applications, however, spatial information corresponding to a desired user is available. For example, a receiver may perform DOA estimation before signal detection. An SDMA system employed to increase system capacity [51]-[54] transmits or receives signal from a certain direction only. In these cases, it is possible to utilize the *a priori* spatial information and avoid the requirement of training sequences. A straightforward idea is to use a GSC for CCI suppression and a blind equalizer for ISI compensation. Unfortunately, the result of this direct cascade is often far from satisfactory. The reasons are stated below.

For computational complexity consideration, the GSC is often implemented with an adaptive structure. As stated in Chapter 2, the adaptive GSC with the LMS algorithm usually converges slowly. The steady-state error signal is large even in the noiseless envi-

ronment. In addition to this, the adaptive GSC is sensitive to constraint mismatch, which is caused by incorrect spatial information. In typical applications, constraint mismatch can easily arise due to multipath channels and spatial signature errors.

It is well-known that the DFE can have much better performance than the linear equalizer in severe fading channels. This statement is also true for blind equalization. However, the blind DFE is difficult to derive and it is rarely reported in the literature. The major contribution in this field is from Labat *et al.*, who proposed an interesting blind DFE in [55]. They used an infinite impulse response (IIR) whitening filter cascaded with a blind finite impulse response (FIR) linear equalizer in the start-up period. After convergence, it switches the cascading order of the IIR and FIR filters yielding a decision feedback structure. At the same time, a decision-directed MMSE training is initiated. For easy reference, we call this blind DFE as the LBDFE hereafter. One inherent problem associated with the DFE is its error propagation effect and this will have even more impact in its adaptive implementation. Since the LBDFE uses the decision-directed training, it is sensitive to error propagation.

In this chapter, we propose an adaptive STE for systems with *a priori* spatial information. The proposed structure comprises a DFGSC, a blind DFE, and a channel estimator. The adaptive DFGSC structure can eliminate the desired signal component from the error signal in the LMS algorithm. As a consequence, it not only improves CCI suppression, but also allows the simple point distortionless constraint robust to multipath channel environments and spatial signature errors. The proposed blind DFE adopts a channel-aided structure yielding better ISI compensation and higher resistance to error propagation. We will demonstrate that the proposed blind DFE performs better than the LBDFE.

The chapter is organized as follows. In Section 3.2, the space-time signal model for the STE is described. The effect of both CCI and ISI to the desired signal is also explained. In Section 3.3, we describe a straightforward approach of a conventional GSC and an LBDFE in a hybrid manner. In Section 3.4, we propose the new adaptive STE and describe the corresponding operation mechanisms in detail. Section 3.5 shows that the proposed STE is robust to general space-time multipath channels and spatial signature errors. Finally,

simulation results and summary are given in Section 3.6 and 3.7, respectively.

3.2 Space-Time Signal Model

Consider a ULA of N antenna elements at the receiver. Let the $N \times 1$ received continuous-time equivalent baseband signal vector denoted by

$$\mathbf{x}(t) = [x_0(t) \;\; x_1(t) \;\; \cdots \;\; x_{N-1}(t)]^T \tag{3.1}$$

where $x_n(t)$ $(0 \leq n \leq N-1)$ is the signal received from the nth antenna element at time t. Define an $N \times 1$ vector $\mathbf{a}(\theta)$ as the spatial signature for the signal from the DOA θ. It is written as $\mathbf{a}(\theta) = [1 \;\; e^{i\zeta_\theta} \;\; e^{i2\zeta_\theta} \;\; \cdots \;\; e^{i(N-1)\zeta_\theta}]^T$, where $i = \sqrt{-1}$ and $\zeta_\theta = (2\pi d/L_\lambda)\sin\theta$, in which d is the element spacing and L_λ is the signal wavelength. Assume that there are M sources (including both desired and interfering sources) coming from M different and distinguishable directions. The transmitted signal waveform of the mth source $s_m(t)$ $(0 \leq m \leq M-1)$ can be written as

$$s_m(t) = \sum_{k=-\infty}^{+\infty} p(t - kT) b_m(k) \tag{3.2}$$

where $b_m(k)$ is the kth information symbol of the mth source, $p(t)$ is the pulse shape of the transmitted symbol, and T is the symbol duration. Let L_m be the number of propagation paths for the mth source, θ_m be the DOA for the mth source, τ_{ml} be the delay time for the lth path $(0 \leq l \leq L_m - 1)$ of the mth source, and $\bar{h}_{ml}$ be the channel coefficient for the lth path of the mth source, respectively. Assume that the channel parameters for different sources are independent and remain constant over the observation period. The received signal vector in (3.1) can then be expressed as

$$\mathbf{x}(t) = \sum_{m=0}^{M-1} \sum_{l=0}^{L_m-1} \bar{h}_{ml} s_m(t - \tau_{ml}) \mathbf{a}(\theta_m) + \mathbf{n}(t) \tag{3.3}$$

where $\mathbf{n}(t)$ is an $N \times 1$ additive white Gaussian noise (AWGN) vector. Each component in the AWGN vector is assumed to be spatially white with a variance of σ_n^2. Substituting (3.2) into (3.3), we have

$$\mathbf{x}(t) = \sum_{m=0}^{M-1} \sum_{k=-\infty}^{+\infty} \mathbf{h}_m(t - kT) b_m(k) + \mathbf{n}(t) \tag{3.4}$$

where $\mathbf{h}_m(t)$ is an $N \times 1$ vector summarizing the total transmission effect of the mth source to the information symbol, and it is given by

$$\mathbf{h}_m(t) = \sum_{l=0}^{L_m-1} \bar{h}_{ml} p(t - \tau_{ml}) \mathbf{a}(\theta_m). \tag{3.5}$$

Without loss of generality, we assume that the first M_0 sources ($M_0 \leq M$) are related to the desired user, i.e., $b_0(k) = b_1(k) = \cdots = b_{M_0-1}(k)$. The DOA θ_0 is the DOA of the desired signal with the strongest signal strength, and θ_1 to θ_{M_0-1} correspond to the DOAs of the desired signal's ISI. Here, we call the ISI comes from θ_0 in different time instants as temporal ISI and the ISI comes from θ_1 to θ_{M_0-1} in different time instants as temporal and spatial ISI. The received equivalent baseband signal is sampled at the symbol rate, i.e., $t = kT$. We assume that the sampling clock is synchronized with the transmission clock. After sampling, the channel effect for all sources is of a finite duration within $[0, (D_m - 1)T]$, where D_m is the channel order of the mth source. Here, we define $\mathbf{d}(k)$ as an $N \times 1$ vector representing the components from the main source of the desired signal ($m = 0$) and $\mathbf{i}(k)$ as an $N \times 1$ vector summing up the components from the temporal and spatial ISI sources ($m = 1, 2, \cdots, M_0 - 1$), and they are given by

$$\mathbf{d}(k) = \sum_{d=0}^{D_0-1} \mathbf{h}_0(d) b_0(k - d) \tag{3.6}$$

and

$$\mathbf{i}(k) = \sum_{m=1}^{M_0-1} \sum_{d=0}^{D_m-1} \mathbf{h}_m(d) b_0(k - d) \tag{3.7}$$

respectively. Also define an $N \times 1$ vector $\mathbf{z}(k)$ representing the uncorrelated CCI-plus-noise components as

$$\mathbf{z}(k) = \sum_{m=M_0}^{M-1} \sum_{d=0}^{D_m-1} \mathbf{h}_m(d) b_m(k - d) + \mathbf{n}(k). \tag{3.8}$$

Then, the expression in (3.4) can be rewritten as

$$\mathbf{x}(k) = \mathbf{d}(k) + \mathbf{i}(k) + \mathbf{z}(k). \tag{3.9}$$

The main task of the STE is to suppress CCI and ISI, and recover the transmitted information symbols. To simplify the notations, we write $b(k)$ instead of $b_0(k)$ for the desired user's information symbols in the following derivations.

3.3 Hybrid of GSC and LBDFE

In some wireless communication systems, such as SDMA applications, the main DOA (and hence the spatial signature) of the desired signal is known as *a priori* or can be estimated. A straightforward approach as mentioned above for CCI and ISI mitigation is the hybrid of a conventional GSC and an LBDFE. Using the spatial signature, the conventional GSC is to suppress CCI and the LBDFE is to compensate ISI. No extra training sequences are required for both processing. The operation and the weakness of this approach will be elaborated below.

3.3.1 GSC

The conventional GSC for CCI suppression is optimized with the LCMV criterion. The LCMV beamformer determines the N-tap weight vector $\mathbf{w}$ through

$$\min_{\mathbf{w}} \quad \mathbf{w}^H \mathbf{R}_{\mathbf{x}} \mathbf{w}, \qquad \text{subject to} \quad \mathbf{C}^H \mathbf{w} = \mathbf{f} \tag{3.10}$$

where $\mathbf{R}_{\mathbf{x}} = \mathrm{E}\{\mathbf{x}(k)\mathbf{x}^H(k)\}$ is the input correlation matrix, $\mathbf{C}$ is an $N \times U$ constraint matrix, and $\mathbf{f}$ is a $U \times 1$ response vector, with U being the number of constraints. The structure of GSC is shown in the left part of Fig. 3.1. Then we have an equivalent spatial filter as $\mathbf{w} = \mathbf{w}_{\mathrm{q}} - \mathbf{B}\mathbf{w}_{\mathrm{a}}$. Ideally, the span of $\mathbf{B}$ is in the null space of $\mathbf{C}^H$. Using the constraint in (3.10), $\mathbf{w}_{\mathrm{q}}$ can be readily found as $\mathbf{w}_{\mathrm{q}} = \mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{f}$, and $\mathbf{w}_{\mathrm{a}}$ is optimized according to the output power of the GSC as

$$J = \mathrm{E}\{|y_{\mathrm{s}}(k)|^2\} = \mathrm{E}\{|(\mathbf{w}_{\mathrm{q}} - \mathbf{B}\mathbf{w}_{\mathrm{a}})^H \mathbf{x}(k)|^2\}. \tag{3.11}$$

The constrained optimization problem in (3.10) can then be changed to

$$\min_{\mathbf{w}_{\mathrm{a}}} \quad J = \min_{\mathbf{w}_{\mathrm{a}}} \quad (\mathbf{w}_{\mathrm{q}} - \mathbf{B}\mathbf{w}_{\mathrm{a}})^H \mathbf{R}_{\mathbf{x}} (\mathbf{w}_{\mathrm{q}} - \mathbf{B}\mathbf{w}_{\mathrm{a}}). \tag{3.12}$$

The optimum $\mathbf{w}_{\mathrm{a}}$ is classically solved to be

$$\mathbf{w}_{\mathrm{a,opt}} = (\mathbf{B}^H \mathbf{R}_{\mathbf{x}} \mathbf{B})^{-1} \mathbf{B}^H \mathbf{R}_{\mathbf{x}} \mathbf{w}_{\mathrm{q}}. \tag{3.13}$$

With the optimum weight vector, the minimum value of J in (3.11), denoted as $J_{\min}$ (which is also the minimum output power, denoted as $P_{\mathrm{o,min}}$, of the conventional GSC),

can be calculated as

$$\begin{aligned} J_{\min} &= P_{\text{o,min}} \\ &= \mathbf{w}_{\text{opt}}^H \mathbf{R}_{\mathbf{x}} \mathbf{w}_{\text{opt}} \qquad (3.14) \\ &= \mathbf{w}_{\text{q}}^H \mathbf{R}_{\mathbf{x}} \mathbf{w}_{\text{opt}} \qquad (3.15) \end{aligned}$$

where we let $\mathbf{w}_{\text{opt}} = \mathbf{w}_{\text{q}} - \mathbf{B}\mathbf{w}_{\text{a,opt}}$. As stated in Chapter 2, we can use the LMS algorithm to approach the optimum weights iteratively. The LMS update equation for $\mathbf{w}_{\text{a}}$ can be written as

$$\mathbf{w}_{\text{a}}(k+1) = \mathbf{w}_{\text{a}}(k) + \mu_{\text{a}} \mathbf{v}(k) e_{\text{s}}^*(k) \qquad (3.16)$$

where $\mathbf{w}_{\text{a}}(k)$ is the estimate of $\mathbf{w}_{\text{a,opt}}$ at the kth iteration, $\mathbf{v}(k) = \mathbf{B}^H \mathbf{x}(k)$ is the filter input vector, μ_{a} is the step size controlling the convergence rate, and $e_{\text{s}}(k)$ is an error signal as $e_{\text{s}}(k) = y_{\text{s}}(k)$. When $\mathbf{w}_{\text{a}}$ is optimized, the error signal $e_{\text{s}}(k)$ will chiefly include the component from the desired signal, i.e., the desired user's transmitted symbols $b(k)$. When the LMS algorithm is applied to estimate $\mathbf{w}_{\text{a,opt}}$, the performance will be affected due to the large error signal used. The other problem with the conventional GSC is its sensitivity to constraint mismatch. Whenever the setting of the constraint matrix $\mathbf{C}$ in (3.10) (or the blocking matrix $\mathbf{B}$ in (3.12)) is not fit for the actual spatial signature of the desired signal, constraint mismatch occurs. Constraint mismatch can easily arise due to multipath channels and spatial signature errors. This seriously degrades the performance of the conventional adaptive GSC.

3.3.2 LBDFE

The output of the GSC is fed into the LBDFE for equalization. The equalizer adaptation process is divided into two periods. In the start-up period, the received signal is prewhitened by an IIR filter and then equalized by a blind FIR linear equalizer. The structure of the LBDFE in the start-up period is shown in the right part of Fig. 3.1. In [55], the constant modulus algorithm (CMA) was used as the blind algorithm for equalization. Recently, a sophisticated blind equalization algorithm called the multimodulus algorithm (MMA) was proposed [56]. Analysis shows that the MMA can provide much stable performance especially with high-order constellation modulation [57]. The cost

function can yield an equalized constellation rotated with a multiple of 90°, eliminating the need for additional constellation phase recovery as that in the CMA. The remaining phase ambiguity problem is classically solved by differential encoding. For these reasons, we use the MMA as our blind equalization algorithm (instead of the CMA) throughout this chapter. For the LBDFE, we denote the length of $\mathbf{w}_{\mathrm{f}}$ and $\mathbf{w}_{\mathrm{b}}$ as α and β, respectively. In the start-up period, let $y_{\mathrm{s}}(k)$ be the output of the GSC and $y_{\mathrm{s}}^{'}(k)$ be the difference between $y_{\mathrm{s}}(k)$ and the prewhitening filter output. Also let $\mathbf{y}^{'}(k)$ be the input vector of $\mathbf{w}_{\mathrm{b}}$ and $\mathbf{y}^{'}(k) = [y_{\mathrm{s}}^{'}(k-1) \;\; y_{\mathrm{s}}^{'}(k-2) \;\; \cdots \;\; y_{\mathrm{s}}^{'}(k-\beta)]^T$. So, we have $y_{\mathrm{s}}^{'}(k) = y_{\mathrm{s}}(k) - \mathbf{w}_{\mathrm{b}}^H \mathbf{y}^{'}(k)$. The filter $\mathbf{w}_{\mathrm{b}}$ is optimized according to the criterion: $\min_{\mathbf{w}_{\mathrm{b}}} \mathrm{E}\{|y_{\mathrm{s}}^{'}(k)|^2\}$. As shown in Fig. 3.1, $y_{\mathrm{s}}^{'}(k)$ serves as the input of the blind FIR filter $\mathbf{w}_{\mathrm{f}}$ as well. We define another vector as $\mathbf{y}^{''}(k) = [y_{\mathrm{s}}^{'}(k) \;\; y_{\mathrm{s}}^{'}(k-1) \;\; \cdots \;\; y_{\mathrm{s}}^{'}(k-\alpha+1)]^T$, and, as a consequence, the output of the LBDFE in the start-up period is $y_{\mathrm{t}}(k) = \mathbf{w}_{\mathrm{f}}^H \mathbf{y}^{''}(k)$. The cost function with the MMA for the optimization of $\mathbf{w}_{\mathrm{f}}$ is written as

$$J_{\mathrm{MMA}} = \mathrm{E}\{e_{\mathrm{t}}(k)\} = \mathrm{E}\{(y_{\mathrm{t,r}}^2(k) - R^2)^2 + (y_{\mathrm{t,i}}^2(k) - R^2)^2\} \tag{3.17}$$

with

$$R^2 = \frac{\mathrm{E}\{b_{\mathrm{r}}^4(k)\}}{\mathrm{E}\{b_{\mathrm{r}}^2(k)\}} = \frac{\mathrm{E}\{b_{\mathrm{i}}^4(k)\}}{\mathrm{E}\{b_{\mathrm{i}}^2(k)\}}. \tag{3.18}$$

Let μ_{f} be the step size controlling the convergence of $\mathbf{w}_{\mathrm{f}}$. From [56], the stochastic gradient algorithm (SGA) for the MMA of $\mathbf{w}_{\mathrm{f}}$ can be written as

$$\mathbf{w}_{\mathrm{f}}(k+1) = \mathbf{w}_{\mathrm{f}}(k) + \mu_{\mathrm{f}} \phi(y_{\mathrm{t}}(k)) \mathbf{y}^{''}(k) \tag{3.19}$$

with

$$\phi(y_{\mathrm{t}}(k)) = y_{\mathrm{t,r}}^3(k) + i y_{\mathrm{t,i}}^3(k) - R^2 y_{\mathrm{t}}(k). \tag{3.20}$$

The start-up period is expected to sufficiently open the eye pattern such that the error rate is low enough to initiate the second (tracking) period. In the tracking period, the cascading order of the IIR prewhitening filter and the FIR linear equalizer is swapped and this turns the whole system into a conventional DFE structure. A decision-directed MMSE tracking operation, similar to that of the conventional DFE, is then activated. This approach may avoid the possible error propagation phenomenon initially and give a smooth transition strategy between the blind and decision-directed equalization. Nevertheless, the stability

of the adaptive IIR filter makes the performance sensitive to the parameters chosen. Also, error propagation may occur during the decision-directed mode. The behavior of the LBDFE becomes not easy to control. We will show empirically in Section 3.6 that the performance of the hybrid of the conventional GSC and the LBDFE is often far from satisfactory in severe channel environments.

3.4 Proposed Hybrid STE

In this section, we propose a new adaptive STE being a hybrid of a DFGSC and a channel-aided blind DFE (CBDFE). Fig. 3.2 is the block diagram of the proposed STE. Note that in Fig. 3.2 a channel estimator is included. Let the coefficients of the channel estimator be denoted as $\mathbf{w}_{\mathrm{h}}$. It is used to model the equivalent temporal channel for the desired signal and its operation will be explained soon. In [15], a channel estimator was also used in a training-based STE such that the corresponding beamformer can achieve better performance. In the STE scheme proposed here, the role of the channel estimator is much more involved. For spatial processing, with the help of the channel estimator, we can formulate the adaptive DFGSC achieving better CCI suppression performance. Besides, the DFGSC can have extra robustness against constraint mismatch. For temporal processing, with the help of the channel estimator, we can formulate the CBDFE such that it is more stable in ISI compensation. This CBDFE is different from those channel-estimation-based DFEs proposed in [58] and [59], where both the feedforward and feedback filters are calculated based on the estimated channel response. In the CBDFE, however, the adaptive structure is remained. The operation of the DFGSC and the CBDFE is presented separately as follows.

3.4.1 DFGSC

Let the channel estimator $\mathbf{w}_{\mathrm{h}}$ have a dimension of $\gamma \times 1$. Here, the value of γ is chosen to be equal to or larger than the maximum value of channel order D_m $(0 \leq m \leq M_0 - 1)$. We first define $\hat{\mathbf{b}}(k)$ as the input vector of $\mathbf{w}_{\mathrm{h}}$, i.e., $\hat{\mathbf{b}}(k) = [\hat{b}(k)\ \ \hat{b}(k-1)\ \cdots\ \hat{b}(k-\gamma+1)]^T$. To optimize the interference cancelling filter $\mathbf{w}_{\mathrm{a}}$ and the channel estimator $\mathbf{w}_{\mathrm{h}}$, we propose a

new cost function as

$$\begin{aligned} J &= \mathrm{E}\{|e_{\mathrm{s}}(k)|^2\} \\ &= \mathrm{E}\{|y_{\mathrm{s}}(k) - \mathbf{w}_{\mathrm{h}}^H \hat{\mathbf{b}}(k)|^2\} \\ &= \mathrm{E}\{|\mathbf{w}_{\mathrm{q}}^H \mathbf{x}(k) - [\mathbf{w}_{\mathrm{a}}^H \quad \mathbf{w}_{\mathrm{h}}^H] \begin{bmatrix} \mathbf{B}^H \mathbf{x}(k) \\ \hat{\mathbf{b}}(k) \end{bmatrix} |^2\}. \end{aligned} \tag{3.21}$$

To understand the operation mechanisms, we first consider a simplified scenario that only temporal ISI is present. In other words, the desired signal and its ISI only come from the main DOA, i.e., $M_0 = 1$ (so $\mathbf{i}(k)$ in (3.7) is a zero vector). We assume that the spatial signature for the desired signal is exactly known and a distortionless constraint is set towards it. In this case, the desired signal will be completely blocked in the lower path of the DFGSC. The general case that temporal and spatial ISI exists and the distortionless constraint may be set improperly will be discussed in the next section. Assuming that the decision is correct, i.e., $\hat{b}(k) = b(k)$, we can rewrite the minimization of the cost function in (3.21) as

$$\begin{aligned} \min_{\mathbf{w}_{\mathrm{a}},\, \mathbf{w}_{\mathrm{h}}} \quad J = \min_{\mathbf{w_c}} \quad & \mathbf{w}_{\mathrm{q}}^H \mathbf{R}_{\mathbf{x}} \mathbf{w}_{\mathrm{q}} - \mathbf{w}_{\mathrm{q}}^H \begin{bmatrix} \mathbf{R}_{\mathbf{x}} \mathbf{B} & \mathbf{R}_{\mathbf{p}} \end{bmatrix} \mathbf{w_c} - \\ & - \mathbf{w_c}^H \begin{bmatrix} \mathbf{B}^H \mathbf{R}_{\mathbf{x}} \\ \mathbf{R}_{\mathbf{p}}^H \end{bmatrix} \mathbf{w}_{\mathrm{q}} + \mathbf{w_c}^H \mathbf{R_c} \mathbf{w_c} \end{aligned} \tag{3.22}$$

where we let

$$\mathbf{w_c} = \begin{bmatrix} \mathbf{w}_{\mathrm{a}} \\ \mathbf{w}_{\mathrm{h}} \end{bmatrix} \tag{3.23}$$

$$\mathbf{R_c} = \mathrm{E}\{ \begin{bmatrix} \mathbf{B}^H \mathbf{x}(k) \\ \hat{\mathbf{b}}(k) \end{bmatrix} \begin{bmatrix} \mathbf{x}^H(k)\mathbf{B} & \hat{\mathbf{b}}^H(k) \end{bmatrix} \} = \begin{bmatrix} \mathbf{B}^H \mathbf{R}_{\mathbf{x}} \mathbf{B} & \mathbf{0} \\ \mathbf{0}^H & \sigma_b^2 \mathbf{I}_\gamma \end{bmatrix} \tag{3.24}$$

and

$$\mathbf{R}_{\mathbf{p}} = \mathrm{E}\{\mathbf{x}(k)\hat{\mathbf{b}}^H(k)\} = \sigma_b^2 [\mathbf{h}_0(0)\ \mathbf{h}_0(1)\ \cdots\ \mathbf{h}_0(\gamma - 1)] \tag{3.25}$$

in which $\mathbf{0}$ denotes a zero matrix with dimension $(N-1) \times \gamma$, $\mathbf{I}_\gamma$ denotes an identity matrix with dimension $\gamma \times \gamma$, and σ_b^2 denotes the power of transmitted symbols. We can set $\mathbf{w}_{\mathrm{q}}$ the same as that in the conventional GSC. Following the same procedure described in Chapter 2, we can obtain the optimum $\mathbf{w_c}$ as

$$\mathbf{w}_{\mathbf{c},\mathrm{opt}} = \mathbf{R_c}^{-1} \begin{bmatrix} \mathbf{B}^H \mathbf{R}_{\mathbf{x}} \\ \mathbf{R}_{\mathbf{p}}^H \end{bmatrix} \mathbf{w}_{\mathrm{q}}. \tag{3.26}$$

With the special structure of $\mathbf{R_c}$, we can decompose $\mathbf{w}_{\mathbf{c},\text{opt}}$ back into the two weights

$$\mathbf{w}_{\text{a,opt}} = (\mathbf{B}^H \mathbf{R}_\mathbf{x} \mathbf{B})^{-1} \mathbf{B}^H \mathbf{R}_\mathbf{x} \mathbf{w}_\text{q} \tag{3.27}$$

$$\mathbf{w}_{\text{h,opt}} = \frac{\mathbf{R}_\mathbf{p}^H}{\sigma_b^2} \mathbf{w}_\text{q} = [\mathbf{h}_0(0)\ \mathbf{h}_0(1)\ \cdots\ \mathbf{h}_0(\gamma-1)]^H \mathbf{w}_\text{q}. \tag{3.28}$$

From (3.27), we observe that the expression of $\mathbf{w}_{\text{a,opt}}$ is the same as that of the conventional GSC given in (3.13). Since $\mathbf{w}_\text{q}$ and $\mathbf{w}_{\text{a,opt}}$ remain the same, the minimum output power $P_{\text{o,min}}$ for the DFGSC is also the same. From (3.28), we observe that $\mathbf{w}_{\text{h,opt}}$ corresponds to the equivalent channel effect from the transmitter to the DFGSC output. With $\mathbf{w}_{\text{a,opt}}$ and $\mathbf{w}_{\text{h,opt}}$ given above, the minimum J in (3.21) for the DFGSC becomes

$$\begin{aligned} J_\text{min} &= \text{E}\{|\mathbf{w}_\text{opt}^H \mathbf{x}(k) - \mathbf{w}_{\text{h,opt}}^H \hat{\mathbf{b}}(k)|^2\} \\ &= \mathbf{w}_\text{q}^H \mathbf{R}_\mathbf{x} \mathbf{w}_\text{opt} - \sigma_b^2 |\mathbf{w}_\text{q}^H [\mathbf{h}_0(0)\ \mathbf{h}_0(1)\ \cdots\ \mathbf{h}_0(\gamma-1)]|^2. \end{aligned} \tag{3.29}$$

We let $\mathbf{R_d} = \text{E}\{\mathbf{d}(k)\mathbf{d}^H(k)\}$ be the input correlation matrix of the desired signal excluding temporal and spatial ISI, and $\mathbf{R_z} = \text{E}\{\mathbf{z}(k)\mathbf{z}^H(k)\}$ be the input correlation matrix of CCI-plus-noise. The first term in (3.29) equals the minimum J of the conventional GSC in (3.15) and the second term in (3.29) is the power of the desired signal in the DFGSC output, which can be defined as

$$\begin{aligned} P_\mathbf{d} &\triangleq \mathbf{w}_\text{opt}^H \mathbf{R_d} \mathbf{w}_\text{opt} \\ &= \sigma_b^2 |\mathbf{w}_\text{q}^H [\mathbf{h}_0(0)\ \mathbf{h}_0(1)\ \cdots\ \mathbf{h}_0(\gamma-1)]|^2. \end{aligned} \tag{3.30}$$

So, J_min in (3.29) can be written as

$$\begin{aligned} J_\text{min} &= \mathbf{w}_\text{q}^H \mathbf{R}_\mathbf{x} \mathbf{w}_\text{opt} - P_\mathbf{d} \\ &= \mathbf{w}_\text{q}^H \mathbf{R_z} \mathbf{w}_\text{opt}. \end{aligned} \tag{3.31}$$

Comparing (3.15) and (3.31), we can see that the desired signal is totally excluded from the above expression with $\mathbf{w}_{\text{a,opt}}$ and $\mathbf{w}_{\text{h,opt}}$. When an adaptive algorithm such as the LMS is used to estimate $\mathbf{w}_{\text{a,opt}}$, the performance of the GSC can be greatly improved by the decision feedback operation. The LMS update equations for the DFGSC can be written as

$$\mathbf{w}_\text{a}(k+1) = \mathbf{w}_\text{a}(k) + \mu_\text{a} \mathbf{v}(k) e_\text{s}^*(k) \tag{3.32}$$

$$\mathbf{w}_\text{h}(k+1) = \mathbf{w}_\text{h}(k) + \mu_\text{h} \hat{\mathbf{b}}(k) e_\text{s}^*(k) \tag{3.33}$$

where $\mu_{\rm a}$ is the step size for $\mathbf{w}_{\rm a}$, $\mu_{\rm h}$ is the step size for $\mathbf{w}_{\rm h}$, $\mathbf{v}(k)$ is the output from the blocking matrix, and $e_{\rm s}(k) = y_{\rm s}(k) - \mathbf{w}_{\rm h}^H(k)\hat{\mathbf{b}}(k)$. Again, unlike the conventional GSC, the steady-state $e_{\rm s}(k)$ will exclude the desired signal and hence can be quite small.

To analyze the steady-state performance for both the conventional adaptive GSC and the proposed adaptive DFGSC, we denote the value of J (both (3.11) and (3.21)) in the steady state as $J(\infty)$. Suppose that the decision is correct and $\mathbf{w}_{\rm h}$ is fixed at optimum. Then

$$J(\infty) = J_{\rm min} + J_{\rm ex}(\infty) \tag{3.34}$$

with

$$J_{\rm ex}(\infty) = J_{\rm min} \sum_{l=1}^{N-1} \frac{\mu_{\rm a}\lambda_l(\mathbf{B}^H\mathbf{R}_{\mathbf{x}}\mathbf{B})}{2 - \mu_{\rm a}\lambda_l(\mathbf{B}^H\mathbf{R}_{\mathbf{x}}\mathbf{B})} \tag{3.35}$$

where $\lambda_l(\mathbf{B}^H\mathbf{R}_{\mathbf{x}}\mathbf{B})$ represents the lth eigenvalue of $\mathbf{B}^H\mathbf{R}_{\mathbf{x}}\mathbf{B}$. If $\mu_{\rm a}$ is small, (3.35) can be approximated as

$$J_{\rm ex}(\infty) \approx \frac{\mu_{\rm a} J_{\rm min}}{2} \sum_{l=1}^{N-1} \lambda_l(\mathbf{B}^H\mathbf{R}_{\mathbf{x}}\mathbf{B}). \tag{3.36}$$

Thus, $J_{\rm ex}(\infty)$ is proportional to $J_{\rm min}$ and the step size $\mu_{\rm a}$. With reference to Chapter 2, the optimum and steady-state output SINR (with the LMS algorithm) can then be calculated as

$$\text{SINR}_{\rm opt} = \frac{P_{\mathbf{d}}}{P_{\rm o,min} - P_{\mathbf{d}}} \tag{3.37}$$

and

$$\text{SINR}_{\rm LMS} = \frac{P_{\mathbf{d}}}{P_{\rm o,min} + J_{\rm ex}(\infty) - P_{\mathbf{d}}} \tag{3.38}$$

respectively. With enough DOFs, CCI is mostly suppressed. For the same reason given in Chapter 2, from (3.15), (3.31), (3.36) and (3.38), we can see that with the same step size, the steady-state output SINR of the adaptive DFGSC will be higher than that of the conventional adaptive GSC. Also note that the step size bound for the LMS algorithm is determined by the eigenvalue spread of the input correlation matrix. The input vectors for the conventional GSC and the DFGSC are the same and so step size bounds for both schemes are the same.

3.4.2 CBDFE

The MMA is conventionally applied to the blind FIR linear equalizer only. Since the DFE is a nonlinear and IIR equalizer, direct application of the MMA may result in severe error propagation in the start-up period. It may make the performance of the blind DFE even worse than that of its FIR linear companion. Here, we make use of the channel estimator $\mathbf{w}_{\mathrm{h}}$ derived previously and propose a new structure, i.e., the CBDFE, to overcome this problem. In FIR linear equalization, the performance of the MMA can be demonstrated to approach that of the training-based MMSE method [57]. Thus, we assume that the weights solved by the MMA is close to those solved by the MMSE criterion. Our approach uses a basic principle of the DFE, i.e., the postcursor response of the channel convolved with the feedforward filter is cancelled by the feedback filter. Let $\mathbf{w}_{\mathrm{f}}$ and $\mathbf{w}_{\mathrm{b}}$ be the feedforward and feedback filter weight vectors of a conventional DFE, with length α and β, respectively. Similar to the previous section, $y_{\mathrm{s}}(k)$ denotes the output of the DFGSC. For notation simplicity, the input of $\mathbf{w}_{\mathrm{f}}$ and $\mathbf{w}_{\mathrm{b}}$ is again written as $\mathbf{y}(k) = [y_{\mathrm{s}}(k)\ \ y_{\mathrm{s}}(k-1)\ \cdots\ y_{\mathrm{s}}(k-\alpha+1)]^T$ and $\hat{\mathbf{b}}(k) = [\hat{b}(k-\kappa-1)\ \ \hat{b}(k-\kappa-2)\ \cdots\ \hat{b}(k-\kappa-\beta)]^T$, where κ is the decision delay. Note that the convolution of the equivalent channel response and the feedforward filter results in a response of length $\alpha+\gamma-1$. Thus, for perfect postcursor cancellation, we must have $\beta \geq \alpha+\gamma-2-\kappa$. Without loss of generality, we let $\beta = \alpha+\gamma-2-\kappa$. We now prove the postcursor cancellation property mentioned above. Assuming that the decision is correct, we have the error signal as

$$\begin{aligned} e_{\mathrm{t}}(k) &= b(k-\kappa) - y_{\mathrm{t}}(k) \\ &= b(k-\kappa) - (\mathbf{w}_{\mathrm{f}}^H \mathbf{y}(k) - \mathbf{w}_{\mathrm{b}}^H \mathbf{b}(k)) \end{aligned} \tag{3.39}$$

where $y_{\mathrm{t}}(k)$ is the equalizer output. Straightforward manipulations give the equalizer output MSE as

$$\begin{aligned} \mathrm{E}\{|e_{\mathrm{t}}(k)|^2\} = {} & \mathbf{w}_{\mathrm{f}}^H \mathbf{R}_{\mathbf{yy}} \mathbf{w}_{\mathrm{f}} - \mathbf{w}_{\mathrm{f}}^H \mathbf{R}_{\mathbf{yb}} \mathbf{w}_{\mathrm{b}} - \mathbf{w}_{\mathrm{f}}^H \mathbf{p}_{\mathbf{y}b} - \mathbf{w}_{\mathrm{b}}^H \mathbf{R}_{\mathbf{yb}}^H \mathbf{w}_{\mathrm{f}} \\ & + \mathbf{w}_{\mathrm{b}}^H \mathbf{R}_{\mathbf{bb}} \mathbf{w}_{\mathrm{b}} + \mathbf{w}_{\mathrm{b}}^H \mathbf{p}_{\mathbf{b}b} - \mathbf{p}_{\mathbf{y}b}^H \mathbf{w}_{\mathrm{f}} + \mathbf{p}_{\mathbf{b}b}^H \mathbf{w}_{\mathrm{b}} + \sigma_b^2 \end{aligned} \tag{3.40}$$

with $\mathbf{R}_{\mathbf{yy}} = \mathrm{E}\{\mathbf{y}(k)\mathbf{y}^H(k)\}$, $\mathbf{R}_{\mathbf{bb}} = \mathrm{E}\{\mathbf{b}(k)\mathbf{b}^H(k)\}$, $\mathbf{R}_{\mathbf{yb}} = \mathrm{E}\{\mathbf{y}(k)\mathbf{b}^H(k)\}$, $\mathbf{p}_{\mathbf{y}b} = \mathrm{E}\{\mathbf{y}(k)b^*(k-\kappa)\}$ and $\mathbf{p}_{\mathbf{b}b} = \mathrm{E}\{\mathbf{b}(k)b^*(k-\kappa)\}$. To obtain the optimum weights with the MMSE cri-

terion, we set the gradient of $\mathrm{E}\{|e_{\mathrm{t}}(k)|^2\}$ with respect to the vectors $\mathbf{w}_{\mathrm{f}}^*$ and $\mathbf{w}_{\mathrm{b}}^*$ to zero. This results in

$$\mathbf{w}_{\mathrm{f,opt}} = (\mathbf{R}_{\mathbf{yy}} - \frac{1}{\sigma_b^2}\mathbf{R}_{\mathbf{yb}}\mathbf{R}_{\mathbf{yb}}^H)^{-1}\mathbf{p}_{\mathbf{y}b} \tag{3.41}$$

$$\mathbf{w}_{\mathrm{b,opt}} = \frac{1}{\sigma_b^2}\mathbf{R}_{\mathbf{yb}}^H\mathbf{w}_{\mathrm{f,opt}}. \tag{3.42}$$

Let $\mathbf{w}_{\mathrm{h,opt}} = [\omega_0 \quad \omega_1 \quad \cdots \quad \omega_{\gamma-1}]^T$ and represent the convolution of $\mathbf{w}_{\mathrm{h,opt}}$ and $\mathbf{w}_{\mathrm{f,opt}}$ as $\mathbf{H}\mathbf{w}_{\mathrm{f,opt}}$, where $\mathbf{H}$ is an $(\alpha+\gamma-1)\times\alpha$ matrix as

$$\mathbf{H} = \begin{bmatrix} \omega_0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \omega_1 & \omega_0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & & \ddots & & & & & \vdots \\ \omega_{\gamma-1} & \omega_{\gamma-2} & \cdots & \omega_0 & 0 & \cdots & \cdots & 0 \\ 0 & \omega_{\gamma-1} & \omega_{\gamma-2} & \cdots & \omega_0 & 0 & \cdots & 0 \\ \vdots & & \ddots & & & \ddots & & \vdots \\ 0 & \cdots & 0 & \omega_{\gamma-1} & \omega_{\gamma-2} & \cdots & \omega_0 & 0 \\ 0 & \cdots & \cdots & 0 & \omega_{\gamma-1} & \omega_{\gamma-2} & \cdots & \omega_0 \\ \vdots & & & & & \ddots & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & \omega_{\gamma-1} & \omega_{\gamma-2} \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \omega_{\gamma-1} \end{bmatrix}. \tag{3.43}$$

We can further partition $\mathbf{H}$ as $\mathbf{H} = [\mathbf{H}_{\mathrm{r}}^T \quad \mathbf{H}_{\mathrm{p}}^T]^T$, where $\mathbf{H}_{\mathrm{r}}$ is of dimension $(\kappa+1)\times\alpha$ and $\mathbf{H}_{\mathrm{p}}$ is of dimension $(\alpha+\gamma-2-\kappa)\times\alpha$. It is simple to see that $\mathbf{H}_{\mathrm{r}}\mathbf{w}_{\mathrm{f,opt}}$ corresponds to the precursor response of $\mathbf{H}\mathbf{w}_{\mathrm{f,opt}}$ while $\mathbf{H}_{\mathrm{p}}\mathbf{w}_{\mathrm{f,opt}}$ the postcursor response. With enough DOFs, the DFGSC output can be modeled as

$$y_{\mathrm{s}}(k) = \mathbf{w}_{\mathrm{q}}^H\mathbf{d}(k) + \nu(k) \tag{3.44}$$

where $\nu(k)$ is a white noise independent to $\mathbf{d}(k)$. With (3.44) and some manipulations, we can derive

$$\frac{1}{\sigma_b^2}\mathbf{R}_{\mathbf{yb}}^H = \mathbf{H}_{\mathrm{p}} \tag{3.45}$$

From (3.42), we then obtain $\mathbf{w}_{\mathrm{b,opt}} = \mathbf{H}_{\mathrm{p}}\mathbf{w}_{\mathrm{f,opt}}$. This result can be re-stated as

$$\mathbf{w}_{\mathrm{b,opt}} = \mathrm{post}\{\mathbf{w}_{\mathrm{h,opt}} \otimes \mathbf{w}_{\mathrm{f,opt}}\}. \tag{3.46}$$

This result suggests an adaptation approach for the training-based MMSE-DFE. Let $\mathbf{w}_{\mathrm{f}}(k)$ and $\mathbf{w}_{\mathrm{b}}(k)$ be the feedforward and feedback weight vectors at time instant k. With

reference to (3.46), we can let

$$\mathbf{w}_{\mathrm{b}}(k) = \mathrm{post}\{\mathbf{w}_{\mathrm{h}}(k) \otimes \mathbf{w}_{\mathrm{f}}(k)\} \tag{3.47}$$

in which $\mathbf{w}_{\mathrm{h}}(k)$ is the channel estimate at time instant k. If $\mathbf{w}_{\mathrm{h}}(k)$ converges to $\mathbf{w}_{\mathrm{h,opt}}$, $\mathbf{w}_{\mathrm{b}}(k)$ will converge to $\mathbf{w}_{\mathrm{b,opt}}$ too. The difference between this approach and the conventional method lies in that only $\mathbf{w}_{\mathrm{f}}(k)$ is adapted (not both $\mathbf{w}_{\mathrm{f}}(k)$ and $\mathbf{w}_{\mathrm{b}}(k)$). Consider a blind DFE scenario that both $\mathbf{w}_{\mathrm{f}}(k)$ and $\mathbf{w}_{\mathrm{b}}(k)$ are adapted. With $y_{\mathrm{t}}(k) = \mathbf{w}_{\mathrm{f}}^{H}(k)\mathbf{y}(k) - \mathbf{w}_{\mathrm{b}}^{H}(k)\hat{\mathbf{b}}(k)$, the update equations for the MMA can be written as

$$\mathbf{w}_{\mathrm{f}}(k+1) = \mathbf{w}_{\mathrm{f}}(k) + \mu_{\mathrm{f}}\phi(y_{\mathrm{t}}(k))\mathbf{y}(k) \tag{3.48}$$

$$\mathbf{w}_{\mathrm{b}}(k+1) = \mathbf{w}_{\mathrm{b}}(k) - \mu_{\mathrm{b}}\phi(y_{\mathrm{t}}(k))\hat{\mathbf{b}}(k) \tag{3.49}$$

where $\phi(y_{\mathrm{t}}(k))$ is the same as that given in (3.20). From (3.48) and (3.49), we see that if there is a decision error, the error will immediately reflect to $\hat{\mathbf{b}}(k)$ and then $\phi(y_{\mathrm{t}}(k))$. Note that the adaptation of $\mathbf{w}_{\mathrm{f}}(k)$ involves erroneous $\phi(y_{\mathrm{t}}(k))$ only while that of $\mathbf{w}_{\mathrm{b}}(k)$ involves both erroneous $\hat{\mathbf{b}}(k)$ and erroneous $\phi(y_{\mathrm{t}}(k))$. The two error sources in (3.49) will make $\mathbf{w}_{\mathrm{b}}(k)$ quite sensitive to decision errors. On the other hand, in the proposed method, only $\mathbf{w}_{\mathrm{f}}(k)$ is adapted as given in (3.48). Although the effect of decision error will also pass to $e_{\mathrm{s}}(k)$ which will perturb the adaptation of $\mathbf{w}_{\mathrm{a}}(k)$ and $\mathbf{w}_{\mathrm{h}}(k)$, the influence is smaller. This is because the convergence of $\mathbf{w}_{\mathrm{a}}(k)$ and $\mathbf{w}_{\mathrm{h}}(k)$ in the DFGSC is much faster and more stable than that of the blind DFE. By using (3.47) to calculate $\mathbf{w}_{\mathrm{b}}(k)$, the feedback filter of the blind DFE will perform much better. In one word, with the proposed operation, the resultant CBDFE will be less sensitive to decision errors.

In the start-up period, decisions may not be trustworthy. Including decisions in the operation of the CBDFE may affect its stability. Unlike the LBDFE's mode switching, we propose a simple mechanism allowing a smooth transition from FIR linear equalization to DFE. We let the feedback filter be multiplied by a time-varying factor $f(k)$ where $f(k) \leq 1$. Initially, k is small and we set $f(k)$ small such that the weighting of the feedback filter is small. In this stage, the proposed equalizer will behave much like an FIR linear equalizer. As k increases and decisions become more reliable, we increase $f(k)$. Eventually, $f(k)$ approaches one and the equalizer becomes a full DFE. As known in [45],

the convergence of an adaptive algorithm exhibits an exponentially decay behavior. Thus, a natural choice for $f(k)$ will be

$$f(k) = 1 - \mathrm{e}^{-\xi k} \tag{3.50}$$

where ξ is a design parameter controlling the increase rate of this factor. With the proposed mechanism, the feedback filter is gradually taken into account and the error propagation effect will be reduced. The CBDFE will approach the optimum MMSE-DFE ultimately.

3.5 Spatial Multipath and Spatial Signature Mismatch

In this section, we will show that the proposed STE is robust to two unfavorable phenomena, spatial multipath propagation and spatial signature mismatch, which are frequently encountered in array signal processing. Let us consider the spatial multipath propagation first. Multipath propagation may cause ISI coming from different directions, which induces both coherent interference [60] and angular spread [61]. In this case, the conventional GSC with the simple point distortionless constraint tends to cancel the desired signal itself and so signal cancellation occurs. In the proposed STE, even with this temporal and spatial ISI, the phenomenon of signal cancellation will not exist. This is due to the decision feedback structure. In the spatial multipath environment, the desired signal will leak to the output of the blocking matrix, producing a correlation between the upper and lower paths in the DFGSC. By extending the weight calculation for the mismatch scenario in Chapter 2, we have

$$\mathbf{w}_{\mathbf{c},\mathrm{opt}} = \begin{bmatrix} \mathbf{w}_{\mathrm{a,opt}} \\ \mathbf{w}_{\mathrm{h,opt}} \end{bmatrix} = \mathbf{R}_{\mathbf{c}}^{-1} \begin{bmatrix} \mathbf{B}^H \mathbf{R}_{\mathbf{x}} \\ \mathbf{R}_{\mathbf{p}}^H \end{bmatrix} \mathbf{w}_{\mathrm{q}} \tag{3.51}$$

where

$$\mathbf{R}_{\mathbf{p}} = \mathrm{E}\{\mathbf{x}(k)\hat{\mathbf{b}}^H(k)\} = \sigma_b^2 [\sum_{m=0}^{M_0-1} \mathbf{h}_m(0) \ \sum_{m=0}^{M_0-1} \mathbf{h}_m(1) \ \cdots \ \sum_{m=0}^{M_0-1} \mathbf{h}_m(\gamma-1)] \tag{3.52}$$

$$\mathbf{R}_{\mathbf{c}} = \begin{bmatrix} \mathbf{B}^H \mathbf{R}_{\mathbf{x}} \mathbf{B} & \mathbf{M} \\ \mathbf{M}^H & \sigma_b^2 \mathbf{I}_\gamma \end{bmatrix} \tag{3.53}$$

with

$$\mathbf{M} = \sigma_b^2 \mathbf{B}^H [\sum_{m=0}^{M_0-1} \mathbf{h}_m(0) \ \sum_{m=0}^{M_0-1} \mathbf{h}_m(1) \ \cdots \ \sum_{m=0}^{M_0-1} \mathbf{h}_m(\gamma-1)] \tag{3.54}$$

which is the correlation matrix between the blocking matrix output $\mathbf{B}^H\mathbf{x}(k)$ and the decision vector $\hat{\mathbf{b}}(k)$. Once more, we can find the inverse of $\mathbf{R_c}$ and so $\mathbf{w}_{\mathbf{c},\text{opt}}$ in (3.51), i.e.,

$$\mathbf{w}_{\mathbf{c},\text{opt}} = \begin{bmatrix} (\mathbf{B}^H\mathbf{R_z}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{R_z} \\ \left[\sum_{m=0}^{M_0-1}\mathbf{h}_m(0) \ \cdots \ \sum_{m=0}^{M_0-1}\mathbf{h}_m(\gamma-1)\right]^H(\mathbf{I}-\mathbf{B}(\mathbf{B}^H\mathbf{R_z}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{R_z}) \end{bmatrix}\mathbf{w}_\text{q} \tag{3.55}$$

and thus the coupled $\mathbf{w}_{\text{a,opt}}$ and $\mathbf{w}_{\text{h,opt}}$ can be written as

$$\mathbf{w}_{\text{a,opt}} = (\mathbf{B}^H\mathbf{R_z}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{R_z}\mathbf{w}_\text{q} \tag{3.56}$$

$$\mathbf{w}_{\text{h,opt}} = [\sum_{m=0}^{M_0-1}\mathbf{h}_m(0) \ \cdots \ \sum_{m=0}^{M_0-1}\mathbf{h}_m(\gamma-1)]^H(\mathbf{w}_\text{q}-\mathbf{B}\mathbf{w}_{\text{a,opt}}). \tag{3.57}$$

The minimum J of the DFGSC as that given in (3.31) can then be solved to be

$$J_\text{min} = \mathbf{w}_\text{q}^H\mathbf{R_z}\mathbf{w}_\text{opt}. \tag{3.58}$$

From (3.56) to (3.58), we see that while $\mathbf{w}_{\text{a,opt}}$ and $\mathbf{w}_{\text{h,opt}}$ are changed, J_min (and so the error signal) still contains no desired signal. The leaky desired signal in the lower path of the DFGSC becomes invisible when optimizing $\mathbf{w}_\text{a}$. This indicates that no signal cancellation occurs.

We then consider the spatial signature mismatch problem. If the information of the desired signal's main DOA has error, $\mathbf{w}_\text{q}$ does not match to the desired signal's spatial signature any more, and $\mathbf{B}$ cannot obstruct the desired signal. In this case, signal cancellation may occur as well. In the case of spatial multipath propagation, we have already proven that for the DFGSC even if there is leaky desired signal in the output of the blocking matrix, $\mathbf{w}_\text{a}$ will not cancel the desired signal. Spatial signature mismatch can be considered as a similar signal leaking problem as that in the spatial multipath propagation. Thus, no signal cancellation will occur in the DFGSC. The derivation details, however, are omitted.

For these scenarios, the expression of the optimum output SINR is the same as that given in (3.37), but now $P_\mathbf{d}$ and $P_{\text{o,min}}$ are changed. The output desired signal power becomes

$$P_\mathbf{d} = \sigma_b^2|(\mathbf{w}_\text{q}-\mathbf{B}\mathbf{w}_{\text{a,opt}})^H[\sum_{m=0}^{M_0-1}\mathbf{h}_m(0) \ \sum_{m=0}^{M_0-1}\mathbf{h}_m(1) \ \cdots \ \sum_{m=0}^{M_0-1}\mathbf{h}_m(\gamma-1)]|^2 \tag{3.59}$$

which is different from (3.30). The calculation of $P_{\mathrm{o,min}}$ should use (3.14) instead of (3.15). Moreover, due to the existence of the desired signal in the input of the interference cancelling filter, the excess MSE yielded by the LMS algorithm will be different too. Let $J_{\mathrm{ex}}^{\mathbf{d}}(\infty)$ denote the excess MSE value for the desired signal in these leaky scenarios. Using (3.35), we can obtain

$$J_{\mathrm{ex}}^{\mathbf{d}}(\infty) = J_{\min}\frac{\mu_{\mathrm{a}}\lambda_1(\mathbf{B}^H\mathbf{R}_{\mathbf{d}}\mathbf{B})}{2-\mu_{\mathrm{a}}\lambda_1(\mathbf{B}^H\mathbf{R}_{\mathbf{d}}\mathbf{B})}. \tag{3.60}$$

The result in (3.60) is due to the fact that only one eigenvalue of $\mathbf{B}^H\mathbf{R}_{\mathbf{d}}\mathbf{B}$ is nonzero. The expression of the steady-state output SINR for the GSC structure with temporal and spatial ISI or spatial signature mismatch is then slightly modified to

$$\mathrm{SINR}_{\mathrm{LMS}} = \frac{P_{\mathbf{d}} + J_{\mathrm{ex}}^{\mathbf{d}}(\infty)}{P_{\mathrm{o,min}} + J_{\mathrm{ex}}(\infty) - (P_{\mathbf{d}} + J_{\mathrm{ex}}^{\mathbf{d}}(\infty))} \tag{3.61}$$

in which $P_{\mathbf{d}}$ and $P_{\mathrm{o,min}}$ should also be changed to use (3.59) and (3.14), respectively.

3.6 Simulations

Computer simulations are conducted to demonstrate the effectiveness of the proposed STE (called DFGSC-CBDFE in this section for clarity) and verify our analytic results. For comparison, we also consider the hybrid of the conventional GSC and the LBDFE (called GSC-LBDFE hereafter), and the hybrid of the DFGSC and the LBDFE (called DFGSC-LBDFE hereafter). Note that the later scheme is used for the comparison of the CBDFE and the LBDFE. In all cases, we assume a ULA with five omnidirectional antennas spaced half a wavelength apart. The main DOA of the desired signal is known as *a priori*. Only the point distortionless constraint to the main DOA is used for the GSC in all three schemes. The received desired signal is corrupted by CCI, ISI and AWGN. In the first part, we only consider channels with temporal ISI. In the second part, we consider more realistic channels where temporal and spatial ISI is present. In all figures, at least 500 simulation runs are averaged to obtain each simulated result.

3.6.1 Channels with Temporal ISI Only

In this set of simulations, we consider channels with temporal ISI only. We compare the performance of DFGSC-CBDFE with GSC-LBDFE and DFGSC-LBDFE. The trans-

mitted symbols are randomly generated and then modulated by QPSK. The co-channel interference-to-noise ratio (CCINR) is set as 25 dB and the SNR as 15 dB. The channel for the desired signal coming from 0^o is [0.407 0.815 0.407] [8, pp. 616] and the channel for CCI coming from -60^o is [1 0 0]. Parameters used for the adaptive GSC, DFGSC, LBDFE and CBDFE are listed in Table 3.1. The decision-directed MMSE training starts after 500 iterations. As mentioned, the factor $f(k)$ for the CBDFE is used to reduce error propagation. In this case, ξ is chosen to be 0.01 and $f(k)$ approaches unity around 500 iterations. Fig. 3.3 shows the GSC output SINR for DFGSC-CBDFE, GSC-LBDFE and DFGSC-LBDFE. Also shown are the optimum SINR calculated with the Wiener solution, and the steady-state SINR obtained with the LMS algorithm. From Fig. 3.3, we see that all schemes are comparable in convergence rate, but the SINR achieved by those schemes with the adaptive DFGSC is higher than that with the conventional adaptive GSC. As expected, the adaptive DFGSC can approach the optimum solution much more closely, which means that the effect of the excess MSE induced by the LMS is smaller. In terms of GSC output SINR and convergence rate, DFGSC-CBDFE provides the best performance. Fig. 3.4 reveals the beam patterns (enlarging the region around the CCI's DOA) of those schemes used in Fig. 3.3 after 5000 iterations. The improvement of the decision feedback operation can be seen clearly as well. Fig. 3.5 shows the equalizer output MSE for all three schemes. From the figure, we see that both schemes with the adaptive DFGSC work well in suppressing ISI. However, DFGSC-CBDFE performs better than DFGSC-LBDFE. This means that the CBDFE is more stable in ISI suppression.

We then show the robustness of the DFGSC against the spatial signature mismatch, we repeat the same experiment for the scenario with a 5^o difference between the actual and presumed main DOA of the desired signal. We plot the resultant GSC output SINR, beam patterns, and equalizer output MSE in Fig. 3.6, Fig. 3.7, and Fig. 3.8, respectively. We see that the spatial processing of DFGSC-CBDFE keeps good performance while that of GSC-LBDFE fails completely due to signal cancellation. Also note that DFGSC-CBDFE outperforms DFGSC-LBDFE and this indicates that the CBDFE is better than the LBDFE as well.

3.6.2 Channels with Temporal and Spatial ISI

In this part, we consider a more general case that ISI comes from different directions and time instants. Again, the CCINR is fixed at 25 dB and the SNR is also 25 dB. The filter length for the blind DFE is the same as given previously. The step sizes and channel settings used are shown in Table 3.2. From the table, we see that ISI comes from different directions and the effect of both coherent interference and angular spread is included. The parameter ξ is selected as 0.005 and then $f(k)$ is close to unity around 1000 iterations. Fig. 3.9 shows the equalizer output MSE with 16 quadrature amplitude modulation (16-QAM) for the three schemes. To observe the convergence behavior of these schemes, we exclude the decision-directed mode and use the MMA blind algorithm all the way first. We see that DFGSC-CBDFE outperforms the other two schemes in terms of both convergence rate and equalizer output MSE. The channel estimator in the proposed scheme really helps to construct a more stable and efficient blind DFE. Note that GSC-LBDFE completely fails again due to signal cancellation.

Finally, we repeat the same experiment with the decision-directed MMSE training in the tracking period (after 1000 iterations). However, some artificial errors are added in the decision process. When an error occurs, we randomly shift the decision to a constellation point near its true value. Ten consecutive errors are added from the 3000th iteration. The result is presented in Fig. 3.10. Again, it shows that DFGSC-CBDFE performs better than DFGSC-LBDFE. The proposed DFGSC-CBDFE can re-converge after the bursty errors but DFGSC-LBDFE diverges. This clearly shows that the CBDFE is less sensitive to decision errors and so makes the whole processing scheme more reliable. We then conclude that DFGSC-CBDFE is the best hybrid scheme for the scenario we consider.

3.7 Summary

In this chapter, a new adaptive STE has been developed for the suppression of both CCI and ISI. The proposed STE introduces a hybrid of an adaptive DFGSC and an adaptive CBDFE. With the main DOA known as *a priori*, training sequences are not required for the adaptation of the whole STE. We show that the included channel estimator can not

only improve the performance of the conventional adaptive GSC but also make the blind DFE more reliable. For spatial processing, the DFGSC improves the CCI suppression capability when implemented with the LMS algorithm. Also, the adaptive DFGSC with the simple point distortionless constraint is robust against multipath propagation environments and spatial signature errors. For temporal processing, the proposed CBDFE can have better performance than the LBDFE. With our special adaptation, the problem of error propagation is reduced. Simulation results verified our analysis and confirmed that the proposed STE can achieve good performance even in severe channel environments.

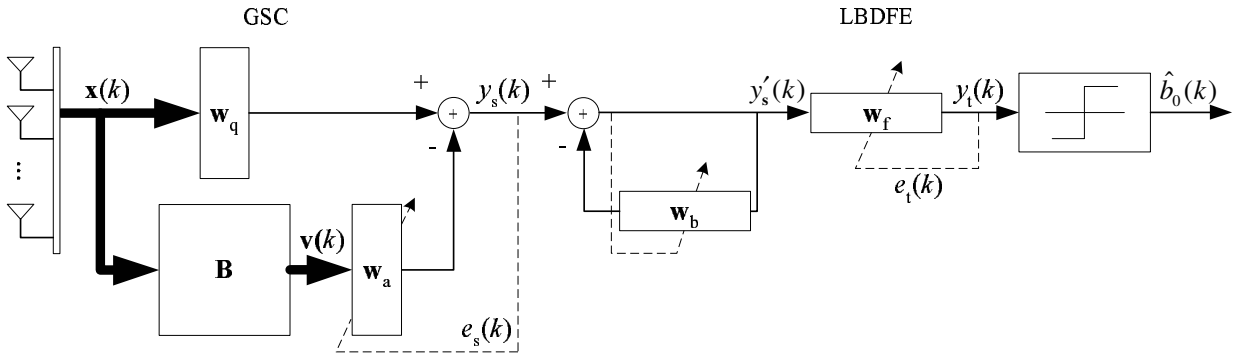

Figure 3.1: Hybrid of GSC and LBDFE (with LBDFE in start-up period).

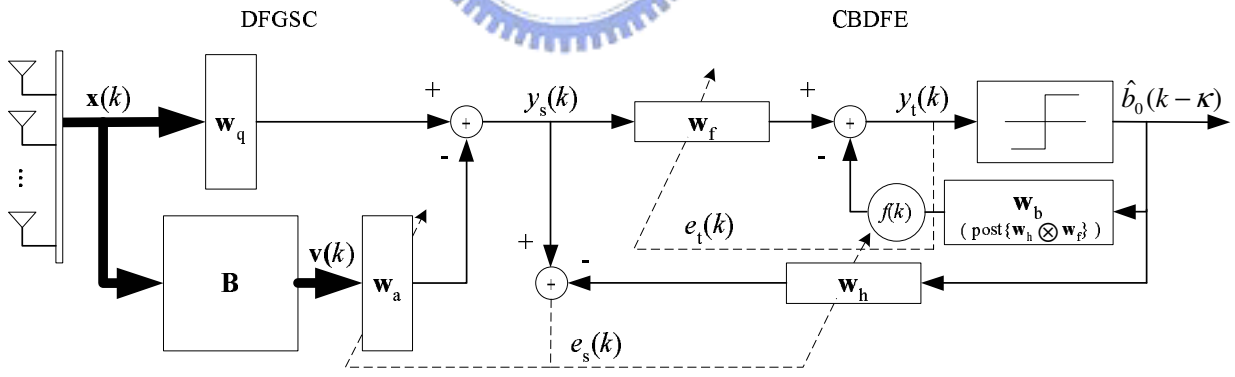

Figure 3.2: Proposed hybrid STE.

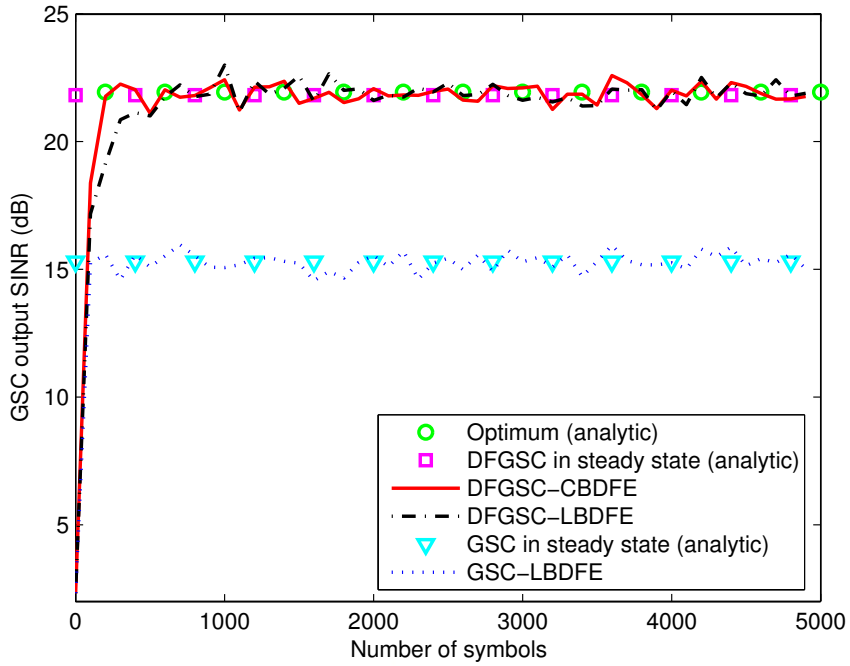

Figure 3.3: Learning curves of GSC output SINR for different schemes in suppressing CCI.

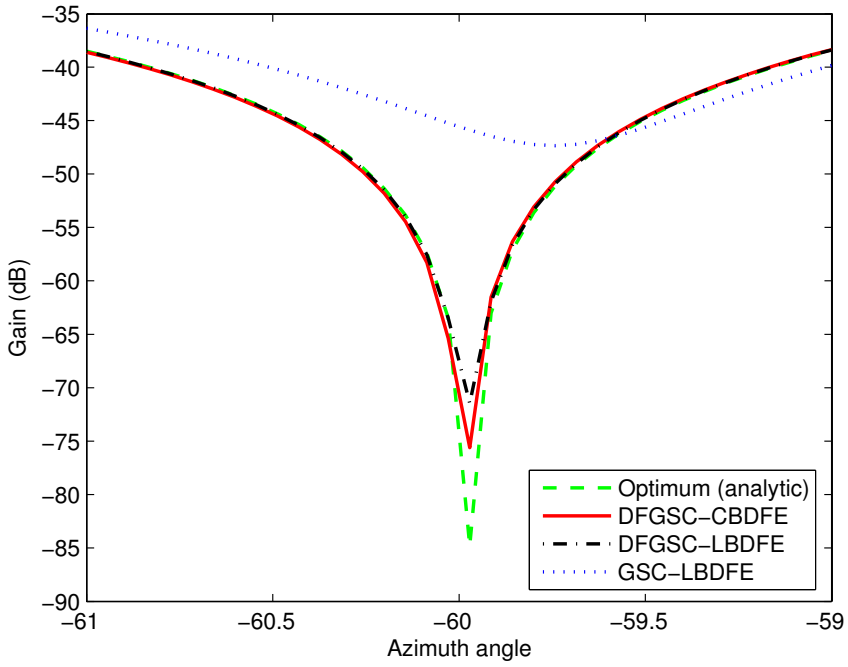

Figure 3.4: Beam patterns (enlarged region around CCI's DOA) of different schemes in Fig. 3.3 after 5000 iterations.

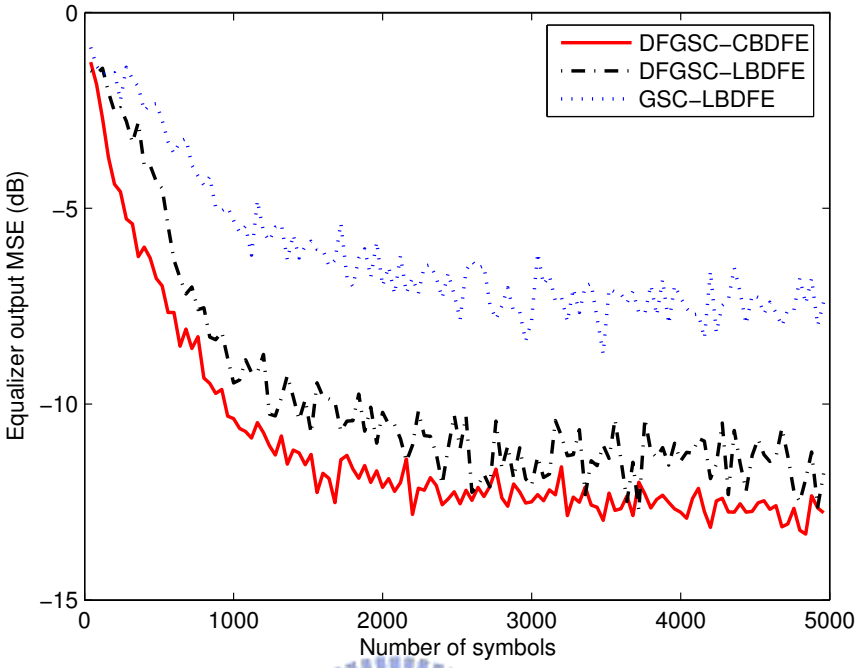

Figure 3.5: Learning curves of equalizer output MSE for different schemes.

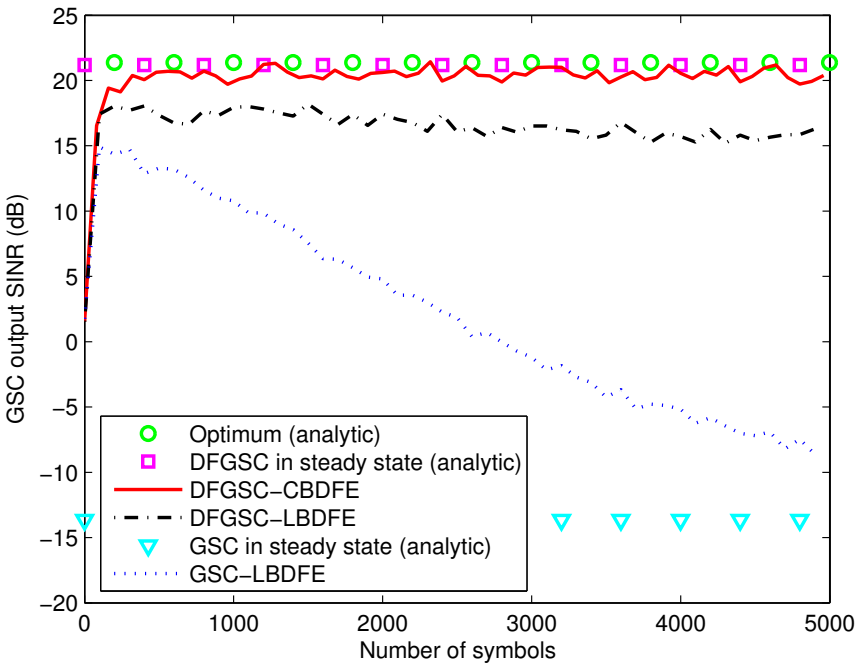

Figure 3.6: Learning curves of GSC output SINR for different schemes with 5° DOA mismatch.

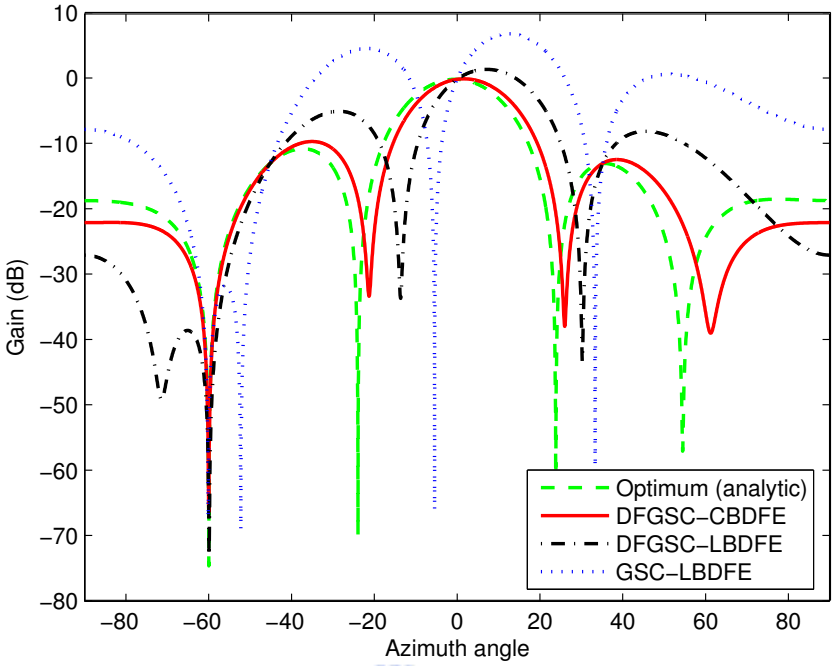

Figure 3.7: Beam patterns of different schemes in Fig. 3.6 after 5000 iterations with 5° DOA mismatch.

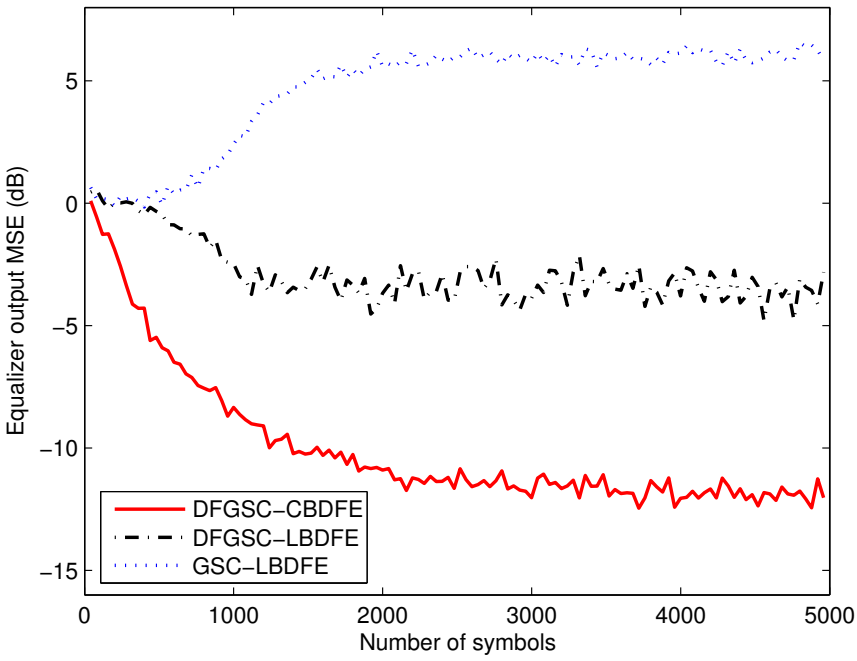

Figure 3.8: Learning curves of equalizer output MSE for different schemes with 5° DOA mismatch.

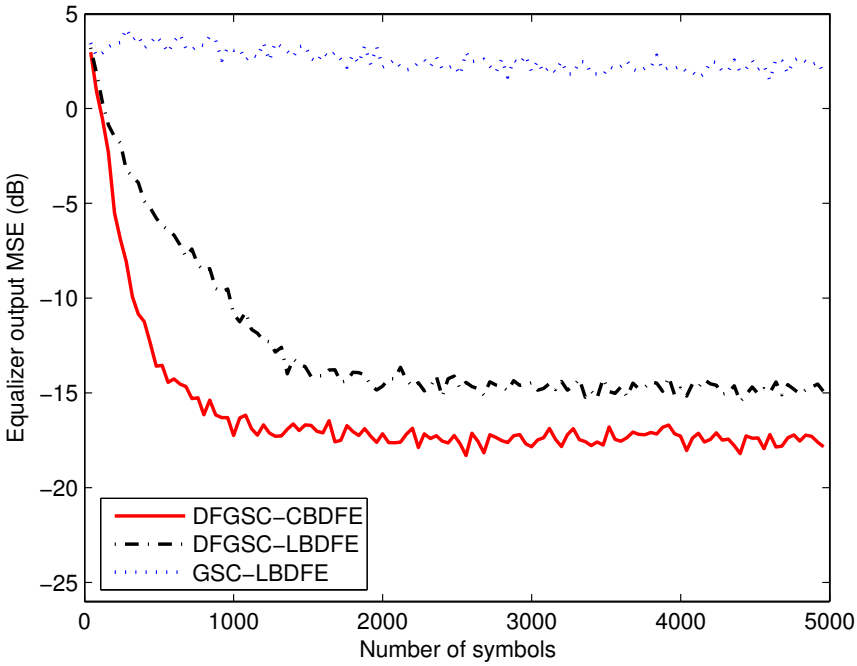

Figure 3.9: Learning curves of equalizer output MSE for different schemes with temporal and spatial ISI.

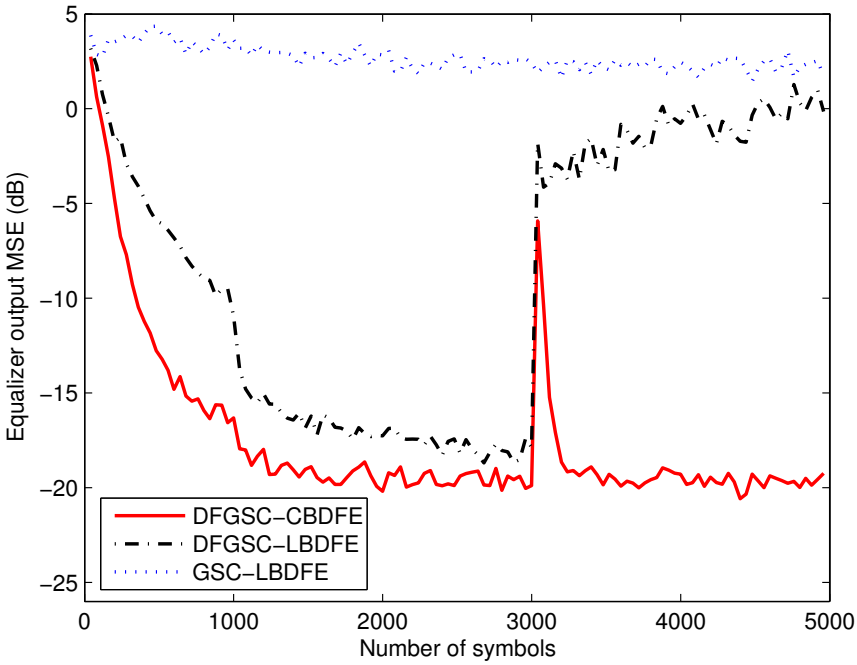

Figure 3.10: Learning curves of equalizer output MSE for different schemes with temporal and spatial ISI (artificial errors added).

Table 3.1: Parameters used for simulations in Section 3.6.1: (a) Filter length, and (b) Step sizes

	$\mathbf{w}_\mathrm{a}$	$\mathbf{w}_\mathrm{f}$	$\mathbf{w}_\mathrm{b}$	$\mathbf{w}_\mathrm{h}$
All schemes	4	7	7	3

(a)

	$\mathbf{w}_\mathrm{a}$	$\mathbf{w}_\mathrm{f}$	$\mathbf{w}_\mathrm{b}$	$\mathbf{w}_\mathrm{h}$
DFGSC-CBDFE (start-up and tracking)	5×10^{-4}	1×10^{-2}	/	2×10^{-3}
GSC-LBDFE (start-up)	5×10^{-4}	1×10^{-4}	2×10^{-4}	/
GSC-LBDFE (tracking)	5×10^{-4}	1×10^{-2}	1×10^{-2}	/
DFGSC-LBDFE (start-up)	5×10^{-4}	1×10^{-4}	2×10^{-4}	2×10^{-3}
DFGSC-LBDFE (tracking)	5×10^{-4}	1×10^{-2}	1×10^{-2}	2×10^{-3}

(b)

Table 3.2: Parameters used for simulations in Section 3.6.2: (a) Step sizes, and (b) Channel settings

	$\mathbf{w}_\mathrm{a}$	$\mathbf{w}_\mathrm{f}$	$\mathbf{w}_\mathrm{b}$	$\mathbf{w}_\mathrm{h}$
DFGSC-CBDFE (start-up)	5×10^{-4}	2×10^{-5}	/	1×10^{-3}
DFGSC-CBDFE (tracking)	5×10^{-4}	3×10^{-3}	/	1×10^{-3}
GSC-LBDFE (start-up)	5×10^{-4}	1×10^{-5}	2×10^{-5}	/
GSC-LBDFE (tracking)	5×10^{-4}	3×10^{-3}	3×10^{-3}	/
DFGSC-LBDFE (start-up)	5×10^{-4}	1×10^{-5}	2×10^{-5}	1×10^{-3}
DFGSC-LBDFE (tracking)	5×10^{-4}	3×10^{-3}	3×10^{-3}	1×10^{-3}

(a)

	DOA	Channel
Desired signal	0^o	[0.1 1.0 0.2]
Temporal and spatial ISI	2^o	[0.1 0.5 0.1]
	30^o	[0.0 0.2 0.7]
	50^o	[0.2 0.6 0.0]
CCI	-60^o	[1.0 0.0 0.0]

(b)

Chapter 4

Adaptive Channel-Aided Decision Feedback Equalization for SISO and MIMO Systems

4.1 Introduction

DFE is a well-known channel equalizer in single-input single-output (SISO) systems [62]-[64]. It has been widely used in digital communications to suppress ISI for over several decades. When the channel spectrum exhibits spectral nulls due to multipath propagation, the DFE performs significantly better than the linear equalizer. Though the MLSE [8] can have better performance than the DFE, the computational complexity is much higher.

A DFE incorporates a feedforward filter operating on the received signal to suppress precursor ISI, and a feedback filter operating on previously detected symbols to suppress postcursor ISI. A DFE uses a nonlinear decision device at the output, and the output represents a noise-free replica of the transmitted symbol assuming that the probability of decision error is small. However, if a symbol is detected incorrectly, the next input to the feedback filter will be in error. As this error advances through the feedback loop, the probability of making an error in the detection of subsequent symbols will be increased. It can result in error propagation that causes bursts of incorrect decisions and a corresponding increase in the decision-error rate [65]. A number of schemes were proposed to reduce error propagation for DFE. A technique combining DFE with partial response precoding and detection was presented in [66]. In [67]-[69], soft decisions and specifically designed constraints were suggested to prevent questionable decisions from being used in

the feedback loop and thereby the probability of error burst was reduced. Besides, a periodic transmission of a short resetting sequence calculated based on a certain steady-state error probability was used to control the error behavior of DFE [70].

Recently, much attention is paid in the development of MIMO systems. For high data-rate transmission, frequency selective fading is present between pairs of transmit and receive antennas. This brings a great design challenge at MIMO receivers. One solution for this problem is to use a MIMO DFE, where both the feedforward filter and the feedback filter are extended to have multiple inputs and multiple outputs, i.e., multi-dimensional feedforward filter and multi-dimensional feedback filter performing multi-dimensional channel equalization [71]-[73]. For the MIMO DFE, the problem of error propagation is even more severe than that of its SISO companion due to the complicated channel configuration and the need to detect signal buried in ISI as well as CCI, in addition to noise.

Since the communication environment may be time-varying, tap weights in the DFE should be updated dynamically for better performance [74]. The LMS adaptive algorithm is often utilized to adapt both the feedforward filter and the feedback filter in SISO DFE systems. It can be shown that the LMS algorithm is also attractive to the adaptive MIMO DFE for dispersive MIMO channels [75], [76]. As described, the error propagation effect will have a greater impact in the adaptive implementation of the DFE. A decision error not only affects the DFE future outputs, but also disturbs the reference signal of the adaptive algorithm. As a result, the DFE will be adapted toward an incorrect direction. In the worst case, error propagation can diverge the DFE adaptation.

The most popular design strategy for channel equalization by far is the use of the MMSE criterion. Its well-accepted theoretical framework and amenability to adaptive implementation make it very attractive for practical usage. Another strategy for equalizer design is to use the minimum bit-error-rate/minimum symbol-error-rate (MBER/MSER) criterion [77]. Various adaptive realizations for the MBER/MSER equalization were proposed in [78]-[81]. Though better results can be obtained in terms of this criterion, there is no guarantee that the global minimum can be reached. In addition, the convergence rate may be slower and the computational complexity may be higher. All these may make

the MBER/MSER equalizer less effective in time-varying channel environments. In the following, we only consider the DFE optimized by the MMSE criterion.

In this chapter, a new LMS-based MMSE-DFE is proposed to reduce the error propagation effect. Inspired by the blind DFE structure given in Chapter 3, we introduce a particularly designed channel estimator to the conventional DFE structure. The resultant adaptive channel-aided (ACA) DFE can perform better than the conventional adaptive DFE and the error propagation effect can be effectively reduced, especially in time-varying channel environments. Only the feedforward filter is adapted with the LMS algorithm, and the feedback filter is obtained from the postcursors of the up-to-date estimated channel convolved with the feedforward filter. We will also show that our SISO ACA-DFE can be extended to a MIMO ACA-DFE.

This chapter is organized as follows. In Section 4.2, the background materials for the MMSE-DFE for both SISO and MIMO systems are described. In Section 4.3, we propose the new ACA-DFE and explain its operation mechanisms. This result can be extended to use in MIMO channels resulting a MIMO ACA-DFE. Finally, simulation results and summary are presented in Section 4.4 and 4.5, respectively.

4.2 Background

4.2.1 Conventional DFE for SISO Systems

Let the tap weights of the feedforward filter and the feedback filter of a DFE be denoted by the column vectors $\mathbf{f}$ with length α and $\mathbf{b}$ with length β, respectively. The complex dispersive channel is modeled by discrete path h^l with $0 \leq l \leq L-1$, in which L is the channel order. We assume that the transmitted symbol $a(k)$ is randomly generated and the noise sample sequence $n(k)$ is zero mean, white, and Gaussian distributed. The received discrete-time equivalent baseband signal at the kth time instant can then be modeled as

$$x(k) = \sum_{l=0}^{L-1} h^l a(k-l) + n(k) = \mathbf{h}^T \mathbf{a}(k) + n(k) \tag{4.1}$$

with $\mathbf{h} = [h^0 \quad h^1 \quad \cdots \quad h^{L-1}]^T$ and $\mathbf{a}(k) = [a(k) \quad a(k-1) \quad \cdots \quad a(k-L+1)]^T$. Let $\mathbf{x}(k)$ be the input vector of the feedforward filter with length α, i.e., $\mathbf{x}(k) = [x(k) \quad x(k-$

$1) \ \cdots \ x(k-\alpha+1)]^T$, and $\hat{\mathbf{a}}(k)$ be the input vector of the feedback filter with length β, i.e., $\hat{\mathbf{a}}(k) = [\hat{a}(k-\kappa-1) \ \ \hat{a}(k-\kappa-2) \ \ \cdots \ \ \hat{a}(k-\kappa-\beta)]^T$, where κ is a suitably chosen decision delay. For the training based MMSE-DFE, the error signal can then be written as

$$e(k) = a(k-\kappa) - (\mathbf{f}^H\mathbf{x}(k) - \mathbf{b}^H\hat{\mathbf{a}}(k)). \tag{4.2}$$

Assuming decisions are correct, i.e., $\hat{a}(k) = a(k)$, we can write the MSE as

$$\begin{aligned} \mathrm{E}\{|e(k)|^2\} = \mathbf{f}^H\mathbf{R}_{\mathbf{xx}}\mathbf{f} - \mathbf{f}^H\mathbf{R}_{\mathbf{xa}}\mathbf{b} - \mathbf{f}^H\mathbf{p}_{\mathbf{x}a} - \mathbf{b}^H\mathbf{R}_{\mathbf{xa}}^H\mathbf{f} \\ +\mathbf{b}^H\mathbf{R}_{\mathbf{aa}}\mathbf{b} + \mathbf{b}^H\mathbf{p}_{\mathbf{a}a} - \mathbf{p}_{\mathbf{x}a}^H\mathbf{f} + \mathbf{p}_{\mathbf{a}a}^H\mathbf{b} + \sigma_a^2 \end{aligned} \tag{4.3}$$

with $\mathbf{R}_{\mathbf{xx}} = \mathrm{E}\{\mathbf{x}(k)\mathbf{x}^H(k)\}$, $\mathbf{R}_{\mathbf{aa}} = \mathrm{E}\{\mathbf{a}(k)\mathbf{a}^H(k)\}$, $\mathbf{R}_{\mathbf{xa}} = \mathrm{E}\{\mathbf{x}(k)\mathbf{a}^H(k)\}$, $\mathbf{p}_{\mathbf{x}a} = \mathrm{E}\{\mathbf{x}(k)a^*(k-\kappa)\}$, $\mathbf{p}_{\mathbf{a}a} = \mathrm{E}\{\mathbf{a}(k)a^*(k-\kappa)\}$, and $\sigma_a^2 = \mathrm{E}\{a(k-\kappa)a^*(k-\kappa)\}$. The optimum choice for $\mathbf{f}$ and $\mathbf{b}$ can be found to be

$$\mathbf{f}_{\mathrm{opt}} = (\mathbf{R}_{\mathbf{xx}} - \frac{1}{\sigma_a^2}\mathbf{R}_{\mathbf{xa}}\mathbf{R}_{\mathbf{xa}}^H)^{-1}\mathbf{p}_{\mathbf{x}a} \tag{4.4}$$

$$\mathbf{b}_{\mathrm{opt}} = \frac{1}{\sigma_a^2}\mathbf{R}_{\mathbf{xa}}^H\mathbf{f}_{\mathrm{opt}}. \tag{4.5}$$

We see that the matrix inverse operation is involved in the calculation of $\mathbf{f}_{\mathrm{opt}}$. A simple alternative to find the optimum tap weights is to use the LMS algorithm. The LMS update equations for $\mathbf{f}$ and $\mathbf{b}$ are expressed as

$$\mathbf{f}(k+1) = \mathbf{f}(k) + \mu_{\mathrm{f}}\mathbf{x}(k)e^*(k) \tag{4.6}$$

$$\mathbf{b}(k+1) = \mathbf{b}(k) - \mu_{\mathrm{b}}\hat{\mathbf{a}}(k)e^*(k) \tag{4.7}$$

where $\mathbf{f}(k)$ and $\mathbf{b}(k)$ are the estimates of $\mathbf{f}_{\mathrm{opt}}$ and $\mathbf{b}_{\mathrm{opt}}$ at the kth time instant, μ_{f} and μ_{b} are the step sizes controlling the convergence rate, and $e(k)$ is the error signal given in (4.2). A typical adaptation process consists of a training mode and a decision-directed mode. Initially, the training mode is launched and sufficient training symbols are transmitted to let both $\mathbf{f}(k)$ and $\mathbf{b}(k)$ converge around the optimum. Then, the DFE switches to the decision-directed mode in which DFE decisions are used as the reference signal and the DFE is continuously adapted. However, DFE decisions may not be always reliable, especially in time-varying channels. Decision errors not only affect the DFE future output, but also disturb the DFE adaptation. In the worst case, the adaptive DFE can diverge, and another training period needs to be re-initiated.

4.2.2 Conventional DFE for MIMO Systems

The SISO DFE can be extended to a MIMO DFE for the equalization of MIMO channels. Here, we use $M \times N$ to signify the configuration with M transmit and N receive antennas, and L to indicate the maximum order for the multi-dimensional channel. Generally, $M \leq N$ is assumed. A sequence of data symbols $a_m(k)$ $(1 \leq m \leq M)$ is transmitted from the mth antenna. We define $\mathbf{a}_m(k) = [a_m(k) \;\; a_m(k-1) \;\; \cdots \;\; a_m(k-L+1)]^T$ as a collection of L successive data symbols from the mth antenna. These data symbols are randomly generated (both in time and space domain) and drawn from the same signal constellation with a variance of σ_a^2. All M data sequences are transmitted over the MIMO channel. The sampled channel response from the mth transmit antenna to the nth receive antenna is given by

$$\mathbf{h}_{nm} = [h_{nm}^0 \;\; h_{nm}^1 \;\; \cdots \;\; h_{nm}^{L-1}]^T \tag{4.8}$$

for $m = 1, 2, \cdots, M$ and $n = 1, 2, \cdots, N$. We can assemble the vectors $\mathbf{h}_{nm}$ into a matrix of size $L \times N$ as

$$\mathbf{H}_m = [\mathbf{h}_{1m} \;\; \mathbf{h}_{2m} \;\; \cdots \;\; \mathbf{h}_{Nm}] \tag{4.9}$$

for $m = 1, 2, \cdots, M$. We also let $\mathbf{n}(k) = [n_1(k) \;\; n_2(k) \;\; \cdots \;\; n_N(k)]^T$ be an N-dimensional noise vector with zero mean, white, and Gaussian distributed elements. With the formulation, M different symbols are simultaneously transmitted through M antennas and received by N antennas to yield the N-dimensional signal vector $\mathbf{x}(k) = [x_1(k) \;\; x_2(k) \;\; \cdots \;\; x_N(k)]^T$. With this premise, the received discrete-time equivalent base-band signal vector can be written as

$$\mathbf{x}(k) = \sum_{m=1}^{M} \mathbf{H}_m^T \mathbf{a}_m(k) + \mathbf{n}(k). \tag{4.10}$$

The formulation of the MIMO DFE is similar to that of the SISO DFE. Nevertheless, M decision devices are employed for M different data sequences. For simplicity, we only consider the most basic form of the MIMO DFE which does not include any SIC action [71]. To be consistent with the previous derivation for the SISO DFE, we first arrange the structure of the feedforward filter into M matrices $\mathbf{F}_m$, for $m = 1, 2, \cdots, M$, with dimension $\alpha \times N$, and the feedback filter into M matrices $\mathbf{B}_m$, for $m = 1, 2, \cdots, M$, with

dimension $\beta \times M$. Both α and β are selected to be long enough to cover the ISI effect in the multi-dimensional channel. The matrices $\mathbf{F}_m$ and $\mathbf{B}_m$ have the forms as

$$\mathbf{F}_m = \begin{bmatrix} f_{m1}^0 & f_{m2}^0 & \cdots & f_{mN}^0 \\ \vdots & \vdots & \ddots & \vdots \\ f_{m1}^{\alpha-1} & f_{m2}^{\alpha-1} & \cdots & f_{mN}^{\alpha-1} \end{bmatrix} = [\mathbf{f}_{m1} \quad \mathbf{f}_{m2} \quad \cdots \quad \mathbf{f}_{mN}] \tag{4.11}$$

$$\mathbf{B}_m = \begin{bmatrix} b_{m1}^0 & b_{m2}^0 & \cdots & b_{mM}^0 \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1}^{\beta-1} & b_{m2}^{\beta-1} & \cdots & b_{mM}^{\beta-1} \end{bmatrix} = [\mathbf{b}_{m1} \quad \mathbf{b}_{m2} \quad \cdots \quad \mathbf{b}_{mM}]. \tag{4.12}$$

To be more compact, we stack the components in the above matrices to form the following vectors

$$\bar{\mathbf{f}}_m = [\mathbf{f}_{m1}^T \quad \mathbf{f}_{m2}^T \quad \cdots \quad \mathbf{f}_{mN}^T]^T \tag{4.13}$$

$$\bar{\mathbf{b}}_m = [\mathbf{b}_{m1}^T \quad \mathbf{b}_{m2}^T \quad \cdots \quad \mathbf{b}_{mM}^T]^T \tag{4.14}$$

for the mth feedforward filter and the mth feedback filter, respectively. Similarly, the successive received signal of (4.10) for the nth antenna can be first grouped as $\mathbf{x}_n(k) = [x_n(k) \quad x_n(k-1) \quad \cdots \quad x_n(k-\alpha+1)]^T$, for $n = 1, 2, \cdots, N$, and then the total received signal vector is described as $\bar{\mathbf{x}}(k) = [\mathbf{x}_1^T(k) \quad \mathbf{x}_2^T(k) \quad \cdots \quad \mathbf{x}_N^T(k)]^T$, which serves as the input to the feedforward filter. The most recent β decisions from the output of the mth decision device after delay κ_m are labeled as $\hat{\mathbf{a}}_m(k) = [\hat{a}_m(k-\kappa_m-1) \quad \hat{a}_m(k-\kappa_m-2) \quad \cdots \quad \hat{a}_m(k-\kappa_m-\beta)]^T$, for $m = 1, 2, \cdots, M$. Here, we assume that all decision delays κ_m are known at the receiver. Then, we can write the overall decision vector as $\bar{\mathbf{a}}(k) = [\mathbf{a}_1^T(k) \quad \mathbf{a}_2^T(k) \quad \cdots \quad \mathbf{a}_M^T(k)]^T$, which constitutes the input to the feedback filter. After that, we may express the estimate in the output of the MIMO DFE prior to the mth decision device at the kth time instant as

$$y_m(k) = \bar{\mathbf{f}}_m^H \bar{\mathbf{x}}(k) - \bar{\mathbf{b}}_m^H \bar{\mathbf{a}}(k) \tag{4.15}$$

and the estimation error for it as

$$e_m(k) = a_m(k-\kappa_m) - y_m(k) \tag{4.16}$$

for $m = 1, 2, \cdots, M$. We see that the error signal given in (4.16) is similar to that of the SISO case except dimension expansion. Architecture-wise, the MIMO DFE can

be treated as a generalization of the SISO DFE, where the scalar delay line, the taps, and the decision are replaced by the vector delay line, the matrix taps, and the decision vector, respectively. With reference to (4.4) and (4.5), for each data sequence and the corresponding decision device, we may have the optimum feedforward filter and feedback filter expressed as

$$\bar{\mathbf{f}}_{m,\text{opt}} = (\mathbf{R}_{\bar{\mathbf{x}}\bar{\mathbf{x}}} - \frac{1}{\sigma_a^2}\mathbf{R}_{\bar{\mathbf{x}}\bar{\mathbf{a}}}\mathbf{R}_{\bar{\mathbf{x}}\bar{\mathbf{a}}}^H)^{-1}\mathbf{p}_{\bar{\mathbf{x}}a_m} \tag{4.17}$$

$$\bar{\mathbf{b}}_{m,\text{opt}} = \frac{1}{\sigma_a^2}\mathbf{R}_{\bar{\mathbf{x}}\bar{\mathbf{a}}}^H\bar{\mathbf{f}}_{m,\text{opt}} \tag{4.18}$$

with the matrix and vector elements defined similar to those for the SISO DFE. Again, to avoid the matrix inverse operation in the feedforward filter calculation, we may adopt the LMS algorithm to find the optimum tap weights recursively. It is not difficult to obtain the update equations for the MIMO DFE as

$$\bar{\mathbf{f}}_m(k+1) = \bar{\mathbf{f}}_m(k) + \mu_\text{f}\bar{\mathbf{x}}(k)e_m^*(k) \tag{4.19}$$

$$\bar{\mathbf{b}}_m(k+1) = \bar{\mathbf{b}}_m(k) - \mu_\text{b}\bar{\mathbf{a}}(k)e_m^*(k) \tag{4.20}$$

for $m = 1, 2, \cdots, M$. Since the received signal is also corrupted by CCI in MIMO channel environments, it tends to make the error signal in the above update equations more noisy. For similar ISI conditions, the performance of the adaptive MIMO DFE is worse than that of the adaptive SISO DFE.

4.3 Proposed Adaptive Channel-Aided DFE

4.3.1 ACA-DFE for SISO Systems

Fig. 4.1 is the block diagram of the proposed ACA-DFE for SISO systems. To obtain the channel response, we first introduce a channel estimator in the DFE structure. Let the coefficients of the included channel estimator be denoted as $\mathbf{q}$ and its dimension is $\gamma \times 1$. The value of γ is chosen to be larger than or equal to that of the channel order L. For convenience, we choose $\gamma = L$. According to Fig. 4.1, the channel estimator $\mathbf{q}$ is tuned by a new error signal $e_\text{q}(k)$, and the cost function for the optimization of $\mathbf{q}$ can be written as

$$\min_{\mathbf{q}} \quad \text{E}\{|e_\text{q}(k)|^2\} = \min_{\mathbf{q}} \quad \text{E}\{|x(k) - \mathbf{q}^H\tilde{\mathbf{a}}(k)|^2\} \tag{4.21}$$

where $x(k)$ is the received signal and $\tilde{\mathbf{a}}(k) = [\hat{a}(k) \quad \hat{a}(k-1) \quad \cdots \quad \hat{a}(k-\gamma+1)]^T$ is the input vector to the channel estimator. Assume that decisions are correct and input data symbols are white. We can then calculate the input correlation matrix for $\mathbf{q}$ as $\mathbf{R}_{\tilde{\mathbf{a}}\tilde{\mathbf{a}}} = \mathrm{E}\{\tilde{\mathbf{a}}(k)\tilde{\mathbf{a}}^H(k)\} = \sigma_a^2\mathbf{I}_\gamma$, where $\mathbf{I}_\gamma$ is a $\gamma \times \gamma$ identity matrix, and the cross-correlation vector for $\mathbf{a}(k)$ and $x(k)$ as $\mathbf{p}_{\tilde{\mathbf{a}}x} = \mathrm{E}\{\tilde{\mathbf{a}}(k)x^*(k)\} = \sigma_a^2\mathbf{h}^*$. From (4.21), the optimum $\mathbf{q}$ solved by the classic Wiener solution is

$$\mathbf{q}_{\mathrm{opt}} = \mathbf{R}_{\tilde{\mathbf{a}}\tilde{\mathbf{a}}}^{-1}\mathbf{p}_{\mathbf{a}x} = \mathbf{h}^*. \tag{4.22}$$

We see that the optimum $\mathbf{q}$ equals the complex conjugate of the channel response. As previous, we can use the LMS algorithm to approach $\mathbf{q}_{\mathrm{opt}}$ recursively. The update equation is stated as

$$\mathbf{q}(k+1) = \mathbf{q}(k) + \mu_{\mathrm{q}}\tilde{\mathbf{a}}(k)e_{\mathrm{q}}^*(k) \tag{4.23}$$

where $\mathbf{q}(k)$ is the estimate of $\mathbf{q}_{\mathrm{opt}}$ at the kth time instant and μ_{q} is the step size for the adaptation. We observe that the channel estimation problem is essentially a system identification problem. For uncorrelated input data symbols, the eigenvalues of the input correlation matrix are all identical, and thus the eigenvalue spread equals unity, which is the minimum possible value. It is well-known that the convergence rate of the LMS algorithm is inversely proportional to the eigenvalue spread [45]. Thus, the convergence of $\mathbf{q}(k)$ is expected to be fast and stable.

Here, we make use of this channel estimator to propose the ACA-DFE. Our approach uses the basic property of the DFE that the postcursors of the channel response convolved with the feedforward filter is cancelled by the feedback filter. With reference to Chapter 3, we can write the result straightforwardly as

$$\mathbf{b}_{\mathrm{opt}} = \mathrm{post}\{\mathbf{q}_{\mathrm{opt}} \otimes \mathbf{f}_{\mathrm{opt}}\}. \tag{4.24}$$

This result suggests an adaptation approach for the training-based MMSE-DFE. Let $\mathbf{f}(k)$ and $\mathbf{b}(k)$ be the feedforward filter and feedback filter at the kth time instant. From (4.24), we can let

$$\mathbf{b}(k) = \mathrm{post}\{\mathbf{q}(k) \otimes \mathbf{f}(k)\} \tag{4.25}$$

in which $\mathbf{q}(k)$ is the channel estimate at the kth time instant. For the conventional adaptive DFE in the decision-directed mode, the scenario is that both $\mathbf{f}(k)$ and $\mathbf{b}(k)$ are

adapted based on the LMS update equations as given in (4.6) and (4.7). We observe that if there is a decision error, the error will immediately reflect to $\hat{\mathbf{a}}(k)$ and then $e(k)$. Note that the adaptation of $\mathbf{f}(k)$ involves erroneous $e(k)$ only while that of $\mathbf{b}(k)$ involves both erroneous $\hat{\mathbf{a}}(k)$ and erroneous $e(k)$. For the same reason given in Chapter 3, by using (4.25) to calculate $\mathbf{b}(k)$, the resultant ACA-DFE will be less sensitive to decision error and the error propagation effect will be reduced, especially in time-varying environments.

4.3.2 ACA-DFE for MIMO Systems

In Section 4.2.2, we have already shown that for the mth decision device in MIMO DFE, the optimum formulation is similar to that for SISO DFE except dimension expansion. This motivates us to extend the idea of the SISO ACA-DFE to MIMO ACA-DFE. The block diagram of the MIMO ACA-DFE is described in Fig. 4.2. First, we define the channel estimators $\mathbf{q}_{nm}$ with dimension $\gamma \times 1$, for $m = 1, 2, \cdots, M$ and $n = 1, 2, \cdots, N$, to estimate $\mathbf{h}_{nm}$ given in (4.8). Following the development presented previously, we can express this task as a system identification problem as

$$\min_{\mathbf{q}_{nm}} \quad \mathrm{E}\{|e_{\mathrm{q},nm}(k)|^2\} = \min_{\mathbf{q}_{nm}} \quad \mathrm{E}\{|x_n(k) - \mathbf{q}_{nm}^H \tilde{\mathbf{a}}_m(k)|^2\} \tag{4.26}$$

where $x_n(k)$ is the received signal from the nth antenna and $\tilde{\mathbf{a}}_m(k) = [\hat{a}_m(k) \quad \hat{a}_m(k-1) \quad \cdots \quad \hat{a}_m(k-\gamma+1)]^T$ is the decision vector from the mth decision device as the input to the corresponding channel estimators. Similar to (4.22), the solution is in the form as

$$\mathbf{q}_{nm,\mathrm{opt}} = \mathbf{h}_{nm}^*. \tag{4.27}$$

Referring to (4.23), we can then use the LMS algorithm to approach $\mathbf{q}_{nm,\mathrm{opt}}$ too. Likewise, the mth optimum feedback filter given in (4.18) can be formulated as

$$\bar{\mathbf{b}}_{m,\mathrm{opt}} = \begin{bmatrix} \sum_{n=1}^{N} \mathrm{post}\{\mathbf{q}_{n1,\mathrm{opt}} \otimes \mathbf{f}_{mn,\mathrm{opt}}\} \\ \sum_{n=1}^{N} \mathrm{post}\{\mathbf{q}_{n2,\mathrm{opt}} \otimes \mathbf{f}_{mn,\mathrm{opt}}\} \\ \vdots \\ \sum_{n=1}^{N} \mathrm{post}\{\mathbf{q}_{nM,\mathrm{opt}} \otimes \mathbf{f}_{mn,\mathrm{opt}}\} \end{bmatrix} \tag{4.28}$$

in which $\mathbf{f}_{mn,\mathrm{opt}}$ is the nth sub-vector in $\bar{\mathbf{f}}_{m,\mathrm{opt}}$, as represented in (4.13). While both $\mathbf{q}_{nm,\mathrm{opt}}$ and $\bar{\mathbf{f}}_{m,\mathrm{opt}}$ are estimated using the LMS algorithm, the estimation of $\bar{\mathbf{b}}_{m,\mathrm{opt}}$ at the kth time

instant can then be calculated as

$$\bar{\mathbf{b}}_m(k) = \begin{bmatrix} \sum_{n=1}^{N} \text{post}\{\mathbf{q}_{n1}(k) \otimes \mathbf{f}_{mn}(k)\} \\ \sum_{n=1}^{N} \text{post}\{\mathbf{q}_{n2}(k) \otimes \mathbf{f}_{mn}(k)\} \\ \vdots \\ \sum_{n=1}^{N} \text{post}\{\mathbf{q}_{nM}(k) \otimes \mathbf{f}_{mn}(k)\} \end{bmatrix} \tag{4.29}$$

Since $\mathbf{q}_{nm}$ can estimate the corresponding channel response $\mathbf{h}_{nm}$, for the same reason described in the SISO case, the proposed operation in (4.29) can enhance the adaptation of the feedback filter. The resultant MIMO ACA-DFE can then improve the robustness against error propagation for MIMO channel equalization.

4.4 Simulations

Computer simulations are conducted to demonstrate the effectiveness of the proposed ACA-DFE and MIMO ACA-DFE. In the first part, we consider SISO channels. In the second part, we consider MIMO channels. All transmitted symbols are randomly generated and then modulated by QPSK. All decision delays are chosen to optimize the performance. In all figures, at least 500 simulation runs are averaged to obtain each simulated result.

4.4.1 SISO channel environments

In this set of simulations, we demonstrate that the proposed ACA-DFE can provide more robust and stable performance than the conventional adaptive DFE against error propagation under severe ISI environments. We first consider a static channel chosen from [8, pp. 616], which is $[0.227\ 0.460\ 0.688\ 0.460\ 0.227]^T$ (Proakis C channel). The parameters α, β and γ for those filters are set to be 9, 9 and 5, respectively. For comparison, we also show the case of ACA-DFE with perfect channel state information (CSI). The SNR is set as 25 dB. Here, $\mu_{\text{f}} = \mu_{\text{b}} = 0.005$, $\mu_{\text{q}} = 0.002$, and the number of training symbols $T_{\text{t}} = 2000$. The decision-directed mode follows immediately after the training mode. Fig. 4.3 gives the learning curves for various equalization schemes. We see that there is no big difference in performance between the proposed ACA-DFE with and without perfect CSI. It implies that the channel estimator works fairly well. The ACA-DFE performs better than the conventional adaptive DFE in the decision-directed mode under

this severe ISI scenario. To demonstrate the merits of the proposed ACA-DFE further, we give the relation between the average SER and the step size used in the feedforward filter (the same step size is used in the feedback filter of the conventional adaptive DFE as well) in Fig. 4.4. This figure reveals that the ACA-DFE always has lower SER than the conventional adaptive DFE with the same step size. There are a couple of things that we can observe from the figure. First, there is an optimum step size for a DFE. As known, for the LMS algorithm, the smaller the step size, the smaller the output MSE in the steady state (possibly the lower the average SER). However, a smaller step size will also make the convergence slower. As a result, there exists an optimum step size balancing these two effects. The optimum μ_{f} giving the lowest SER is around 0.005 for both schemes, and the SER improvement with the ACA-DFE is almost an order of magnitude. Second, the ACA-DFE outperforms the conventional adaptive DFE for any step size. Last, given a target SER, the applicable step-size range of the ACA-DFE is wider than that of the conventional adaptive DFE. We then consider a time-varying channel constructed from Proakis C channel used previously. The first, second, forth and fifth paths in the channel now undergo fading according to Jakes' model [82], and are upper bounded by their corresponding path magnitudes. The normalized Doppler frequency $f_{\mathrm{d}}T_{\mathrm{s}}$ equals 5×10^{-4}, in which f_{d} is the Doppler frequency and T_{s} is the symbol duration. For example, it is to say that the data rate used is about 4.6×10^5 symbols/s. The vehicle speed is set as 100 km/h and the carrier frequency is 2.4 GHz. This time-varying channel is normalized to keep the SNR at a constant level. Fig. 4.5 shows the average SER versus the step size used in the DFE. We observe that the SER improvement for this time-varying case is quite significant, and is more than an order of magnitude most of the time. In practice, it is difficult to know the exact channel variation pattern and the optimum step size. For the ACA-DFE, since the applicable step-size range is wider and the resultant SER is always lower, making the choice of the step size becomes much easier. This enables the ACA-DFE to work adequately in general time-varying environments.

4.4.2 MIMO channel environments

In this part, we consider the MIMO ACA-DFE for dispersive MIMO channels. First, we use the static 2×2 MIMO channel given in [71]: $\mathbf{h}_{11} = [0.781\ 0.625]^T$, $\mathbf{h}_{12} = [0.781\ -0.625]^T$, $\mathbf{h}_{21} = [0.895\ \ -0.447]^T$, and $\mathbf{h}_{22} = [0.958\ 0.287]^T$. The parameters used are $T_{\mathrm{t}} = 200$, SNR = 20 dB, $\mu_{\mathrm{f}} = \mu_{\mathrm{b}} = 0.005$, and $\mu_{\mathrm{q}} = 0.002$. The learning curves for the MIMO ACA-DFE and the conventional adaptive MIMO DFE are shown in Fig. 4.6. We see that the MIMO ACA-DFE can achieve an MSE lower than that of the conventional adaptive MIMO DFE in the decision-directed mode. Similarly, Fig. 4.7 presents the relationship between the average SER and the step size used in the MIMO DFEs. The MIMO ACA-DFE generally achieves lower SER than the conventional adaptive MIMO DFE. Next, we conduct the experiment under time-varying channel environments, in which we let the second channel tap in $\mathbf{h}_{nm}$ (n=1,2 and m=1,2) be varied with Jakes' model and upper bounded by its corresponding path magnitude. The normalized Doppler frequency is now changed to 2×10^{-4}. It is equivalent to say that the vehicle speed is reduced to about 40 km/h. Fig. 4.8 shows the average SER versus the step size used for the time-varying channel. We can see that the MIMO ACA-DFE still provides better performance. In this scenario, the optimum step size for both DFE schemes is around 0.006, and the SER improvement with the MIMO ACA-DFE is more than an order of magnitude. However, note that the performance obtained with the proposed method is not as good as that in the SISO scenario. It is because the MIMO environment induces CCI for each transmitted sequence and this lowers the input SNR. Also, the channel becomes multi-dimensional and is more difficult to estimate.

4.5 Summary

In this chapter, we have developed the ACA-DFE for SISO systems and MIMO systems. With the additional channel estimator(s) and the special operation for the feedback filter, the stability and robustness against error propagation are improved. Simulation results confirm the usefulness of these proposed schemes.

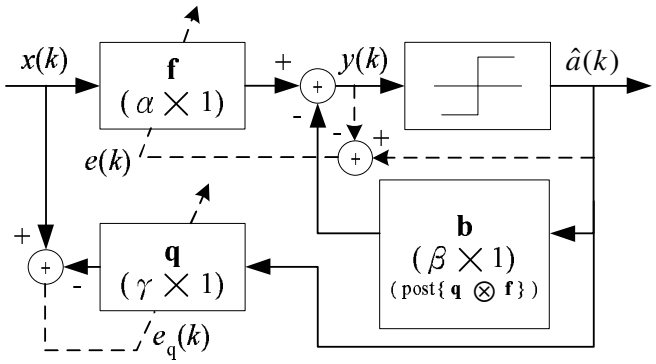

Figure 4.1: ACA-DFE in decision-directed mode for SISO systems.

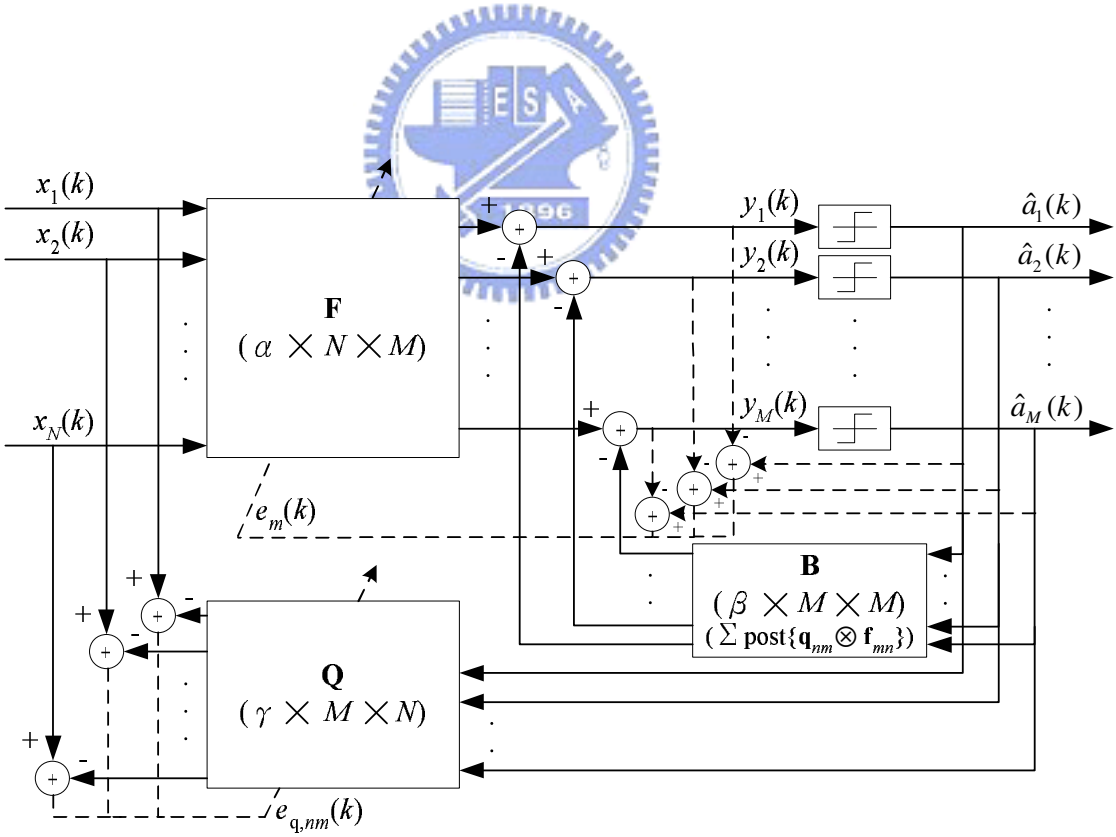

Figure 4.2: ACA-DFE in decision-directed mode for MIMO systems.

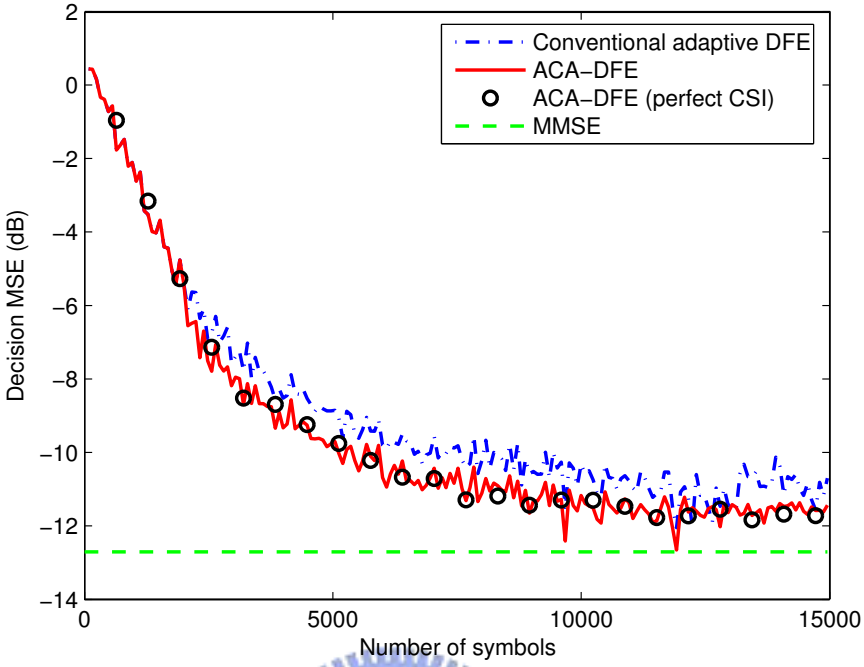

Figure 4.3: MSE learning curves for static Proakis C channel.

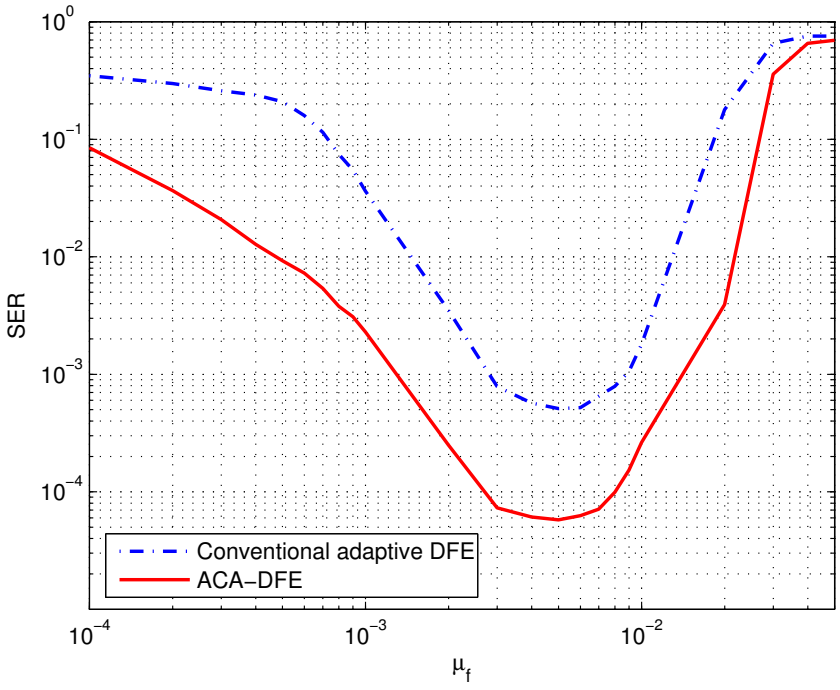

Figure 4.4: SER for static Proakis C channel with different step sizes.

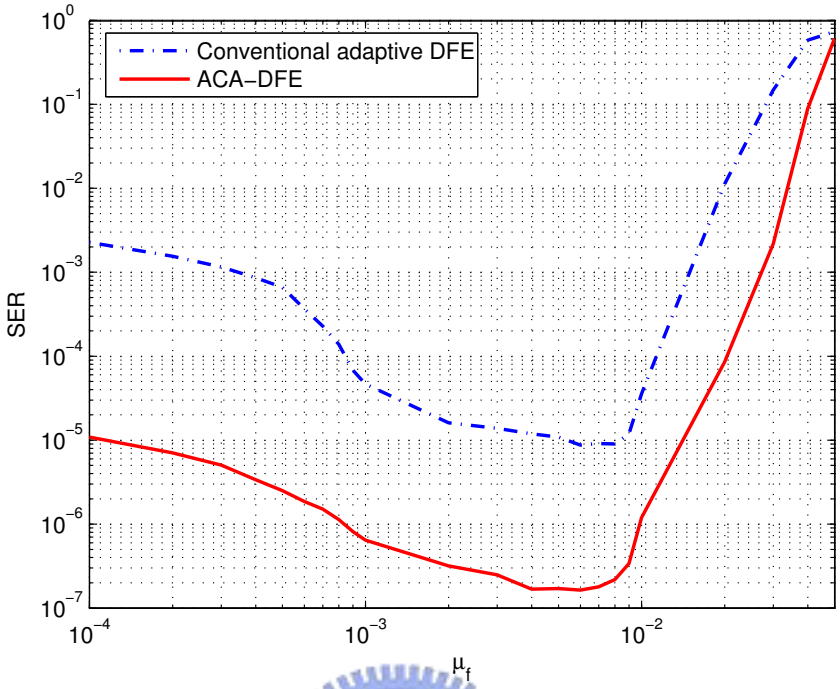

Figure 4.5: SER for time-varying Proakis C channel ($f_d T_s = 5 \times 10^{-4}$) with different step sizes.

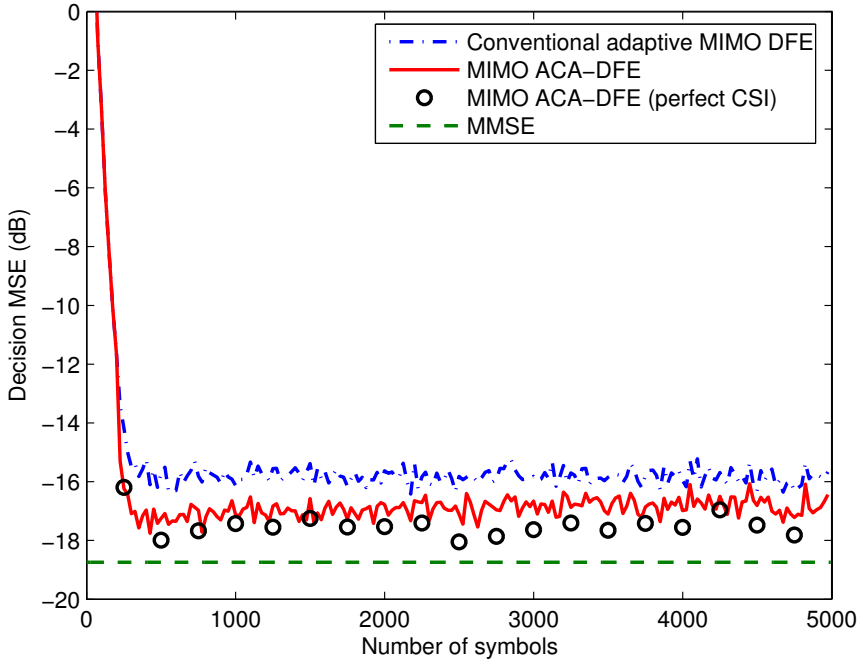

Figure 4.6: MSE learning curves for static MIMO channel.

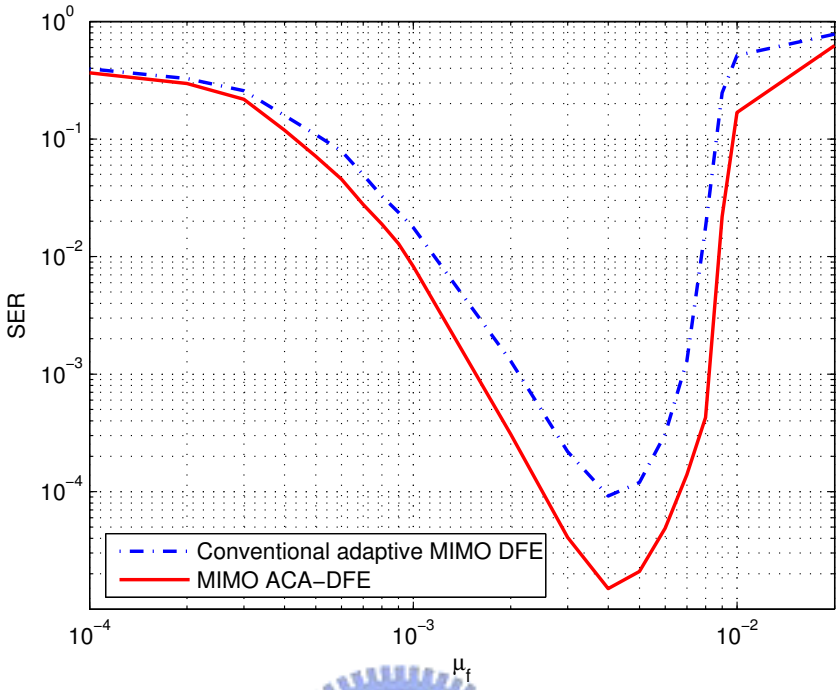

Figure 4.7: SER for static MIMO channel with different step sizes.

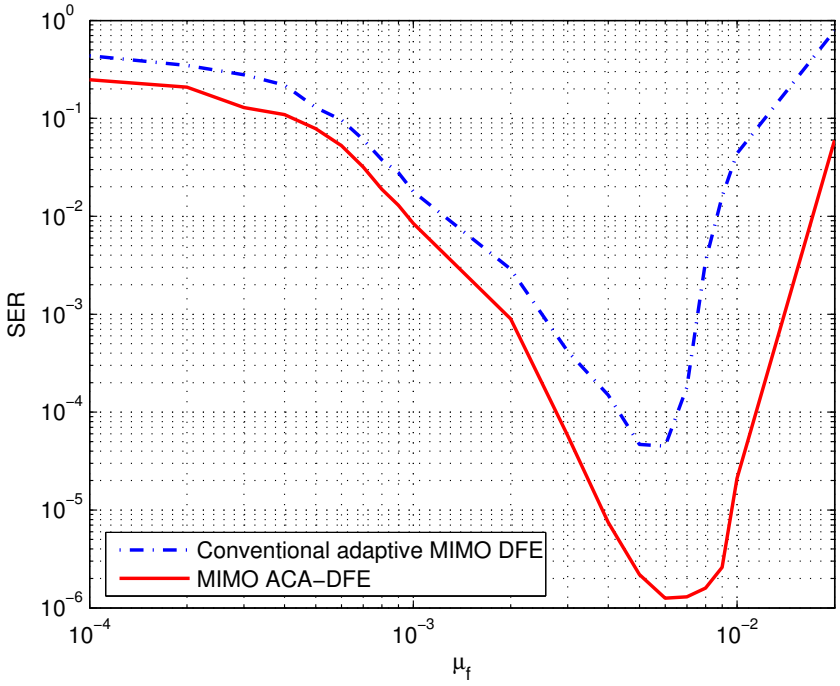

Figure 4.8: SER for time-varying MIMO channel ($f_{\mathrm{d}}T_{\mathrm{s}} = 2 \times 10^{-4}$) with different step sizes.

Chapter 5

Adaptive Two-Stage GSC-Based PIC Detection for Time-Varying MIMO Channels

5.1 Introduction

In MIMO systems, multiple antennas at both transmitter and receiver are used to increase the spectral efficiency. Recently, PIC detection schemes were proposed for signal detection in this communication scenario. In [21]-[25], PIC was realized with zero-forcing (ZF) or MMSE nulling. In noisy environments, the PIC with MMSE nulling (called the MMSE-PIC hereafter) usually provides better performance. However, finding the optimum weights requires high computational complexity. Since decisions in early stages tend to have higher error rates, the idea of partial PIC was then developed [23]. The conventional PIC can provide satisfactory performance in time-invariant channel environments, but its performance can be significantly degraded in time-variant channel environments. This is due to the error propagation effect inherent in the multistage PIC scheme. Also, it is known that the MMSE criterion cannot lead to the MBER performance.

In this chapter, we propose a new adaptive two-stage PIC detection scheme to improve the performance of the MMSE-PIC. The optimization is based on the MV criterion [83]. It is well-known that the MV detection can lead to maximum output SNR and then the MBER. The MV detector, originated from array beamforming, can be effectively realized with the GSC structure. However, the conventional GSC cannot provide best performance when implemented with an adaptive algorithm. Also, it is sensitive to model mismatch. In

time-variant channel environments, direct application of the adaptive GSC in PIC detection may lead to divergence. The DFGSC proposed in Chapter 2 can solve the problems successfully. Here, we extend the use of the DFGSC in MIMO PIC detection problem. We propose an adaptive DFGSC-based PIC structure that can effectively enhance the detection performance, particularly in time-varying channels. All adaptations are based on the LMS algorithm. This will keep the overall computational complexity at a low level and make the proposed scheme feasible for real-world applications. For easy reference, we call our scheme as the DFGSC-PIC hereafter. We also derive the optimum solutions for the DFGSC-PIC and analyze its convergence behavior. Simulations show that the DFGSC-PIC detection scheme can significantly outperform the conventional MMSE-PIC detection scheme in fast-variant channel environments.

The remainder of this chapter is organized as follows. In Section 5.2, the background materials for MIMO signal model and the basics of PIC are briefly described. In Section 5.3, we propose the new adaptive two-stage DFGSC-PIC detection scheme and explain its operation in detail. Section 5.4 gives convergence analysis of the proposed algorithm. Finally, simulation results and summary are presented in Section 5.5 and 5.6, respectively.

5.2 MIMO Signal Model and Parallel Interference Cancellation

Consider a wireless communication system with M antennas at the transmitter and N antennas at the receiver, assuming $M \leq N$. Transmission is over independent time-varying flat Rayleigh fading channels. We define $h_{nm}(k)$ as the complex channel response from transmit antenna m to receive antenna n at time instant k, with $1 \leq m \leq M$ and $1 \leq n \leq N$. Spatial multiplexing is used in this MIMO communication environment. In other words, there are M data streams transmitted at the same time. At time instant k, the M transmitted data symbols can be collected into an $M \times 1$ vector, denoted as $\mathbf{d}(k) = [d_1(k) \;\; d_2(k) \;\; \cdots \;\; d_M(k)]^T$. Let $\mathbf{h}_m(k)$ be a vector containing the channel response for $d_m(k)$, i.e.,

$$\mathbf{h}_m(k) = [h_{1m}(k) \;\; h_{2m}(k) \;\; \cdots \;\; h_{Nm}(k)]^T. \tag{5.1}$$

Then the complete $N \times M$ channel matrix for the transmission of $\mathbf{d}(k)$ can be presented as

$$\mathbf{H}(k) = [\mathbf{h}_1(k) \;\; \mathbf{h}_2(k) \;\; \cdots \;\; \mathbf{h}_M(k)]. \tag{5.2}$$

The data symbols are simultaneously transmitted through the M antennas and received by the N antennas to yield an $N \times 1$ received signal vector $\mathbf{r}(k)$, given by

$$\mathbf{r}(k) = \mathbf{H}(k)\mathbf{d}(k) + \mathbf{n}(k) \tag{5.3}$$

where $\mathbf{n}(k)$ is an $N \times 1$ complex Gaussian noise vector with zero mean and equal variance in each dimension. In addition, noise components at receive antennas are assumed to be independent.

To detect the transmitted data symbols, the multistage PIC is applied. Here, we focus on the development of a two-stage structure. Although more stages can be used, the performance tends to saturate quickly. It has been found that a two-stage processing can have the best trade-off between performance and complexity [84]. The block diagram of a general two-stage PIC detection scheme for MIMO systems is illustrated in Fig. 5.1. As shown, the estimate for the mth element of $\mathbf{d}(k)$ in the first-stage output is expressed as $\tilde{d}_m(k) = \mathrm{Dec}\{\tilde{\mathbf{w}}_m^H(k)\mathbf{r}(k)\}$, where $\tilde{\mathbf{w}}_m(k)$ is a first-stage $N \times 1$ weight vector for the estimation of $d_m(k)$. The input vector to the mth branch in the second stage, denoted as $\hat{\mathbf{r}}_m(k)$, is constructed as

$$\hat{\mathbf{r}}_m(k) = \mathbf{r}(k) - \sum_{j=1, j\neq m}^{M} \mathbf{h}_j(k)\tilde{d}_j(k). \tag{5.4}$$

Note that we treat those data symbols $d_j(k)$ for $j \neq m$ as interference to $d_m(k)$. If the decisions from the first stage are mostly correct, interference from other transmit antennas can be eliminated in $\hat{\mathbf{r}}_m(k)$, and thus the estimation performance in the second stage can be improved. We write the estimate for the mth element of $\mathbf{d}(k)$ in the second-stage output as $\hat{d}_m(k) = \mathrm{Dec}\{\hat{\mathbf{w}}_m^H(k)\hat{\mathbf{r}}_m(k)\}$, where $\hat{\mathbf{w}}_m(k)$ is a second-stage $N \times 1$ weight vector for the estimation of $d_m(k)$. Conventionally, to avoid noise enhancement, the weight vectors $\tilde{\mathbf{w}}_m$ and $\hat{\mathbf{w}}_m$ are optimized through the MMSE criterion [8]. We define the MMSE cost function for the mth weight vector in the first stage as

$$\min_{\tilde{\mathbf{w}}_m(k)} \; \mathrm{E}\{|e(k)|^2\} = \min_{\tilde{\mathbf{w}}_m(k)} \; \mathrm{E}\{|\tilde{d}_m(k) - \tilde{\mathbf{w}}_m^H(k)\mathbf{r}(k)|^2\}. \tag{5.5}$$

It is well-known that the optimum solution for (5.5) involves the matrix inversion, which is not desirable in real-world applications. This problem becomes more troublesome when channels are time-variant. As an alternative, adaptive methods are adopted to obtain the weight vector recursively. The simple yet effective LMS algorithm is used here as well. The update equation for $\tilde{\mathbf{w}}_m(k)$ with the LMS algorithm is given by

$$\tilde{\mathbf{w}}_m(k+1) = \tilde{\mathbf{w}}_m(k) + \mu \mathbf{r}(k) e^*(k) \tag{5.6}$$

where μ is the step size controlling the convergence rate, and $e(k)$ is an error signal as that given in (5.5). The operation of the MMSE-PIC can be divided into two modes, training and decision-directed. In the training mode, training sequences are transmitted for the initial adaptation of $\tilde{\mathbf{w}}_m(k)$. Once convergence is reached, the decision-directed mode is activated in which detected symbols are fed back as reference signals and the PIC weights are updated continuously. The adaptation of $\hat{\mathbf{w}}_m(k)$ in the second stage is identical to that given above. However, due to the satisfactory result usually provided in the first-stage processing, the adaptation in the second stage can be omitted. Here, we simply let the PIC weights for a data stream match to the corresponding channel response (in the second stage), i.e.,

$$\hat{\mathbf{w}}_m(k) = (\mathbf{h}_m^H(k)\mathbf{h}_m(k))^{-1}\mathbf{h}_m(k) \tag{5.7}$$

in which we assume that $\mathbf{h}_m(k)$ is known or can be estimated. With extensive simulations, we found that this approach gives little performance degradation. We then use this simple way for the second-stage PIC weights throughout this chapter.

One problem associated with the decision-directed mode is the error propagation effect. It can seriously affect the MMSE-PIC training. In the worst case, the receiver may lose track of the time-varying MIMO channel. When this occurs, the training mode has to be re-initiated. It is well-known that the transmission of training sequences will reduce the bandwidth utilization efficiency. In the following, we will develop the DFGSC-PIC detection scheme having much higher resistance to error propagation. It can continuously track the fast-variant MIMO channel and thus keep the high transmission efficiency.

5.3 DFGSC-PIC and Its Adaptive Realization

In this section, an adaptive two-stage DFGSC-PIC detection scheme for MIMO systems is proposed. In the scheme, a channel estimator is included. The estimated channel information is employed to assist the detection process. With the up-to-date channel information, we will show that the DFGSC-PIC can greatly enhance the detection performance in fast-variant channel environments. As mentioned, the DFGSC-PIC is only applied to the first-stage processing.

Since the structure and the corresponding processing for each data stream are identical in the first stage, for notation simplicity, we will drop the subscript m, and use $\tilde{\mathbf{w}}(k)$ to denote the weight vector and $\tilde{d}(k)$ the output decision for each data stream in the first stage. Consequently, the detection process for each data stream in the first stage, as shown in Fig. 5.1, can be depicted as the left part of Fig. 5.2. Similarly, the subscript m for the channel vector $\mathbf{h}_m(k)$ given in (5.1) is also omitted. We propose to use the MV criterion for the optimization of $\tilde{\mathbf{w}}(k)$. As described, the MMSE criterion cannot provide the MBER performance. Also, it suffers from the error propagation effect when implemented with adaptive structure. It will be shown below that the PIC optimized with the MV criterion is much more robust and resistant to error propagation. The MV criterion determines the weight vector $\tilde{\mathbf{w}}(k)$ for each data stream with the following minimization:

$$\min_{\tilde{\mathbf{w}}(k)} \ \tilde{\mathbf{w}}^H(k)\mathbf{R}(k)\tilde{\mathbf{w}}(k), \qquad \text{subject to } \mathbf{h}^H(k)\tilde{\mathbf{w}}(k) = 1 \tag{5.8}$$

where $\mathbf{R}(k) = \mathrm{E}\{\mathbf{r}(k)\mathbf{r}^H(k)\}$ is the input correlation matrix. The idea behind (5.8) is to minimize the total output power provided that the data symbol $d(k)$ from a particular stream is kept at a constant level (with distortionless response in this case). If we treat the difference between $d(k)$ and the actual output as noise, this MV criterion will give the maximum SNR performance. It is well-known that the GSC is an alternative formulation for the MV criterion. Originally, the GSC was proposed for beamforming in adaptive array processing. With a designated look direction, GSC cancels interference from other directions. For spatial multiplexing in MIMO systems considered here, the channel response $\mathbf{h}(k)$ can be viewed as a spatial direction for $d(k)$. As a result, GSC can be used to

perform MV optimization in MIMO systems. The structure of GSC is depicted within the square marked as "Conventional GSC" in the right part of Fig. 5.2. We first ignore other components illustrated in the figure and concentrate on the operation of the conventional GSC. As shown in the figure, the upper path includes an N-tap signal matched filter $\mathbf{w}_{\mathrm{q}}(k)$. The lower path includes an $N \times (N-1)$ blocking matrix $\mathbf{B}(k)$ and an $(N-1)$-tap interference cancelling filter $\mathbf{w}_{\mathrm{a}}(k)$. The span of $\mathbf{B}(k)$ is designed to fall into the null space of $\mathbf{h}^H(k)$ and so $\mathbf{B}(k)$ can block the desired data symbol with channel $\mathbf{h}(k)$ entering the lower path. We then have an equivalent spatial filter as $\mathbf{w}(k) = \mathbf{w}_{\mathrm{q}}(k) - \mathbf{B}(k)\mathbf{w}_{\mathrm{a}}(k)$. Using the constraint in (5.8), $\mathbf{w}_{\mathrm{q}}(k)$ can be readily found as $\mathbf{w}_{\mathrm{q}}(k) = (\mathbf{h}^H(k)\mathbf{h}(k))^{-1}\mathbf{h}(k)$, and $\mathbf{w}_{\mathrm{a}}(k)$ is optimized according to the output power of the GSC as

$$\begin{aligned} J(k) &= \mathrm{E}\{|y(k)|^2\} \\ &= \mathrm{E}\{|(\mathbf{w}_{\mathrm{q}}(k) - \mathbf{B}(k)\mathbf{w}_{\mathrm{a}}(k))^H \mathbf{r}(k)|^2\}. \end{aligned} \tag{5.9}$$

The constrained optimization problem in (5.8) can then be rewritten as the following unconstrained one.

$$\min_{\mathbf{w}_{\mathrm{a}}(k)} \quad J(k) = \min_{\mathbf{w}_{\mathrm{a}}(k)} \quad (\mathbf{w}_{\mathrm{q}}(k) - \mathbf{B}(k)\mathbf{w}_{\mathrm{a}}(k))^H \mathbf{R}(k)(\mathbf{w}_{\mathrm{q}}(k) - \mathbf{B}(k)\mathbf{w}_{\mathrm{a}}(k)). \tag{5.10}$$

The optimum $\mathbf{w}_{\mathrm{a}}(k)$ is classically solved to be

$$\mathbf{w}_{\mathrm{a,opt}}(k) = (\mathbf{B}(k)^H \mathbf{R}(k)\mathbf{B}(k))^{-1}\mathbf{B}^H(k)\mathbf{R}(k)\mathbf{w}_{\mathrm{q}}(k). \tag{5.11}$$

Let $\mathbf{w}_{\mathrm{opt}}(k) = \mathbf{w}_{\mathrm{q}}(k) - \mathbf{B}(k)\mathbf{w}_{\mathrm{a,opt}}(k)$. With the optimum weight vector, the minimum value of $J(k)$ in (5.9), denoted as $J_{\min}(k)$, can be calculated as

$$J_{\min}(k) = \mathbf{w}_{\mathrm{q}}^H(k)\mathbf{R}(k)\mathbf{w}_{\mathrm{opt}}(k). \tag{5.12}$$

For the case of $M \leq N$, there are enough DOFs for interference cancellation. Thus, the minimum value for (5.12) will be dominated by the power of the desired data symbol, i.e.,

$$J_{\min}(k) \simeq \sigma_d^2 |\mathbf{w}_{\mathrm{q}}^H(k)\mathbf{h}(k)|^2 = \sigma_d^2 \tag{5.13}$$

where σ_d^2 denotes the variance of the desired data symbol. Again, the LMS algorithm is chosen to approach $\mathbf{w}_{\mathrm{a,opt}}(k)$ recursively, i.e.,

$$\mathbf{w}_{\mathrm{a}}(k+1) = \mathbf{w}_{\mathrm{a}}(k) + \mu_{\mathrm{a}}\mathbf{v}(k)e^*(k) \tag{5.14}$$

where $\mathbf{w}_{\mathrm{a}}(k)$ is the estimate of $\mathbf{w}_{\mathrm{a,opt}}(k)$ at the kth time instant, $\mathbf{v}(k) = \mathbf{B}^H(k)\mathbf{x}(k)$ is the filter input vector, μ_{a} is the step size controlling the convergence rate, and $e(k)$ is an error signal. Here, $e(k) = y(k)$. When $\mathbf{w}_{\mathrm{a}}(k)$ is optimized, the error signal $e(k)$ will chiefly consist of the desired data symbol $d(k)$. Due to the large error signal, the step size in the LMS algorithm must be small and this will reduce the tracking capability of the adaptive GSC. The other problem with the conventional GSC is that it is sensitive to model mismatch. In the case of model mismatch, signal cancellation will occur. This can seriously degrade the performance of the GSC. With reference to Chapter 2, a decision feedback operation was introduced to overcome the problems, and the resultant scheme was referred to as the DFGSC. Here, we use this DFGSC for the adaptive realization of the MIMO MV detector. The architecture of DFGSC is shown in the right part of Fig. 5.2. As seen, a feedback tap $w_{\mathrm{b}}(k)$ is added. A new cost function for the optimization of both $\mathbf{w}_{\mathrm{a}}(k)$ and $w_{\mathrm{b}}(k)$ is given by

$$\begin{aligned} J(k) &= \mathrm{E}\{|e(k)|^2\} \\ &= \mathrm{E}\{|y(k) - w_{\mathrm{b}}^*(k)\tilde{d}(k)|^2\} \\ &= \mathrm{E}\{|\mathbf{w}_{\mathrm{q}}^H(k)\mathbf{r}(k) - [\mathbf{w}_{\mathrm{a}}^H(k) \;\; w_{\mathrm{b}}^*(k)] \begin{bmatrix} \mathbf{B}^H(k)\mathbf{r}(k) \\ \tilde{d}(k) \end{bmatrix} |^2\}. \end{aligned} \tag{5.15}$$

For simplicity, we assume that the decision for the desired data symbol is correct, i.e., $\tilde{d}(k) = d(k)$, in the derivation of the optimum $\mathbf{w}_{\mathrm{a}}(k)$ and $w_{\mathrm{b}}(k)$. Following the similar procedure in Chapter 2, we can derive

$$\mathbf{w}_{\mathrm{a,opt}}(k) = (\mathbf{B}^H(k)\mathbf{R}(k)\mathbf{B}(k))^{-1}\mathbf{B}^H(k)\mathbf{R}(k)\mathbf{w}_{\mathrm{q}}(k) \tag{5.16}$$

$$w_{\mathrm{b,opt}}(k) = \mathbf{h}^H(k)\mathbf{w}_{\mathrm{q}}(k) = 1. \tag{5.17}$$

Immediately, we can find that the expression of $\mathbf{w}_{\mathrm{a,opt}}(k)$ in (5.16) is the same as that of the conventional GSC. With $\mathbf{w}_{\mathrm{a,opt}}(k)$ and $w_{\mathrm{b,opt}}(k)$ given above, the minimum $J(k)$ of (5.15) for the DFGSC becomes

$$J_{\min}(k) = \mathbf{w}_{\mathrm{q}}^H(k)\mathbf{R}(k)\mathbf{w}_{\mathrm{opt}}(k) - \sigma_d^2. \tag{5.18}$$

The first term in (5.18) equals the minimum $J(k)$ of the conventional GSC given in (5.12) and the second term in (5.18) is the variance of the desired data symbol. The input

correlation matrix can be decomposed as $\mathbf{R}(k) = \mathbf{R_d}(k) + \mathbf{R_z}(k)$, in which $\mathbf{R_d}(k)$ is the input correlation matrix of the desired data symbol, and $\mathbf{R_z}(k)$ is the input correlation matrix of interference-plus-noise. Then, $J_{\min}(k)$ in (5.18) can be rewritten as

$$J_{\min}(k) = \mathbf{w}_{\mathrm{q}}^{H}(k)\mathbf{R_z}(k)\mathbf{w}_{\mathrm{opt}}(k). \tag{5.19}$$

As we can see, the desired data symbol is totally excluded from $J_{\min}(k)$. When an adaptive algorithm such as the LMS is applied to approach $\mathbf{w}_{\mathrm{a,opt}}(k)$, the performance can be greatly improved by the feedback operation. This is because the error signal, given as $e(k) = y(k) - w_{\mathrm{b}}^{*}(k)\tilde{d}(k)$, does not contain the desired data symbol and its value is much smaller than that of the conventional GSC. As a result, the stochastic gradient is small and a large step size can be used to enhance the tracking capability of the GSC. The LMS update equations for the DFGSC can be expressed as

$$\mathbf{w}_{\mathrm{a}}(k+1) = \mathbf{w}_{\mathrm{a}}(k) + \mu_{\mathrm{a}}\mathbf{v}(k)e^{*}(k) \tag{5.20}$$

$$w_{\mathrm{b}}(k+1) = w_{\mathrm{b}}(k) + \mu_{\mathrm{b}}\tilde{d}(k)e^{*}(k) \tag{5.21}$$

where μ_{a} is the step size for $\mathbf{w}_{\mathrm{a}}(k)$, μ_{b} is the step size for $w_{\mathrm{b}}(k)$, $\mathbf{v}(k)$ is the input vector for $\mathbf{w}_{\mathrm{a}}(k)$, and $e(k)$ is the error signal given above.

The feedback operation can also make the DFGSC robust to model mismatch. For example, in time-variant channel environments, $\mathbf{w}_{\mathrm{q}}(k)$ may not always match to $\mathbf{h}(k)$. In this case, $\mathbf{B}(k)$ cannot obstruct the desired data symbol. If there are enough DOFs, $\mathbf{w}_{\mathrm{a}}(k)$ will be adapted to cancel the desired data symbol and the performance of the conventional GSC may be degraded severely. This phenomenon is referred to as signal cancellation. Also shown in Chapter 2, the signal cancellation phenomenon can be avoided in the DFGSC. Now, the optimum solutions for $\mathbf{w}_{\mathrm{a}}(k)$ and $w_{\mathrm{b}}(k)$ in the DFGSC are different from those in conventional GSC. Following the procedure in Chapter 2, we can derive

$$\mathbf{w}_{\mathrm{a,opt}}(k) = (\bar{\mathbf{B}}^{H}(k)\mathbf{R_z}(k)\bar{\mathbf{B}}(k))^{-1}\bar{\mathbf{B}}^{H}(k)\mathbf{R_z}(k)\bar{\mathbf{w}}_{\mathrm{q}}(k) \tag{5.22}$$

$$w_{\mathrm{b,opt}}(k) = \mathbf{h}^{H}(k)(\bar{\mathbf{w}}_{\mathrm{q}}(k) - \bar{\mathbf{B}}(k)\mathbf{w}_{\mathrm{a,opt}}(k)). \tag{5.23}$$

where $\bar{\mathbf{w}}_{\mathrm{q}}(k)$ and $\bar{\mathbf{B}}(k)$ correspond to $\mathbf{w}_{\mathrm{q}}(k)$ and $\mathbf{B}(k)$ calculated based on the mismatched channel estimate $\bar{\mathbf{h}}(k)$, respectively. The minimum $J(k)$ of the DFGSC in this case can

be calculated as

$$J_{\min}(k) = \bar{\mathbf{w}}_{\mathrm{q}}^{H}(k)\mathbf{R}_{\mathbf{z}}(k)\mathbf{w}_{\mathrm{opt}}(k) \tag{5.24}$$

with $\mathbf{w}_{\mathrm{opt}}(k) = \bar{\mathbf{w}}_{\mathrm{q}}(k) - \bar{\mathbf{B}}(k)\mathbf{w}_{\mathrm{a,opt}}(k)$. From (5.22) to (5.24), we can see that $J_{\min}(k)$ (and the error signal) still contains no desired data symbol. Although the signal leakage is observed in the DFGSC, it becomes invisible in $\mathbf{w}_{\mathrm{a}}(k)$ adaptation. This is the reason why no signal cancellation occurs.

We have already seen that the DFGSC can resist model mismatch. Thus, signal cancellation will not occur when the channel response changes. However, in this mismatch scenario, $\mathbf{w}_{\mathrm{q}}(k)$ cannot match to the channel and it will reduce the signal power in the upper path of the DFGSC. Also, $\mathbf{B}(k)$ cannot block the desired data symbol and this leaky component will increase the interference power. As a result, the overall performance of the DFGSC will be degraded. The degradation can be ignored when the DFGSC is operated in time-varying MIMO channels within a short period of time. To keep the good performance over a long-term period, however, $\mathbf{w}_{\mathrm{q}}(k)$ and $\mathbf{B}(k)$ have to be properly updated. Here, we propose a simple method to do the job. First, define a diagonal matrix as

$$\mathbf{P}(k) = \mathrm{Diag}\{(\mathbf{h}^{H}(k)\mathbf{h}(k))^{-1}\mathbf{h}(k)\}. \tag{5.25}$$

This matrix, called the steering matrix, is used to *pre-steer* the look direction of the DFGSC. In other words, the received signal vector is first multiplied with the matrix. With this operation, $\mathbf{w}_{\mathrm{q}}(k)$ will become a time-invariant vector with components of all ones, and $\mathbf{B}(k)$ can be any time-invariant orthogonal matrix, denoted as $\mathbf{B}$. With $\mathbf{B}$ chosen carefully, the computational complexity for signal blocking can be significantly reduced. For the case of $N = 2^l$, where l is any nonnegative integer, a simple choice for $\mathbf{B}$ is the Hadamard matrix (excluding the first column). For the case that $N \neq 2^l$, $\mathbf{B}$ can still be designed to achieve low complexity as reported in [85]. With this architecture, only $\mathbf{P}(k)$ has to be updated.

Another problem is how to acquire the up-to-date channel response. With feedback decisions, this can be easily solved using a channel estimator. Let the coefficients of the channel estimator be denoted as $\mathbf{q}(k)$, which is an $N \times 1$ vector. It can be tuned by a

new error signal vector $\mathbf{e}_{\mathrm{q}}(k)$, and the optimization problem can be written as

$$\min_{\mathbf{q}(k)} \mathrm{E}\{\|\mathbf{e}_{\mathrm{q}}(k)\|^2\} = \min_{\mathbf{q}(k)} \mathrm{E}\{\|\mathbf{r}(k) - \mathbf{q}(k)\hat{d}(k)\|^2\}. \tag{5.26}$$

To have better performance, we use the output in the second stage of the PIC, $\hat{d}(k)$, as the input to the channel estimator. Also note that the estimation exists over all parallel channels. With the assumption of correct decisions, it is not difficult to see that the optimum $\mathbf{q}(k)$, denoted as $\mathbf{q}_{\mathrm{opt}}(k)$, will be equal to $\mathbf{h}(k)$. We still use the LMS algorithm to approach $\mathbf{q}_{\mathrm{opt}}(k)$ recursively. The update equation is given as

$$\mathbf{q}(k+1) = \mathbf{q}(k) + \mu_{\mathrm{q}}\hat{d}^*(k)\mathbf{e}_{\mathrm{q}}(k) \tag{5.27}$$

where μ_{q} is the step size for the adaptation. We observe that the input to $\mathbf{q}(k)$ is a white sequence, which owns the smallest eigenvalue spread of the input correlation matrix. Since the convergence rate of (5.27) is inversely proportional to the eigenvalue spread [45], the adaptation of $\mathbf{q}(k)$ is expected to be fast and stable. In Section 5.5, we will show empirically that this estimation mechanism can successfully track the channel variations.

With the application of the steering matrix, the adaptive DFGSC can be operated in the time-variant channel environment. However, since the steering matrix and $\mathbf{w}_{\mathrm{a}}(k)$ are connected in series, continuous update of the matrix may yield always un-convergent $\mathbf{w}_{\mathrm{a}}(k)$. Since DFGSC can resist channel mismatch, there is no need to update the steering matrix continuously; periodic update is more appropriate. There is also one problem associated with periodic update. This is because the abrupt change of the steering matrix will make $\mathbf{w}_{\mathrm{a}}(k)$ deviate from its optimum state instantaneously. As a result, the performance of the system will be degraded until $\mathbf{w}_{\mathrm{a}}(k)$ re-converges. To solve this problem, we proposed the use of a dual-DFGSC configuration, as illustrated in Fig. 5.3, for each branch in the first stage. These two adaptive DFGSCs are operated simultaneously and complementarily. We let $\mathbf{P}(k)$ and $\tilde{d}(k)$ in the first DFGSC be donated as $\mathbf{P}_1(k)$ and $\tilde{d}_1(k)$, and those in the second DFGSC as $\mathbf{P}_2(k)$ and $\tilde{d}_2(k)$, respectively. Both $\mathbf{P}_1(k)$ and $\mathbf{P}_2(k)$ are updated periodically in different time instants, and only one of $\tilde{d}_1(k)$ and $\tilde{d}_2(k)$ is selected as the decision passed to the second stage and used as the reference signal for the dual-DFGSC adaptation. Let the time origin be zero, L be the update period and

$k = 2jL + l$, where j is a nonnegative integer and $0 \leq l \leq 2L - 1$. Then, $\mathbf{P}_1(k)$ and $\mathbf{P}_2(k)$ can be express as

$$\mathbf{P}_1(2jL + l) = \mathbf{P}(2jL) \tag{5.28}$$

$$\mathbf{P}_2(2jL + L + l) = \mathbf{P}(2jL + L) \tag{5.29}$$

where $\mathbf{P}(k)$ is as that in (5.25). Also, $\mathbf{P}_2(l) = \mathbf{P}(0)$ for $0 \leq l \leq L - 1$. The decision passed to the second stage and used for the dual-DFGSC adaptation can be expressed as

$$\tilde{d}(2jL + l) = \begin{cases} \tilde{d}_2(2jL + l), & \text{if } 2jL \leq l \leq 2jL + L - 1 \\ \tilde{d}_1(2jL + l), & \text{if } 2jL + L \leq l \leq 2jL + 2L - 1 \end{cases}. \tag{5.30}$$

Fig. 5.4 shows the update and decision time relation for the dual-DFGSC structure. With this approach, the DFGSC-PIC can keep good track to the channel variations and provide smooth operation for channel updating. In the next section, we will provide a guideline for the determination of L. As we can see, the dual structure will increase the computational complexity. Fortunately, as mentioned, the structure is only applied in the first stage. The weight vector $\hat{\mathbf{w}}(k)$ for each branch in the second stage (the subscript m is omitted also) only performs the matching operation, i.e.,

$$\hat{\mathbf{w}}(k) = (\mathbf{q}^H(k)\mathbf{q}(k))^{-1}\mathbf{q}(k) \tag{5.31}$$

where $\mathbf{q}(k)$ is the up-to-date channel estimate for the corresponding channel.

5.4 Convergence Analysis

In this section, we will analyze the convergence behavior of the proposed DFGSC operated in time-varying channel environments. The analysis is difficult and has not been done before even for the conventional GSC. To let our analysis mathematically traceable, we have to make some assumptions. First, we assume that each coefficient in the MIMO channel varies independently with Jakes' model and the channel vector is normalized (made possible by the automatic gain control in practice). Other assumptions will be given later whenever necessary. Recall that the correlation function of a normalized time-varying channel vector $\mathbf{h}(k)$ can be modeled as a zeroth-order Bessel function of the first

kind [86], namely,

$$\begin{aligned} r_{\mathrm{hh}}(\tau) &\triangleq \mathrm{E}\{\mathbf{h}^H(k)\mathbf{h}(k-\tau)\} \\ &= \mathfrak{J}_0(2\pi f_{\mathrm{d}} T_{\mathrm{s}} \tau) \end{aligned} \tag{5.32}$$

in which τ is the time lag, f_{d} is the Doppler frequency and T_{s} is the symbol duration, and the function $\mathfrak{J}_0(\cdot)$ is the zeroth-order Bessel function of the first kind, which is defined by

$$\mathfrak{J}_0(z) \triangleq \frac{1}{\pi}\int_0^{\pi} \cos(z\sin\theta)d\theta. \tag{5.33}$$

We assume that the channel variations are small in a short period of time and the channel vector can be described by a random walk process as

$$\mathbf{h}(k+1) = \mathbf{h}(k) + \boldsymbol{\eta}(k) \tag{5.34}$$

where $\boldsymbol{\eta}(k)$ denotes a white noise vector with the variance of each component as σ_η^2. The variance can be calculated as

$$\begin{aligned} N\sigma_\eta^2 &= \mathrm{E}\{\| \mathbf{h}(k) - \mathbf{h}(k+1) \|^2\} \\ &= \mathrm{E}\{\| \mathbf{h}(k) \|^2\} + \mathrm{E}\{\| \mathbf{h}(k+1) \|^2\} - 2\mathrm{E}\{\mathbf{h}^H(k)\mathbf{h}(k+1)\} \\ &= 2 - 2r_{\mathrm{hh}}(1). \end{aligned} \tag{5.35}$$

Thus, $\sigma_\eta^2 = (2 - 2r_{\mathrm{hh}}(1))/N$. For notation simplicity, we drop the time index for $\mathbf{w}_{\mathrm{q}}(k)$ and $\mathbf{B}(k)$ since they are constants during an update period. Let $\mathbf{B} = [\mathbf{b}_1 \;\; \mathbf{b}_2 \;\; \cdots \;\; \mathbf{b}_{N-1}]$ and $\mathbf{w}_{\mathrm{a,opt}}(k) = [w_1(k) \;\; w_2(k) \;\; \cdots \;\; w_{N-1}(k)]^T$ in the following analysis. Due to the decision feedback operation and small channel variations, the desired data symbol can be assumed to be totally eliminated in the optimization. Also, the leaky component in the output of $\mathbf{B}$ is small and can be ignored. To simplify the problem, we first consider a noiseless environment with only one interference, in which $\mathbf{h}(k)$ denotes the interfering channel vector. For perfect interference cancellation, we have

$$\begin{aligned} \mathbf{w}_{\mathrm{q}}^H\mathbf{h}(k) &= \mathbf{w}_{\mathrm{a,opt}}^H(k)\mathbf{B}^H\mathbf{h}(k) \\ &= w_1^*(k)\mathbf{b}_1^H\mathbf{h}(k) + w_2^*(k)\mathbf{b}_2^H\mathbf{h}(k) + \cdots + w_{N-1}^*(k)\mathbf{b}_{N-1}^H\mathbf{h}(k). \end{aligned} \tag{5.36}$$

At time instant $k+1$, $\mathbf{h}(k+1)$ is also perfectly cancelled. We shall have

$$\mathbf{w}_{\mathrm{q}}^H\mathbf{h}(k+1) = \mathbf{w}_{\mathrm{a,opt}}^H(k+1)\mathbf{B}^H\mathbf{h}(k+1) \tag{5.37}$$

Substituting (5.34) and (5.36) into (5.37), we have

$$\begin{aligned} & w_1^*(k)\mathbf{b}_1^H\mathbf{h}(k) + w_2^*(k)\mathbf{b}_2^H\mathbf{h}(k) + \cdots + w_{N-1}^*(k)\mathbf{b}_{N-1}^H\mathbf{h}(k) + \mathbf{w}_{\mathrm{q}}^H\boldsymbol{\eta}(k) \\ = & \, w_1^*(k+1)\mathbf{b}_1^H(\mathbf{h}(k)+\boldsymbol{\eta}(k)) + w_2^*(k+1)\mathbf{b}_2^H(\mathbf{h}(k)+\boldsymbol{\eta}(k)) + \cdots \\ & + w_{N-1}^*(k+1)\mathbf{b}_{N-1}^H(\mathbf{h}(k)+\boldsymbol{\eta}(k)). \end{aligned} \tag{5.38}$$

Now, assume that the variation term $\mathbf{w}_{\mathrm{q}}^H\boldsymbol{\eta}(k)$ in (5.38) can be evenly cancelled by each weight element in $\mathbf{w}_{\mathrm{a,opt}}(k)$. We then have

$$w_n^*(k)\mathbf{b}_n^H\mathbf{h}(k) + \frac{1}{N-1}\mathbf{w}_{\mathrm{q}}^H\boldsymbol{\eta}(k) = w_n^*(k+1)\mathbf{b}_n^H(\mathbf{h}(k)+\boldsymbol{\eta}(k)) \tag{5.39}$$

for $1 \le n \le N-1$. From (5.39), we can obtain

$$\begin{aligned} w_n^*(k+1) &= w_n^*(k)\frac{\mathbf{b}_n^H\mathbf{h}(k)}{\mathbf{b}_n^H(\mathbf{h}(k)+\boldsymbol{\eta}(k))} + \frac{1}{N-1}\frac{\mathbf{w}_{\mathrm{q}}^H\boldsymbol{\eta}(k)}{\mathbf{b}_n^H(\mathbf{h}(k)+\boldsymbol{\eta}(k))} \\ &\simeq w_n^*(k) + \frac{1}{N-1}\frac{\mathbf{w}_{\mathrm{q}}^H\boldsymbol{\eta}(k)}{\mathbf{b}_n^H\mathbf{h}(k)}. \end{aligned} \tag{5.40}$$

Thus, $\mathbf{w}_{\mathrm{a,opt}}(k)$ can be approximated as

$$\mathbf{w}_{\mathrm{a,opt}}(k+1) = \mathbf{w}_{\mathrm{a,opt}}(k) + \boldsymbol{\omega}(k) \tag{5.41}$$

where $\boldsymbol{\omega}(k)$ is the process noise vector with its elements defined as the second term in (5.40). From (5.41), we can see that the optimum weight vector $\mathbf{w}_{\mathrm{a,opt}}(k)$ can also be modeled as a random walk process. Note that $\mathrm{E}\{|\mathbf{w}_{\mathrm{q}}^H\boldsymbol{\eta}(k)|^2\} = \sigma_\eta^2$ and $\mathrm{E}\{|\mathbf{b}_n^H\mathbf{h}(k)|^2\} = 1/N$. Thus, the variance of each component in $\boldsymbol{\omega}(k)$ is then $N\sigma_\eta^2/(N-1)^2$. For the scenario of multiple interfering streams, (5.36) and (5.37) can be modified to include multiple channel vectors. With the similar method, we can derive the same result. Using the model in (5.41), we can formulate the adaptation of $\mathbf{w}_{\mathrm{a}}(k)$ as a time-variant system identification problem. As assumed, the desired data symbol is totally eliminated by the feedback operation. Thus, we can deem that the whole DFGSC as an interference only system. Now, the input to the system is $\mathbf{B}^H\acute{\mathbf{H}}(k)\mathbf{i}(k)$, where $\acute{\mathbf{H}}(k)$ is the channel matrix excluding the column corresponding to the desired data symbol and $\mathbf{i}(k)$ is the vector consisting of interfering data streams. The output of the system is then $\mathbf{w}_{\mathrm{a,opt}}^H(k)\mathbf{B}^H\acute{\mathbf{H}}(k)\mathbf{i}(k) + z(k)$, where $z(k)$ is white noise with zero mean and variance σ_z^2. Then, the filter $\mathbf{w}_{\mathrm{a}}(k)$ is used to identify $\mathbf{w}_{\mathrm{a,opt}}(k)$. Also assume that $\boldsymbol{\omega}(k)$ is independent to both the input vector

$\mathbf{B}^H\acute{\mathbf{H}}(k)\mathbf{i}(k)$ and the noise $z(k)$. With these settings, the error signal $e(k)$, involved in the adaptation of $\mathbf{w}_{\mathrm{a}}(k)$, can be stated as

$$e(k) = \mathbf{w}_{\mathrm{a,opt}}^H(k)\mathbf{B}^H\acute{\mathbf{H}}(k)\mathbf{i}(k) + z(k) - \mathbf{w}_{\mathrm{a}}^H(k)\mathbf{B}^H\acute{\mathbf{H}}(k)\mathbf{i}(k). \tag{5.42}$$

Recall that the update equation for $\mathbf{w}_{\mathrm{a}}(k)$ is given in (5.20). We let $\mathbf{v}(k) = \mathbf{B}^H\acute{\mathbf{H}}(k)\mathbf{i}(k)$ and substitute (5.42) for $e(k)$ in (5.20), we may reformulate the weight update in the expanded form as

$$\mathbf{w}_{\mathrm{a}}(k+1) = (\mathbf{I} - \mu_{\mathrm{a}}\mathbf{v}(k)\mathbf{v}^H(k))\mathbf{w}_{\mathrm{a}}(k) + \mu_{\mathrm{a}}\mathbf{v}(k)\mathbf{v}^H(k)\mathbf{w}_{\mathrm{a,opt}}(k) + \mu_{\mathrm{a}}\mathbf{v}(k)z^*(k). \tag{5.43}$$

Also define the weight-error vector as $\boldsymbol{\epsilon}(k) = \mathbf{w}_{\mathrm{a}}(k) - \mathbf{w}_{\mathrm{a,opt}}(k)$. Using the expression in (5.41), we may write

$$\boldsymbol{\epsilon}(k+1) = (\mathbf{I} - \mu_{\mathrm{a}}\mathbf{v}(k)\mathbf{v}^H(k))\boldsymbol{\epsilon}(k) + \mu_{\mathrm{a}}\mathbf{v}(k)z^*(k) - \boldsymbol{\omega}(k). \tag{5.44}$$

With the direct-averaging method [45], we replace the matrix $(\mathbf{I} - \mu_{\mathrm{a}}\mathbf{v}(k)\mathbf{v}^H(k))$ with its ensemble average $(\mathbf{I} - \mu_{\mathrm{a}}\mathbf{R}_{\mathbf{v}})$, where $\mathbf{R}_{\mathbf{v}}$ is the correlation matrix of the input vector $\mathbf{v}(k)$. The stochastic difference equation can be expressed as

$$\boldsymbol{\epsilon}(k+1) = (\mathbf{I} - \mu_{\mathrm{a}}\mathbf{R}_{\mathbf{v}})\boldsymbol{\epsilon}(k) + \mu_{\mathrm{a}}\mathbf{v}(k)z^*(k) - \boldsymbol{\omega}(k). \tag{5.45}$$

Invoking the independence assumption [45], the weight-error correlation matrix, i.e., $\mathbf{K}(k) = \mathrm{E}\{\boldsymbol{\epsilon}(k)\boldsymbol{\epsilon}^H(k)\}$, is readily determined from (5.45) as

$$\mathbf{K}(k+1) = (\mathbf{I} - \mu_{\mathrm{a}}\mathbf{R}_{\mathbf{v}})\mathbf{K}(k)(\mathbf{I} - \mu_{\mathrm{a}}\mathbf{R}_{\mathbf{v}}) + \mu_{\mathrm{a}}^2\sigma_z^2\mathbf{R}_{\mathbf{v}} + \boldsymbol{\Omega} \tag{5.46}$$

in which $\boldsymbol{\Omega}$ is the covariance matrix of $\boldsymbol{\omega}(k)$ given by

$$\boldsymbol{\Omega} = \frac{N}{(N-1)^2}\sigma_\eta^2\mathbf{I}. \tag{5.47}$$

Assume that μ_{a} is small such that (5.46) will not diverge. In this case, we can see that $\mathbf{K}(k)$ will reach a steady state and (5.46) can be reduced to

$$\mathbf{R}_{\mathbf{v}}\mathbf{K}(k) + \mathbf{K}(k)\mathbf{R}_{\mathbf{v}} \simeq \mu_{\mathrm{a}}\sigma_z^2\mathbf{R}_{\mathbf{v}} + \frac{1}{\mu_{\mathrm{a}}}\boldsymbol{\Omega}. \tag{5.48}$$

One of the commonly used figure of merit for adaptive algorithms is the misadjustment $\mathfrak{M}(k)$, which can be defined by

$$\mathfrak{M}(k) = \frac{\mathrm{Tr}\{\mathbf{R}_{\mathbf{v}}\mathbf{K}(k)\}}{\sigma_z^2}. \tag{5.49}$$

Using (5.48), we can then have

$$\mathrm{Tr}\{\mathbf{R_v}\mathbf{K}(k)\} + \mathrm{Tr}\{\mathbf{K}(k)\mathbf{R_v}\} \simeq \mu_\mathrm{a}\sigma_z^2\mathrm{Tr}\{\mathbf{R_v}\} + \frac{1}{\mu_\mathrm{a}}\mathrm{Tr}\{\mathbf{\Omega}\} \tag{5.50}$$

and then the misadjustment of the adaptive DFGSC is approximated as

$$\mathfrak{M}(k) \simeq \frac{\mu_\mathrm{a}}{2}\mathrm{Tr}\{\mathbf{R_v}\} + \frac{1}{2\mu_\mathrm{a}\sigma_z^2}\mathrm{Tr}\{\mathbf{\Omega}\} \tag{5.51}$$

for large k. Taking the derivative of (5.51) with respect to μ_a and setting the result to zero, we readily find the optimum step size as

$$\mu_{\mathrm{a,opt}} \simeq \sqrt{\frac{\mathrm{Tr}\{\mathbf{\Omega}\}}{\sigma_z^2\mathrm{Tr}\{\mathbf{R_v}\}}}. \tag{5.52}$$

With the step size provided in (5.52), the misadjustment can be minimized.

Now, we discuss the determination of the updating period L. The period should be long enough for $\mathbf{w}_\mathrm{a}(k)$ to converge, and short enough for channel mismatch to remain small. The settling time of the LMS algorithm is proportional to the average time constant, which is approximated as [45]

$$\tau_\mathrm{av} \approx \frac{1}{2\mu_\mathrm{a}\lambda_\mathrm{av}} \tag{5.53}$$

where λ_av is the average eigenvalue for the underlying correlation matrix, i.e., $\mathbf{R_v}$ in our scenario. As a rule of thumb, we choose two times of this average time constant for a good trade-off between the convergence of $\mathbf{w}_\mathrm{a}(k)$ and channel mismatch.

5.5 Simulations

Computer simulations are conducted to demonstrate the effectiveness of the proposed adaptive two-stage DFGSC-PIC detection scheme for time-varying MIMO channels. We first examine the validity of the optimum step size derived in (5.52). For an adaptive DFGSC, we experimentally calculate the interference cancellation MSE defined as

$$\mathrm{E}\{|(\mathbf{w}_\mathrm{q} - \mathbf{B}\mathbf{w}_\mathrm{a}(k))^H(\acute{\mathbf{H}}(k)\mathbf{i}(k) + \mathbf{n}(k))|^2\}. \tag{5.54}$$

For the calculation of $\mu_{\mathrm{a,opt}}$ in (5.52), we set $z(k)$ in (5.42) to be the noise in the upper path, i.e., $\mathbf{w}_\mathrm{q}^H\mathbf{n}(k)$. The MIMO channel is drawn from independent flat fading. The normalized

Doppler frequencies $f_{\rm d}T_{\rm s}$ equal 1×10^{-4} and 5×10^{-4}, respectively. For example, it is to say that the data rate used is about 4.6×10^5 symbols/s and the carrier frequency is 2.4 GHz. The vehicle speed is set as about 20 km/h and 100 km/h, respectively. Assume that the channel response is known at the beginning. The number of antennas is set at $M = 2$ and $N = 4$. All transmitted symbols are randomly drawn and then modulated by 16-QAM. The SNR is 30 dB. We have $\mu_{\rm a,opt} \simeq 1.2 \times 10^{-3}$ for $f_{\rm d}T_{\rm s} = 1 \times 10^{-4}$ and $\mu_{\rm a,opt} \simeq 6.0 \times 10^{-3}$ for $f_{\rm d}T_{\rm s} = 5 \times 10^{-4}$. In addition, $L = 2\tau_{\rm av} \simeq 85$ for $\mu_{\rm a,opt} \simeq 1.2 \times 10^{-3}$, and $L = 2\tau_{\rm av} \simeq 17$ for $\mu_{\rm a,opt} \simeq 6.0 \times 10^{-3}$. Fig. 5.5 shows the interference cancellation MSE versus different $\mu_{\rm a}$. As we can see, the MSE yielded by the optimum step size is close to the minimum.

We then investigate the convergence behavior of the DFGSC-PIC and show the learning curves of the output SINR first. For comparison, we also consider the MMSE-PIC detector. Again, the channels are independent flat fading and generated by Jakes' model. The transmitted data symbols are grouped into blocks and the block size is 1024. The channel response is assumed to be known at the beginning of each block. Other settings for the proposed DFGSC-PIC, including the step size and the update period, are the same as those given in Fig. 5.5. All tracking operations are based on the LMS algorithm. The same channel tracking mechanism as given in Section 5.3 is used for the MMSE-PIC. The step sizes in LMS adaptation for the MMSE-PIC are chosen to optimize the performance as well. Fig. 5.6 shows the learning curves of the output SINR in the first and second stages of the DFGSC-PIC and the MMSE-PIC with $f_{\rm d}T_{\rm s} = 1 \times 10^{-4}$. In terms of SINR, we see that the proposed DFGSC-PIC detector can usually perform better than the MMSE-PIC detector. For the DFGSC-PIC in the first stage, it is clear that the update of the steering matrix does not affect the learning behavior. Although the channel varies over time (relatively slowly), both PIC schemes keep the good performance throughout the block. However, as shown in Fig. 5.7, when $f_{\rm d}T_{\rm s}$ is increased to 5×10^{-4}, it becomes difficult for the MMSE-PIC to track this rapid changing environment. The SINR values for both stages of the MMSE-PIC degrades continuously with time. For the proposed DFGSC-PIC, the output SINRs in both stages remain the same. This confirms that the our DFGSC-PIC is much more robust to fast time-varying MIMO channels.

We next show the SER performance for the environments described previously. We compare the DFGSC-PIC with the MMSE-PIC and the V-BLAST system with parameters updated blockwisely. The V-BLAST system re-calculates the detection order and weight vectors with the MMSE criterion at the beginning of each block. Hence, the smaller the block length is, the better the performance is expected, but the higher the computational complexity it requires. Usually, the V-BLAST system can perform very well provided that the channel is exactly known. However, if the V-BLAST system only updates in a blockwise manner in time-varying channel environments, its performance degrades significantly, especially when the update interval is large. In general, the computational complexity for calculating both the detection order and weight vectors is on the order of $O(N^3)$ [87]. While the DFGSC-PIC is more complex than the MMSE-PIC, the computational complexity of both PIC schemes keeps on the order of $O(N)$. Fig. 5.8 shows the resultant SER against different SNR values with $f_{\mathrm{d}}T_{\mathrm{s}} = 1 \times 10^{-4}$. As demonstrated, the proposed DFGSC-PIC and the MMSE-PIC performs almost the same. We see that the V-BLAST system needs to reduce the block length to 150 to have the comparable performance of both PIC schemes. In Fig. 5.9, we change $f_{\mathrm{d}}T_{\mathrm{s}}$ to 5×10^{-4} and re-run the simulations. This time the performance of the MMSE-PIC becomes poor, but our DFGSC-PIC still performs quite well. In order for the V-BLAST system to come close to the performance of the proposed DFGSC-PIC, the block length should be further shortened to 50.

Finally, Fig. 5.10 presents the SER against different $f_{\mathrm{d}}T_{\mathrm{s}}$ values when the SNR is fixed at 30 dB. We observe that the DFGSC-PIC outperforms the MMSE-PIC from moderate to high $f_{\mathrm{d}}T_{\mathrm{s}}$ regions. It also outperforms the V-BLAST system with block length 50 in high $f_{\mathrm{d}}T_{\mathrm{s}}$ region.

5.6 Summary

In this chapter, a new adaptive two-stage DFGSC-PIC detection scheme has been developed for time-variant MIMO systems. The proposed scheme utilizes the MV criterion for PIC optimization. Using a specially designed dual-DFGSC structure, we are able to realize the adaptive MV detector. The proposed scheme can have better tracking capa-

bility than the MMSE-PIC detector, and also have the resistance to model mismatch. This provides an effective yet low-complexity solution for spatial multiplexing in MIMO systems. Simulations show that the proposed DFGSC-PIC can significantly outperform the MMSE-PIC in fast-variant channel environments. We also compare the DFGSC-PIC with the blockwisely updated V-BLAST system. It is shown that in fast-variant environments, the update frequency of the V-BLAST scheme has to be high in order to have the performance comparable to that of the DFGSC-PIC. The high computational complexity requirement often makes the V-BLAST difficult to apply in real-world implementations.

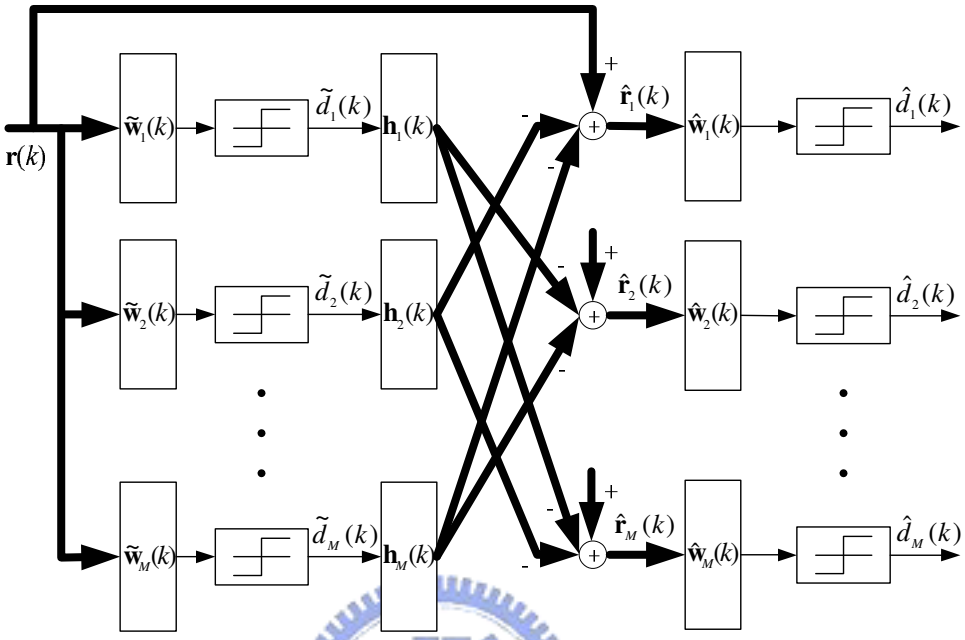

Figure 5.1: General two-stage PIC detection scheme for MIMO systems.

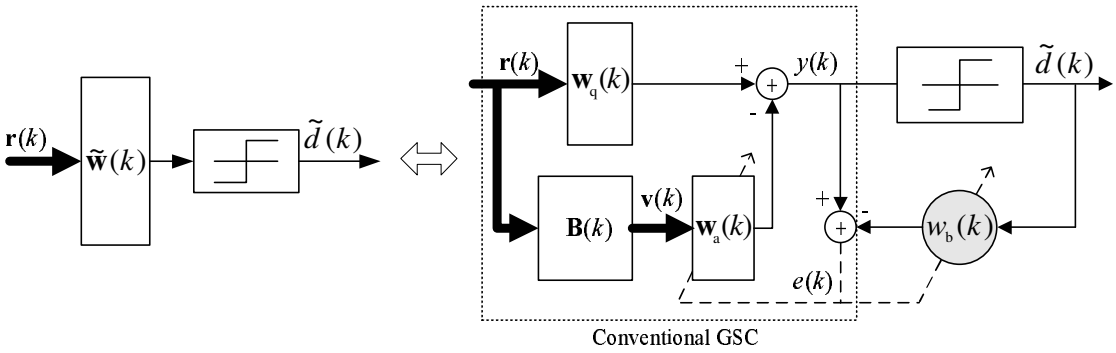

Figure 5.2: Adaptive DFGSC for realization of weight vector in the first stage of PIC.

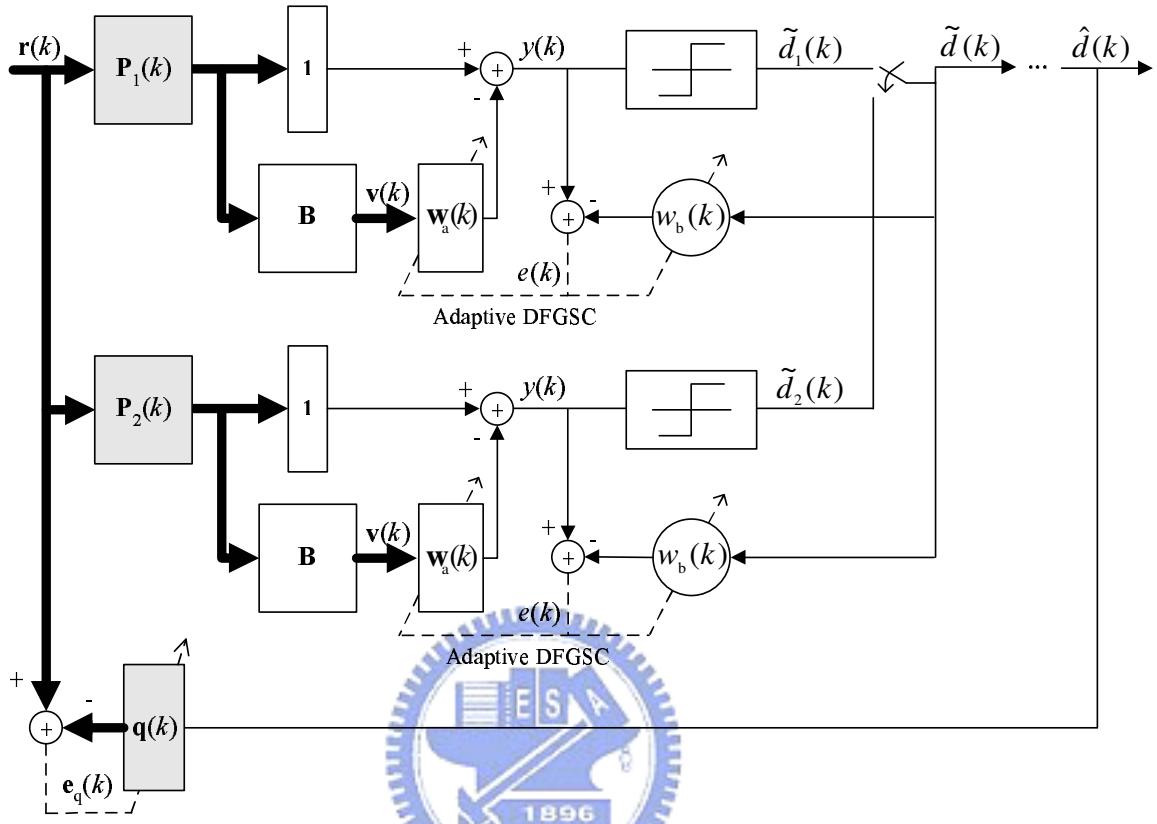

Figure 5.3: Adaptive DFGSC with steering matrix $\mathbf{P}(k)$ and channel estimator $\mathbf{q}(k)$. Dual-DFGSC works complementarily for each branch in the first stage of DFGSC-PIC.

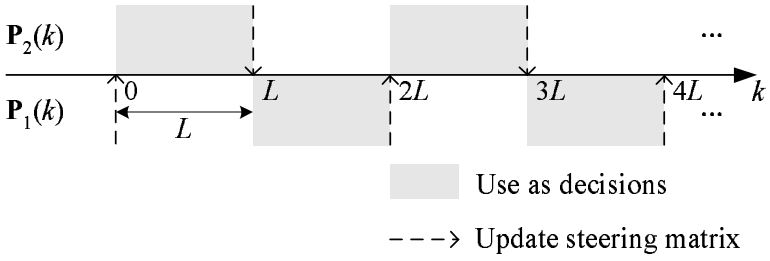

Figure 5.4: Update and decision time relation for dual-DFGSC.

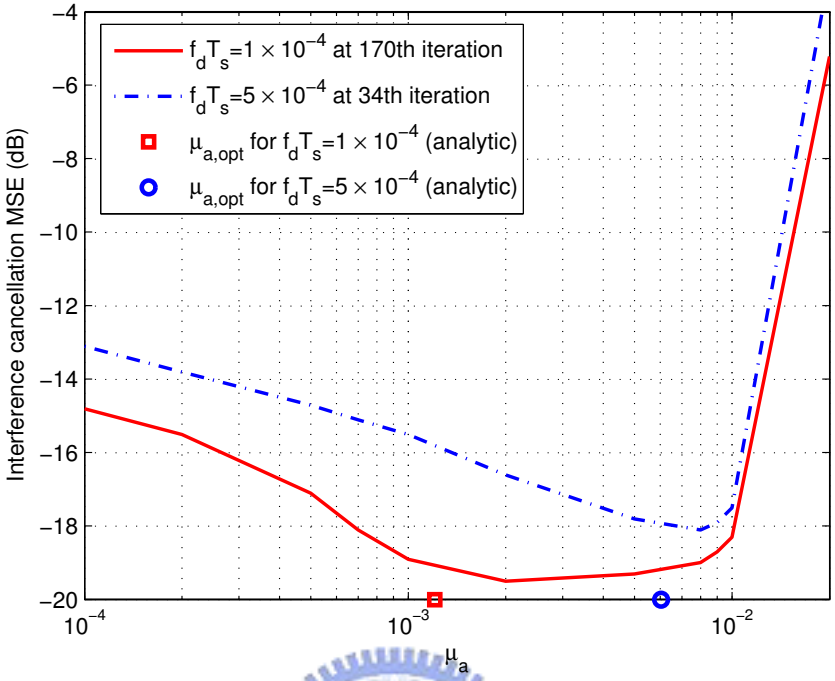

Figure 5.5: Interference cancellation MSE with different μ_{a}.

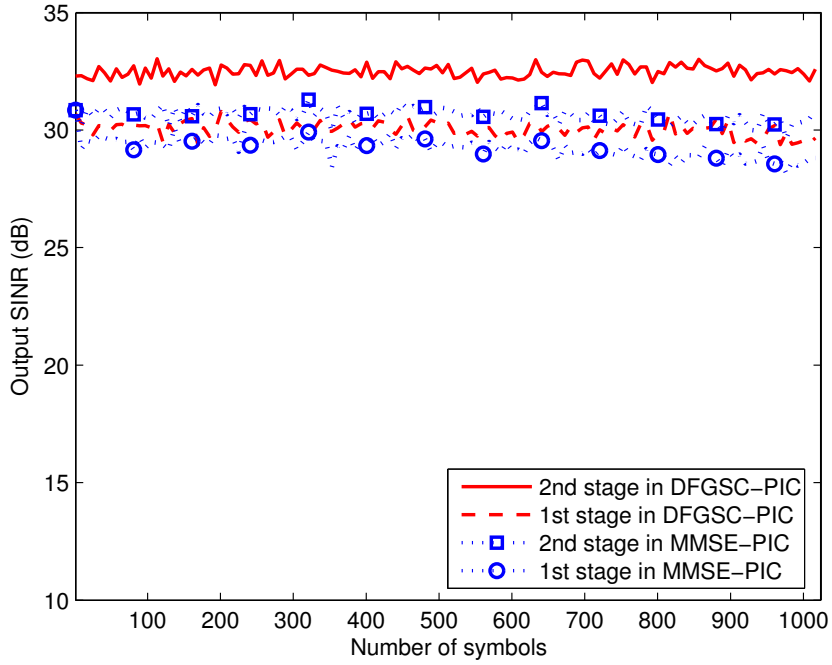

Figure 5.6: Learning curves of PIC output SINR with $f_{\mathrm{d}}T_{\mathrm{s}} = 1 \times 10^{-4}$ and SNR = 30 dB.

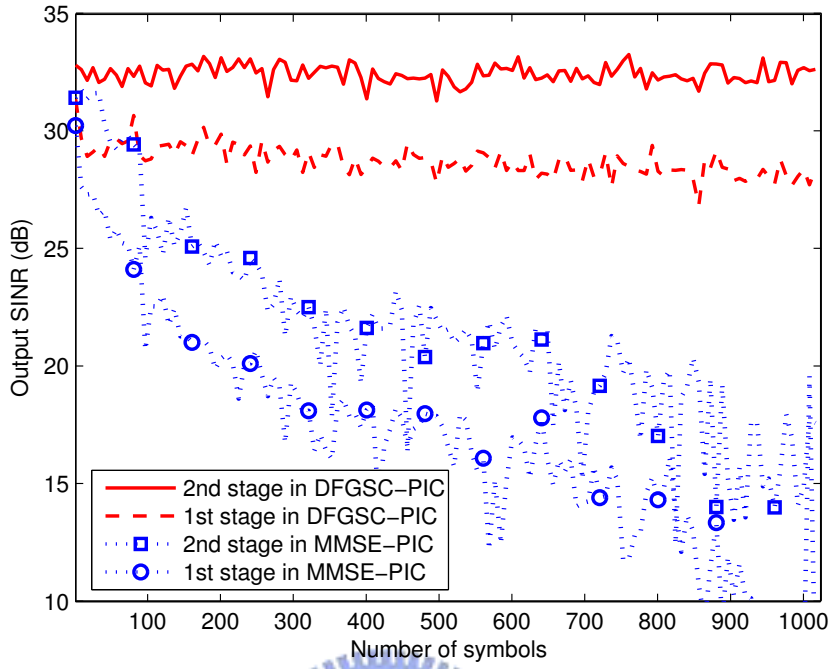

Figure 5.7: Learning curves of PIC output SINR with $f_{\mathrm{d}}T_{\mathrm{s}} = 5 \times 10^{-4}$ and SNR = 30 dB.

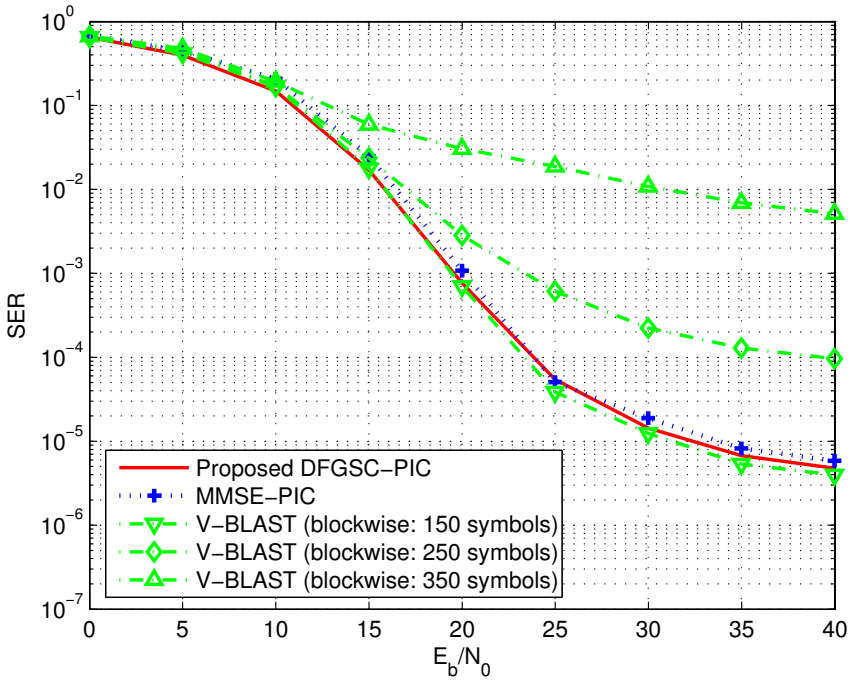

Figure 5.8: SER for various SNR values with $f_{\mathrm{d}}T_{\mathrm{s}} = 1 \times 10^{-4}$.

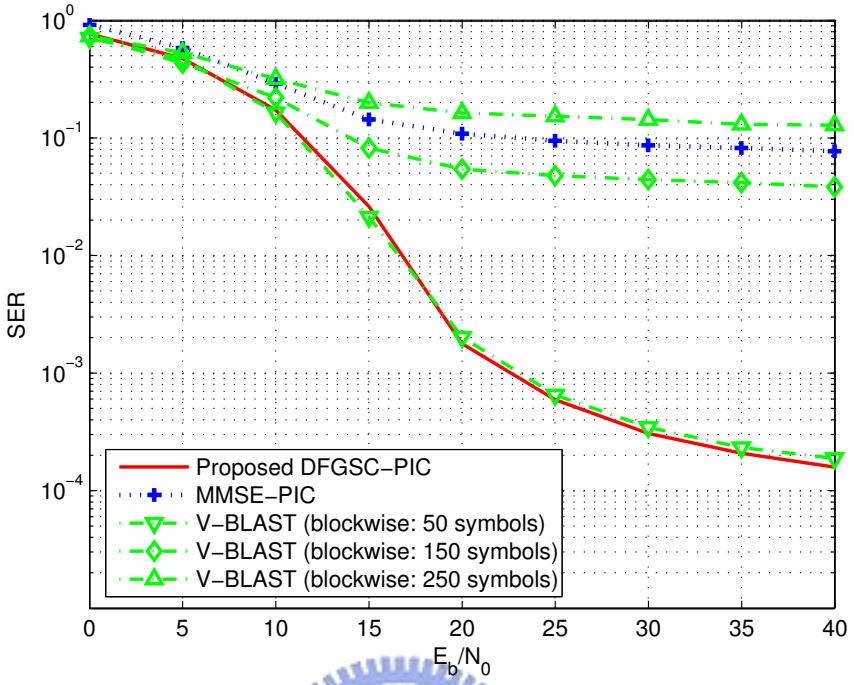

Figure 5.9: SER for various SNR values with $f_{\mathrm{d}}T_{\mathrm{s}} = 5 \times 10^{-4}$.

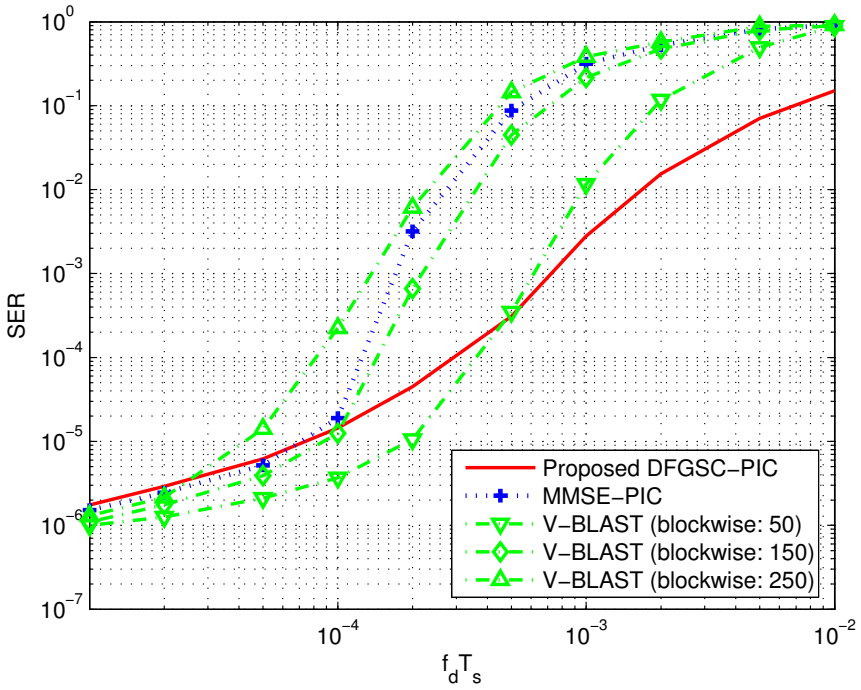

Figure 5.10: SER for various $f_{\mathrm{d}}T_{\mathrm{s}}$ values with SNR = 30 dB.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

Multiple antenna systems have been an important research area for future wireless communications because of their potential for interference suppression, high capacity, and increased diversity. Most current wireless standardization bodies (3GPP, IEEE 802.11, 802.16, and 802.20) are discussing or have already included multiple antenna techniques. This shows that the techniques are becoming sufficiently mature from both a theoretical and implementation point of view. It is clear that there are still many interesting problems to be considered. In this dissertation, we have introduced the adaptive DFGSC and its applications for various wireless communication scenarios. The adaptive DFGSC and the corresponding decision feedback operation are successfully applied to STE, DFE, and PIC spatial multiplexing detection in MIMO systems, and the performance of these systems is greatly enhanced.

In Chapter 2, the adaptive DFGSC based on the LMS algorithm has been proposed. The DFGSC contains a single-tap decision-directed equalizer and a single-tap feedback filter in the GSC structure. It can avoid the signal cancellation problem and provide extra robustness against a DOA mismatch. When the optimum weights are estimated by the LMS algorithm, the convergence of the adaptive DFGSC is faster than that of the conventional adaptive GSC.

In Chapter 3, based on the DFGSC developed in Chapter 2, a new adaptive STE has been proposed for the suppression of both CCI and ISI. In addition to the adaptive DFGSC, a particularly designed CBDFE is also investigated. With the main DOA

known as *a priori*, training sequences are not required for the adaptation of the whole STE. For spatial processing, the DFGSC improves the CCI suppression capability when implemented with the LMS algorithm. Moreover, the adaptive DFGSC with the simple point distortionless constraint is robust against multipath propagation environments and spatial signature errors. For temporal processing, the proposed CBDFE can have better performance than the LBDFE. With our special adaptation, the problem of error propagation is reduced.

In Chapter 4, using the feedback operation proposed in Chapter 3, we have developed the ACA-DFE for SISO systems and MIMO systems to reduce the error propagation effect, especially in time-varying channel environments. With the special channel-aided structure and the special operation for the feedback filter, the equalization performance is improved.

In Chapter 5, an adaptive two-stage PIC detection scheme has been developed for MIMO systems in time-varying environments. The DFGSC with the dual-DFGSC configuration is utilized for adaptive realization. All tracking operations are based on the LMS algorithm. This provides a low complexity solution for spatial multiplexing in multiple antenna systems. The DFGSC-PIC performs quite well even in fast time-varying environments.

6.2 Future Work

In concluding this dissertation, we suggest following topics for further research.

- First, research efforts can be placed on further improving the performance of the adaptive DFGSC. For rapid changing environments, the convergence rate of the LMS algorithm may not be fast enough. We can then apply a fast-convergent adaptive algorithm such as the RLS or self-orthogonalization algorithms. This may serve as a potential topic for further studies.

- In CDMA systems, it is known that multiple access interference (MAI) is the main obstacle for capacity enhancement. The adaptive GSC detector has been applied to MAI suppression. However, signal cancellation has been an unsolved problem.

To alleviate the problem, a constraint is usually posed on the filter weights. As a straightforward extension, the adaptive DFGSC can be applied in this application and signal cancellation can be completely avoided.

- Since the GSC structure is also widely used in microphone array, the adaptive DFGSC may be applied in the area of acoustic signal processing. However, the speech signal is continuous and so the concept of decisions does not exist. Note that there exists a strong correlation between a continuous and its quantized signal. Whether or not we can still apply the decision feedback operation to the speech signal deserves further investigation.

- In Chapter 3, the CBDFE has been used to enhance the performance of the blind DFE. Although it is not difficult to see where the improvement stems from, theoretical analysis is missing in this part. It is worth to examine the whole behavior of the CBDFE and the effect of error propagation in great detail.

- Note that the DFE considered in Chapter 4 for MIMO systems is generally referred to as a one-stage PIC scheme in which decisions are made for all recovered bit streams simultaneously. To further enhance the performance in MIMO systems, we can apply the SIC technique where the decision for each bit stream is made sequentially. The decision sequence is determined according to the decision-error probabilities in the recovered bit streams. Combining SIC and the proposed MIMO ACA-DFE, we can reduce the error propagation effect even more effectively.

- Low-complexity algorithms for spatial multiplexing with convincing performance are still missing. Furthermore, if the channel varies with time, the design challenge will increase significantly. We only consider the scenario of flat fading MIMO channels for the two-stage PIC detection scheme in Chapter 5. For high data-rate transmission, frequency selective fading is present between pairs of transmit and receive antennas. Orthogonal frequency division multiplexing (OFDM) is a popular technique for transmission of signals over wireless channels. It converts a frequency-selective channel into a parallel collection of frequency flat subchannels, which makes

the receiver simpler. The combination of the two powerful techniques, MIMO and OFDM, is very attractive, and has become the most promising broadband wireless access scheme. Extending the proposed DFGSC-PIC to use with OFDM in general multipath MIMO channels remains to be investigated.

Bibliography

[1] L. C. Godara, "Applications of antenna arrays to mobile communications, Part I: Performance improvement, feasibility, and system considerations," *Proc. IEEE*, vol. 85, pp. 1031–1060, Jul. 1997.

[2] L. C. Godara, "Application of antenna arrays to mobile communications, Part II: Beam-forming and direction-of-arrival considerations," *Proc. IEEE*, vol. 85, pp. 1195–1245, Aug. 1997.

[3] G. V. Tsoulos (Editor), *Adaptive Antennas for Wireless Communications*. Wiley-IEEE Press, Inc., 2001.

[4] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*. New York: Wiley, 1980.

[5] O. L. Frost III, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, pp. 926–935, Aug. 1972.

[6] L. J. Griffiths and C. W. Jim, "Alternative approach to linear constrained adaptive beamforming," *IEEE Trans. Antennas Propagat.*, vol. AP-30, pp. 27–34, Jan. 1982.

[7] B. Widrow, K. M. Duvall, R. P. Gooch, and W. C. Newman, "Signal cancellation phenomena in adaptive antennas: Causes and curse," *IEEE Trans. Antennas Propagat.*, vol. AP-30, pp. 469–478, May 1982.

[8] J. G. Proakis, *Digital Communications*, 3-rd Ed. New York: McGraw-Hill, 1995.

[9] R. B. Ertel, P. Cardieri, K. W. Sowerby, T. S. Rappaport, and J. H. Reed, "Overview of spatial channel models for antenna array communication systems," *IEEE Personal Commun.*, pp. 10–22, Feb. 1998.

[10] R. Kohno, "Spatial and temporal communication theory using adaptive antenna array," *IEEE Personal Commun.*, pp. 28–35, Feb. 1998.

[11] A. J. Paulraj and B. C. Ng "Space-time modems for wireless personal communications," *IEEE Personal Commun.*, pp. 36–48, Feb. 1998.

[12] J. W. Liang, J. T. Chen, and A. J. Paulraj, "A two-stage hybrid approach for CCI/ISI reduction with space-time processing," *IEEE Commun. Letter*, vol. 6, pp. 163–165, Nov. 1997.

[13] N. Ishii and R. Kohno, "Spatial and temporal equalization based on an adaptive tapped delay line array antenna," *IEICE Trans. Commun.*, vol. E78-B, pp. 1162–1169, Aug. 1995.

[14] M. Chiani and A. Zanella, "Spatial and temporal equalization for broadband wireless indoor networks at millimeter waves," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1725–1734, Oct. 1999.

[15] M. L. Leou, C. C. Yeh, and H. J. Li, "A novel hybrid of adaptive array and equalizer for mobile communications," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 1–10, Jan. 2000.

[16] K. Hayashi and S. Hara, "A new spatio-temporal equalization method based on estimated channel response," *IEEE Trans. Veh. Technol.*, vol. 50, pp. 1250–1259, Sep. 2001.

[17] G. J. Foschini and M. J. Gans, "On the limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 315–335, Mar. 1998.

[18] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Woniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electron. Lett.*, vol. 35, no. 1, pp. 14–16, Jan. 1999.

[19] Q. Sun and D. C. Cox, "Training-based channel estimation for continuous flat fading BLAST," *Proc. IEEE ICC 2002*, Helsinki, Finland, pp. 325–329, Jun. 2002.

[20] J. Choi, H. Yu, and Y. H. Lee, "Adaptive MIMO decision feedback equalization receivers with time-varying channels," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4295–4303, Nov. 2005.

[21] W. H. Chin, A. G. Constantinides, and D. B. Ward, "Parallel multistage detection for multiple antenna wireless systems," *Electron. Lett.*, vol. 38, no. 12, pp. 597–599, Jun. 2002.

[22] H. Dai, A. F. Molisch, and H. V. Poor, "Downlink capacity of interference-limited MIMO systems with joint detection," *IEEE Trans. Wireless Commun.*, vol. 3, no. 2, pp. 442–453, Mar. 2004.

[23] S. M. Tseng and H. L. Lee, "An adaptive partial parallel multistage detection for MIMO systems," *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 587–591, Apr. 2005.

[24] F. C. Zheng and A. G. Burr, "Signal detection for orthogonal space-time block coding over time-selective fading channels: A PIC approach for the $\mathcal{G}_i$ systems," *IEEE Trans. Commun.*, vol. 53, no. 6, pp. 587–591, Jun. 2005.

[25] F. C. Zheng and A. G. Burr, "Signal detection for orthogonal space-time block coding over time-selective fading channels: the $\mathcal{H}_i$ systems," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 40–46, Jan. 2006.

[26] S. Moshavi, "Multi-user detection for DS-CDMA communications," *IEEE Commun. Mag.*, pp. 124–136, Oct. 1996.

[27] S. Verdú, *Multiuser Detection.* Cambridge, U.K.: Cambridge Univ. Press, 1998.

[28] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Acoust., Speech, Signal Processing Mag.*, vol. 5, pp. 4–24, Apr. 1988.

[29] T. K. Sarkar, M. C. Wicks, M. S. Palma, and R. J. Bonneau, *Smart Antennas.* Wiley-IEEE Press, Inc., 2003.

[30] A. B. Gershman, "Robust adaptive beamforming in sensor arrays," *Int. Journ. Electronics and Communications*, vol. 53, pp. 305–314, Dec. 1999.

[31] T. Takao and C. S. Boon, "Importance of the exclusion of the desired signal from the control of a generalised sidelobe canceller," *IEE Processings-F*, vol. 139, pp. 256–272, Aug. 1992.

[32] D. D. Feldman and L. J. Griffiths, "A projection approach to robust adaptive beamforming," *IEEE Trans. Signal Processing*, vol. 42, pp. 867–876, Apr. 1994.

[33] D. D. Feldman, "An analysis of the projection method for robust adaptive beamforming," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1023–1030, Jul. 1996.

[34] K. C. Huarng and C. C. Yeh, "Performance analysis of derivative constraint adaptive arrays with pointing errors," *IEEE Trans. Antennas Propagat.*, vol. 40, pp. 975–981, Aug. 1992.

[35] S. Park and T. K. Sarkar, "Prevention of signal cancellation in adaptive nulling problem," *Digital Signal Processing*, vol. 8, Issue: 2, pp. 95–102, Apr. 1998.

[36] Y. Chu and W. H. Fang, "A novel wavelet-based generalized sidelobe canceller," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 1485–1494, Sep. 1999.

[37] H. Cox, R. M. Zeskind, and M. H. Owen, "Robust adaptive beamforming," *IEEE Trans. Signal Processing*, vol. 35, pp. 1365–1376, Oct. 1987.

[38] J. R. Guerci, "Theory and application of covariance matrix tapers for robust adaptive beamforming," *IEEE Trans. Signal Processing*, vol. 47, pp. 977–985, Apr. 1999.

[39] H. L. Van Trees, *Detection, Estimation, and Modulation Theory: Optimum Array Processing, Pt. 4.* John Wiley & Sons, Inc., 2002.

[40] M. H. Er, "A unified approach to the design of robust narrow-band antenna array processors," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 17–23, Jan. 1990.

[41] A. Cantoni, X. G. Lin, and K. L. Teo, "A new approach to the optimization of robust antenna array processors," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 403–411, Apr. 1993.

[42] J. H. Lee and Y. T. Lee, "Robust adaptive array beamforming for cyclostationary signals under cycle frequency error," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 233–241, Feb. 1999.

[43] S. A. Vorobyov, A. B. Gershman, and Z. Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Processing*, vol. 51, pp. 313–324, Feb. 2003.

[44] S. Shahbazpanahi, A. B. Gershman, Z. Q. Luo, and K. M. Wong, "Robust adaptive beamforming for general-rank signal models," *IEEE Trans. Signal Processing*, vol. 51, pp. 2257–2269, Sep. 2003.

[45] S. Haykin, *Adaptive Filter Theory*, 3-rd Ed. Prentice-Hall, Inc., 1996.

[46] T. K. Sarkar, E. Arvas, and S. M. Rao, "Application of FFT and the conjugate gradient method for the solution of electromagnetic radiation from electrically large and small conducting bodies," *IEEE Trans. Antennas Propagat.*, vol. AP-34 , pp. 635–640, May 1986.

[47] T. K. Sarkar and N. Sangruji, "An adaptive nulling system for a narrow-band signal with a look-direction constraint utilizing the conjugate gradient method," *IEEE Trans. Antennas Propagat.*, vol. 37 , pp. 940–944, Jul. 1989.

[48] T. K. Sarkar, J. Koh, R. Adve, R. A. Schneible, M. C. Wicks, S. Choi, and M. S. Palma, "A pragmatic approach to adaptive antennas," *IEEE Antennas Propagat. Mag.*, vol. 42 , pp. 39–55, Apr. 2000.

[49] H. Cox, "Resolving power and sensitivity to mismatch of optimum array processors," *J. Acoust. Soc. Am.*, vol. ASSP–54, pp. 771–785, Sep. 1973.

[50] G. H. Golub and C. F. Van Loan, *Matrix computation*, 3-rd Ed. The Johns Hopkins University Press, 1996.

[51] G. V. Tsoulos, "Smart antennas for mobile communication systems: Benefits and challenges," *Elect. and Commun. Eng. J.*, vol. 11, pp. 120–130, Apr. 1999.

[52] J. C. Liberti and T. S. Rappaport, *Smart Antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications.* Englewood Cliffs, NJ: Prentice Hall, 1999.

[53] B. Suard, G. Xu, H. Liu, and T. Kailath, "Uplink channel capacity of space-division-multiple-access schemes," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1468–1476, Jul. 1998.

[54] G. M. Galvan-Tejada and J. G. Gardiner, "Theoretical model to determine the blocking probability for SDMA systems," *IEEE Trans. Veh. Technol.*, vol. 50, pp. 1279–1288, Sep. 2001.

[55] J. Labat, O. Macchi, and C. Laot, "Adaptive decision feedback equalization: Can you skip the training period?" *IEEE Trans. Commun.*, vol. 46, pp. 921–930, Jul. 1998.

[56] J. Yang, J. J. Werner, and G. A. Dumont, "The multimodulus blind equalization and its generalized algorithms," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 997–1015, Jun. 2002.

[57] L. M. Garth, "A dynamic convergence analysis of blind equalization algorithms," *IEEE Trans. Commun.*, vol. 49, pp. 624–634, Apr. 2001.

[58] P. K. Shukla and L. F. Turner, "Channel-estimation-based adaptive DFE for fading multipath radio channels,", *IEE Proc.-I*, vol. 138, pp. 525–543, Dec. 1991.

[59] A. A. Rontogiannis and K. Berberidis, "Effficient decision feedback equalization for sparse wireless channels," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 570–581, May 2003.

[60] T. J. Shan and T. Kailath, "Adaptive beamforming for coherent signals and interference," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 527–536, Jun. 1985.

[61] S. Valaee, B. Champagne, and P. Kabal, "Parametric localization of distributed sources," *IEEE Trans. Signal Processing*, vol. 43, pp. 2144–2153, Sep. 1995.

[62] C. Belfiore and J. Park, Jr., "Decision feedback equalization," *Proc. IEEE*, vol. 67, pp. 1143–1156, Aug. 1979.

[63] N. Al-Dhahir and J. M. Cioffi, "MMSE decision-feedback equalizers: Finite-length results," *IEEE Trans. Inform. Theory*, vol. 41, pp. 961–975, Jul. 1995.

[64] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney, "MMSE decision-feedback equalizers and coding. I. Equalization results," *IEEE Trans. Commun.*, vol. 43, pp. 2582–2594, Oct. 1995.

[65] J. E. Smee and N. C. Beaulieu, "Error-rate evaluation of linear equalization and decision feedback equalization with error propagation," *IEEE Trans. Commun.*, Vol. 46, pp. 656–665, May 1998.

[66] M. Russell and W. M. Bergmans, "A technique to reduce error propagation in M-ary decision feedback equalization," *IEEE Trans. Commun.*, Vol. 43, pp. 2878–2881, Dec. 1995.

[67] M. Chiani, "Introducing erasures in decision-feedback equalization to reduce error propagation," *IEEE Trans. Commun.*, Vol. 45, pp. 757–760, Jul. 1997.

[68] A. Fertner, "Improvement of bit-error-rate in decision feedback equalizer by preventing decision error propagation," *IEEE Trans. Signal Processing*, vol. 46, pp. 1872–1877, Jul. 1998.

[69] M. Reuter, J. C. Allen, J. R. Zeidler, and R. C. North, "Mitigating error propagation effects in a decision feedback equalizer," *IEEE Trans. Commun.*, Vol. 49, pp. 2028–2041, Nov. 2001.

[70] D. Wulich and A. Geva "On a periodic training sequence in DFE to reduce the steady-state error probability," *IEEE Trans. Commun.*, Vol. 47, pp. 1288–1292, Sep. 1999.

[71] N. Al-Dhahir and A. H. Sayed, "The finite-length multi-input multi-output MMSE-DFE," *IEEE Trans. Signal Processing*, vol. 48, pp. 2921–2936, Oct. 2000.

[72] C. Tidestav, A. Ahlén, and M. Sternad, "Realizable MIMO decision feedback equalizers: Structure and design," *IEEE Trans. Signal Processing*, vol. 49, pp. 121–133, Jan. 2001.

[73] R. Fischer, J. Huber, and C. Windpassinger, "Signal processing in decision-feedback equalization of intersymbol-interference and multiple-input/multiple-output channels: A unified view," *Signal Processing*, 83 pp. 1633–1642, Aug. 2003.

[74] S. U. H. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol. 73, pp. 1349–1387, Sep. 1985.

[75] J. F. Frigon and B. Daneshrad, "A multiple input-multiple output (MIMO) adaptive decision feedback equalizer (DFE) with cancellation for wideband space-time communications," *Int. J. Wireless Info. Networks*, vol. 9, pp. 13–23, Jan. 2002.

[76] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Multi-input multiple-output fading channel tracking and equalization using Kalman estimation," *IEEE Trans. Signal Processing*, vol. 50, pp. 1065–1076, May 2002.

[77] M. R. Aaron and D. W. Tufts, "Intersymbol interference and error probability," *IEEE Trans. Inform. Theory*, vol. 1, pp. 26–34, Dec. 1966.

[78] S. Chen, B. Mulgrew, E. S. Chng, and G. J. Gibson, "Space translation properties and the minimum-BER linear-combiner DFE," *IEE Proc. Commun.*, vol. 145, pp. 316–322, Oct. 1998.

[79] S. Chen, L. Hanzo, and B. Mulgrew, "Adaptive minimum symbol-error-rate decision feedback equalization for multilevel pulse-amplitude modulation," *IEEE Trans. Signal Process.*, vol. 52, pp. 2092–2101, Jul. 2004.

[80] C. C. Yeh and J. R. Barry, "Adaptive minimum bit-error rate equalization for binary signaling," *IEEE Trans. Commun.*, vol. 48, pp. 1226–1235, Jul. 2000.

[81] C. C. Yeh and J. R. Barry, "Adaptive minimum symbol-error rate equalization for quadrature-amplitude modulation," *IEEE Trans. Signal Process.*, vol. 51, pp. 3263–3269, Dec. 2003.

[82] P. Dent, G. E. Bottomley, and T. Croft, "Jakes fading model revisited," *Electron. Lett.*, vol. 29, pp. 1162–1163, Jun. 1993.

[83] S. Shahbazpanahi, M. Beheshti, A. B. Gershman, M. Gharavi-Alkhansari, and K. M. Wong, "Minimum variance linear receivers for multiaccess MIMO wireless systems with space-time block coding," *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3306–3312, Dec. 2004.

[84] Y. T. Hsieh and W. R. Wu, "Optimal two-stage decoupled partial PIC receivers for multiuser detection," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 112–127, Jan. 2005.

[85] J. Q. Trelewicz, J. L. Mitchell, and M. T. Brady, "Vectorized transforms in scalar processors," *IEEE Signal Process. Mag.*, vol. 19, pp. 22-31, July 2002.

[86] W. C. Jakes, *Microwave Mobile Communications.* New York: Wiley, 1974

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