Switch-Mode Blind Equalizers for Single-User Systems

Wee Gin Lim

B.Eng.(Hons) (University College London)

December 2003

A thesis submitted for the degree of Doctor of Philosophy
of The Australian National University

Department of Telecommunications Engineering
Research School of Information Sciences and Engineering
The Australian National University

Declaration

The contents of this thesis are the results of original research and have not been submitted for a higher degree to any other university or institution.

Much of the work in this thesis has been accepted or has been submitted for publication as journal papers, conference papers or patents. These papers are:

- W. G. Lim, T. D. Abhayapala, and R. A. Kennedy, "Reliability based soft transition between blind startup and decision directed algorithms," *submitted to IEEE Trans. Commun. for publication in Oct 2003.*

- W. G. Lim, R. A. Kennedy, and T. D. Abhayapala, "Fast-convergence decision feedback equalizer under parallel adaptation," *submitted to IEEE Trans. Commun. for publication in Dec 2003.*

Papers accepted and submitted in Conference Proceedings, but in some cases containing material overlapping with the above are:

- W. G. Lim, T. D. Abhayapala, and R. A. Kennedy, "Reliability based soft transition technique for dual-mode blind equalizers," *submitted to IEEE International Conference on Communications, Paris, 2004, for publication.*

- W. G. Lim, "Torque based rotational study of several blind equalization algorithms for QAM signals," *Accepted to Australian Communications Theory Workshop (AusCTW), Newcastle, Australia, Feb 2004.*

Patent disclosures connected with research contract work for News Technology, Los Angeles, CA and Philips Research USA, Briarchiff Manor, NY:

- W. G. Lim, R. A. Kennedy, and T. D. Abhayapala, "Parallel adaptation strategy for robust and rapid convergence to the decision directed mode in a blind adaptive decision feedback equalizer," Jun 2001.

- W. G. Lim, "Tri mode equalizer (TRIME) for time varying channels," Jun 2001.

The research represented in this thesis has been performed jointly with Prof. Rodney A. Kennedy and Dr Thushara D. Abhayapala.

Wee Gin Lim
Department of Telecommunications
Research School of Information Sciences and Engineering
The Australian National University
Canberra ACT 0200
Australia

Acknowledgements

The work presented in this thesis would not have been possible without the support of a number of individuals and organizations, and they are gratefully acknowledged below:

- My supervisors Prof Rod Kennedy and Dr Thushara Abhayapala who guided and nurtured me throughout my years of PhD in Australia and eventually, producing this thesis. I am considered very blessed to be able to work under these two very distinguished and outstanding researchers both in my field of adaptive equalization and several other emerging fields.

- My fellow students in the Department of Telecommunication Engineering, for their friendships. Ms Lesley Cox, administrator of the Department of Telecommunication Engineering, for simply being the very best administrator I have ever known. Ms Rita Murray, business manager of the Research School of Information Sciences and Engineering, for all her kind assistance in making my stay in Australia a pleasant, memorable and problem-free one.

- The Australian National University Malaysian Alumni body for providing me with the PhD scholarship.

- The industrial collaboration with News Technology, Los Angeles, CA and Philips Research USA, Briarchiff Manor, NY, that gave me the opportunity to be involved in designs for real-life practical telecommunication systems.

- My parents and brother for all their heart-felt love, unceasing encouragement, guidance and prayers. Most of all, for being a wonderful family that is always there whenever I needed them.

Abstract

This thesis considers the design of equalizers which need to operate in various modes (two or more modes) depending on the difficulty of the channel and the performance of the equalizer. We consider the formulation of such switch-mode equalizers where it is usually broken into an acquisition mode and a tracking mode. As a broad goal, our objectives include speeding up the convergence rate of blind adaptive algorithms which are well-known for their slow rate of convergence, achieving low steady state errors as well as reducing switching transients during the switch-over between operation modes.

A novel concept based on the reliability of the equalizer output is developed and presented where unlike conventional algorithms which utilize only explicit information regarding the equalizer output, the reliability measure is calculated as a function of both the equalizer output and its estimated statistical distribution. Two separate algorithms are developed based on this concept. The first is a switch-mode algorithm that uses the reliability measure to combine the acquisition algorithm and the tracking algorithm to achieve not only a smooth transition between modes, but also increased rate of convergence and lower steady state errors. While the first algorithm that computes the reliability measure requires an estimate of the variance of the residual intersymbol-interference (ISI) and noise term, the second algorithm uses a simple technique of combining that computes the probability of the equalizer output being close to the constellation data points instead. The second algorithm extends and simplifies the first algorithm and is shown to achieve fast convergence, low steady state errors and smooth transition at a significantly lower computational cost.

We further propose modifications that extend a fast-convergence equalization scheme proposed by Labat, Macchi and Laot in 1998. The main accomplishment is in the design of a new strategy that successfully reduces the switching transients that occur during the switch-over between the linear acquisition equalizer and the non-linear tracking decision feedback equalizer (DFE). This is achieved by combining both the linear and the non-linear equalizers in a parallel fashion while adapting the tap parameters using combining techniques designed for dual-mode type algorithms such as the Benveniste-Goursat algorithm and the reliability-based algorithm developed in this thesis. In addition, we also extended the original equalization scheme to incorporate a linear transversal whitening filter in place of a recursive whitening filter to overcome some deficiencies inherent in recursive filtering structures. Thirdly, we introduced a novel DFE scheme that uses two

DFE's to accomplish the robust tracking of time-varying channel statistics and to retain low steady state errors simultaneously by a novel interplay of the equalizer structure and the transfer of equalizer tap parameters between the filter blocks.

In addition, we also analyzed the "rotational" behavior of several blind algorithms using a torque concept, where we assess the susceptibility of the equalizer output constellation to converge to stable but undesirable rotated solutions.This new analysis tool is used to investigate various blind algorithms and we reveal the multi-modulus algorithm (MMA) has superior convergence properties.

Contents

Declaration i

Acknowledgements iii

List of Figures xv

List of Tables xvii

1 Introduction 1

1.1 Overview of Thesis 1

1.2 Contributions 3

2 Background on Single-User Blind Equalizers 7

2.1 Model of a Digital Communications System 7

2.1.1 Digitally Modulated Source Signals 8

2.1.2 Inter-Symbol Interference and Eye Diagram 10

2.1.3 Performance Measures 12

2.2 Equalizer Structures 13

2.2.1 Linear transversal equalizers 13

2.2.2 Non-linear equalizers 14

2.3 Criteria That Lead To ISI Cancellation 18

2.3.1 Zero Forcing Algorithm 19

2.3.2 Least Mean Squared (LMS) Algorithm 20

2.3.3 Unsupervised (Blind) Algorithms 21

2.4 Overview of Non-Bussgang Techniques 22

2.4.1 Historical Notes 22

2.4.2 Classifications of Blind Deconvolution Algorithms 23

2.4.3 Polyspectra algorithms 24

2.4.4 Algorithms based on cyclostationary statistics 27

2.4.5 Probabilistic algorithms 28

2.5 Bussgang Algorithms 28
2.5.1 Maximum-Level-Error (MLE) algorithm 29
2.5.2 Sato algorithm 31
2.5.3 Godard algorithm 34
2.5.4 Shalvi and Weinstein algorithm 36
2.5.5 Multi-Modulus Algorithm 37
2.5.6 Bussgang algorithms for multiple-modulus constellation . . . 39
2.5.7 Decision directed algorithms 40
2.6 Principal Aims Of Thesis 41

3 Reliability Based Technique For Switch-Mode Blind Algorithms 43
3.1 Motivation 44
3.2 System Model 46
3.2.1 Expressions for error function of switch-mode algorithms . . 47
3.3 Review of traditional switch-mode algorithms 49
3.3.1 Benveniste-Goursat (BG) algorithm 49
3.3.2 Stop-And-Go (SAG) decision directed algorithm 49
3.3.3 Dual-Mode Generalized Sato Algorithm (DMGSA) and Dual-Mode Godard Algorithm (DMGA) 50
3.3.4 Hilal-Duhamel (HD) algorithm 51
3.3.5 Diagrams of error functions 52
3.4 A Novel Reliability Based Switch-Mode Algorithm 54
3.4.1 Computation of Reliability Measure $\alpha(k)$ 55
3.5 Discussions on convergence and approximation 59
3.5.1 Convergence 59
3.5.2 Gaussian assumption of effective noise 61
3.6 Simulation Results 63
3.7 Conclusions 66

4 Probabilistic-Based Switching Technique For Switch-Mode Algorithms 67
4.1 Problem Statement 68
4.2 Graphical Illustration of Switching Difficulty in Switch-Mode Algorithms 69
4.2.1 Non-Constant Modulus Constellation 71
4.2.2 Constant Modulus Constellation 71
4.3 A Novel Probabilistic-Based Switching Technique 72
4.4 Algorithm For Joint Blind Equalization And Phase Recovery 75

4.4.1 System Model 75
4.4.2 Development of Novel Tri-Mode Algorithm 76
4.4.3 Performance Improvements via Pre-Whitening 80
4.5 Simulation Results 81
4.5.1 Results for Stationary Channel Without Phase Errors 84
4.5.2 Results for Joint Equalization and Phase Recovery 89
4.6 Conclusions 91

5 Fast Convergence Switch-Mode DFE Schemes 93
5.1 Introduction 93
5.1.1 Channel Decomposition Property 94
5.2 Design Objectives 96
5.2.1 System Model 97
5.3 Development of Alternative Fast-Convergence DFE 97
5.4 Parallel Adaptation Strategy for Dual Mode Equalization Schemes . 101
5.5 Alternative Fast-Convergence Predictive DFE Scheme 103
5.5.1 Switching Strategy 104
5.6 Novel Dual Decision Feedback Equalizer 105
5.6.1 Problem Statement 105
5.6.2 Development 105
5.6.3 Operation Details 107
5.7 Simulation Results 109
5.7.1 Switch-Mode Conventional DFE Scheme of Section 5.3 – Channel Equalization Results 109
5.7.2 Switch-Mode Conventional DFE Scheme of Section 5.3 – Channel Identification Results 112
5.7.3 Switch-Mode Conventional DFE Scheme Under Parallel Adaptation Strategy 114
5.7.4 New Switch-Mode DFE Schemes of Sections 5.3, 5.4 and 5.5 – Channel Equalization Results 115
5.7.5 Switch-Mode Dual DFE Scheme of Section 5.6 – Channel Equalization Results 117
5.8 Conclusions 118

6 Rotational Analysis On Several Blind Equalization Algorithms 121
6.1 Introduction 121
6.2 System Setup and Assumptions 122
6.3 Torque Analysis of Undesirable Rotated Solutions 124

6.3.1 Some Wrong Rotated Solutions 124
6.3.2 Preliminaries on Torque Concepts and Some Definitions . . . 125
6.3.3 Rotational Behavior of the CMA 126
6.3.4 Rotational Behaviors of RCA and MMA 127
6.4 Conclusions 130

7 Conclusions **133**
7.1 Executive Summary 133
7.2 Future Work 134
7.2.1 Direct Extensions 134
7.2.2 Generalizations 135

Appendix A Computation of mean of absolute value of error functions, $\mathrm{E}\{|\epsilon(k)|\}$ 137

Appendix B Validity of (5.18) due to parallel adaptation 143

List of Figures

2.1 A typical communication system. 8

2.2 Constellation diagram of modulation formats for (A) 16-QAM, (B) 4-PAM, (C) 8-PSK, (D) 16-APK (V29-CCITT). 9

2.3 (A) Linear noisy baseband equivalent channel model. (B) Discrete-time channel impulse response obtained from the continuous-time response. 10

2.4 Eye diagram. 12

2.5 Linear Transversal Filter. 14

2.6 Conventional decision feedback equalizer. 16

2.7 Predictive decision feedback equalizer. 17

2.8 Linear Transversal Bussgang Equalizer. 29

2.9 Maximum Level Error (MLE) Algorithm: Constellation diagram for 16-QAM with dotted-line boundary connecting the edge data points. The equalizer is only updated if the output is exceeds the boundary depicted by the dotted lines. 30

2.10 Reduced Constellation Algorithm (RCA): Constellation diagram for 16-QAM and its reduced constellation whose 4 data points are represented by 'o' at coordinates $[\gamma_S, \gamma_S]$, where γ_S is defined in (2.44). 33

2.11 Godard algorithm or Constant Modulus Algorithm (CMA): It penalizes deviations from a constant modulus represented by the dotted circles for (A) 8-PSK which is a constant modulus source, and (B) 16-QAM which is a non-constant modulus source. 34

2.12 Multi Modulus Algorithm (MMA): It penalizes deviations from γ_M for real and complex components separately. 38

2.13 MMA2 and the DAMA algorithms : The 9 radii of the 64-QAM constellation. 40

3.1 A typical baseband equivalent channel and a linear equalizer that employs the switch-mode algorithm. 45

3.2 The error functions of the switch-mode algorithms outlined in the above subsections for 8-PAM signalling. Their shapes vary with the equalizer output $z(k)$ only. (A) Benveniste-Goursat (BG) algorithm [15], with parameters $k_1 = 1, k_2 = 0.25$. (B) Stop-And-Go (SAG) algorithm [116]. (C) Dual-Mode Generalized Sato Algorithm (DMGSA) [145], with the decision region $D(k) = d_i \pm 0.5,\ \forall i$, where $\{d_i\}$ is the set of the source constellation. (D) Dual-Mode Godard Algorithm (DMGA) [145] with the decision region $D(k) = d_i \pm 0.5, \forall i$. 53

3.3 Contour plots of the reliability measure of (3.29) for a 4-PAM data. 56

3.4 Relationship between the DD MSE and σ_v for M-ary PAM with $M = 2, 4, \cdots, 64$. The range of values of σ_v is from 0.1 to 1.4 presented in the log scale. 58

3.5 The error function of the novel switch-mode algorithm of Equation (3.21). Here $\gamma = 14$. 60

3.6 Cross sections of the 3-D graph of the error function of the new dual-mode algorithm for 8-PAM signalling and $\gamma = 100$. (A) $\sigma_v = 1.2$ and (B) $\sigma_v = 0.7$. The error functions of the CMA and the LMS algorithms are also plotted as dashed and dotted lines respectively. 61

3.7 A comparison of several dual-mode algorithms equalizing channels (A) $\mathbf{h}'$, and (B) $\mathbf{h}''$, both at an SNR of 25dB for 4-PAM signalling. . 63

3.8 A comparison of several dual-mode algorithms equalizing channels $H_1(z)$ with an SNR of (A) 20 dB, and (B) 15 dB for 4-PAM signalling. 64

3.9 Similar to Fig. 3.7, this graph also includes the new equalizer that is prefixed with a recursive whitening filter where lower steady state MSE is achieved due to the additional whitening filter which increases the effective length of the equalizer, thereby being able to estimate the channel more accurately. The SNR is 25 dB, and a 4-PAM source is used. 65

4.1 The plots of $|\epsilon(k)|$ for (A) CMA-1, (B) CMA-2, (C) DD algorithm, (D) DAMA algorithm, for the non-constant modulus 16-QAM source signals. 70

4.2 The plots of $|\epsilon(k)|$ for (A) DD algorithm, (B) DAMA algorithm, for the constant modulus 8-PSK source signals. The plots of $|\epsilon(k)|$ for CMA-1 and CMA-2 are identical to Fig. 4.1-A,B, except that $\gamma_C^{(p)}$ has to be readjusted. 71

4.3 The new switch-mode algorithm featuring adaptation with the acquisition algorithm when $z(k)$ is outside the individual square regions surrounding each data point, and with the tracking algorithm otherwise. 72

4.4 Baseband QAM system model for joint blind equalization and phase recovery. 75

4.5 The new tri-mode blind equalization algorithm. The reliable regions $D(k) = \cup D_p(k), p = 1, 2, \cdots, 32$ and $D_{\text{out}}(k)$ where the DD algorithm and the MMA are respectively employed for the 32-QAM constellation are illustrated. The CMA is employed outside these regions. 78

4.6 Figures illustrating the technique for joint equalization and phase recovery for 16-QAM signals in the presence of phase errors. The emphasis is on the variable-size reliable region, $D(k)$. 79

4.7 Comparing various algorithms using the channel $\mathbf{h}'$ under SNR levels of 15dB and 25dB only. The modulation format is 16-QAM. Equalizer has 31 taps. 84

4.8 Comparing various algorithms using the channel $\mathbf{h}''$ under SNR level of 15dB. The modulation format is 16-QAM. Equalizer has 41 taps. 86

4.9 Comparing various algorithms using the channel $\mathbf{h}'''$ under SNR level of 17dB. The modulation format is 16-QAM. Equalizer has 41 taps. 87

4.10 Comparing various algorithms using the channel $\mathbf{h}'$ under SNR levels of 25dB and 30dB. The modulation format is 32-QAM. 88

4.11 Comparing various algorithms using the channels $\mathbf{h}'''$ under SNR level of 30dB. The modulation format is 32-QAM. 89

4.12 Comparing various algorithms in different equalizer setups. 92

5.1 Decomposition of a transversal non-minimum phase channel via the channel decomposition property with $H(z) = \mathcal{G}H_{\min}(z)H_{\text{ap}}(z)$, assuming the complex gain, $\mathcal{G}$, is unity. 94

5.2 Novel blind linear equalizer setup [75] as a cascade of four devices used to acquire the "decomposed" channel. 95

5.3 Linear equalizer of acquisition mode of the alternative DFE scheme to that of [75], featuring a non-recursive $\mathcal{W}$. 98

5.4 The non-linear equalizer (DFE) of the tracking mode. 100

5.5 New DFE scheme under parallel adaptation. The taps are adapted and shared among filters from "top" and "bottom" paths of the new DFE. 101
5.6 The new alternative fast-convergence predictive DFE. 103
5.7 The dual DFE scheme which features a conventional DFE that is prefixed by a linear recursive filter whose taps are driven by a predictive DFE which is not in the equalizer path (illustrated by the dotted lines) but it minimizes the residual error at the soft output of the conventional DFE, $y(k)$. 106
5.8 (A) and (B) are comparisons of three DFE schemes for the first channel, $\mathbf{h}'$, while (C) and (D) are for the second channel, $\mathbf{h}''$, using 16-QAM data signals. The first equalizer is the DFE of [75]. The second and third equalizers are the DFE described in Section 5.3 without and with parallel adaptation strategy, respectively. The parallel adaptation strategy employs the Benveniste-Goursat type combination parameters. 110
5.9 The 100 most dominant taps of the inverse response of channel $\mathbf{h}'''$. 112
5.10 Channel identification results using new switch-mode DFE scheme proposed in Section 5.3. These results are obtained using (A) 300 symbols, (B) 800 symbols, (C) 2500 symbols. The equalizer is in its acquisition mode for both (A) and (B). As for (C), the equalizer has settled into its tracking (DD) mode. 113
5.11 The average MSE and SER of 20 runs of fast convergence DFE scheme of Section 5.3, with (solid line) and without (dotted line) parallel adaptation strategy on 8-PAM signals at 30 dB for channel $\mathbf{h}'''$. 114
5.12 Results of two new alternative DFE's compared with trained equalizers. 115
5.13 Plot of the equalizer output of various selected schemes extracted at the end of the data stream. (A) CMA equalizer. (B) Trained LMS equalizer. (C) Fast converging blind predictive DFE. (D) Our FIR scheme and the trained DFE, which yields similar MSE. 116
5.14 Comparing three equalizers: 1) the linear acquisition equalizer (LE) of [75], 2) the DFE scheme of [75], and 3) the dual-DFE that extends the DFE in [75]. 118

6.1 Baseband QAM system model. 122

6.2 Examples of undesirable rotated solutions for (A) 16-QAM at 21°, obtained by RCA and CMA; (B) 32-QAM at 45°, obtained by RCA and CMA; (C) 16-QAM under a frequency offset, obtained by CMA. 124

6.3 Rotational behavior of RCA for square-QAM constellations. The non-flat region in Fig. (B) indicates that there is an undesirable but rotationally stable equilibrium point which is at the vicinity of 21°. This is also reflected in Fig. (C) where a shallow local minimum is observed for $\sigma_v \approx 0$ and $\Delta\theta(k) \approx 21°$. 127

6.4 Rotational behavior of RCA for cross-QAM constellations. The large non-flat regions in Figures (B) and (E) at the vicinity of 45° indicate that the RCA is very susceptible to the '45°' wrong solutions. This is also reflected in the cost figures (C), (F) where a local minimum point clearly exists at 45°. There are two more stable points for 32-QAM as shown in Figure (B) at high SNR and around 13° and 21°. 128

6.5 Rotational behavior of MMA for 16-QAM constellation. Identical-shaped figures are observed for 32-, 64- and 128-QAM, so they are not plotted. Figure (B) shows no signs of a negative mean torque and this is true also for 32-, 64- and 128-QAM. This phenomenon is also reflected in the cost function which does not have any undesirable local minimum. 130

A.1 The probability density functions (p.d.f.) of the error functions of the MOE, CMA, and the LMS algorithms. The solid lines represent the c.d.f. for $\sigma_v^2 = 0.28$ and the dotted lines for $\sigma_v^2 = 0.69$. 138

A.2 The cumulative distribution functions (c.d.f.) of the error functions of the MOE, CMA, and the LMS algorithms. The solid lines represent the c.d.f. for $\sigma_v^2 = 0.28$ and the dotted lines for $\sigma_v^2 = 0.69$. The horizontal dotted lines show $D(\epsilon(k)) = 0.75$. 141

List of Tables

3.1 Lookup table for $\mathrm{MSE}_{\mathrm{DD}}(\sigma_v)$ and variance of effective noise, σ_v^2 for M-PAM signals 59

3.2 Summary of results in Fig. 3.7 including failure rates 66

4.1 Characteristics of blind algorithms 77

4.2 Progressive stages in new algorithm and associated step sizes employed in our simulations in the presence of phase errors 82

Chapter 1

Introduction

This thesis deals with switch-mode blind equalizers for single-user digital communication systems. A switch-mode equalizer is a combination of two or more equalizer modes where it usually starts with a linear equalizer that employs an acquisition algorithm and later switches over to a tracking decision directed algorithm after the data estimates have become sufficiently reliable. In terms of the filtering structure, a non-linear decision feedback equalizer may be used in the tracking mode. In the acquisition mode, the equalizer is designed to be robust with respect to the channel's initial conditions but may yield high steady state errors even when its convergence is achieved. The tracking equalizer, on the other hand, yields lower steady state errors at convergence but cannot be used reliably to acquire severe channels (which do not correspond to initial open-eye conditions [92, 99]). Therefore, switch-mode blind equalizers have become very popular in practical digital communication systems because the distinct advantages of the acquisition and tracking modes are complementary.

In this introductory chapter, we present an overview of the contents of the thesis and provide a list of the contributions. It features several novel switch-mode algorithms as well as switch-mode equalization schemes. Other contributions relating to the global topic of switch-mode equalization are also presented.

1.1 Overview of Thesis

Chapter 2 – In this background chapter we give a brief overview of the domain of digital transmission as well as reception over linear channels. The problem of intersymbol interference (ISI) is presented and the notion of equalization is defined. Equalizer structures including linear transversal equalizers, maximum-likelihood equalizers, and decision feedback equalizers (DFE) are

presented. Subsequently, a concise overview of blind equalization techniques is given, the emphasis being on Bussgang-type algorithms. Non-Bussgang algorithms such as the polyspectra algorithms, the cyclostationary-statistics algorithms and the probabilistic algorithms are also briefly described. Lastly, we present an analysis of the undesirable solutions (due to the non-convexity of cost functions) of the most widely used algorithms such as the Sato and Godard algorithms.

Chapter 3 – This chapter addresses the problems induced by abruptly switching between acquisition and tracking algorithms. We use a novel reliability measure to combine the two algorithms to provide a soft transition and avoid having to accurately determine the open-eye threshold for the tracking algorithm to be reliably employed. We derive the reliability measure from Bayes theorem and compute the posterior probability of correctly detecting the equalizer output. From the derivation we find that in order to effectively provide a soft transition between algorithms, not only the equalizer output needs to be utilized, but the statistical distribution of the residual intersymbol-interference (ISI) plus noise term is also required. Once the reliability measure is accurately estimated, it leads to significant improvements over conventional algorithms such as the Benveniste-Goursat algorithm [15] as well as the Stop-And-Go algorithm [116].

Chapter 4 – This chapter further considers the problem of smooth switching between acquisition and tracking algorithms. In the same spirit as the reliability-based algorithm proposed in Chapter 3, we design an alternative switch-mode algorithm that also utilizes information regarding the equalizer output and its estimated distribution to ensure a smooth transition between algorithms. This new algorithm is based on a simple concept that employs the tracking algorithm when the equalizer output is found in reliable regions in the constellation space. Otherwise, an acquisition algorithm is employed. The novelty lies in the way the size of these reliability regions is varied according to a new probabilistic measure regarding the reliability of the equalizer output.

In addition, this chapter also presents an algorithm capable of joint blind equalization and phase recovery by using the new switch-mode technique.

Chapter 5 – This chapter presents three novel modifications of the equalizer *structures* in addition to the adaptation algorithms as is the focus of Chap-

ters 3 and 4. Overall, this chapter addresses the speed of convergence and improved robustness while retaining the simplicity of an equalization scheme proposed by Labat, Macchi and Laot [75]. We first address the problem of the undesirable use of a recursive linear filter structure in the acquisition mode by proposing a new transversal filter while preserving the ability of direct transfer of the taps to the DFE in the tracking mode. Secondly, we address the problem of eliminating switching transients of dual-mode equalizers during the switch-over between acquisition and tracking modes where the positions of several filter blocks need to be rearranged. We propose a novel strategy that combines the acquisition linear equalizer and the tracking DFE in a parallel fashion to reduce the switching transients effectively. Lastly, the chapter addresses the problem of tracking time-varying channels with a new dual-DFE equalization scheme that features two DFE in an equalizer setup and involves a novel interplay of the equalizer structures and the transfer of filter coefficients between the filter blocks.

Chapter 6 – We analyze the phase and/or frequency locking (rotational) convergence behavior of several blind algorithms, namely the CMA, the reduced constellation algorithm (RCA) and the multi-modulus algorithm (MMA)[1]. The principle objective is to assess their susceptibilities to wrong solutions (stable undesirable equiibria) which correspond to an output constellation that is a undesirable rotated version of the original data constellation. In our analysis, we define a "torque" that is the cross product of the equalizer output and the "force" which is due to the error function acting on the equalizer output symbol. A net torque will correspond to a rotational behavior of the equalizer output in either the clockwise direction or the anti-clockwise direction.

Chapter 7 – Some concluding remarks and future work are presented in this chapter.

1.2 Contributions

At this point, a list of major contributions of this thesis is given:

- **New treatment of switching for enhanced equalizer performance** –

[1]The MMA we consider in this Chapter is proposed by Oh and Chin [105], and Yang *et al* [152]. It should not be confused with another algorithm previously proposed by Sethares *et al* [126] under the same name.

Smooth switching algorithms are designed to avoid the difficult task of having to accurately determine an open-eye condition so that the tracking algorithm may be employed reliably. The major contribution drawn from Chapters 3 and 4 is in the design of superior smooth switching switch-mode algorithms. While conventional techniques usually tradeoff convergence speed and steady state performance for smoothness in the switch-over, our techniques actually *accelerate* convergence in addition to ensuring smooth transitions between modes.

- **Dual-parameter measure for estimating reliability of the equalizer output** –
  The reliability-based switch-mode algorithm offers effective smooth-switching that enhances convergence rate and preserves low steady state errors based on a concept that utilizes explicit and implicit information in the equalizer output signals. In contrast to conventional smooth switching algorithms that employ a single-parameter[2] measure for combining the switch-mode algorithm, our new *dual-parameter* reliability measure as proposed in Chapter 3 results in not only a smooth transition between the acquisition and tracking algorithms, but also in a faster, more reliable convergence as well as lower steady state errors.

- **Simple algorithm for rapid acquisition** –
  The new measure for combining the switch-mode algorithm as proposed in Chapter 4 is designed to be simple and effective in ensuring a smooth switching between algorithms. While conventional algorithms usually trade off equalizer performance, e.g., convergence speed and steady state errors, with computational complexity, this algorithm is shown to outperform several well-known switch-mode algorithms at a *reduced* cost. This is because the novel measure can be used to accurately reflect an open-eye condition without having to collect many data samples. The switch-over therefore occurs promptly as soon as the data estimates become sufficiently reliable.

- **Joint blind equalization and automatic phase recovery** –
  By coupling the simple switching technique described above and the results in Chapter 6 regarding phase recovery properties of the MMA, a new switch-mode algorithm is proposed to achieve joint blind equalization and phase recovery effectively. The phase offsets can be corrected by the MMA and the frequency offsets can be handled through the use of the smooth switching technique of Chapter 4.

[2]Conventional single-parameter measures rely solely on either the equalizer output, e.g., [15], or the signs of the error functions, e.g., [116].

- **New fast convergence switch-mode decision feedback equalization (DFE) scheme** –
  We extended a fast convergence equalization scheme proposed by Labat, Macchi and Laot in 1998 to incorporate a simple linear transversal acquisition equalizer which can be switched to a DFE when the error rate is sufficiently low. Deficiencies inherent in the recursive filtering structure in [75] can be overcome with this modification. After a transformation in the transversal filter, the direct transfer of the taps from the acquisition equalizer to the tracking equalizer (which involves a recursive feedback filter) is also retained.

- **Soft transition strategy for new switch-mode DFE scheme** –
  A parallel adaptation strategy that combines the linear equalizer and the tracking DFE in a parallel fashion can reduce the switching transients that are due to the rearrangement of the filter blocks as well as a switch in the adaptation algorithms. We showed significant improvement in terms of smooth switching between the linear equalizer and the DFE as well as rate of convergence with this strategy, even when a simple Benveniste-Goursat combination technique is used.

- **Dual DFE scheme for tracking varying channels** –
  The dual-DFE, unlike conventional equalizers which usually implement just a single DFE, may exhibit enhanced tracking capabilities in non-stationary environments and low steady state MSE in stationary environments. The complexity incurred which is less than twice of that of a single DFE is reasonable for its achieved improvements.

- **Rotational analysis of blind equalization algorithms** –
  We develop a novel analysis technique based on a torque concept to study the phase/frequency locking behavior of blind equalizer algorithms. Through analysis of the RCA and the MMA, we reveal that the MMA is superior to its RCA counterpart in terms of immunity to converge to incorrectly rotated solutions. This revelation motivates the use of the MMA over the RCA for the task of joint equalization and phase recovery.

Chapter 2

Background on Single-User Blind Equalizers

For bandwidth-efficient communication systems operating in high inter-symbol interference (ISI) environments, adaptive equalizers have become a necessary component of the receiver architecture. An accurate estimate of the amplitude and phase distortion introduced by the channel is essential to achieve high data rates with low error probabilities. An adaptive equalizer provides a simple practical device that is capable of both learning and inverting the distorting effects of the channel. Conventional equalizers rely on the transmission of a training sequence known to both the transmitter and receiver. The blind equalizer, on the other hand, does not require a training sequence to be sent for start-up or restart. Rather, the blind equalization algorithms exploit *a priori* knowledge regarding the statistics of the transmitted data sequence as opposed to an exact set of symbols known to the transceiver, thereby achieving improved bandwidth utilization. Blind equalization is also desirable when the communication system must cope with multipoint broadcast environments and unpredictable channel changes [147].

In this chapter, a concise background on the communications systems is provided. We recommend the following references [12, 38, 58, 117–119, 141] for an excellent tutorial overview of the field. Our focus will be to provide sufficient self-contained background on the subjects which the thesis deals with predominantly, i.e., a variety of equalizer structures and popular blind algorithms.

2.1 Model of a Digital Communications System

Throughout this thesis we will be dealing with aspects related to the physical layer of a communications system, i.e., the transmission layer. A typical communica-

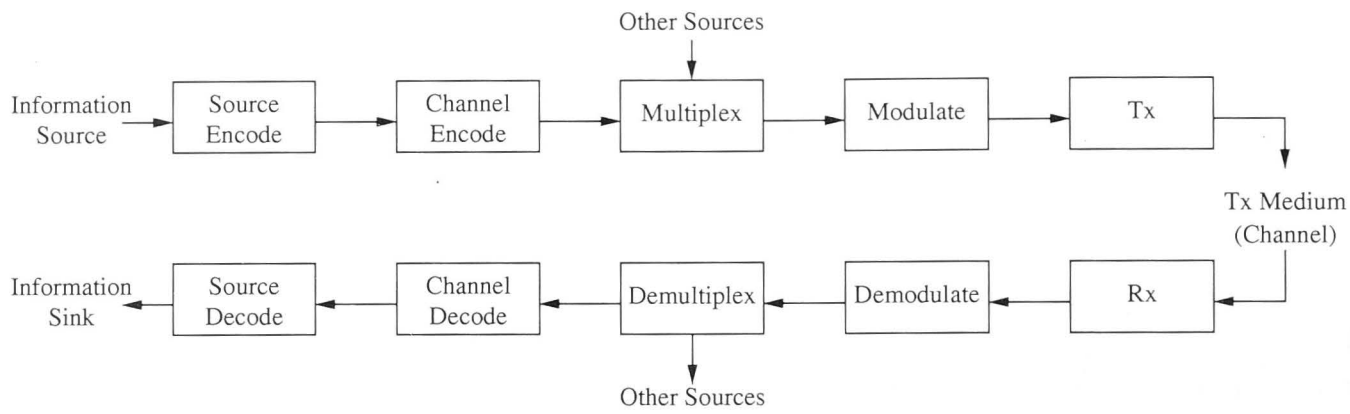

Figure 2.1: A typical communication system.

tions system is illustrated in Fig. 2.1. The complete cycle of the transmit-receive process begins with the original information being formatted in a digital form in order to be represented by a series of bits. These digits are then source-encoded whereby the message is represented in a succinct form and the redundancy is being removed as much as possible. The resulting sequence is then channel-encoded where redundancy is now introduced into the data stream in a deliberate but controlled manner for the purpose of detecting and possibly correcting the errors due to transmission. This is also known as error control coding. One technique of error control coding is to introduce frequent symbol changes so that the symbol clock speed can be easily recovered in the receiver. Also, often a number of bits are encoded into one or more symbols. At this point, multiplexing with sequences from other sources may be performed to efficiently utilize the limited bandwidth of communication channels. Finally, the sequence is converted by the modulator into continuous-time passband waveforms, in accordance with modulation schemes that are suitable for transmission over the channel. Examples of communications channel include coaxial, fiber optic, or twisted-pair cables in wired communications; and the atmosphere or ocean in wireless communications, or some combination of these media. This resulting waveform is then delivered to the receiver at the other end of the channel.

At the receiver, the processes prior to transmission need to be "inverted" in order to recover the original transmitted information. The received waveform is first demodulated and demultiplexed. Subsequently, the sequence is equalized and decoded in order to give the final digital output. Equalization is necessary to combat the impairments of the channel that may have caused signal distortions.

2.1.1 Digitally Modulated Source Signals

The information bearing bits are usually collected in blocks of say K consecutive binary digits. Thus there may be 2^K different binary sequences, and each sequence

would correspond to a symbol. The set of all possible symbols is called the alphabet set $\mathbb{A}$ of the transmitted symbols of a particular modulation format. Each of these possible symbols is then mapped to one of 2^K different continuous-time lowpass waveforms by attributing to the set of symbols a discrete set of amplitudes, frequencies and/or phases, depending on the modulation format of the specific system. For example, let we consider Quadrature Amplitude Modulation (QAM) and let there be $K = 4$ binary digits in each block. Thus, there are $2^4 = 16$ different sequences as follows:

$$[0000, 0001, 0010, 0011, \cdots, 1110, 1111]. \tag{2.1}$$

The QAM alphabet set can be represented by the set of sequences in (2.1) where each sequence is uniquely mapped to one of 16 different constellation points in the complex plane as in Fig. 2.2. The resulting 16-QAM alphabet set is thus

$$\mathbb{A} = \{\pm 1 \pm j, \pm 1 \pm 3j, \pm 3 \pm j, \pm 3 \pm 3j\}. \tag{2.2}$$

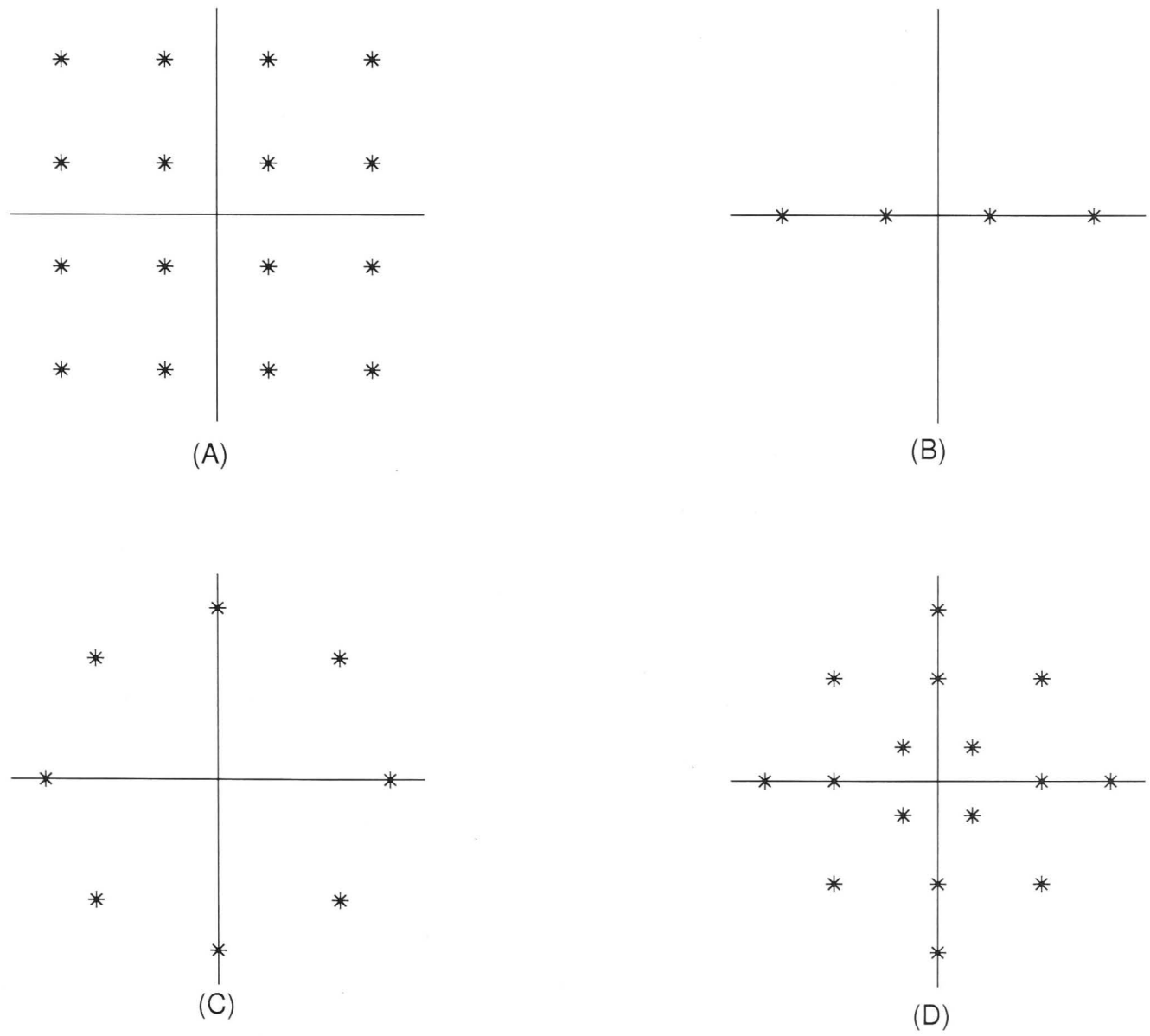

Figure 2.2: Constellation diagram of modulation formats for (A) 16-QAM, (B) 4-PAM, (C) 8-PSK, (D) 16-APK (V29-CCITT).

For other modulation formats such as the Pulse Amplitude Modulation (PAM) and the Phase Shift Keying (PSK) formats, they differ from the QAM example by only the mapping operator that maps each of the sequence in 2.1 to one of the alphabet in 2.2. The constellation points of 16-QAM, 4-PAM, 8-PSK, and the 16-APK (Amplitude-Phase Keying, also widely known as 16 point V29-CCITT for 9,600 bits per second modems). The choice of modulation formats usually trades off between bandwidth efficiency and probability of error, where larger constellations may utilize the bandwidth more efficiently but suffers higher probability of error. More recently, some new techniques that vary the modulation constellation size according to the time-varying channel conditions in order to maintain a specified target reliability in the receiver output have been proposed [52, 61].

2.1.2 Inter-Symbol Interference and Eye Diagram

In a nutshell, inter-symbol interference (ISI) in a digital transmission system is the distortion of the received signal that is manifested in the temporal spreading. The overlap of individual pulses and their adjacent pulses would usually result in the receiver being unable to reliably distinguish between individual signal elements.

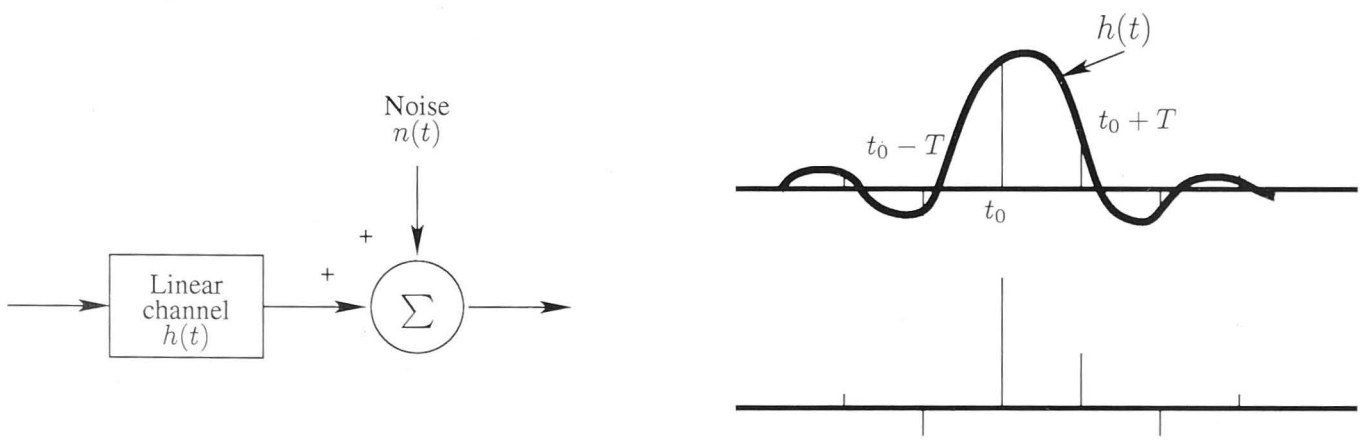

Figure 2.3: (A) Linear noisy baseband equivalent channel model. (B) Discrete-time channel impulse response obtained from the continuous-time response.

Consider a linear time invariant (LTI) baseband equivalent channel[1] whose impulse response $h(t)$ is shown in Fig. 2.3. Then the received signal $r(t)$ is obtained by convolving the input signal $a(t)$ with the channel's impulse response $h(t)$, giving

$$r(t) = a(t) \otimes h(t) + n(t) \tag{2.3a}$$

$$= \sum_j a_j h(t - jT) + n(t) \tag{2.3b}$$

[1]In this model, the channel includes the effects of the transmitter filter, the modulator, the transmission medium, the receiver filter and the demodulator.

where $\otimes$ denotes convolution, T seconds is the signalling interval and $n(t)$ is the additive white Gaussian noise (AWGN). When the continuous-time signal is sampled at $kT + t_0$, where t_0 accounts for the channel delay and sampler phase, the sampled signal becomes

$$r(t_0 + kT) = a_k h(t_0) + \sum_{j \neq k} a_j h(t_0 + kT - jT) + n(t_0 + kT) \tag{2.4a}$$

$$= a_k h(t_0) + v(t_0 + kT). \tag{2.4b}$$

The first term on the right hand side (RHS) of (2.4a) is the desired term scaled by $h(t_0)$ since it can be used to identify the transmitted signal level. The middle and last terms on the RHS of (2.4a) are the ISI and AWGN respectively. The sum of ISI and noise is also known as the *effective noise* and is denoted by $v(k) \triangleq v(t_0 + kT)$.

In the frequency domain, the received signal may be represented by the following multiplicative process

$$R(f) = A(f)H(f) \tag{2.5}$$

where $H(f)$ is the Fourier transform of $h(t)$:

$$H(f) = \mathcal{F}\big(h(t)\big) = \int_{-\infty}^{\infty} h(t) e^{-j2\pi ft} dt, \tag{2.6}$$

and $R(f)$, $A(f)$ are similarly defined as $\mathcal{F}\big(r(t)\big)$ and $\mathcal{F}\big(a(t)\big)$, respectively. A sufficient condition for zero ISI is for the folded spectrum to have a flat magnitude at all frequencies for $|f| \leq 1/2T$ and a linear phase. This is also known as the Nyquist criterion. Most channels that are considered in the thesis are ISI channels which linearly distort the transmit signals and the aim objective of equalization is to combat such distortions.

The presence of ISI and noise can be easily visualized using the eye diagram shown in Fig. 2.4. The figure is generated assuming an ideal raised cosine filter was used with no added noise. The sampling should be carried out at a time offset that corresponds to the two well defined points in the middle where the eye opening is the largest in the vertical direction. Therefore, when and if the eye is open, the symbol can be accurately detected at the receiver by a simple decision device with an appropriate timing. When the signal-to-noise-ratio (SNR) drops due to the addition of the AWGN and/or ISI, then the eye will begin to close, resulting in an increase of the probability of an error.

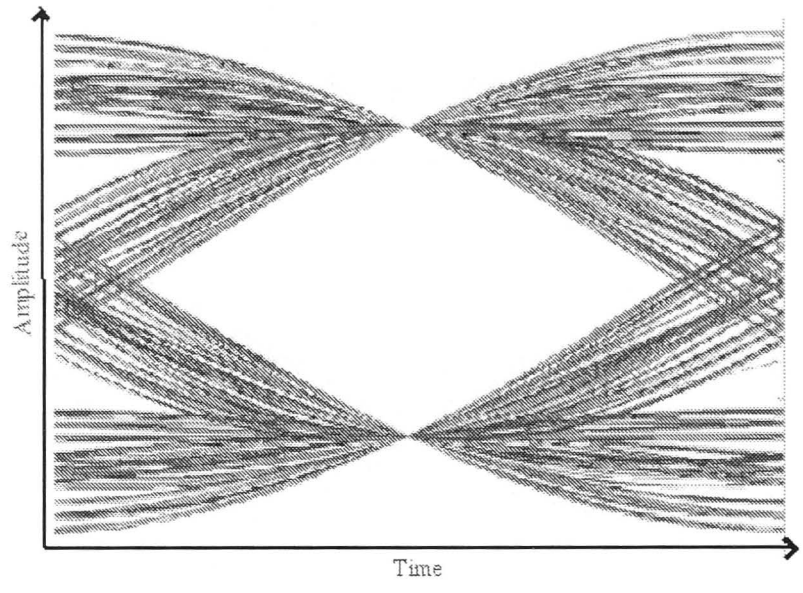

Figure 2.4: Eye diagram.

2.1.3 Performance Measures

The ultimate aim of the receiver is to produce a sequence of data estimates with a minimum probability of error. When such a measure is difficult to estimate or impractical, there are other simpler performance indices that are closely related to the bit-error-rate. Three commonly used indices are:

1. The closed eye measure (CLEM)

$$\text{CLEM} = \frac{\sum_i |s_i| - \max_i(|s_i|)}{\max_i(|s_i|)}, \tag{2.7}$$

where the sequence with element s_i is the combined channel-equalizer response. It is the ratio of the noise margin level to the total signal level. A small value of the CLEM indicates that the channel eye is sufficiently "open". Usually, if the CLEM is less than unity, then the decision directed algorithm can be reliably employed and the convergence to undesirable local equilibria can be avoided [92], [57, Ch. 3].

2. The ISI measure

$$J_{\text{ISI}} = \frac{\sum_i |s_i|^2 - \max_i(|s_i|^2)}{\max_i(|s_i|^2)}. \tag{2.8}$$

This measure indicates to what extent the equalized signal is influenced by adjacent samples of the same signal. It is the ratio of the power noise margin of to the combined channel-equalizer power.

3. Decision directed mean squared error (DD-MSE)

$$\mathrm{MSE}_{\mathrm{DD}}(k) = \mathrm{E}\Big\{ \big| z(k) - \mathrm{Q}(z(k)) \big|^2 \Big\} \tag{2.9}$$

where $\mathrm{Q}(z(k))$ is the nearest neighbor quantization operator appended to the equalizer output $z(k)$. A small value of $\mathrm{MSE}_{\mathrm{DD}}(k)$ indicates the raw (unquantized) equalizer output is close to a symbol value.

Other measures that can be used include the Kolgomorov-Smirnov as well as the Chi-Squared goodness-of-fit tests, with a suitably chosen null hypothesis, to determine if the distribution of the equalizer output corresponds to an open eye condition. An example is found in [8].

Among the three above mentioned performance indices, only the $\mathrm{MSE}_{\mathrm{DD}}(k)$ can be practically used under blind equalization because it does not require knowledge of both the unavailable channel and transmitted sequence. The CLEM and ISI measure can be used for theoretical analysis but not for practical real-time digital communication systems.

2.2 Equalizer Structures

We now describe several equalizer structures which can be used to mitigate the effects of ISI. The choice of the equalizer structure will largely determine the speed of recovery of the linearly distorted data sequence (convergence of the equalizer) and the error probability of the equalizer output. A linear transversal equalizer can be used to invert the channel and hence recover the transmitted input data. Non-linear equalizers such as the maximum-likelihood sequence estimator and the decision feedback equalizer exploit the discrete nature of the input symbols to cancel ISI and noise in a more effective manner.

2.2.1 Linear transversal equalizers

In order for the receiver to reliably detect the original transmitted signal, the effective noise (i.e., the sum of ISI and AWGN) of (2.4) must be reduced to lower decision errors. The most common and simplest channel equalizer structure is the linear transversal equalizer, or otherwise known as the tapped-delay line or non-recursive equalizer as shown in Fig. 2.5. In digital implementation, the sampled received signal at time kT is given as $r(k)$, where the symbol-rate samples are stored in digital shift registers or memory. The equalizer output samples are computed

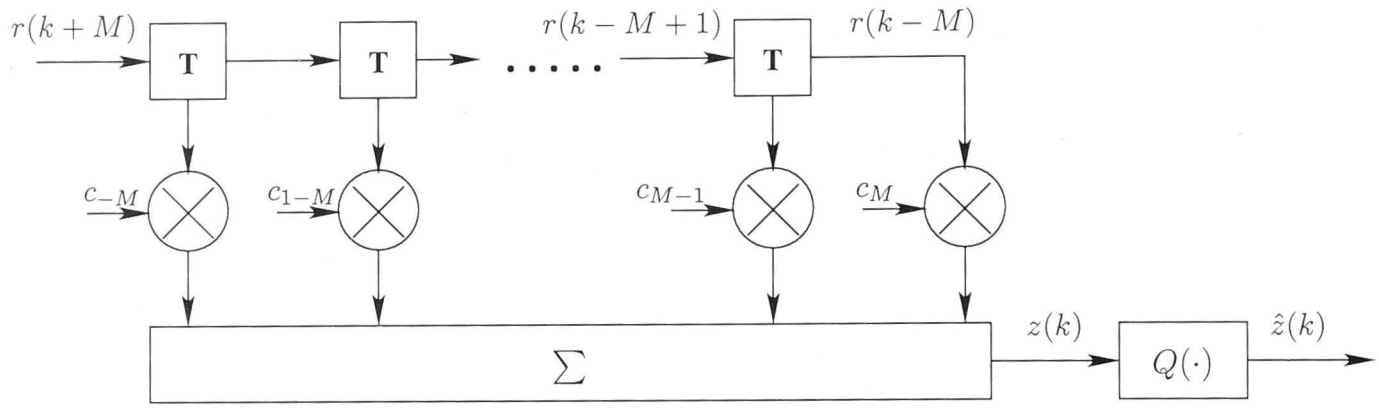

Figure 2.5: Linear Transversal Filter.

according to

$$z(k) = \sum_{i=-M}^{M} c_i r(k-i) \tag{2.10}$$

where $\{c_i\}$ are the $(2M+1)$ tap weight coefficients of the equalizer. Linear equalizers designed on the basis of the baud-rate sampled received signal are quite sensitive to symbol timing errors. Therefore, fractionally spaced linear equalizers are widely used to mitigate the equalizer's sensitivity to symbol timing errors. A fractionally spaced equalizer (FSE) in the linear transversal structure has the output

$$z(k) = \sum_{i=-M}^{M} \mathbf{c}_i^T \mathbf{r}(k-i) \tag{2.11}$$

where $\mathbf{c}_i^T$ and $\mathbf{r}(i)$ are column vectors with more than one sample per symbol according to the oversampling factor, while $\{\mathbf{c}_i\}$ are the $(2M+1)$ *vector* tap weight coefficients of the FSE. The linear equalizers can also be implemented as a lattice filter [118] which is known for its improved convergence and numerical properties.

2.2.2 Non-linear equalizers

In digital communications, the actual goal of equalization is to recover the input which is discrete and not to invert the channel. Hence despite the channel being linear the equalizer needs not be linear. In fact, by most criteria the best equalizer structures should be non-linear.

Linear equalizers generally do not perform satisfactorily when the channel is severely distorted especially when it has deep spectral nulls in the passband. Non-linear equalizers can deal with such channels more effectively. In the realm of non-linear equalizers, two well known approaches are:

Maximum-Likelihood (ML) Sequence Estimator It has been shown be Se-

shadri [125] that it is possible to perform joint data and channel estimation using the Viterbi algorithm [44]. However, unlike conventional maximum likelihood sequence estimators (MLSE) where the Viterbi algorithm operates on only one trellis, for joint data and channel estimation, the VA operates on many trellises where each of them corresponds to a hypothesized estimate of the channel. The number of trellises grows exponentially with the length of the channel impulse response. The MLSE operates based on the following concept:

The MLSE estimates the information sequence to maximize the joint probability of the received sequence conditioned on the information sequence. The linear FIR channel output is

$$r(k) = \sum_{i=0}^{L} h_i a(k-i) + n(k) \tag{2.12}$$

where $L+1$ is the channel length and $n(k)$ is the channel noise. Since $n(k)$ is often white Gaussian, the ML estimate of the channel input $a(k)$ based on a sequence of channel output $r(k)$ can be obtained if the channel impulse response is known or has been estimated. This is done by maximizing the likelihood function, or equivalently, by minimizing

$$\sum_{k=L}^{\infty} \left| r(k) - \sum_{i=0}^{L} h_i a(k-i) \right|^2 . \tag{2.13}$$

If the size of the symbol alphabet is M, then there are M^L different possible states or transmitted sequences. The principle of ML receivers is then to choose, among these M^L sequences the most likely one to have produced the received sequence $\{r(k)\}$. Finding the most likely sequence involves exhaustive calculation of all the M^L metrics, which can be quite complex if the number of states M^L is large. A substantially lower computational complexity can be achieved by employing the Viterbi algorithm. Reduced state Viterbi algorithms that provide good compromises between complexity and performance by assuming some past decisions are correct have been proposed for channels with long but small tails [37].

Due to the high computational cost of the MLSE which is impractical in real-time high-data-rates digital communication systems, alternative structures such as the DFE are usually preferred.

Decision Feedback Equalizer The philosophy of a decision feedback equalizer (DFE) is to employ previously detected symbols to reduce the impact of ISI and noise due to those symbols that have distorted the current symbol. The nonlinearity that arises from the use of a decision device significantly improves the performance of the DFE because it exploits the discrete properties of the transmitted sequence. However, it suffers from the error propagation phenomenon whenever a decision error is made [39, 71]. This phenomenon is due to incorrect decisions traversing the feedback delay line, resulting in enhanced probability of error until the states have been flushed out completely.

In this thesis, we deal with two types of DFE extensively in Chapter 5. These DFE's have been commonly used in the literature and are briefly described below:

Conventional Decision Feedback Equalizer

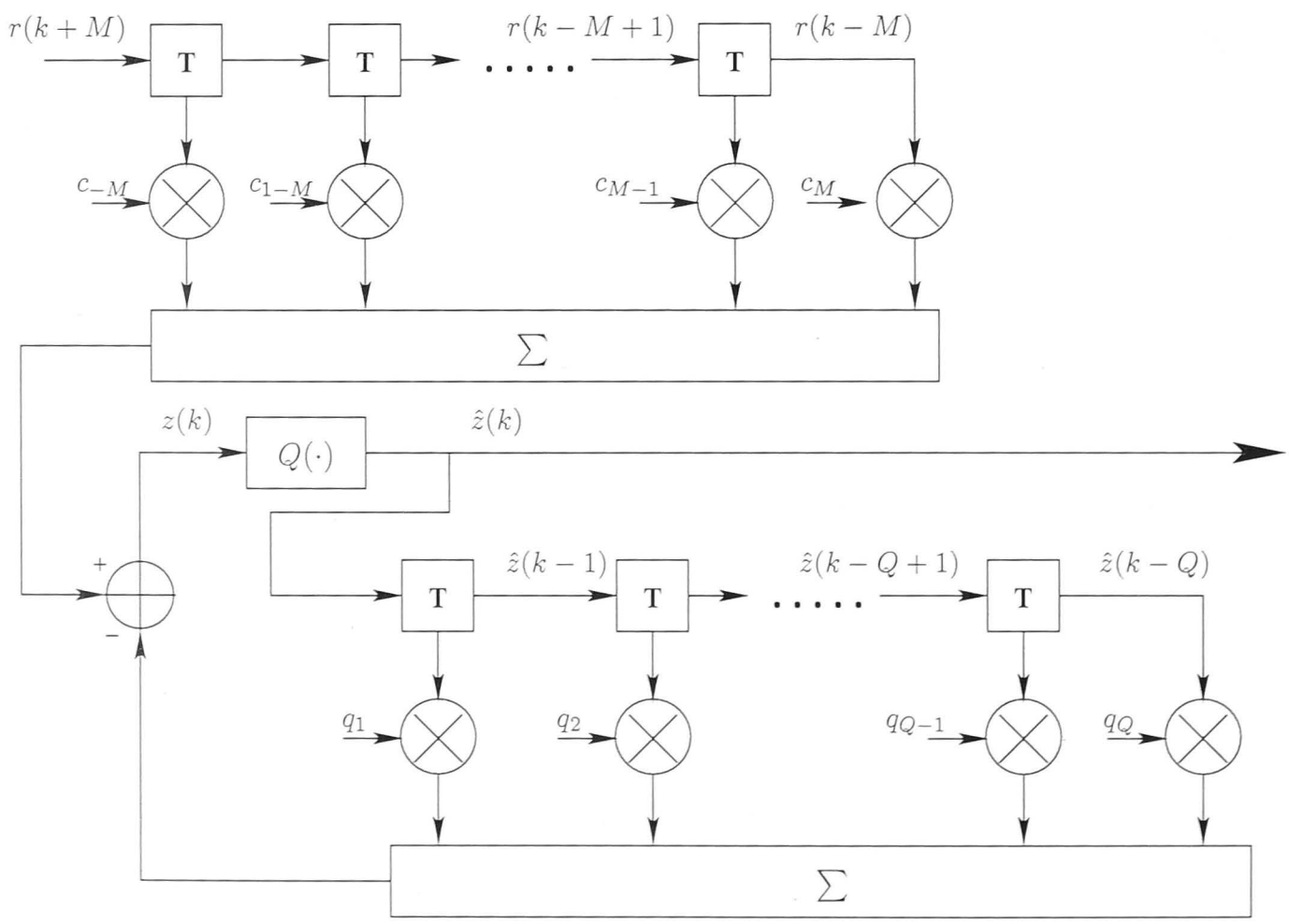

Figure 2.6: Conventional decision feedback equalizer.

The DFE was first developed by Austin [7]. It was intended to combat the ISI due to channels with severe amplitude distortions, using decision feedback to cancel the interference from symbols which have already been detected. It is implemented as a cascade of a forward filter and a feedback filter that is equipped with a memoryless decision device. Fig. 2.6 shows a symbol-rate DFE. The equalized signal is the sum of the outputs of the forward and feedback parts of the equalizer,

given by

$$z(k) = \sum_{i=-M}^{M} c_i r(k-i) - \sum_{j=1}^{Q} q_j \hat{a}(k-j) \tag{2.14}$$

where q_j is the tap weight coefficient of the feedback filter with length Q, and $\hat{a}(k)$ is the output of the decision device. The forward filter is similar to the linear equalizer that may be symbol-rate or fractionally spaced. The feedback filter however must have tap spacing that equals the symbol interval.

Predictive Decision Feedback Equalizer

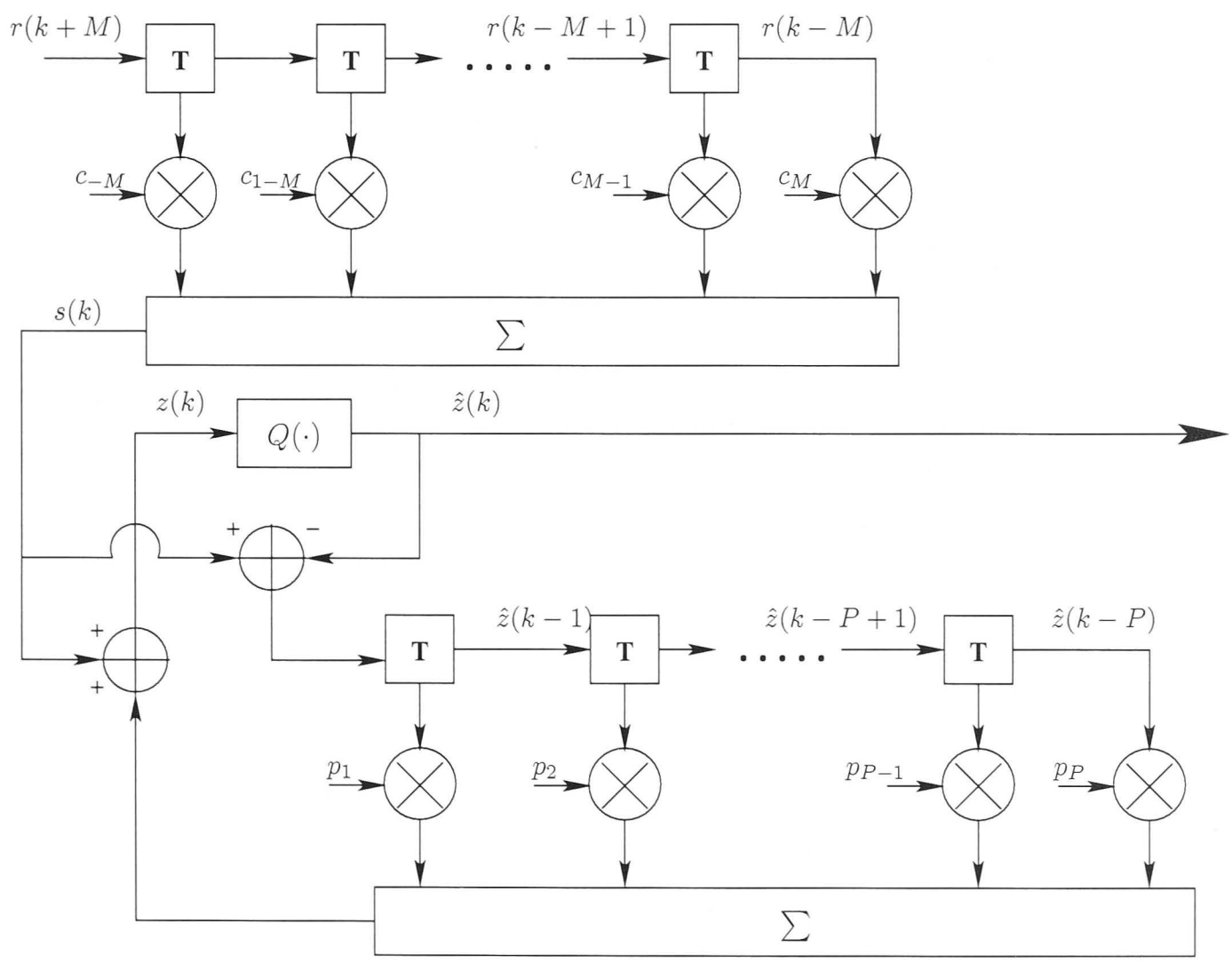

Figure 2.7: Predictive decision feedback equalizer.

A nonlinear equalizer, shown in Fig. 2.7, that is equivalent to the conventional DFE under the condition that the forward filter has an infinite number of taps was proposed by Belfiore and Park [10]. It is known as the predictive DFE and is developed as follows. Given the MMSE forward filter, say an infinite-length FSE, then the sequence of symbol-rate signals at the output of this forward filter forms a set of sufficient statistics for estimating the transmitting sequence. Thus by implementing the linear predictor as a feedback filter shown in the configuration given in Fig. 2.7, the MSE may be further reduced by the memoryless decision

device that eliminates the contribution of ISI and noise to the equalized signal.

The output of the MMSE forward filter is given in the usual form of (2.11). The predictive DFE output, which is the input to the decision device, is given by

$$z(k) = \sum_{i=-M}^{M} c_i r(k-i) - \sum_{j=1}^{P} p_j \{ s(k-j) - \hat{a}(k-j) \} \tag{2.15}$$

where p_j is the tap weight coefficient of the predictive feedback filter with length P, and $s(k)$ is the output of the forward filter.

Remarks on conventional DFE and predictive DFE :

1. Both conventional DFE and predictive DFE have identical expressions of minimum achievable MSE when the lengths of the forward filters of the equalizers are unconstrained, even when the feedback filter is reduced to a finite length. The output of the infinite length linear predictor is a white noise sequence with corresponding minimum MSE that is identical to that of the conventional DFE. A proof of their equivalence can be found in [119], [118].

2. In contrast to the conventional DFE whose forward filter depends on the number of feedback coefficients, the forward filter of the predictive DFE is independent of the predictor coefficients. Optimizations of the forward filter and the feedback predictor in the predictive DFE can be done separately. Hence, its MSE is at least as large as that of the conventional DFE. In spite of this sub-optimality of the predictive DFE, it is more suited for trellis-coded signals than the conventional DFE [40, 143].

3. Usually the conventional DFE and the predictive DFE are not used simultaneously because they are essentially equivalent [10]. However, we found a new way to incorporate both DFE's in Chapter 5.6 to enhance the tracking capability of the equalizer in time-varying channels.

2.3 Criteria That Lead To ISI Cancellation

Throughout Section 2.2, commonly used linear and non-linear equalizer structures were presented. The transmitted sequence is estimated by filtering the channel output with a set of filter coefficients that ideally produce an output sequence with minimum average probability of error. Unfortunately, the probability of error is a highly non-linear function of the tap coefficients and cannot be easily computed

without a known training sequence. Consequently other criteria in the spirit of lowering the error probability have been proposed. We will review two criteria in this section, namely the peak distortion criterion and the mean-square-error criterion. The former is used to invert the channel directly while the latter minimizes the error between the raw equalizer output data with its quantized output. In digital communications where the input is discrete, the latter criterion is usually preferred.

2.3.1 Zero Forcing Algorithm

The zero forcing (ZF) equalizer minimizes the so-called *peak distortion* criterion [118], which is in essence the residual ISI (without the AWGN term) given by

$$J_{\mathrm{ZF}} = \sum_{k=-\infty, k\neq 0}^{\infty} |s(k)| \tag{2.16a}$$

$$= \sum_{k=-\infty, k\neq 0}^{\infty} \left| \sum_{j=-\infty}^{\infty} c_j(k) h(k-j) \right| \tag{2.16b}$$

where $\{s(k)\}$ is the combined channel-equalizer impulse response, i.e., the convolution of the channel impulse response, $\{c_j\}$, and the equalizer coefficients, $\{h_j\}$. The ZF cost is the peak value of the interference term *minus* the term corresponding to the $(k=0)^{\mathrm{th}}$ tap and the noise term, divided by $s(0)$ which is often normalized to be unity. With an infinitely long equalizer, J_{ZF} can be zeroed. Effectively, this corresponds to the complete elimination of ISI. The values of the taps are chosen to yield non-zero overall channel-equalizer response at decision time, i.e., $k=0$, and zero elsewhere. In other words this corresponds to choosing equalizer coefficients so that the following condition is satisfied:

$$s(k) = \sum_{j=-\infty}^{\infty} c_j(k) h(k-j) = \delta_{k,0} \tag{2.17}$$

where

$$\delta_{k,0} \triangleq \begin{cases} 1 & k=0 \\ 0 & \text{otherwise} \end{cases}$$

denotes the Kronecker delta. In the z domain, this leads to the direct inversion of the channel (assuming it is invertible):

$$C(z) = \frac{1}{H(z)} \tag{2.18}$$

where $C(z)$ and $H(z)$ denote the z transform of the $\{c_j\}$ and $\{h_j\}$, respectively. Even though the zero-ISI condition generally requires infinite-length symbol-rate ZF equalizers, it can be shown that zero-ISI ZF equalizers of finite length generally exist for fractionally-spaced receivers [118].

Minimizing the peak distortion using a *finite* length symbol-rate linear equalizers cannot completely eliminate the ISI at the output of the equalizer. However, it can at the least be shown to be a convex function of the equalizer coefficients [88]. A steepest-descent recursive algorithm for adjusting the $(2K+1)$ equalizer coefficients would therefore be

$$c_j(k+1) = c_j(k) + \mu\epsilon(k))a^*(k-j), \qquad j = -K, \cdots, -1, 0, 1, \cdots, K \tag{2.19}$$

where $c_j(k)$ is the value of the j^{th} coefficient at time $t = kT$, $\epsilon(k) = a(k) - z(k)$ is the error signal at time $t = kT$ and μ is the adaptation step size. When the equalizer coefficients have converged to their global minima, the error is orthogonal to the transmitted sequence, giving

$$\begin{aligned}
\mathrm{E}(\epsilon(k)a^*(k-j)) &= \mathrm{E}\big[(a(k) - z(k))a^*(k-j)\big] && \text{(2.20a)}\\
&= \mathrm{E}\big[a(k)a^*(k-j)\big] - \mathrm{E}\big[z(k)a^*(k-j)\big] && \text{(2.20b)}\\
&= \delta_{j,0} - s(j) && \text{(2.20c)}\\
&= 0 \qquad\qquad j = -K, \cdots, K && \text{(2.20d)}
\end{aligned}$$

where we have assumed the transmitted sequence to be i.i.d. and the equalizer output is correct, so that $\mathrm{Q}(z(k)) = a(k)$. The last two lines follow because the optimal tap coefficients will give the zero-ISI condition, i.e., $s(0) = 1$ and $s(k) = 0$ for $1 \leq |k| \leq K$.

2.3.2 Least Mean Squared (LMS) Algorithm

The ZF equalizer minimizes the ISI without explicitly taking into account the effects of channel noise. For channels that exhibit deep spectral nulls, the channel inversion property of the ZF algorithm may amplify the noise component at frequencies corresponding to the nulls in the spectrum, resulting in significant deterioration

of the bit-error-rates of the equalizer output. The minimum mean squared error (MMSE) criterion on the other hand ameliorates this problem. Its philosophy is rather than trying to completely eliminate the ISI, the MMSE equalizer minimizes a balanced contribution from the ISI and the channel noise. The instantaneous mean squared value of the error is

$$\epsilon(k) = a(k) - z(k) \tag{2.21}$$

which is the derivative of the MMSE cost function with respective to the equalizer output. Thus, the MMSE cost is

$$J_{\text{MMSE}} = \mathrm{E}\Big(\big|a(k) - z(k)\big|^2\Big) \tag{2.22}$$

$$= \mathrm{E}\Big(\big|a(k) - \mathbf{r}^H(k)\mathbf{c}(k)\big|^2\Big) \tag{2.23}$$

where $\mathbf{r}(k)$ is the column vector of the channel output and $\mathbf{c}(k)$ is the column vector of the tap coefficients of the transversal equalizer. The solution to the above criterion is the MMSE equalizer which is given in the z domain by

$$C(z) = \frac{1}{H(z) + N_0} \tag{2.24}$$

where the additive noise at the channel output is assumed white with power spectral density N_0. The minimization of the cost leads to the well known Wiener filter whose (optimal) tap weights $\mathbf{c}^*$ are determined by the following Wiener-Hopf equation

$$\mathbf{c}^* = \mathbf{R}^{-1}\mathbf{p} \tag{2.25}$$

where $\mathbf{R}$ is the autocorrelation matrix of the equalizer input and $\mathbf{p}$ denotes the cross-correlation of the equalizer input and the desired signals [58].

2.3.3 Unsupervised (Blind) Algorithms

The ZF and MMSE equalizers are usually implemented with a training sequence that is known to both the transmitter and the receiver. When the transmission of these sequences is impractical or too costly, an alternative solution is to rely on blind equalization techniques. The Bussgang-type techniques estimate the training signal based only on the received signal and certain *a priori* statistical information regarding the transmitted signals. We devote the following sections to discuss

certain blind algorithms in the literature as this thesis deals primarily with the class of blind algorithms. We first give a brief overview of non-Bussgang techniques before discussing in more detail the class of Bussgang algorithms which is used extensively in this thesis.

2.4 Overview of Non-Bussgang Techniques

2.4.1 Historical Notes

To understand a science it is necessary to know its history.
Auguste Comte (1798–1857)

The problem of blind deconvolution is the subject of thorough research over the past few decades; under different names and for various applications. The homomorphic filter was introduced to remove effects of non uniform illumination of images [107, Ch. 10]. An observed image is usually described as a convolution of the object brightness distribution (OBD) by a Point Spread Function (PSF), and blind deconvolution is used to recover the OBD when no reliable information regarding the PSF is available. The name "blind deconvolution" was first used by Stockham *et al* for the restoration of old records [132]. Later "minimum entropy deconvolution" in seismic data analysis was introduced [149] to find the inverse of the channel that maximizes the kurtosis of the deconvolved data. The same concept was later applied to synthetic aperture radar focusing to estimate the rate of change of the Doppler shift of radar echoes [77]. It is also used for image reconstruction to remove the effects of the blur induced on astronomical plates by short term variations of the refraction index of the atmosphere [154]. At the same time, similar concepts were applied to multilevel blind data transmission over telephone and radio channels.

The search for cost-function-based blind algorithms started with the seminal work of Sato and Godard [49,124]. Their algorithms were generalized by Benveniste *et al* [14] and Sethares *et al* [126], respectively. Later the Bussgang methods for blind equalization, which subsume all the algorithms in this section, were discovered by Bellini [11]. Other methods include CMA with multiple radii [120, 126], sign algorithms [144], convex cost functions [68], super exponential methods [127] and the so-called multimodulus-algorithm (MMA) [105, 106, 151, 152]. Alternatively, second order statistics-based algorithms that emerged in the early 90's due to Gardner and Tong *et al* [45,135] have received a lot of attention. Some earlier works by Tong *et al* include [134,136,155]. Higher-order statistics-based algorithms which

include cumulant matching algorithms were proposed [56,140], as well as maximum likelihood estimation using the Viterbi algorithm [47,125], [57, Ch. 6].

Much of the research over these two decades on blind equalization centered on the analysis of particular blind algorithms. Generally, researchers are concern with the convergence behaviors of these algorithms under both ideal conditions [14,99] and also the case when they are violated, such as finite length filter effects [26,27,35]. Apart from the work related to analysis, significant amount of research is dedicated to extension of original equalization algorithms and structures. For example, Benveniste *et al* [15] and Picchi *et al* (Stop-And-Go) [116] proposed dual-mode algorithms that cater for smooth switching between the Sato and the decision directed (DD) algorithms. In addition, some new algorithms that enable a soft transition between the CMA and the DD algorithm have also been proposed [60,145]. In the domain of equalizer structures, both blind decision feedback equalization [5,17,66,75,112,133] and neural network implementations [43], [58, Ch. 19] have also been considered more recently.

2.4.2 Classifications of Blind Deconvolution Algorithms

We have broadly identified four families of blind deconvolution algorithms. They are:

1. The Bussgang algorithms where a non-linear operation is performed on the *output* of the adaptive equalization filter.

2. The polyspectra algorithms where a non-linear operation is performed on the *input* of the equalizer.

3. Probabilistic algorithms, where the non-linearity is in the data detection process.

4. The cyclostationary statistics-based algorithms which exploit the second-order cyclostationary statistics of the received signal. The property of cyclostationarity is known to arise in a modulated signal that results from varying the amplitude, phase, or frequency of a sinusoid carrier, which is basic to the electrical communications process.

For the remaining parts of this section, the last three types of algorithms will be briefly described, namely the polyspectra type, cyclostationary statistics based and the probabilistic algorithms. As the focus of this thesis is blind Bussgang algorithms, a whole section is dedicated to reviewing the rich but essential literature in this area.

2.4.3 Polyspectra algorithms

One method of blind deconvolution that has attracted a lot of attention in the late 80's and early 90's is the polyspectra algorithm. It explicitly exploits higher order moments of the channel output in order to recover both magnitude and phase information regarding the channel, where second order statistics based algorithms fail. One important assumption is that the transmitted signal is not Gaussian because otherwise phase identification is not possible (since for a Gaussian signal, its higher order moments, i.e., order three and above, are zero). This assumption is easily met in practical digital communications system.

The motivation behind using higher order cumulants or their Fourier transforms called polyspectra for blind equalization include:

i. Equalization techniques based on higher order statistics are independent of a particular transmitted sequence (the training sequence) since digital communications signals are data sequences with common statistical properties. They are also more robust with respect to the channel input probability density functions provided they are not too close to being Gaussian, which is typically the case because the distribution is discrete.

ii. Polyspectra have the ability to identify non-minimum phase communications channels from output measurements by virtue of their ability to be sensitive to phase and magnitude information in the channel output.

iii. All polyspectra of Gaussian processes of order greater than two are zero. Consequently, polyspectra techniques will not be affected by additive Gaussian noise which might be present in communications systems, assuming they are independent of the data.

iv. Global convergence to the desired solution is guaranteed in approaches of the equation error type (see below).

As noted in [141], given the mathematical model, there are two broad classes of approaches to channel estimation and equalization. They are distinguished only by their respective choices of optimization criteria:

Equation Error The approaches of the Tricepstrum Equalization Algorithm (TEA) and its extensions by Hatzinakos *et al* [56] and in [48] minimize an "equation error" in some equation which is satisfied ideally. They showed equalization of a non-minimum phase channel (which we now assume is FIR for illustration,

under the assumption that no zeros lie on the unit circle) given by

$$H(z) = \mathcal{K} I(z) O(z^{-1}) \tag{2.26a}$$

$$I(z) = \sum_{n=1}^{N_1} (1 - a_n z^{-1}), \qquad |a_n| < 1 \tag{2.26b}$$

$$O(z^{-1}) = \sum_{n=1}^{N_2} (1 - b_n z), \qquad |b_n| < 1 \tag{2.26c}$$

where $I(z)$, $O(z^{-1})$ are minimum phase and maximum phase polynomials respectively, can be identified and used for sequence estimation. Here we give the briefest outline of the equation error approach whilst the details can be found in [57, Ch. 5], [58, Ch. 18] and [56].

Let $c_4(\tau_1, \tau_2, \tau_3)$ denote the fourth-order cumulant of the channel output $r(k)$. The tricepstrum of $r(k)$ is thus

$$C_4(\omega_1, \omega_2, \omega_3) = F\big[c_4(\tau_1, \tau_2, \tau_3)\big] \tag{2.27}$$

where $F[\cdot]$ denotes the three-dimensional discrete Fourier transformation. Define also the so-called tricepstrum of the process $r(k)$ as

$$\kappa_4(\omega_1, \omega_2, \omega_3) = F^{-1}\big[\ln C_4(\omega_1, \omega_2, \omega_3)\big] \tag{2.28}$$

where F^{-1} denotes the inverse three-dimensional discrete Fourier transformation and ln is $\log_e$. The following relation between $c_4(\tau_1, \tau_2, \tau_3)$ and the assumed i.i.d. channel input sequence holds [58, Ch. 3.7], [100]

$$c_4(\tau_1, \tau_2, \tau_3) = \gamma_4 \sum_{i=0}^{\infty} h_i h_{i+\tau_1} h_{i+\tau_2} h_{i+\tau_3}. \tag{2.29}$$

It can be shown that $\kappa_4(\omega_1, \omega_2, \omega_3)$ can be expressed in terms of the zeros and

poles of $H(z)$ in the following form:

$$\kappa_4(\omega_1,\omega_2,\omega_3) = \begin{cases} \ln \mathcal{K} + 3\ln\gamma_4, & \tau_1=\tau_2=\tau_3=0 \\ -\tau_1^{-1}A^{(\tau_1)}, & \tau_1>0, \tau_2=\tau_3=0 \\ -\tau_2^{-1}A^{(\tau_2)}, & \tau_2>0, \tau_1=\tau_3=0 \\ -\tau_3^{-1}A^{(\tau_3)}, & \tau_3>0, \tau_1=\tau_2=0 \\ \tau_1^{-1}B^{(\tau_1)}, & \tau_1<0, \tau_2=\tau_3=0 \\ \tau_1^{-1}B^{(\tau_2)}, & \tau_2<0, \tau_1=\tau_3=0 \\ \tau_1^{-1}B^{(\tau_3)}, & \tau_3<0, \tau_1=\tau_2=0 \\ -\tau_2^{-1}B^{(\tau_2)}, & \tau_1=\tau_2=\tau_3>0 \\ \tau_2^{-1}A^{(\tau_2)}, & \tau_1=\tau_2=\tau_3<0 \\ 0 & \text{otherwise} \end{cases} \tag{2.30}$$

where

$$A^{(m)} = \sum_{n=1}^{N_1} a_n^m, \qquad B^{(m)} = \sum_{n=1}^{N_2} b_n^m. \tag{2.31}$$

Note that $A^{(m)}$, $B^{(m)}$ contain information regarding the minimum phase and maximum phase roots of $H(z)$ in (2.26). The key equation that connects the fourth-order cumulant to the tricepstrum is [110]

$$\sum_{r=-\infty}^{\infty}\sum_{s=-\infty}^{\infty}\sum_{t=-\infty}^{\infty} r\kappa_4(r,s,t)c_4(\tau_1-r,\tau_2-s,\tau_3-t) = -\tau_1 c_4(\tau_1,\tau_2,\tau_3). \tag{2.32}$$

Substituting (2.30) into (2.32) yields

$$\begin{aligned} &\sum_{m=1}^{p}\Big(A^{(m)}\big[c_4(\tau_1-m,\tau_2,\tau_3) - c_4(\tau_1+m,\tau_2+m,\tau_3+m)\big]\Big) \\ &+\sum_{m=1}^{q}\Big(B^{(m)}\big[c_4(\tau_1-m,\tau_2-m,\tau_3-m) - c_4(\tau_1+m,\tau_2,\tau_3)\big]\Big) \\ &= -\tau_1 c_4(\tau_1,\tau_2,\tau_3). \end{aligned} \tag{2.33}$$

Representing (2.33) in matrix form as an over-determined linear system of equations we get

$$\mathbf{C}\,\mathbf{a} = \mathbf{p} \tag{2.34}$$

where we wish to determine $\mathbf{a}$ to identify the non-minimum phase channel,

and $\mathbf{C}$ is a matrix with entries of the form $\{c_4(\tau_1, \tau_2, \tau_3) - c_4(\tau_1', \tau_2', \tau_3')\}$ and dimensions given in [58, Ch. 18], $\mathbf{p}$ is a vector with entries of the form $\{\tau_1 c_4(\tau_1, \tau_2, \tau_3)\}$, and $\mathbf{a}$ is a $(p+q)$-by-1 coefficient vector in terms of $A^{(m)}$ and $B^{(m)}$ given by

$$\mathbf{a} = \left[A^{(1)}, A^{(2)}, \cdots, A^{(p)}, B^{(1)}, B^{(2)}, \cdots, B^{(q)}\right]^T. \tag{2.35}$$

By estimating $\mathbf{a}$ it can be shown that the channel may be identified and a corresponding equalizer may be constructed.

Fitting Error This so-called cumulant matching technique that is adopted in [140] matches the model-based higher order statistics to the estimated data-based statistic in a least-squares sense to estimate the channel impulse response. Unlike the "equation error" approach which results in a zero-forcing equalizer, this approach allows consideration of noisy observations. In general, it is a less robust approach since it requires a good initial guess to prevent the convergence of the channel estimator to undesirable local minima.

2.4.4 Algorithms based on cyclostationary statistics

For linear time-invariant frequency-selective deterministic channels, when the received waveforms are over-sampled at p times the baud-rate, the discrete signal displays scalar cyclostationarity properties that may be exploited for channel identification and equalization. Depending on its application, sequence estimation using this technique is viewed as a single-input-multiple-output (SIMO) equalization where a discrete-time vector of stationary sequences is available at the equalizer input, or fractionally-spaced equalization. The former considers the channel output as

$$r(k) = \sum_i \mathbf{h}_i a(k-i) + \mathbf{n}(k) \tag{2.36}$$

where $\mathbf{h}_i$ and $\mathbf{n}(k)$ are p-by-1 vectors. When the received signal is oversampled, the fractionally-spaced equalizer that employs the constant modulus algorithm (CMA) has been shown to converge to complete ISI removal under noiseless models [25, 79]. However this is achieved at the expense of higher complexity relative to baud-rate equalizers.

A subspace method that exhibits low complexity that leads to closed-form solutions is given in [86]. Consider the collection of m over-sampled received sequences

$\mathbf{r}(k)$ (multiple snapshots of $\mathbf{r}(k)$)

$$\mathcal{H}\big(\mathbf{r}(k)\big) \triangleq [\mathbf{r}_k, \mathbf{r}_{k-1}, \cdots, \mathbf{r}_{k-m}] = \mathcal{T}\big(\mathbf{h}\big)\big[\mathbf{a}_k, \mathbf{a}_{k-1}, \cdots, \mathbf{a}_{k-m}\big] \tag{2.37}$$

where $\mathcal{T}\big(\mathbf{h}\big)$ is the so-called filtering matrix, $\mathbf{h}$ is the channel vector, $\mathbf{a}_k$ is the input source vector. As long as $\mathcal{T}(\mathbf{h})$ has full column rank, the span of $\mathcal{H}(\mathbf{r})$ is identical to that of $\mathcal{H}(\mathbf{a})$. Then when $\mathcal{H}(\mathbf{a})$ has more modes than the rank of $\mathcal{T}(\mathbf{h})$, $\mathcal{H}(\mathbf{a})$ can be uniquely determined from the nullspace of $\mathcal{H}(\mathbf{r})$ and the transmitted sequence can be found by solving the set of linear equations

$$U^H(k)\mathcal{H}(\mathbf{a}) = 0 \tag{2.38}$$

where $U(k)$ is the nullspace of $\mathcal{H}(\mathbf{r})$ and ^{H} denotes the Hermitian transpose operator.

2.4.5 Probabilistic algorithms

An overview of this technique has been previously discussed in Section 2.2.2 under the maximum-likelihood sequence estimation section.

2.5 Bussgang Algorithms

This section reviews the class of Bussgang-type blind algorithms that is used extensively in this thesis. The Bussgang technique [11, 57] for blind equalization is to obtain an estimate of the input signal by sending the equalizer output through a zero-memory non-linear device $g(\cdot)$, and then the filter taps are updated using the LMS algorithm by minimizing the error between the equalizer output and the estimate of the input signal. The cost function of the Bussgang algorithm is thus

$$J_{\text{Bussgang}} = \mathrm{E}\left\{\big|z(k) - g\big(z(k)\big)\big|^2\right\}. \tag{2.39}$$

The equalizer model is assumed to be linear transversal with $2N+1$ symbol-rate taps, whose tap weight vector at sampling time kT is denoted as $\mathbf{w}(k)$. This same model is used for the rest of this section when dealing with Bussgang blind equalization techniques. The tap update equation of the equalizer is

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \mu\epsilon\big(z(k)\big)\mathbf{X}^*(k) \tag{2.40a}$$

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \mu\left(z(k) - g\big(z(k)\big)\right)\mathbf{X}^*(k), \tag{2.40b}$$

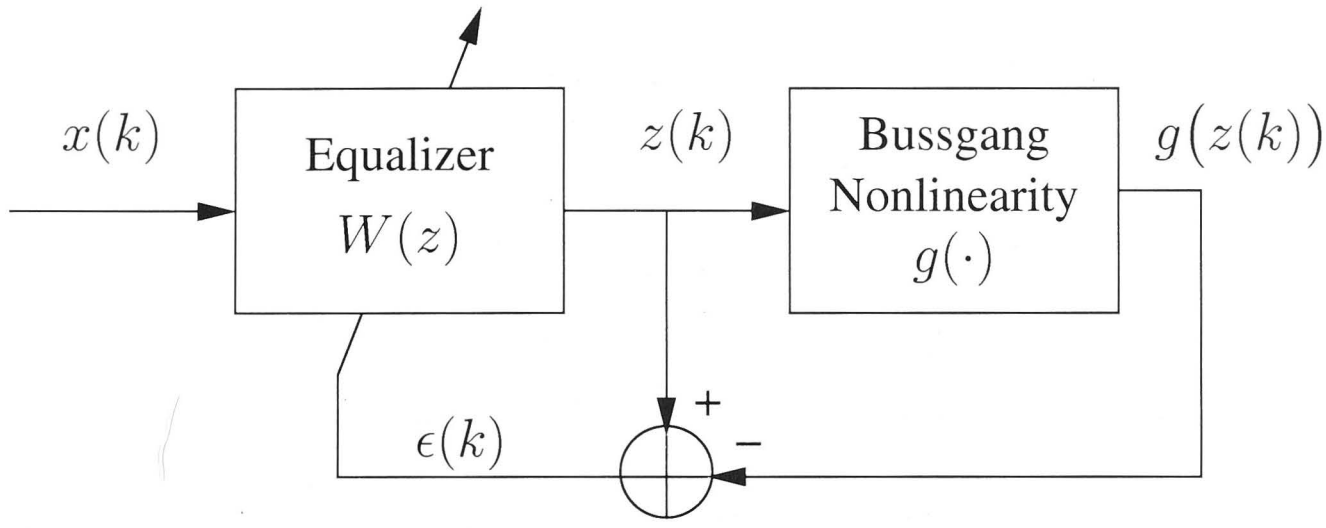

Figure 2.8: Linear Transversal Bussgang Equalizer.

where $\mathbf{X}(k)$ is the regressor vector of the linear equalizer that employs the Bussgang algorithm and $\epsilon(k)$ is the LMS error function implicit in (2.39). This approach was proposed by Godfrey and Rocca [51] for deconvolution of seismic traces and the term Bussgang algorithm was coined by Bellini [11] after he revealed that the algorithms based on higher order statistics implicitly are indeed identical to the methods described by Barrett and Lampard [9]. In the approach, the estimated input signal is cross-correlated with the equalizer output to obtain an approximate inverse filter.

The very first truly blind algorithm for multilevel digital communication systems was proposed by Sato back in 1975. However reliable blind digital transmissions has been sought after even before the revolutionary pioneering work of Sato. In this section, we go one step before the invention of the Sato algorithm to describe a simple blind algorithm called the Maximum Level Error (MLE) algorithm before describing some common blind algorithms that have enjoyed widespread popularity in today's digital receivers. In fact, the concept of the MLE algorithm has been partially exploited in the design of our switch-mode algorithm as described in Chapter 4.

2.5.1 Maximum-Level-Error (MLE) algorithm

The MLE algorithm [153] is usually not regarded as a truly blind algorithm. It is briefly explained here just to show how the evolution of blind equalization techniques started. The algorithm was proposed by Yatsuboshi et al based on a very simple idea. They observed that when the equalized signals are found beyond the largest constellation points, there is a higher probability that these signals are correct regardless of the SNR of the residual equalizer output. Therefore, the LMS algorithm which is decision directed, may be employed with a relatively higher rate

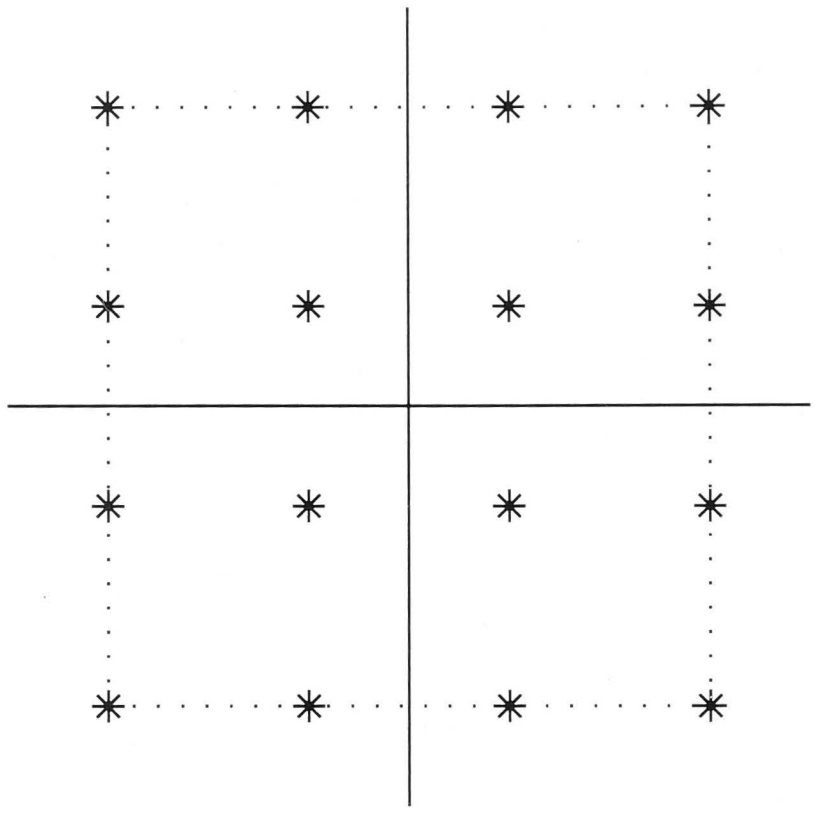

Figure 2.9: Maximum Level Error (MLE) Algorithm: Constellation diagram for 16-QAM with dotted-line boundary connecting the edge data points. The equalizer is only updated if the output is exceeds the boundary depicted by the dotted lines.

of detecting a correct symbol when the signal exceeds the boundary that joins the edge symbols, as depicted in Fig. 2.9. The filter taps are updated only when at least one of the following conditions

$$|\mathrm{Re}(z(k))| \geq \max_{a \in \mathbb{A}} |\mathrm{Re}(a(k))| \tag{2.41a}$$

$$|\mathrm{Im}(z(k))| \geq \max_{a \in \mathbb{A}} |\mathrm{Im}(a(k))| \tag{2.41b}$$

are met. The stochastic gradient update equation of the MLE equalizer is

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \mu\delta(k)\big(z(k) - \mathrm{Q}(z(k))\big)\mathbf{X}^*(k) \tag{2.42}$$

where $\delta(k)$ is the Kronecker delta that yields a value of 1 when at least one condition in (2.41) is satisfied, and 0 otherwise. Due to the low probability of satisfying the conditions in (2.41), very slow convergence of the equalizer is expected especially for large constellation sizes.

However, its concept has been exploited by several equalization and phase recovery techniques. In Chapter 4, we designed a switch-mode algorithm that performs joint equalization and phase recovery when one of the conditions in (2.41) is satisfied because such signals are generally more reliable. This concept is also similar to the "four-corner" phase recovery technique of [63, 139] where a phase discrimination is only performed when signals are "far" away from the origin, for example

when they are close to the "four-corners" of the QAM constellation.

2.5.2 Sato algorithm

The first blind algorithm was proposed by Sato [124] for multilevel digital transmission of real one-dimensional signals. His work has been widely recognized as a major breakthrough in blind digital communication systems. Sato's M-ary PAM blind equalization algorithm consists of minimizing a non-convex cost function

$$J_{\text{Sato}} = \mathrm{E}\left\{\left(z(k) - \gamma_{\text{S}} \cdot \operatorname{sgn}\big[z(k)\big]\right)^2\right\} \tag{2.43}$$

with the scaling coefficient that sets the gain of the equalizer defined by

$$\gamma_{\text{S}} \triangleq \frac{\mathrm{E}\{a^2(k)\}}{\mathrm{E}\{|a(k)|\}} \tag{2.44}$$

where $a(k)$ is the channel input sample. Sato's error function is thus

$$\epsilon_{\text{Sato}}(k) = z(k) - \gamma_{\text{S}} \cdot \operatorname{sgn}\big[z(k)\big]. \tag{2.45}$$

In what follows, we provide a summary of the work of Sato in [124]. He used the example of 8-PAM whereby the signal levels are given by

$$V\left(\frac{1}{2}d_1 + \frac{1}{4}d_2 + \frac{1}{8}d_3\right) \tag{2.46}$$

where d_1, d_2 and d_3 are binary random variables and V is a constant factor. The Sato algorithm estimates only the most significant digit d_1 while treating the d_2 and d_3 as random noise superimposed at the data source (see Fig. 1 in [124]). The concept of Sato's algorithm is thus to treat the multilevel digital signal as a "binary" signal. The algorithm then uses the results of this preliminary step to modify the error signal obtained from a conventional decision directed algorithm. The algorithm only makes a distinction between the signs of the equalizer output, subject only to a gain factor, γ_{S}. In short, rather than "pushing" the equalizer output towards the nearest data point as would conventional decision directed algorithms, the algorithm "pushes" the signals towards the half plane where they are found.

The original work of Sato made three strong and rather "impractical" assumptions for the algorithm to be globally convergent:

1. The PAM modulation admits an infinity number of levels, i.e.

$$V \cdot \left(\frac{1}{2} d_1 + \frac{1}{4} d_2 + \frac{1}{8} d_3 + \frac{1}{16} d_4 + \cdots \right)$$

where $d_i = \pm 1$ for all i, with a continuous uniform distribution in a symmetric interval of $(-V, V)$;

2. The combined channel-equalizer impulse response $\{s_i\}$ is doubly infinite in length;

3. The channel eye is initially open, i.e. CLEM < 1, where CLEM is defined in (2.7).

Under the above assumptions, the Sato cost function may be rewritten as

$$J_{\text{Sato}} = \frac{V^2}{3} \sum s_i^2 - \gamma_{\text{S}} \cdot V \left(s_0 + \frac{\sum_{i \neq 0} s_i^2}{s_0} \right) + \gamma_{\text{S}}^2. \tag{2.47}$$

J_{Sato} is shown to be a *convex*[2] function of the tap gains $\{c_i\}$ in Appendix B of [124] under the above assumptions. This convexity result guarantees that it has a unique minimum point that corresponds to a zero forcing equalizer (ZFE). These assumptions are indeed sufficient but *not necessary*. According to the Benveniste-Goursat-Ruget theorem [14], global convergence of the Sato algorithm can be achieved provided that the probability density function of the transmitted data sequence is a sub-Gaussian function such as the uniform distribution. The condition for doubly infinite equalizer is required however to ensure its global convergence. Deviations from this ideal behavior have been reported in the literature [26, 92, 99].

Perhaps the most important asymptotic result regarding the convergence of the BGR algorithm is that it can be shown to be globally convergent in the zero-forcing sense when the distribution of the input sequence $\{a(k)\}$ is sub-Gaussian when the equalizer length is unconstrained. The subsequent investigations carried out by Ding *et al* in [26, 27, 78] yield the conclusion that the violation of the ideal and impractical conditions can result in ill-convergence of the BGR algorithms to undesirable local minima.

Benveniste *et al* developed important results regarding the convergence of the Sato algorithm and its generalizations for sub-Gaussian and super-Gaussian inputs, where the input is continuous instead of discrete. The work of Kennedy, Ding *et*

[2] $f(x)$ is said to be a convex function of x if for any $0 < \alpha < 1$, the following holds: $f(\alpha x_1 + (1-\alpha)x_2 \leq \alpha f(x_1) + (1-\alpha) f(x_2)$, for any $x_1 < x_2$ in the domain of definition of f.

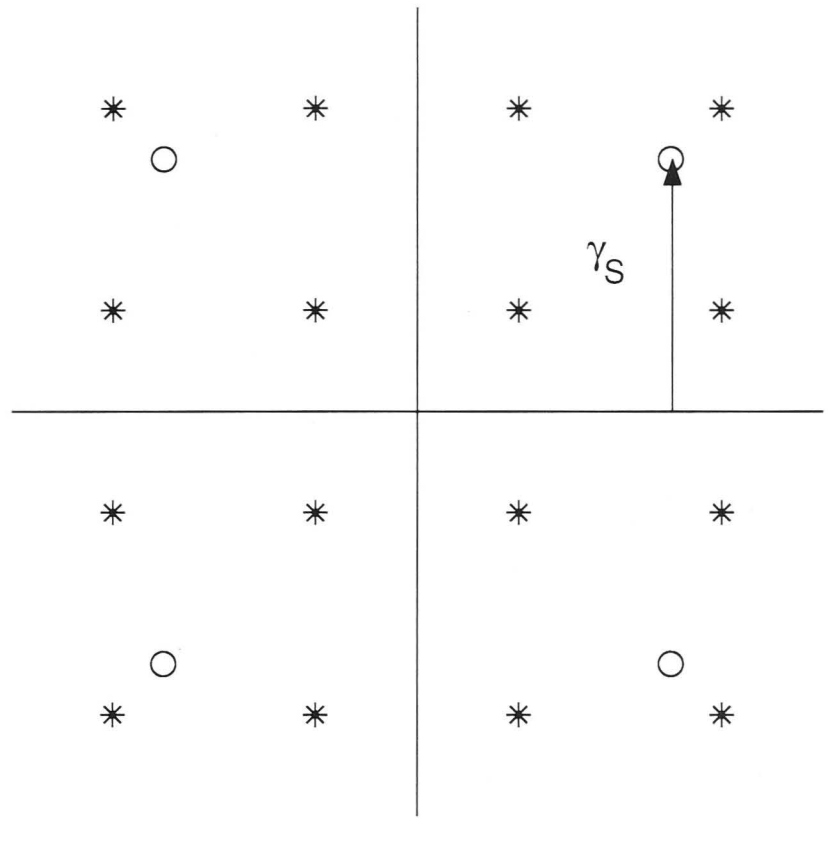

Figure 2.10: Reduced Constellation Algorithm (RCA): Constellation diagram for 16-QAM and its reduced constellation whose 4 data points are represented by 'o' at coordinates $[\gamma_S, \gamma_S]$, where γ_S is defined in (2.44).

al later revealed that undesirable local minima do exist in Sato equalizers under practical constraints of finite equalizer length and discrete inputs [26, 67].

Generalizations of Sato algorithm

The Sato algorithm, which is originally designed for real signals, is subsequently generalized by Benveniste, Goursat and Ruget [14] into a class of error functions, which can be used for complex signals, and is given by

$$\epsilon_{\mathrm{GSA}}\big(z(k)\big) = \tilde{\epsilon}_b(z(k)) - \gamma_b \cdot \mathrm{sgn}\big(z(k)\big) \tag{2.48a}$$

$$\gamma_b \triangleq \frac{\mathrm{E}\{\tilde{\epsilon}_b\big(a(k)\big)a(k)\}}{\mathrm{E}\{|a(k)|\}} \tag{2.48b}$$

where $\tilde{\epsilon}_b(\cdot)$ is an odd function whose second derivative is non-negative for $z \geq 0$, such that $\tilde{\epsilon}_b(-z) = -\tilde{\epsilon}_b(z)$, and $\epsilon_b''(z) = \tilde{\epsilon}_b''(z) \geq 0$ for all $z \geq 0$. This is sometimes known as the Generalized Sato Algorithm since it subsumes the special case of the Sato algorithm which is obtained when one defines $\tilde{\epsilon}_b(z) = z$.

Extension to the complex signalling, for example QAM signalling, has been proposed by Godard and Thirion [50] and Benveniste and Goursat [15]. It is commonly referred to as the Reduced Constellation Algorithm (RCA) because the algorithm, just like the original Sato algorithm, treats the real and imaginary parts of the QAM signals are "binary" signals. Thus it represents a reduced constellation

in the constellation space. The error function is thus

$$\epsilon_{\mathrm{RCA}}\big(z(k)\big) = z(k) - \gamma_{\mathrm{S}} \cdot \mathrm{csgn}\big(z(k)\big) \tag{2.49}$$

where $\mathrm{csgn}(\cdot)$ is the complex signum operator defined as $\mathrm{csgn}(z) = \mathrm{sgn}\big(\mathrm{Re}(z)\big) + \mathrm{sgn}\big(\mathrm{Im}(z)\big)$. Even though the RCA largely resembles the Sato algorithm conceptually, in another sense, it differs from the Sato algorithm because unlike the Sato algorithm, the RCA would alter the *phase* of the equalizer output in addition to adjusting its magnitude (see Fig. 2.10).

2.5.3 Godard algorithm

The arguably most important class of Bussgang algorithm is the Godard algorithm [49] which consists of a family of *constant modulus* blind equalization algorithms. It was also independently proposed by Treichler and Agee [137] under the name, Constant Modulus Algorithm (CMA). The Godard algorithm penalizes the deviation of the equalized signal from a constant modulus via the "property restoral" philosophy, a term coined by Agee in [3]. The deviations from this constant modulus is depicted by the dotted circles in the constellation space as shown in Fig. 2.11. It was initially developed for constellations that exhibit a constant modulus property such as the 4-QAM and PSK formats (see Fig. 2.11-A for the 8-PSK format). However, it exhibits exceptional robustness with respect a variety of other non-constant modulus data formats such as larger QAM formats and the V29-CCITT formats.

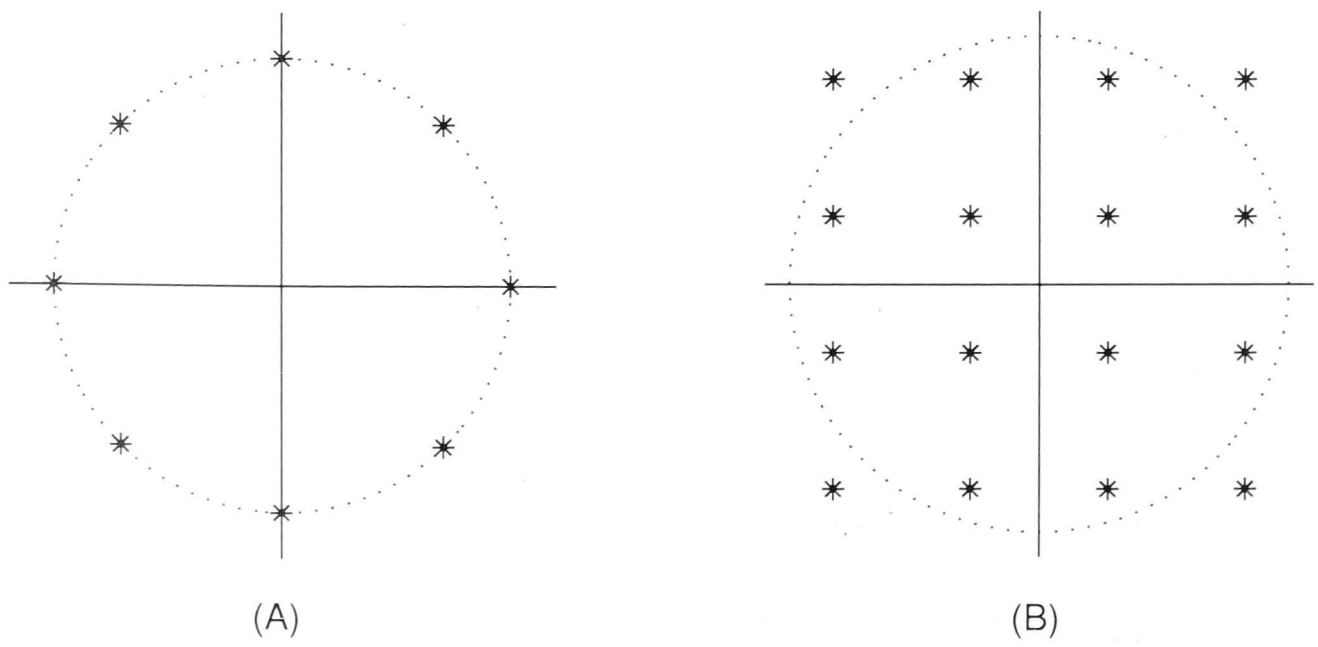

Figure 2.11: Godard algorithm or Constant Modulus Algorithm (CMA): It penalizes deviations from a constant modulus represented by the dotted circles for (A) 8-PSK which is a constant modulus source, and (B) 16-QAM which is a non-constant modulus source.

The Godard algorithm minimizes a non-convex cost function parameterized by a positive integer p given as

$$J_{\mathrm{CMA}} = \mathrm{E}\{(|z(k)|^p - \gamma_{\mathrm{C}}^p)^2\} \tag{2.50}$$

where γ_{C}^p is a positive real constant called the dispersion constant which is given by

$$\gamma_{\mathrm{C}}^p \triangleq \frac{\mathrm{E}\{|a(k)|^{2p}\}}{\mathrm{E}\{|a(k)|^p\}}. \tag{2.51}$$

When $p = 1$, the CMA is identical to the Sato algorithm for real systems. The cost functions and error functions of the Godard algorithm for $p = 1$ and $p = 2$ are respectively given by

$p = 1$:

$$J_{\mathrm{CMA}} = \mathrm{E}\left\{\left(|z(k)| - \gamma_{\mathrm{C}}\right)^2\right\} \tag{2.52a}$$

$$\epsilon_{\mathrm{CMA}} = z(k)\left(1 - \frac{\gamma_{\mathrm{C}}}{|z(k)|}\right) \tag{2.52b}$$

$p = 2$:

$$J_{\mathrm{CMA}}(k) = \mathrm{E}\left\{\left(|z(k)|^2 - \gamma_{\mathrm{C}}^2\right)^2\right\} \tag{2.53a}$$

$$\epsilon_{\mathrm{CMA}} = z(k)\left(|z(k)|^2 - \gamma_{\mathrm{C}}^2\right). \tag{2.53b}$$

The second special case where $p = 2$ is referred to in the literature as the constant modulus algorithm (CMA) and it is the most widely used algorithm in the family of Godard algorithms.

The Godard algorithm is considered the most successful and the most extensively researched Bussgang algorithm. It is supported by some comparative studies performed by Shynk *et al* [130] and Jablon [63]. It can be shown to exhibit the carrier phase blind property [81] that is particularly advantageous in the design of practical digital transmission systems. Therefore it is more robust than other Bussgang algorithms with respect to carrier phase offsets as the equalizer parameters can be adapted independently but simultaneously with the operation of the carrier recovery system, leading to the notion of a decoupled equalizer and phase estimator. Under steady state conditions, the Godard algorithm attains a mean-squared error (MSE) that is lower than other Bussgang algorithm [58].

As the cost function is non-convex, initialization of the equalizer tap parameters is crucial in order to avoid convergence to spurious local minima. Even though convergence to global minima cannot be guaranteed for all cases, practical guidelines do exist to ensure desirable convergence have been suggested in [49, 79]. In contrary, *bad* initialization strategies that should be avoided have been formulated in [35] where it is stated that the initialization with an edge tap set to a non-zero constant may very likely yield convergence to undesirable local minima that do not correspond to the desired ISI removal.

2.5.4 Shalvi and Weinstein algorithm

The methods of Shalvi-Weinstein [128] generalize CMA, but unlike conventional Bussgang algorithms that implicitly exploit higher order statistics of the equalizer output, the Shalvi-Weinstein algorithm (SWA) are explicitly based on the higher order statistics of the equalizer output. The SWA maximizes the absolute value of the kurtosis of the equalizer output $\left|K\big(z(k)\big)\right|$, i.e.,

$$\max \left|K\big(z(k)\big)\right| \tag{2.54a}$$

$$\text{subject to}$$

$$\mathrm{E}\{|z(k)|^2\} = \mathrm{E}\{|a(k)|^2\}, \tag{2.54b}$$

where the kurtosis of $z(k)$ is defined as

$$\left|K\big(z(k)\big)\right| \triangleq \mathrm{E}\big\{|z(k)^4|\big\} - 2\mathrm{E}^2\big\{|z(k)^2|\big\} - \left|\mathrm{E}\big\{z(k)^2\big\}\right|^2. \tag{2.55}$$

For convenience, we denote $K_a = \left|K\big(a(k)\big)\right|$ and $K_z = \left|K\big(z(k)\big)\right|$. Shalvi and Weinstein showed that when (2.54b) is true, then the following must hold [128]:

$$|K_z| \leq |K_a| \tag{2.56a}$$

$$|K_z| = |K_a| \quad \text{if and only if} \quad \mathbf{s} = e^{j\theta}(\cdots, 0, 0, \mathcal{C}, 0, 0, \cdots)^T \tag{2.56b}$$

where $\mathbf{s}$ is the vector of the combined channel-equalizer response, i.e.,

$$\mathbf{s}(k) = [s_1(k), s_2(k), \cdots]^T, \quad s_i(k) = \sum_l h_{i-l} w_l(k) \tag{2.57}$$

and h_i is the channel impulse response, $w_l(k)$ is the equalizer impulse response, and $\mathcal{C}$ is a non-zero constant. The result of (2.56b) is of particular interest as it corresponds to a zero-forcing solution. It is obtained as follows [128]:

Let the l_2-norm $\|s\|_2^2$ be absolutely summable, i.e., $\sum_l |s_l(k)|^2 < \infty$. Then,

$$\sum_l \left|s_l(k)\right|^4 \leq \left(\sum_l \left|s_l(k)\right|^2\right)^2 \tag{2.58}$$

where equality holds if and only if $\mathbf{s}$ has at most one non-zero component, which would correspond to the zero-forcing solution. This completes the proof. The necessary and sufficient conditions for equalization are therefore

$$\mathrm{E}\{|z(k)|^2\} = \mathrm{E}\{|a(k)|^2\} \quad \text{and} \quad \left|K_z\right| = \left|K_a\right|. \tag{2.59}$$

It can also be shown for typical digital modulation, that the condition $|K_z| = |K_a|$ may be replaced by the simpler $\mathrm{E}\{|z(k)|^4\} = \mathrm{E}\{|a(k)|^4\}$.

According to the criterion in (2.54), the cost function may be expressed as

$$J_{\mathrm{SW}} = \left|K_z\right| \tag{2.60}$$

$$= \mathrm{sgn}\,(K_z) \cdot K_z \tag{2.61}$$

$$= \mathrm{sgn}\,(K_z)\left\{\mathrm{E}\{|z(k)|^4\} - 2\big(\mathrm{E}\{|z(k)|^2\}\big)^2 - \big(\mathrm{E}\{z(k)^2\}\big)^2\right\} \tag{2.62}$$

where it can be shown that the second and third terms on the right hand side of (2.60) do not influence the maximum. The criterion thus becomes

$$J_{\mathrm{SW}} = \mathrm{sgn}\big(K_z\big)\mathrm{E}\{|z(k)|^4\}. \tag{2.63}$$

The update equation of the equalizer parameters is

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \mu\ \mathrm{sgn}(K_a) z(k)\mathbf{X}^*(k)|z(k)|^2 \tag{2.64}$$

where μ is the step size parameter.

2.5.5 Multi-Modulus Algorithm

The CMA for multiple moduli constellations is later extended and redeveloped to incorporate joint phase constellation recovery together with equalization [105, 152]. It is called the multi-modulus algorithm (MMA) by Yang, Werner and Darth [152]. The rotational behavior of this algorithm will be studied in Chapter 6 and is compared with the Reduced Constellation Algorithm (RCA) previously described

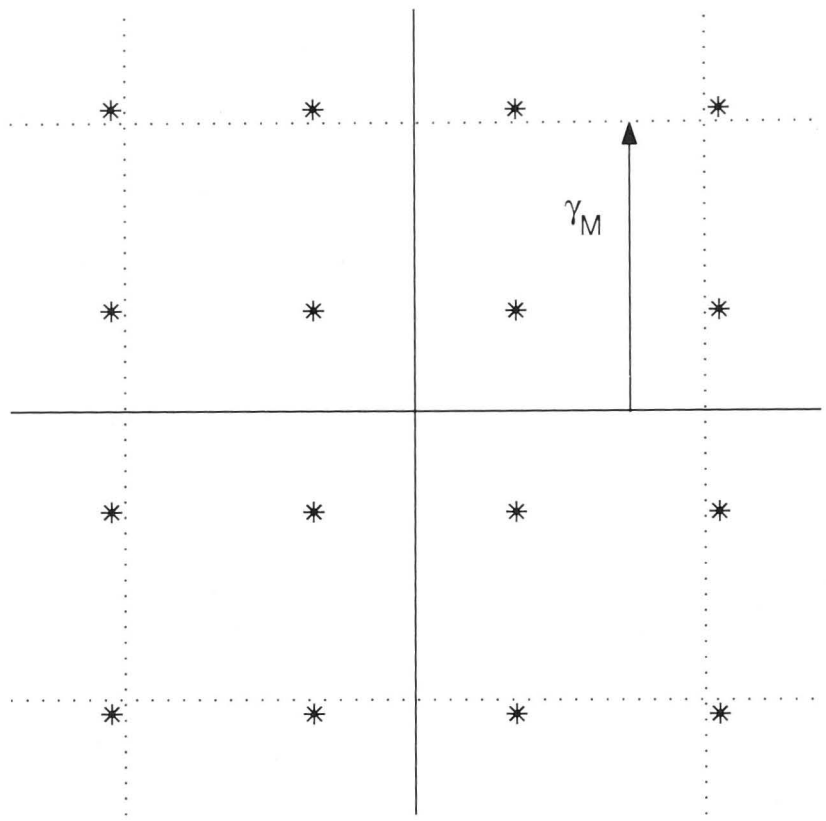

Figure 2.12: Multi Modulus Algorithm (MMA): It penalizes deviations from γ_M for real and complex components separately.

above. The algorithm minimizes the cost function

$$J_{\mathrm{MMA}} = \mathrm{E}\Big\{ \big[\mathrm{Re}(z(k))^2 - \gamma_{\mathrm{M}}^2\big]^2 + \big[\mathrm{Im}(z(k))^2 - \gamma_{\mathrm{M}}^2\big]^2 \Big\} \tag{2.65}$$

where

$$\gamma_{\mathrm{M}}^2 = \frac{\mathrm{E}\big[\mathrm{Re}(a(k))^4 + \mathrm{Im}(a(k))^4\big]}{\mathrm{E}\Big[\big|\mathrm{Re}(a(k))\big|^2 + \big|\mathrm{Im}(a(k))\big|^2\Big]} = \frac{\mathrm{E}\big[a(k)^4\big]}{\mathrm{E}\Big[\big|a(k)\big|^2\Big]} \tag{2.66}$$

for square QAM constellations. For cross QAM constellations, γ_{M}^2 should be assigned different values for different data points in the constellation, where details can be found in [152]. Its concept is to "push" the equalizer output signals towards the four boundaries that form a square of width $2\gamma_{\mathrm{M}}$ as depicted in Fig. 2.12. This cost function is somewhat similar to that of the CMA except that the in-phase and quadrature components are separated and the cross term between the components have been eliminated as we now show. Recall that the cost function of the CMA 2-2 algorithm that has been expanded is

$$J_{\mathrm{CMA}} = \mathrm{E}\Big\{ \big[|z(k)|^2 - \gamma_{\mathrm{C}}^2\big]^2 \Big\} \tag{2.67a}$$

$$= \mathrm{E}\Big\{ \big[\mathrm{Re}(z(k))^2 - \gamma_{\mathrm{C}}^2\big]^2 + \big[\mathrm{Im}(z(k))^2 - \gamma_{\mathrm{C}}^2\big]^2 - \big[2\mathrm{Re}\big(z(k)\big)\mathrm{Im}\big(z(k)\big) - \gamma_{\mathrm{C}}^2\big]^2 \Big\} \tag{2.67b}$$

where γ_C^2 is the dispersion constant of the CMA (2.51). Comparing (2.65) and (2.67b) reveals that the cross term has been dropped in the MMA cost function. The error functions of the MMA for the in-phase (i) and the quadrature (q) filters are therefore

$$\epsilon_{\text{MMA}}^{(i)} = \text{Re}\big(z(k)\big)\big[\text{Re}\big(z(k)\big)^2 - \gamma_M^2\big] \tag{2.68a}$$

$$\epsilon_{\text{MMA}}^{(q)} = \text{Im}\big(z(k)\big)\big[\text{Im}\big(z(k)\big)^2 - \gamma_M^2\big]. \tag{2.68b}$$

Because of the disjoint adaptation of the in-phase and quadrature phase filters, minimization of this cost, unlike the CMA, will lead to phase-aware solutions, i.e., with automatic constellation phase recovery (up to ambiguities related to symmetries in the constellation) simultaneously achieved with the desired ISI removal [81, 105, 152].

2.5.6 Bussgang algorithms for multiple-modulus constellation

The CMA was proposed with the intention of equalizing signals whose source constellations display a constant modulus. It was later shown to be robust with respect to the modulation format, but would on the other hand yield larger excess noise [42, 65]. Recognizing this fact, the CMA is extended for signal formats with multiple radii such as 64-QAM and 8-PAM in [8, 120, 126, 145]. Two algorithms were proposed by Sethares *et al* [126] where the first is also called the multi-modulus algorithm (same as [152]) and the second is the decision adjusted modulus algorithm (DAMA). To avoid confusion with the MMA we described earlier in Chapter 2.5.5, we will abbreviate the first algorithm of [126] as MMA2. The cost function of the MMA2 is

$$J_{\text{MMA2}} = \big[z(k)^2 - r_1^2\big]^2\big[z(k)^2 - r_2^2\big]^2 \cdots \big[z(k)^2 - r_m^2\big]^2 \tag{2.69}$$

where $r_1, \cdots, r_m$ are the m radii that circles the data points as illustrated in Fig. 2.13.

The second algorithm that was jointly proposed by Sethares *et al* with the MMA1 in [126] is called the decision adjusted modulus algorithm (DAMA). It is also independently proposed by Ready and Gooch under the name Radius Directed

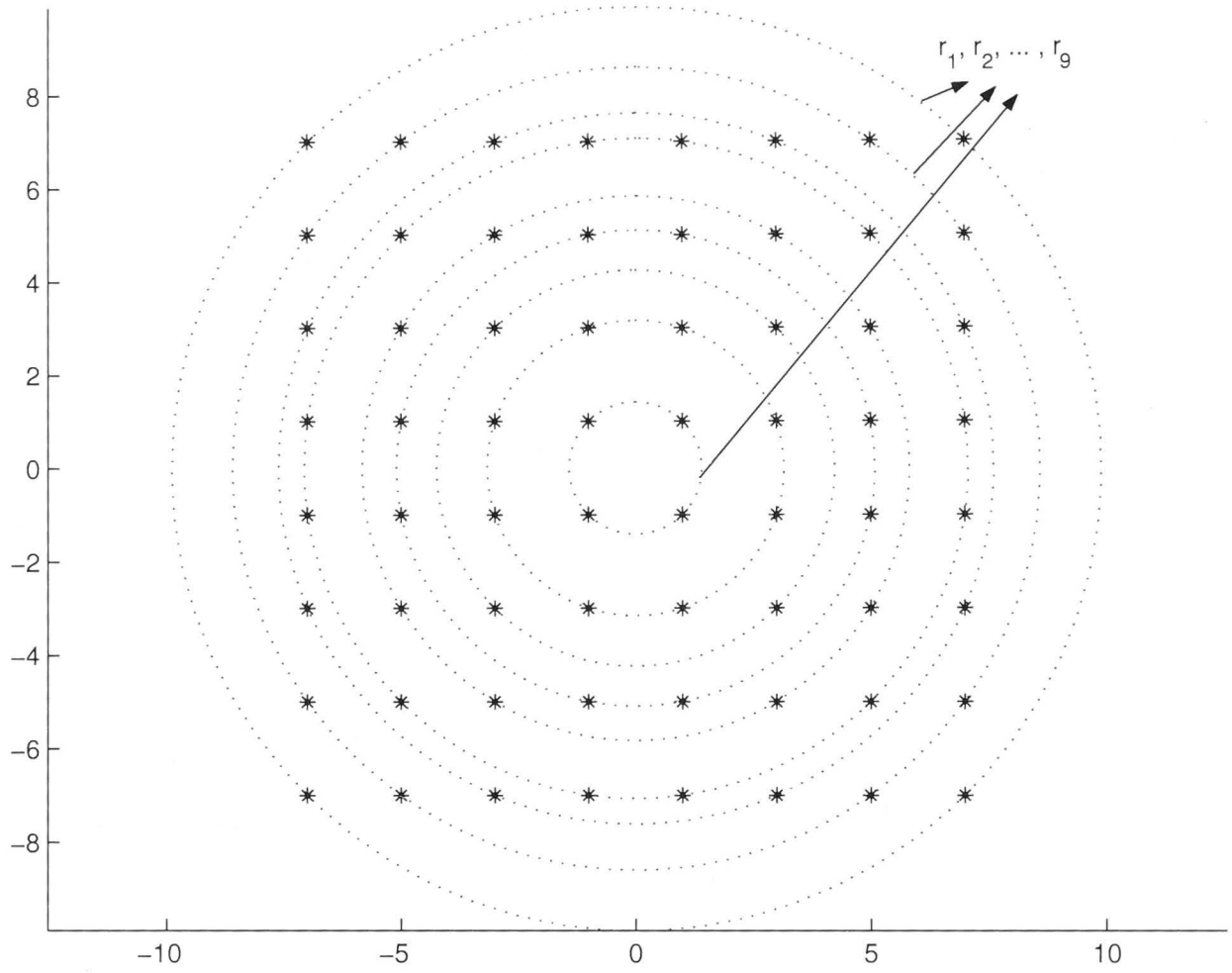

Figure 2.13: MMA2 and the DAMA algorithms : The 9 radii of the 64-QAM constellation.

Equalization (RDE) algorithm [120]. It minimizes the cost function

$$J_{\text{DAMA}} = \text{E}\Big\{ \min_i \big|z(k)^2 - r_i^2\big|^2 \Big\} \tag{2.70a}$$

$$= \text{E}\Big\{ \big|z(k)^2 - \text{Q}_R(z(k))^2\big|^2 \Big\}. \tag{2.70b}$$

DAMA is conceptually identical to a decision directed algorithm except for the decision device, $\text{Q}_R(\cdot)$ that they used. The decision device is a radius discriminator which is defined as

$$\text{Q}_R\big(z(k)\big) \triangleq \arg\min_i \big(\big||z(k)| - r_i\big|\big). \tag{2.71}$$

At convergence, the MMA2 and the DAMA yields lower MSE than the CMA due to the exploitation of the multi-modulus property of the constellation.

2.5.7 Decision directed algorithms

The decision directed algorithm [89] requires no training and is the simplest blind algorithm. Its cost function is

$$J_{\text{LMS}} = \text{E}\Big\{ \Big|z(k) - \text{Q}\big(z(k)\big)\Big|^2 \Big\}. \tag{2.72}$$

where in principle, it minimizes the mean square error between the equalizer output and the quantizer output. The performance of this algorithm can be indicated by the closed-eye-measure, CLEM of (2.7), which is an indication of how (un)reliable the equalizer output is. Global convergence to the optimum equalizer settings is guaranteed is CLEM < 1 [99]. In general, if the initial parameter values cause significant number of decision errors, then local convergence to undesirable minima is highly likely [92].

2.6 Principal Aims Of Thesis

In the remainder of the thesis we will develop several new switch-mode algorithms as well as switch-mode equalization schemes for blind equalization of linear channels. The motivation for the development of each new method is due to the deficiencies and disadvantages of existing algorithms and schemes in the following aspects:

- Their low convergence speed.
- Their high steady-state errors.
- Their abrupt switching which may result in the ill-convergence of the adaptive equalizer.
- Their high computational costs.
- Their poor estimation of open-eye condition that directly determines the switch from the acquisition to the tracking equalizer mode.

Chapter 3

Reliability Based Technique For Switch-Mode Blind Algorithms

Blind acquisition algorithms such as the Sato and the Godard algorithms can be used for the reliable acquisition of an unknown channel driven by a data sequence of known distribution [14, 49, 124]. Unfortunately, in general such algorithms yield non-zero error (noisy) solutions, typically described through the excess mean square error, even when the equalizer parameters have nominally reached their optimal settings, especially when dealing with non-constant modulus data formats [42, 65]. On the other hand, blind tracking algorithms such as the decision directed (DD) algorithm and the decision adjusted modulus algorithm (DAMA) [89, 126] can be used after the channel eye is open to yield a low excess mean square error (dominated only be the channel noise) when desired convergence is achieved, even for non-constant modulus data formats [58, 119]. Unfortunately these decision based algorithms strictly require a sufficiently low level of decision errors to ensure convergence [92, 99]. Since the acquisition and tracking algorithms exactly complement each other's deficiencies, it is common to employ the former during the initial acquisition and subsequently switch to the latter when the error rate is sufficiently low. This combination which pairs up the above blind algorithms is thus called a switch-mode algorithm. The Sato, Godard, LMS and DAMA algorithms have been previously introduced in Chapter 2.

Switching between modes can be done in one of two ways. Firstly, it can take place *instantaneously* as soon as a measure of the equalizer performance surpasses a pre-defined value that is chosen to correspond to low error rates. This is the so-called "hard-switching" technique [58, 82]. Otherwise, switching may occur *smoothly* in the sense that the transition between algorithms is soft and is usually automatic [15, 60, 116, 145]. Unfortunately, smooth switching techniques usually

slow down the convergence speed of the equalizer significantly, in addition to compromising low steady state errors. The following questions are therefore raised:

1) In determining the combination parameters of smooth switching approaches how well does the switching process accurately reflect the error rates?

2) In designing for a sufficiently smooth transition in a switching scheme, is the convergence speed too adversely affected and has the steady state error been compromised as well?

We propose a new smooth switching technique in this chapter that reflects the current error rate more accurately and less conservatively. It involves the computation of the probability of a correctly detected symbol at the equalizer output which has been derived using Bayes theorem. As a result, we observe fast convergence and low steady state errors without the need to manually control the parameters that govern the convergence speed and excess noise.

3.1 Motivation

In a trained digital communications system, the LMS algorithm can be employed with satisfactory convergence and steady state performance. With the availability of a reference signal, its cost function is the error between the equalizer output and the reference signal. Therefore minimizing the cost should lead to the desired recovery of the transmit signal. In blind transmissions, however, the cost function of the DD algorithm is the error between the equalizer output and its quantized value. When the output is unreliable, the quantized output will be erroneous with high probability, and may therefore result in the ill-convergence of the equalizer. A sufficient condition for the DD convergence is for the channel to have an initially open-eye condition [92, 99]. Therefore the goal of blind acquisition algorithms is to achieve an open-eye condition. The critical problem can be posed as:

> *"How can we measure the open eye condition, or at least estimate it as accurately as possible, given only the observed channel output with only knowledge of the statistics of the transmitted sequence?"*

The conventional technique is to perform switching instantaneously when a performance measure, say the estimated MSE of the equalizer output, drops below a pre-defined threshold value. Even though it is simple to implement, this technique requires the user to pre-specify the threshold value on the MSE to trigger the

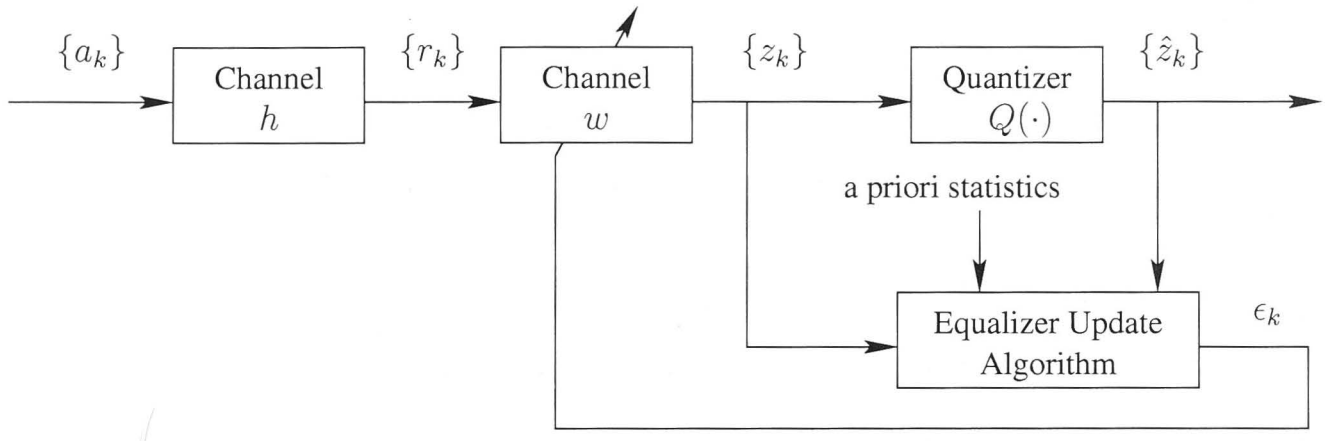

Figure 3.1: A typical baseband equivalent channel and a linear equalizer that employs the switch-mode algorithm.

switch-over. Strictly speaking this can only be done effectively if there is knowledge of the channel, which is not available. Switching too early when there are many errors may result in the ill-convergence of the DD algorithm; switching too late may result in a slow rate of convergence and possibly fail to acquire the channel when the channel is time-varying. There are other instances when the estimated MSE is not low enough even though the eye is clearly sufficiently "open", such as noisy channels and high excess MSE due to equalization of non-constant modulus data using the Sato or Godard algorithm.

A few smooth switching techniques have been suggested in the literature to overcome the lost of a reference (training) sequence. Some of them suggeste combination techniques for the acquisition and the tracking algorithms to be reflective of the current distribution of the equalizer output based on the equalizer output only. For example, if the equalizer output is close to one of the constellation data points, then the tracking algorithm is preferred over the acquisition algorithm because it is "believed" to be resulted from a well converged equalizer. On the contrary, when the equalizer output is far from the data points, the acquisition algorithm is preferred over the tracking algorithm. Thus, their concept is such that the error rate is implicit in this combination parameter which measures the distance of the equalizer output from its data points. The algorithms that employ this concept are those in [15,60,145]. Another concept is the so-called "stop-and-go" algorithm, whereby the algorithm is in its tracking mode at all times, but an update of the equalizer parameters is only performed when the error functions of both acquisition and tracking algorithms agree in sign [54,55,116]. However, the slow convergence of this technique has been reported [82].

3.2 System Model

In this chapter, we largely treat a communication system involving real signals given they usually have obvious extensions to the complex case. Consider the combined channel-equalizer system depicted in Fig. 3.1. Let $\mathbf{h} \triangleq [h_0, h_1, \cdots, h_L]^T$ denote the coefficients of the channel filter of length $L+1$. The channel is assumed stationary, possibly non-minimum phase, but unknown. Let the source data vector be $\mathbf{a}(k) \triangleq [a(k), a(k-1), \cdots, a(k-L)]^T$ at a time index k, with elements drawn from the alphabet set

$$\mathbb{A} \triangleq \{d_1, d_2, \cdots, d_M\} = \{\pm 1, \pm 3, \cdots, \pm(M-1)\} \tag{3.1}$$

representing M-ary PAM signaling. Then the input signal to the equalizer is

$$r(k) = \mathbf{h}^T \mathbf{a}(k) + n(k) \tag{3.2}$$

where $n(k)$ is the additive noise. Also let $\mathbf{w}(k) \triangleq [w_{-N,k}, \cdots, w_{0,k}, \cdots, w_{N,k}]^T$ be the $(2N+1)$ equalizer tap coefficients. We set

$$w_{p,0} = \begin{cases} 1 & p = 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

to have a center-tap initialization, which allows the causal development of approximate inverse filters for non-minimum phase systems. Then the equalizer output is given by

$$z(k) = \sum_{n=-N}^{N} w_n r(k-n) \tag{3.4a}$$

$$= s_0 a(k) + v(k) \tag{3.4b}$$

where

$$v(k) = \sum_{j \neq 0, j=-N}^{N+L} s_j a(k-j) + \sum_{n=-N}^{N} w_n \eta(k-n) \tag{3.5}$$

is the so-called *effective noise* with a variance of σ_v^2, and $\{s_j\}, j = -N, \cdots, N+L$ is the set of coefficients of the combined channel-equalizer filter. Without loss of generality, the coefficient s_0 is defined as unity. We further assume:

($\mathcal{H}1$) the M-PAM source symbols are identically and independently distributed

(i.i.d.);

($\mathcal{H}2$) $v(k)$ is white and Gaussian with zero mean and variance σ_v^2;

($\mathcal{H}3$) the impulse response of the combined channel-equalizer filter is absolutely summable, i.e., $\sum_j |s_j| < \infty$.

Justification for the second assumption is provided by the central limit theorem (see Section 3.5.2 for further justification).

3.2.1 Expressions for error function of switch-mode algorithms

Before we proceed any further, we provide first the different expressions to describe a switch-mode algorithm:

Blind adaptive equalization algorithms are often designed as stochastic gradient descent schemes to update the parameter vector by minimizing some cost functions that do not involve the use of the original input $a(k)$ but reflect the current level of ISI in the equalizer output. Define the mean cost function as

$$J(\mathbf{w}(k)) \triangleq \frac{1}{2}\mathrm{E}\{\epsilon^2(z(k))\} \tag{3.6}$$

where $\epsilon^2(\cdot) : \mathbb{R} \to \mathbb{R}$ is a scalar cost function and $z(k)$ is the equalizer output. Denote the weight vector of the equalizer as $\mathbf{w}(k)$ and $\mathbf{r}(k) = [r(k-N), \cdots, r(k), \cdots, r(k+N)]^T$ is the regressor vector of the equalizer consisting of samples from the channel output. The stochastic gradient descent minimization algorithm is well-known to be

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \mu \frac{\partial}{\partial \mathbf{w}(k)} \frac{1}{2}\epsilon^2(z(k)) \tag{3.7}$$

$$= \mathbf{w}(k) - \mu\epsilon(z(k))\mathbf{r}(k)^T. \tag{3.8}$$

Hence the blind algorithm can either be defined by the cost function or equivalently through $\epsilon(\cdot)$ which we call the *error function* since it replaces the prediction error in the LMS algorithm.

Switch-mode algorithms can be broadly categorized under three distinct groups. Slight modifications are possible as long as the modified form largely resembles the parent form of (3.9) – (3.11) as shown below. Let $\epsilon_{\text{acq}}(k)$ and $\epsilon_{\text{tr}}(k)$ be the error functions of the blind algorithms of the acquisition mode and the tracking mode,

respectively. Then the error function of a switch-mode algorithm can be expressed in one of the following three forms:

($\mathcal{F}1$) A weighted error function:

$$\begin{aligned}\epsilon(k) &= \beta_1(k)\epsilon_{\text{tr}}(k) + \beta_2(k)\epsilon_{\text{acq}}(k) \\ &= \beta(k)\left[\epsilon_{\text{tr}}(k), \epsilon_{\text{acq}}(k)\right]^T\end{aligned} \tag{3.9}$$

where $\beta_1(k)$ and $\beta_2(k)$ are positive functions and we define for convenience $\beta(k) \triangleq [\beta_1(k), \beta_2(k)]$. They can be functions of the equalized output $z(k)$ but generally are taken as positive constants. The desired properties of $\beta(k)$ are such that $\frac{\beta_1(k)}{\beta_2(k)} \ll 1$ when the equalizer is far from convergence, and $\frac{\beta_1(k)}{\beta_2(k)} \gg 1$ when the equalizer is close to convergence.

($\mathcal{F}2$) A conditional update type error function:

$$\epsilon(k) = \begin{cases} [1,0]\left[\epsilon_{\text{tr}}(k), \epsilon_{\text{acq}}(k)\right]^T = \epsilon_{\text{tr}}(k) & \text{Condition 1} \\ [0,1]\left[\epsilon_{\text{tr}}(k), \epsilon_{\text{acq}}(k)\right]^T = \epsilon_{\text{acq}}(k) & \text{Condition 2} \\ [0,0]\left[\epsilon_{\text{tr}}(k), \epsilon_{\text{acq}}(k)\right]^T = 0 & \text{Condition 3.} \end{cases} \tag{3.10}$$

Depending on the conditions, only one of the error functions will be used at any one time, or none at all when Condition 3 is satisfied. Condition 1 is usually a condition that indicates high probability of correct equalized signals, or an event corresponding to low noise levels. Conditions 2 and 3 hold when there is uncertainty in the quality of the equalizer output with condition 3 taking the most conservative approach (no update).

($\mathcal{F}3$) A Bussgang-type error function:

$$\epsilon(k) = z(k) - g(z(k)) = z(k) - g(\epsilon_{\text{acq}}(k), \epsilon_{\text{tr}}(k)) \tag{3.11}$$

where $g(\cdot)$ is a Bussgang memoryless nonlinear function. As mentioned in Chapter 2, Bellini [11] pointed out that most of the blind adaptive algorithms are Bussgang-type algorithms. They include the Sato, the Godard, and even the DD algorithms. The Bussgang-type algorithm resembles closely the LMS type algorithm since the error function of the LMS algorithm is $z(k) - \mathrm{Q}(z(k))$, where $\mathrm{Q}(\cdot)$ is the quantization operator and $\mathrm{Q}(z(k))$ is taken as the reference. The Bussgang error replaces $\mathrm{Q}(z(k))$ with $g(z(k))$ as is shown above. The combination of the respective error functions, given in (3.11), is performed within the nonlinear function $g(z(k))$.

Several approaches to the switching problem have been given in the references [8, 15, 53–55, 60, 75, 87, 116, 130, 145]. We will review four popular switch-mode algorithms in the following subsection because they present novel concepts that enable a soft transition between the acquisition and tracking algorithms.

3.3 Review of traditional switch-mode algorithms

3.3.1 Benveniste-Goursat (BG) algorithm

The very first known switch-mode algorithm in the literature to address the switching problem is proposed by Benveniste and Goursat [15]. They proposed to minimize an error function that combines the Sato and the DD algorithm in the following manner:

$$\epsilon_{\mathrm{BG}}(k) = k_1 \epsilon_{\mathrm{DD}}(k) + k_2 |\epsilon_{\mathrm{DD}}(k)| \epsilon_{\mathrm{Sato}}(k) \tag{3.12}$$

where k_1, k_2 are positive constants. It is expressed in the form of (3.9). In their algorithm, they used the magnitude of the LMS error function as the weight of $\epsilon_{\mathrm{Sato}}(k)$. When the equalizer is far from convergence, then the equalized signals are usually far away from the constellation points. Hence $|\epsilon_{\mathrm{DD}}(k)|$ is large and $\epsilon_{\mathrm{Sato}}(k)$ is largely used for the update. However, when the equalizer is close to convergence, $|\epsilon_{\mathrm{DD}}(k)|$ reduces. It is interesting to note that at perfect equalization, $\epsilon_{\mathrm{DD}}(k) = 0$, and so will $\epsilon_{\mathrm{BG}}(k)$. Therefore, the BG algorithm can achieve the zero cost condition given ideal convergence and no noise. Deshpande [23] reports still better performance if the Sato error is replaced by the CMA error.

3.3.2 Stop-And-Go (SAG) decision directed algorithm

The standard DD algorithm for joint MSE equalization and carrier recovery, which is normally utilized in the open-eye condition, can be turned into an algorithm providing effective blind convergence in the MSE sense and is usable in the closed-eye acquisition phase. This "stop-and-go" (SAG) algorithm [54, 55, 116] uses a binary-valued flag to indicate to both the equalizer and the synchronizer whether the equalizer output is sufficiently reliable. The flag shows "Go", i.e., an update event, when the signs of $\epsilon_{\mathrm{tr}}(k)$ and $\epsilon_{\mathrm{acq}}(k)$ agree. When they disagree, adaptation is stopped for that iteration.

Consider the original algorithm by Picchi and Prati. They employed the Sato algorithm and the DD LMS algorithm whose error functions are $\epsilon_{\mathrm{Sato}}(k)$ and $\epsilon_{\mathrm{DD}}(k)$,

respectively. Their error functions can be expressed in the forms of (3.9), (3.10) in a modified manner and (3.11), as shown:

($\mathcal{F}1$) Weighted error function

$$\epsilon_{\mathrm{SAG}}(k) = 0.5\epsilon_{\mathrm{DD}}(k) + 0.5|\epsilon_{\mathrm{DD}}(k)|\mathrm{sgn}(\epsilon_{\mathrm{Sato}}(k)). \tag{3.13}$$

($\mathcal{F}2$) Conditional update type error function

$$\epsilon_{\mathrm{SAG}}(k) = \begin{cases} \epsilon_{\mathrm{DD}}(k) & \mathrm{sgn}(\epsilon_{\mathrm{DD}}(k)) = \mathrm{sgn}(\epsilon_{\mathrm{Sato}}(k)) \\ 0 & \mathrm{sgn}(\epsilon_{\mathrm{DD}}(k)) \neq \mathrm{sgn}(\epsilon_{\mathrm{Sato}}(k)). \end{cases} \tag{3.14}$$

($\mathcal{F}3$) Bussgang type error function

$$\epsilon_{\mathrm{SAG}}(k) = z(k) - \{0.5(\mathrm{Q}(z(k)) + z(k)) - 0.5|\epsilon_{\mathrm{DD}}(k)|\mathrm{sgn}(\epsilon_{\mathrm{Sato}}(k))\}. \tag{3.15}$$

This is a demonstration that a switch-mode algorithm can be expressed in various forms if required.

3.3.3 Dual-Mode Generalized Sato Algorithm (DMGSA) and Dual-Mode Godard Algorithm (DMGA)

Weekackody and Kassam [130, 145] proposed two schemes that separately use the generalized Sato algorithm (GSA) and the Godard algorithm as the acquisition algorithms, and the DD LMS and the decision adjusted modulus algorithm (DAMA) [126] as their tracking algorithms, respectively. Scheme 1 is a combination of the GSA and the DD LMS algorithm, named dual-mode generalized Sato algorithm (DMGSA). Scheme 2 is a combination of the Godard algorithm and the DAMA, named dual-mode Godard algorithm (DMGA).

The conditions of switching are as follows. For both DMGSA and DMGA, condition 1 is when the equalizer output is within a predefined region $D(k)$. Condition 2 is simply its opposite, i.e., when the equalizer output is not inside $D(k)$. For DMGSA, $D(k)$ is a box around the constellation alphabet. For DMGA, $D(k)$ is an annular region around the moduli of the (non-CM) constellation. So, the error functions expressed in the form of (3.10) are

1. Dual-Mode Generalized Sato Algorithm (DMGSA)

$$\epsilon_{\text{DMGSA}}(k) = \begin{cases} [1,0]\,[\epsilon_{\text{DD}}(k), \epsilon_{\text{Sato}}(k)]^T & z(k) \in D_i \\ [0,1]\,[\epsilon_{\text{DD}}(k), \epsilon_{\text{Sato}}(k)]^T & z(k) \notin D_i, \forall i; \end{cases} \tag{3.16}$$

where

$$\begin{aligned} \epsilon_{\text{DD}}(k) =& (\text{Re}(z(k)) - \text{Q}(\text{Re}(z(k)))) + j(\text{Im}(z(k)) - \text{Q}(\text{Im}(z(k)))) \\ \epsilon_{\text{Sato}}(k) =& (\text{Re}(z(k)) - \gamma \text{Re}(z(k))) + j(\text{Im}(z(k)) - \gamma \text{Im}(z(k))) \end{aligned}$$

where γ is a constant [124].

2. Dual-Mode Godard Algorithm (DMGA)

$$\epsilon_{\text{DMGA}}(k) = \begin{cases} [1,0]\,[\epsilon_{\text{DAMA}}(k), \epsilon_{\text{CMA}}(k)]^T & z(k) \in D_i \\ [0,0]\,[\epsilon_{\text{DAMA}}(k), \epsilon_{\text{CMA}}(k)]^T & z(k) \notin D_i, \forall i; \end{cases} \tag{3.17}$$

where

$$\begin{aligned} \epsilon_{\text{DAMA}}(k) =& (|z(k)|^2 - R_i) z(k) \\ \epsilon_{\text{CMA}}(k) =& (|z(k)|^2 - R) z(k) \end{aligned}$$

where R is the dispersion constant [49] and $\{R_i\}$ is the set of the radii square of the source constellation. For further details of the DAMA algorithm and the DMGA, please refer to [126] and [130, 145] respectively.

3.3.4 Hilal-Duhamel (HD) algorithm

Hilal and Duhamel [60] proposed a dual-mode algorithm for blind equalization of PSK modulated signals. Their algorithm combines the CMA and the DD algorithm in the manner as shown in (3.11), i.e., a Bussgang-type algorithm. The combining method is using a memoryless proximity measure of the actual output $z(k)$ from each of the constellation points, so that the further $z(k)$ is from a constellation point, the smaller is this proximity measure. The concept is similar to the Benveniste-Goursat algorithm, where it can be shown that the closer $z(k)$ is to a constellation point, the more the DD algorithm will dominate. When it is far away, the algorithm will be dominated by the CMA.

The algorithm is as follows. Let $d_i, i = \{1, 2, \cdots, M\}$ be the M-PSK constellation points. Then the proposed Bussgang memoryless nonlinearity yields the

output

$$\epsilon_{\mathrm{HD}}(k) = z(k) - \exp\left\{ j.\arg\left(\sum_{i=0}^{M-1} \mathrm{Prox}(z(k) : d_i).\overrightarrow{Od_i} \right) \right\} \tag{3.18}$$

where the notations are defined as follows. The so-called proximity measure is defined as

$$\mathrm{Prox}(z(k) : d_i) = \frac{\breve{\alpha}}{\|\overrightarrow{z(k)d_i}\|} \equiv \frac{1/(\sum_{i=0}^{M-1} \frac{1}{\|\overrightarrow{z(k)d_i}\|})}{\|\overrightarrow{z(k)d_i}\|}. \tag{3.19}$$

where $\breve{\alpha}$ is a normalization constant that ensures that the sum of all proximities be equal to 1, and $\overrightarrow{Od_i} = z(k) + \overrightarrow{z(k)d_i}$. Then (3.18) becomes

$$\epsilon_{\mathrm{HD}}(k) = z(k) - \exp\left\{ j.\arg\left(z(k) + \sum_{i=0}^{M-1} \breve{\alpha} \frac{\overrightarrow{z(k)d_i}}{\|\overrightarrow{z(k)d_i}\|} \right) \right\}. \tag{3.20}$$

The behaviour of the proximity measure can be categorized under three scenarios:

a) $z(k) \approx 0$ **(Equalizer far from convergence):** The distances $\|\overrightarrow{z(k)d_n}\|$ are almost equal for all n due to the symmetry of the constellation, and $\sum_{i=1}^{M} \alpha \frac{\overrightarrow{z(k)d_i}}{\|\overrightarrow{z(k)d_i}\|}$ is almost a null vector, yields $g^{\mathrm{HD}}(z(k)) \approx \exp(j.\arg(z(k)))$.

b) $|z(k)| \gg \max_n(|d_n|)$ **(Equalizer far from convergence):** This corresponds to the case when $z(k)$ is very far from the constellation points. Then $\sum_{i=1}^{M} \alpha \frac{\overrightarrow{z(k)d_i}}{\|\overrightarrow{z(k)d_i}\|} = \overrightarrow{z(k)O} + \delta$ has the same argument as $z(k)$, δ is a small insignificant value.

c) $z(k) \approx a(k)$ **(Equalizer close to convergence):** It is shown [60] that $\sum_{i=1}^{M} \alpha \frac{\overrightarrow{z(k)d_i}}{\|\overrightarrow{z(k)d_i}\|} = \overrightarrow{z(k)d^*}$, where d^* is the constellation point of the transmitted symbol. Thus, $z(k) + \overrightarrow{z(k)d^*} = \overrightarrow{Od^*}$, which is the decision directed algorithm.

3.3.5 Diagrams of error functions

The diagrams of the error functions of some of the above mentioned switch-mode algorithms are shown in Fig. 3.2. The shapes of the error functions are reflective of a combination of an acquisition and a tracking algorithm which is proportioned according to a parameter that is a function of the equalizer output. We observe the following properties in the above figures:

1. The shape of the error function is determined only by the position of the equalizer output regardless of how reliable these data are. This is the main

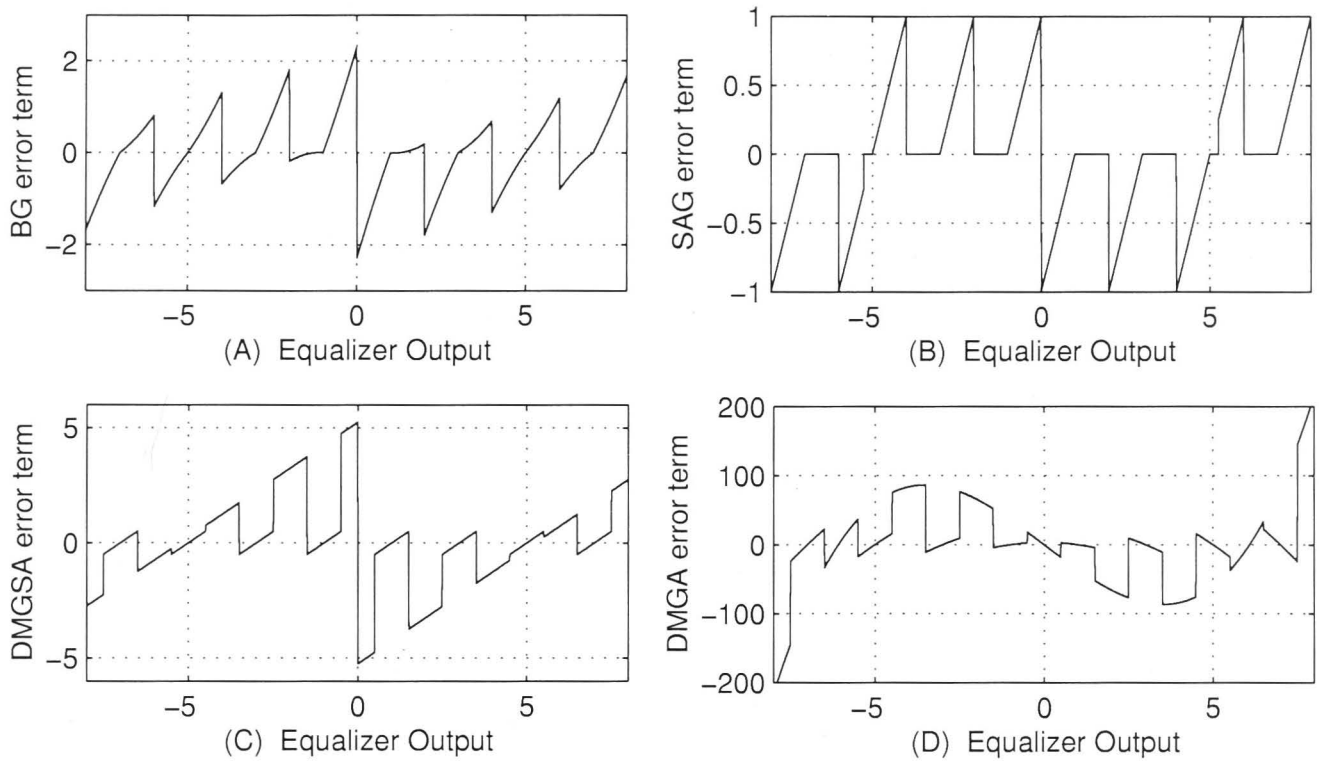

Figure 3.2: The error functions of the switch-mode algorithms outlined in the above subsections for 8-PAM signalling. Their shapes vary with the equalizer output $z(k)$ only. (A) Benveniste-Goursat (BG) algorithm [15], with parameters $k_1 = 1, k_2 = 0.25$. (B) Stop-And-Go (SAG) algorithm [116]. (C) Dual-Mode Generalized Sato Algorithm (DMGSA) [145], with the decision region $D(k) = d_i \pm 0.5, \ \forall i$, where $\{d_i\}$ is the set of the source constellation. (D) Dual-Mode Godard Algorithm (DMGA) [145] with the decision region $D(k) = d_i \pm 0.5, \forall i$.

"deficiency" in conventional algorithms that our switch-mode algorithm overcomes.

2. Whenever the equalizer output is close to a data symbol, the shape of the error function resembles that of the tracking algorithm more. Otherwise when the equalizer output is close to the bisector of the data symbols, its shape will resemble that of the acquisition algorithm.

3. All the above switch-mode algorithms yield zero error when the ideal convergence to optimal equalizer parameters is achieved. They exploit the discrete nature of the input data as would decision directed algorithms.

4. The well-known slow convergence of the "Stop-And-Go" (SAG) algorithm can be explained (see Fig. 3.2-(B)) by observing that almost half of the values of the equalizer output corresponds to zero update conditions. This happens at the values where the signs of both acquisition and tracking error functions do not agree.

3.4 A Novel Reliability Based Switch-Mode Algorithm

Let $\epsilon_{\text{acq}}(k)$ and $\epsilon_{\text{tr}}(k)$ be the error functions of the blind algorithms of the acquisition mode and the tracking mode, respectively. We propose a switch-mode algorithm with an error function

$$\epsilon(k) = \alpha(k)\gamma\epsilon_{\text{tr}}(k) + (1 - \alpha(k))\epsilon_{\text{acq}}(k) \tag{3.21}$$

which represents a convex combination between $\gamma\epsilon_{\text{tr}}(k)$ and $\epsilon_{\text{acq}}(k)$ with convex parameter $\alpha(k)$, and γ is chosen to compensate for the differences in the variances of the respective error functions where it is sensible to assign $\gamma \triangleq \mathrm{E}\{|\epsilon_{\text{acq}}(k)|\}/\mathrm{E}\{|\epsilon_{\text{tr}}(k)|\}$[1]. The formulation in (3.21) is according to (3.9). The principle design issue is how to determine $\alpha(k)$ as a function of signals available at the receiver. In what follows we identify $\alpha(k)$ with a measure of reliability[2].

Let P_C be the probability of the output of the quantizer being correct given the output of the equalizer $z(k)$. Then our proposed reliability measure can be compactly expressed as

$$\alpha(k) = \begin{cases} 1 & \sigma_v^2 < \sigma_{\text{thr}}^2 \\ 2P_C - 1 & 0.5 \leq P_C < 1 \text{ and } \sigma_v^2 \geq \sigma_{\text{thr}}^2 \\ 0 & \text{otherwise} \end{cases} \tag{3.22}$$

which we now explain. Given that $\alpha(k)$ is a convex parameter we can relate the extreme values of $\alpha(k)$ to P_C in the following way, where at one extreme, $\alpha(k) = 0$ when the probability of the equalizer detecting an incorrect symbol exceeds that of the correct symbol, i.e., $P_C \leq 1 - P_C$ or $P_C \leq 0.5$. At the other extreme, $\alpha(k) = 1$ when the probability of the symbol is high enough to allow the DD algorithm to converge to its global minima. As empirical results often show that a bit-error-rate of less than 10% is usually sufficient reliable, we will set $\alpha(k) = 1$ whenever $\sigma_v^2 < \sigma_{\text{thr}}^2$ where σ_{thr}^2 is a suitable threshold. Subsequently we need to decide on the relationship between $\alpha(k)$ and P_C whenever $0.5 \leq P_C < 1$ and $\sigma_v^2 \geq \sigma_{\text{thr}}^2$, which is the crucial portion that governs the need for smooth switching. We propose a

[1] γ is meant to be computed infrequently say at the end of each frame. When we consider the CMA and the LMS algorithm for example, we find that γ is approximately constant for a particular constellation size during the transitional period between acquisition and steady state, i.e., at the vicinity of an open eye condition. For the 4-PAM source, for example, it can be shown empirically that $\gamma \approx 14$.

[2] The definition of reliability is *the ability of a system or component to perform its required functions under stated conditions for a specified period of time* [62].

simple linear relationship between $\alpha(k)$ and P_C which is given by $\alpha(k) = 2P_C - 1$. Other mappings are possible as long as $\alpha(k)$ is a monotonically increasing function of P_C. This concludes our explanation of the assignment of $\alpha(k)$.

The probability P_C can be interpreted as the posterior probability of the event of a correctly detected symbol given certain measurements. Such an event can be mathematically expressed as $\{A_{j^*} : \mathrm{Q}(z(k)) = d_{j^*} = a(k)\}$, where $\mathrm{Q}(\cdot)$ is the quantization operator and j^* is the index of the alphabet symbol that corresponds to a correct quantizer decision. All other incorrect events are $\{A_j : \mathrm{Q}(z(k)) = d_j, \forall j \neq j^*\}$. Thus, we define P_C as the posterior conditional probability

$$P_C \triangleq P(A_{j^*}|z(k)). \tag{3.23}$$

3.4.1 Computation of Reliability Measure $\alpha(k)$

To compute $\alpha(k)$ we need to compute P_C and σ_v^2. The former can be calculated by applying Bayes theorem[3] and the law of total probability:

$$P_C = P(A_{j^*}|z(k)) = \frac{p(z(k)|A_{j^*})P(A_{j^*})}{p(z(k))} \tag{3.25}$$

$$= \frac{p(z(k)|A_{j^*})P(A_{j^*})}{p(z(k)|A_{j^*})P(A_{j^*}) + \sum_{j \neq j^*} p(z(k)|A_j)P(A_j)}, \tag{3.26}$$

[3]The Bayes Theorem expresses the posterior probability as

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{3.24}$$

where $P(B|A)$ is known as the likelihood function and $P(A)$, $P(B)$ are known as the priors. The Law of Total Probability defines

$$P(B) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A})$$

where $\bar{A}$ is the complement event of A. The likelihood function can be obtained if we have prior knowledge of the distribution of the continuous random variable B given event A has occurred:

$$P(B|A) = \frac{p_A(B)}{\int_{B \in \mathbb{B}} p_A(B) dB}.$$

Note that the upper case $P(\cdot)$ is used to denote the probability function, whereas the lower case $p(\cdot)$ is used to denote the probability density function. If $p_A(B)$ is *normalized*, i.e. $\int_{B \in \mathbb{B}} p_A(B) dB = 1$ then the conditional posterior probability becomes $P(B|A) = p_A(B)$. Similarly, we have $P(B|\bar{A}) = p_{\bar{A}}(B)$. So (3.24) becomes

$$P(A|B) = \frac{p_A(B)P(A)}{p_A(B)P(A) + p_{\bar{A}}(B)P(\bar{A})}.$$

where $p(\cdot)$ denotes the probability density function (pdf). From assumption $(\mathcal{H}1)$, we get $P(A_j) = 1/M, \forall j$. From assumption $(\mathcal{H}2)$, $p(z(k)|A_j), \forall j$ is the pdf of a normalized Gaussian distribution, i.e.,

$$p(z(k)|A_j) = \frac{1}{\sigma_v\sqrt{2\pi}}\exp\left(\frac{-(z(k)-d_j)^2}{(2\sigma_v^2)}\right). \tag{3.27}$$

Therefore, due to assumptions $(\mathcal{H}1)$ and $(\mathcal{H}2)$, P_C of (3.25) becomes a function of both $z(k)$ and σ_v^2 so that

$$P_C = \frac{\exp\left(\frac{-(z(k)-\mathrm{Q}(z(k)))^2}{2\sigma_v^2}\right)}{\sum_{j=1}^{M}\exp\left(\frac{-(z(k)-d_j)^2}{2\sigma_v^2}\right)}. \tag{3.28}$$

Then $\alpha(k)$ of (3.22) becomes

$$\alpha(k) = \begin{cases} 1 & \sigma_v^2 < \sigma_{\text{thr}}^2 \\ \dfrac{2\exp\left(\frac{-(z(k)-\mathrm{Q}(z(k)))^2}{2\sigma_v^2}\right)}{\sum_{j=1}^{M}\exp\left(\frac{-(z(k)-d_j)^2}{2\sigma_v^2}\right)} - 1 & 0.5 \le P_C < 1 \ \text{ and } \ \sigma_v^2 \ge \sigma_{\text{thr}}^2 \\ 0 & \text{otherwise.} \end{cases} \tag{3.29}$$

The contour plot of $\alpha(k)$ w.r.t. changes in $z(k)$ and σ_v can be illustrated as in Fig. 3.3. Here σ_{thr} is chosen to be 0.2.

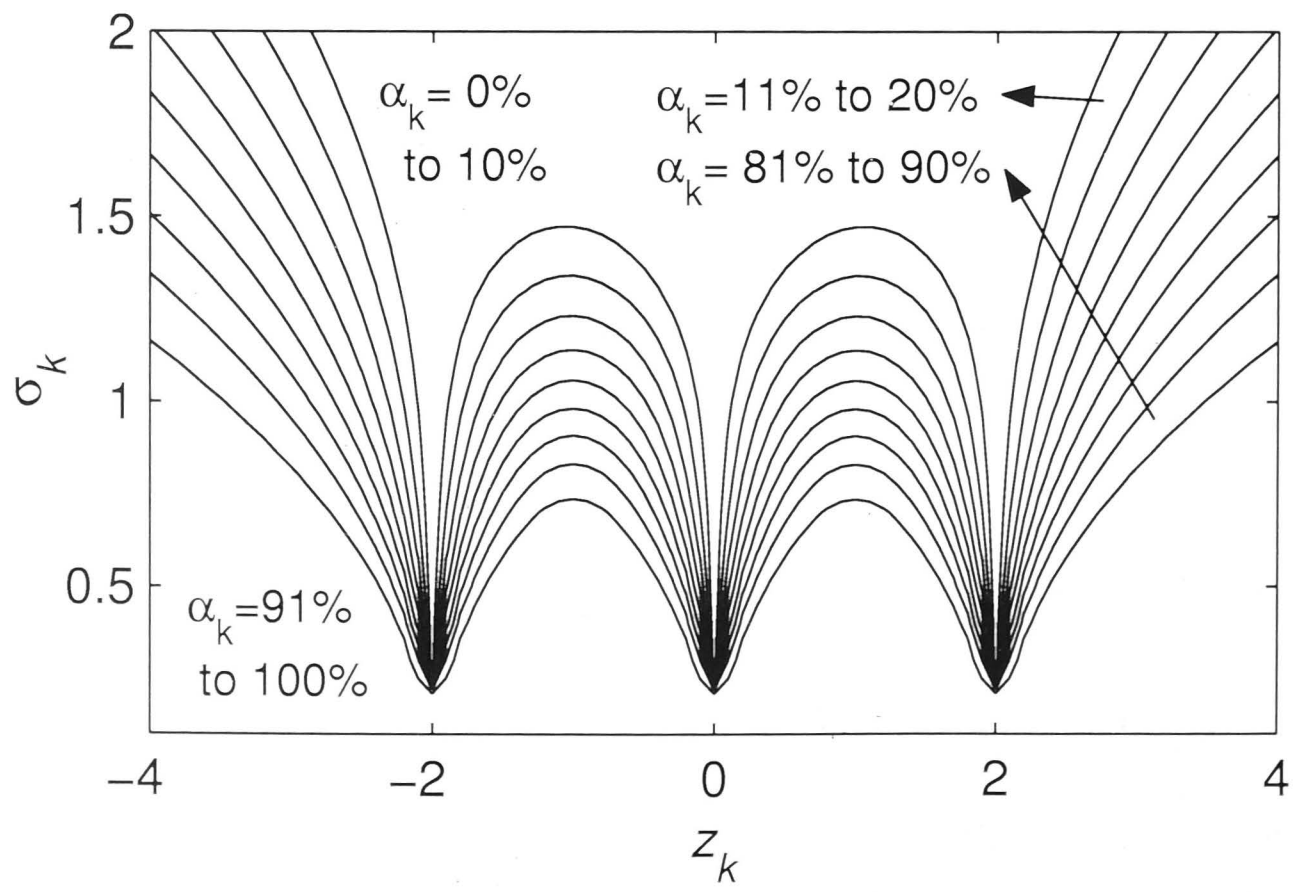

Figure 3.3: Contour plots of the reliability measure of (3.29) for a 4-PAM data.

To calculate the reliability measure $\alpha(k)$ in (3.27) and P_C in (3.36), it is neces-

sary to estimate the variance of the effective noise σ_v^2. However, estimation of this noise level is not a simple task because of the unknown channel. In our situation, we will rely on the effective noise being Gaussian due to the central limit theorem, i.e., the assumption of $(\mathcal{H}2)$. Recall that the conditions of the central limit theorem are such that there should be *many independent* terms in the impulse response of the effective noise which is the combined channel-equalizer response excluding the cursor s_0. In other words, the terms in $\{s_j\}, \forall j \neq 0$ should be many and independent of one another. When the channel eye is almost open, the impulse response $\{s_j\}, \forall j \neq 0$ should contain many small terms whose cross-correlation is small [57, ch. 2] if the equalizer has successfully minimized the cost function. It is actually appropriate to restrict the region of our consideration to the region when the channel eye is almost open as this is when switching usually occurs and the calculation of $\alpha(k)$ is required.

There are two known methods to obtain σ_v^2, namely the decision directed MSE and the signal-to-noise ratio (SNR) moments estimator approaches.

SNR Moments Estimator

The variance of the effective noise can be obtained by solving simultaneously the equations of the second (M_2) and fourth (M_4) order moments of $z(k)$ [115]

$$M_2 \triangleq S + N \tag{3.30}$$

$$M_4 \triangleq k_a S^2 + 6SN + k_v N^2 \tag{3.31}$$

where S, N are the power scaling factors of the unit variance signal and noise respectively, and $k_a \triangleq \mathrm{E}\{|a(k)|^4\}/\mathrm{E}\{|a(k)|^2\}^2$ and $k_v \triangleq \mathrm{E}\{|v(k)|^4\}/\mathrm{E}\{|v(k)|^2\}^2$ are the kurtoses of the signal and the noise, respectively. Let the so-called excess kurtoses of $a(k)$ and $v(k)$ be $G_a \triangleq k_a - 3$ and $G_v \triangleq k_v - 3$. Then solving for S and N in terms of M_2 and M_4, we get

$$S = \frac{M_2 G_v \pm \sqrt{(M_2 G_v)^2 - (G_a + G_v)(M_2^2(G_v + 3) - M_4)}}{G_a + G_v} \tag{3.32}$$

$$N = M_2 - S. \tag{3.33}$$

Further we assume the effective noise is Gaussian so that $G_v = 0$. Thus the last step before $\sigma_v^2 \triangleq N$ is obtained is to solve for S and N of (3.32) and (3.35) simultaneously, yielding

$$\sigma_v^2 \triangleq N = M_2 \pm \sqrt{G_a^{-1}(M_4 - 3M_2^2)}. \tag{3.34}$$

The second and fourth order moments M_2, M_4, can be estimated recursively in time k by

$$\widehat{M_2}(k+1) = \rho\widehat{M_2}(k) + (1-\rho)|z(k)|^2 \tag{3.35}$$

$$\widehat{M_4}(k+1) = \rho\widehat{M_4}(k) + (1-\rho)|z(k)|^4 \tag{3.36}$$

where ρ is a forgetting factor that is close to 1. The initial values of $\widehat{M_2}(0)$ and $\widehat{M_4}(0)$ are set to zero. Equation (3.36) needs to be computed only once in every frame[4]. The estimated variance is then substituted into P_C of (3.28) to acquire the reliability measure $\alpha(k)$.

Decision Directed MSE Method

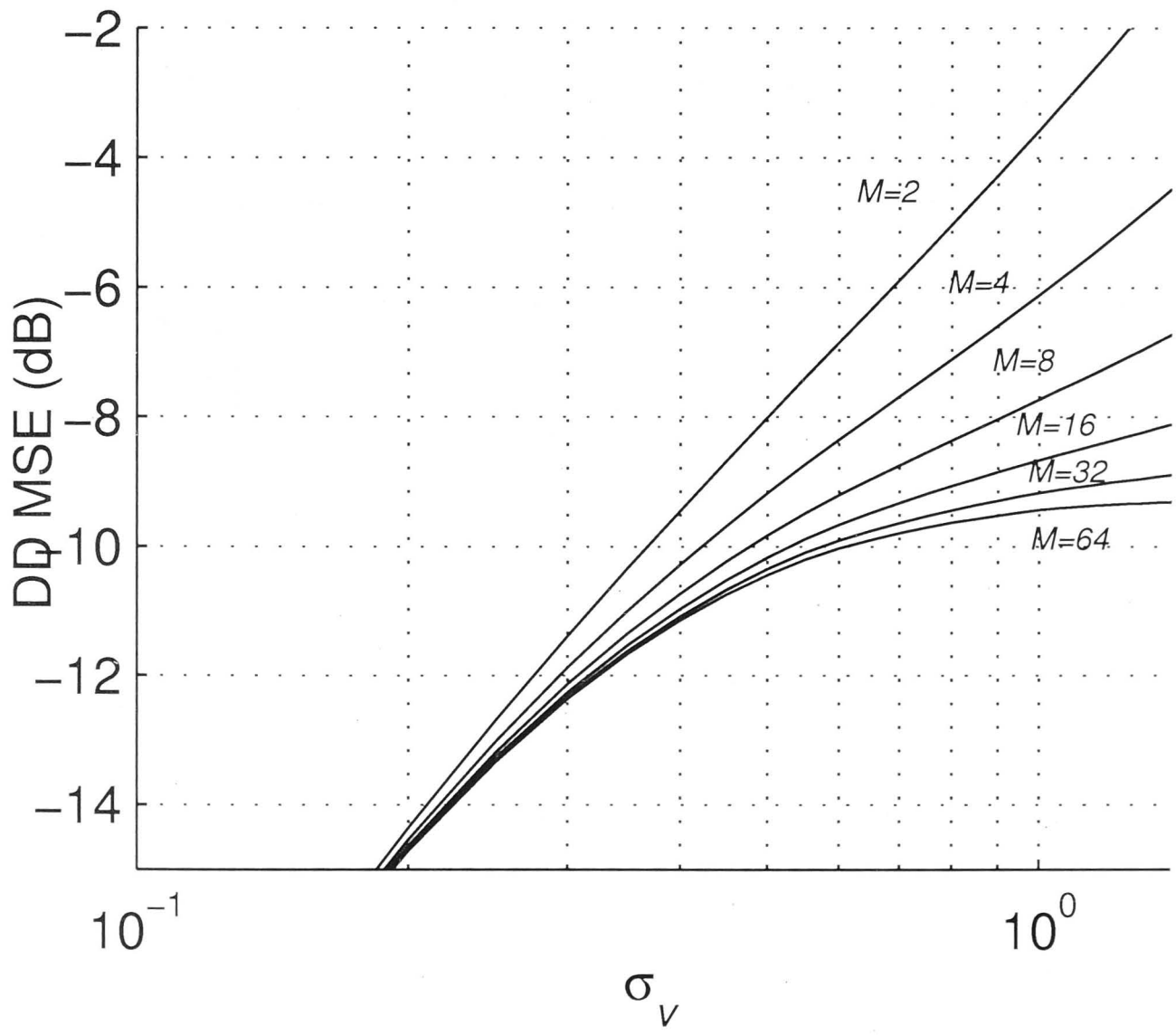

Figure 3.4: Relationship between the DD MSE and σ_v for M-ary PAM with $M = 2, 4, \cdots, 64$. The range of values of σ_v is from 0.1 to 1.4 presented in the log scale.

Once the pdf of $v(k)$ is assumed Gaussian, there is a straight forward relationship between σ_v^2 and the DD MSE. This method involves a lookup table to be tabulated which can be computed offline. The value of σ_v is varied from say 0.4 to

[4]In our simulations using 4-PAM data, one frame consists of 100 symbols.

Table 3.1: Lookup table for $\mathrm{MSE}_{\mathrm{DD}}(\sigma_v)$ and variance of effective noise, σ_v^2 for M-PAM signals

$\sigma_{\mathbf{v}}$	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$\sigma_{\mathbf{v}}^{\mathbf{2}}$	0.16	0.25	0.36	0.49	0.64	0.81	1.0
M-PAM	$\mathrm{MSE}_{\mathrm{DD}}(\sigma_v)$						
2-	0.157	0.232	0.311	0.392	0.476	0.565	0.663
4-	0.155	0.224	0.287	0.344	0.396	0.444	0.495
8-	0.154	0.219	0.275	0.32	0.354	0.383	0.411
16-	0.154	0.217	0.269	0.308	0.334	0.353	0.369
32-	0.154	0.216	0.266	0.302	0.324	0.338	0.348
64-	0.154	0.216	0.265	0.3	0.319	0.33	0.337

1.0 at an interval of 0.1 and a large sample of $z(k)$ is generated for each σ_v. Let the DD MSE of these sample points be $\mathrm{MSE}_{\mathrm{DD}}(\sigma_v)$, where it is a function of σ_v. Then the DD MSE can be computed according to

$$\mathrm{MSE}_{\mathrm{DD}}(\sigma_v) = \frac{\sum_{i=1}^{L} |z(k) - \mathrm{Q}(z(k))|^2}{L} \tag{3.37}$$

where L is the sample size used. The lookup table for M-PAM, $M = 2, 4, 8, 16, 32, 64$ for a range of σ_v values is tabulated in Table 3.1.

One advantage of the DD MSE approach over the SNR moments estimator, apart from being less computationally intensive, is that it is a direct and unbiased estimator. In contrast the SNR moments estimator may suffer a large variance given higher order moments need to be estimated. On the other hand, the DD MSE approach is less robust for greater constellation sizes. For M-PAM, it may not be reliably used for constellation sizes larger than 16 as the relationship becomes increasingly nonlinear (see Fig. 3.4).

3.5 Discussions on convergence and approximation

3.5.1 Convergence

The error function of (3.21) is a function of both $z(k)$ and σ_v^2. We plotted the error function that combines the CMA and the DD LMS algorithm. Figure 3.5 depicts the error function for an 4-PAM source for the range of $z(k)$ values between -4 and 4, and for the range of σ_v values between 0 and 2. At high noise levels,

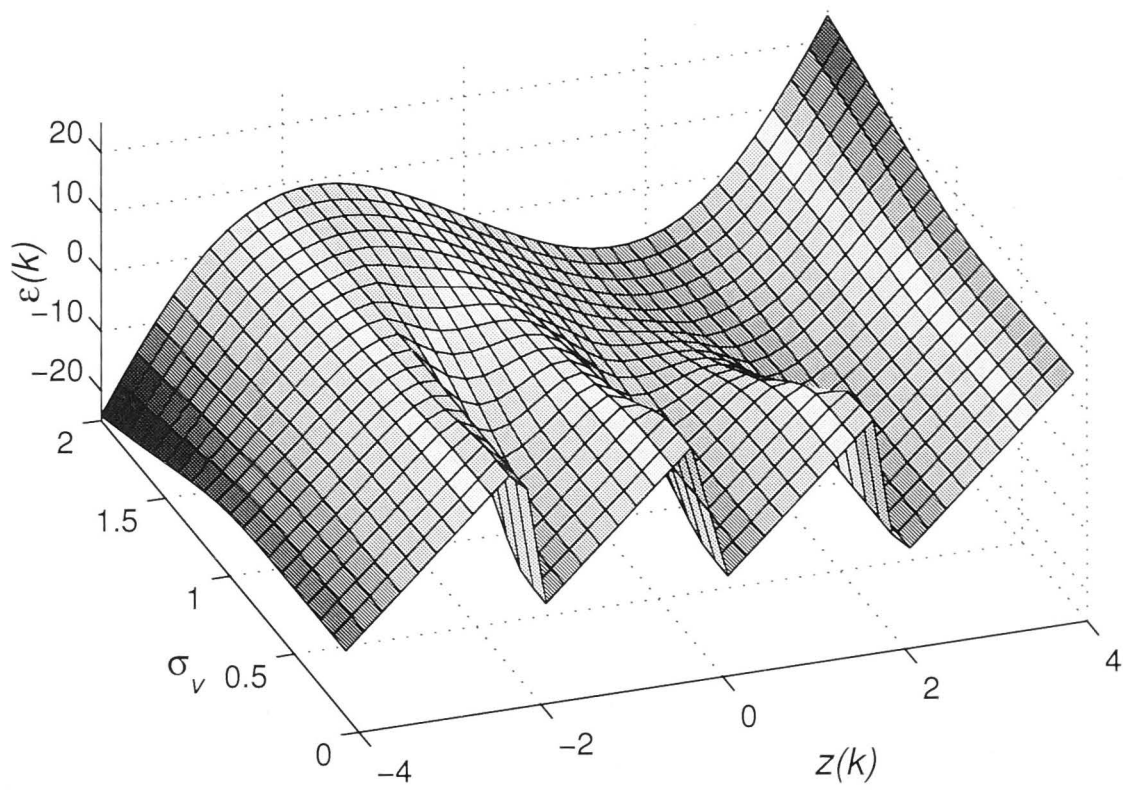

Figure 3.5: The error function of the novel switch-mode algorithm of Equation (3.21). Here $\gamma = 14$.

say when $\sigma_v \approx 2$, the error function is clearly dominated by $\epsilon_{\text{CMA}}(k)$ (recall that $\epsilon_{\text{CMA}}(k)$ has the shape of a cubic, a third order polynomial). At low noise levels, say $\sigma_v \approx 0$, the error function is clearly dominated by $\epsilon_{\text{tr}}(k)$. The transition between $\epsilon_{\text{CMA}}(k)$ and $\epsilon_{\text{tr}}(k)$ which happens when σ_v reduces from 2 to 0 may be observed as smooth. Thus the convergence of the switch-mode algorithm is such that it follows the characteristics of the CMA during the initial acquisition phase and then the characteristics of the DD algorithm at steady state.

A look at the cross section of the new $\epsilon(k)$

Consider further a switch-mode algorithm that combines $\epsilon_{\text{CMA}}(k)$ and $\epsilon_{\text{tr}}(k)$ for an 8-PAM source. If we were to take the cross sections of the error function at $\sigma_v = 1.2$ and 0.7, then the resulting error function evaluated at $\sigma_v = 1.2$ is predominantly $\epsilon_{\text{CMA}}(k)$ and $\epsilon_{\text{tr}}(k)$ for $\sigma_v = 0.7$. See Fig. 3.6. This switch-mode algorithm is not only robust at high noise levels but the error also goes to zero at perfect (and noiseless) equalization. The 3-D characteristic of this error function can be contrasted with the 2-D error functions in Fig. 3.2, typified by Benveniste-Goursat and other researchers.

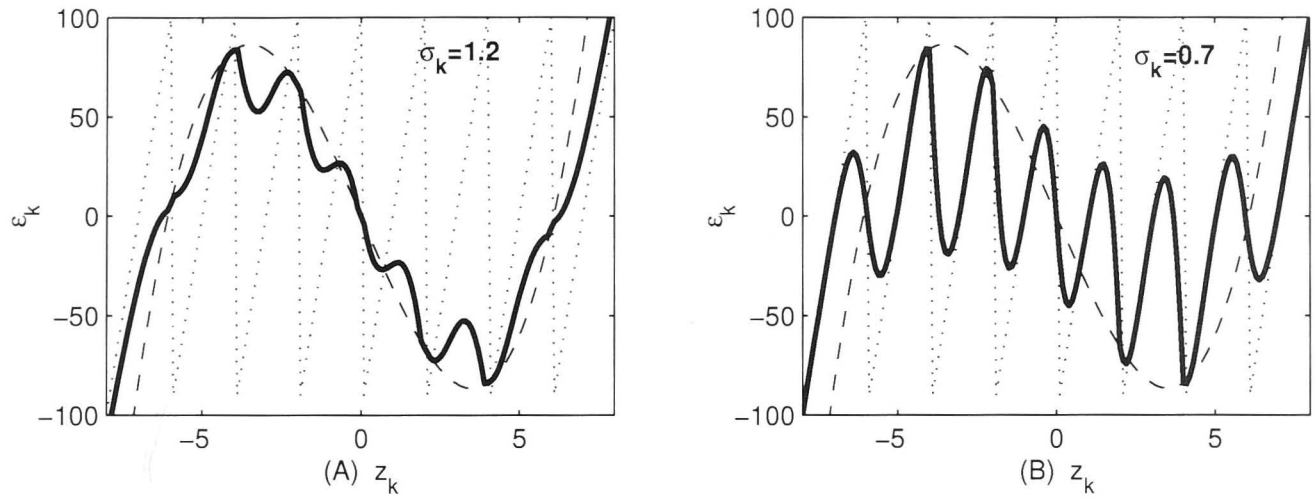

Figure 3.6: Cross sections of the 3-D graph of the error function of the new dual-mode algorithm for 8-PAM signalling and $\gamma = 100$. (A) $\sigma_v = 1.2$ and (B) $\sigma_v = 0.7$. The error functions of the CMA and the LMS algorithms are also plotted as dashed and dotted lines respectively.

3.5.2 Gaussian assumption of effective noise

This subsection is intended for discussing the Gaussian assumption of the effective noise which we have used to estimate the reliability measure. While the assumption that the assumption that the effective noise, $v(k)$, is Gaussian is generally untrue for most dispersive channels and hence it is difficult to derive a universal expression for the distribution of $v(k)$ due to the unknown channel, the central limit theorem may simplify matters significantly. Recall that the conditions of the CLT are that there are *many*, *independent* random variables where none of them is especially dominant. From (3.5) that $v(k)$ is the sum of the sequence $s_j a(k-j)$, $-N \leq j \leq N+L$, $j \neq 0$. Therefore for CLT to be valid, $\{s_j\}$, $-N \leq j \leq N+L$, $j \neq 0$ should be independent of each other and the order of the impulse response, $2N+L$, should be sufficiently large. In this discussion we will cover two scenarios where the first validates the CLT and the second considers the case when $2N+L$ is very large.

1) If $2N+L$ is sufficiently large (usually at least 15 will yield well approximated Gaussian distribution), then the correlation between $v(k)$ and $a(k)$ is small. The normalized correlation coefficients of $v(k)$ and $a(k)$ are [57, ch. 2]

$$\frac{\mathrm{E}\{v(k)v_j\}}{\sqrt{\mathrm{E}\{v(k)^2\}\mathrm{E}\{v_j^2\}}} = \frac{\sum_l s_l s_{l+j-k}}{\sum_l s_l^2} \tag{3.38}$$

and

$$\frac{\mathrm{E}\{a(k)v_j\}}{\sqrt{\mathrm{E}\{a(k)^2\}\mathrm{E}\{v_j^2\}}} = \frac{s_{j-k}}{\sqrt{\sum_l s_l^2}}. \tag{3.39}$$

If $2N + L$ is a large positive integer, then the above coefficients which are on the order of the inverse of the square root of the terms s_j that are different from zero will be small. This results in small cross- and autocorrelation of $v(k)$ and $a(k)$ which validates the CLT. Supportive experimental results were presented by Metzger [101] whereby the pdf's of the effective noise in multilevel digital transmissions look reasonably Gaussian, and becomes increasingly Gaussian-like with increasing constellation sizes.

One remaining issue regards the possibility of several taps with relatively larger values than the rest dominating the overall distribution of $v(k)$. A simple example is when the equalization just begins, i.e., $w(k) = [0, \cdots, 0, 1, 0, \cdots, 0]$, and the channel has just two non-zero taps, $h = [1, \lambda]$, where λ is a non-zero value. Then $v(k)$ cannot be (approximately) a Gaussian random variable. A general conclusion is that the more accomplished equalization is, the more Gaussian-like $v(k)$ will become.

2) Consider $2N + L \to \infty$. Even though the derivation of the CLT assumes the limiting case when the number of terms approaches infinity, it can be shown that this is actually undesirable. As pointed out by Ding *et al* [57, ch. 3], one important condition for the CLT is for the *variance* of $v(k)$ to be finite and non-zero, i.e.,

$$0 < \lambda_1 \leq \mathrm{E}\{|s_j a(k-j)|^2\} \leq \lambda_2 < \infty, \qquad \forall j \tag{3.40}$$

where λ_1 and λ_2 are some finite positive constants.

In order to justify the assumption that $v(k)$ is Gaussian, we restrict our region of consideration to the region where the eye is almost open. By then the condition that s_j are independent random variables is satisfied. Further, we propose to prefix the usual equalizer with a recursive whitening filter [14, 75] in order to satisfy the CLT conditions that $2N+L$ be large but finite. A recursive filter ensures the length of the combined system to be large regardless of the order of the channel. Besides, prefixing the recursive filter also simultaneously satisfies the second condition of $2N+L < \infty$ since its (transversal) impulse response is exponentially decaying with increasing tap length [58]. This is why $2N + L$ is less than infinity in the practical sense. So long as the Gaussian assumption of $v(k)$ is valid, so would $\alpha(k)$ and assumption $(\mathcal{H}2)$.

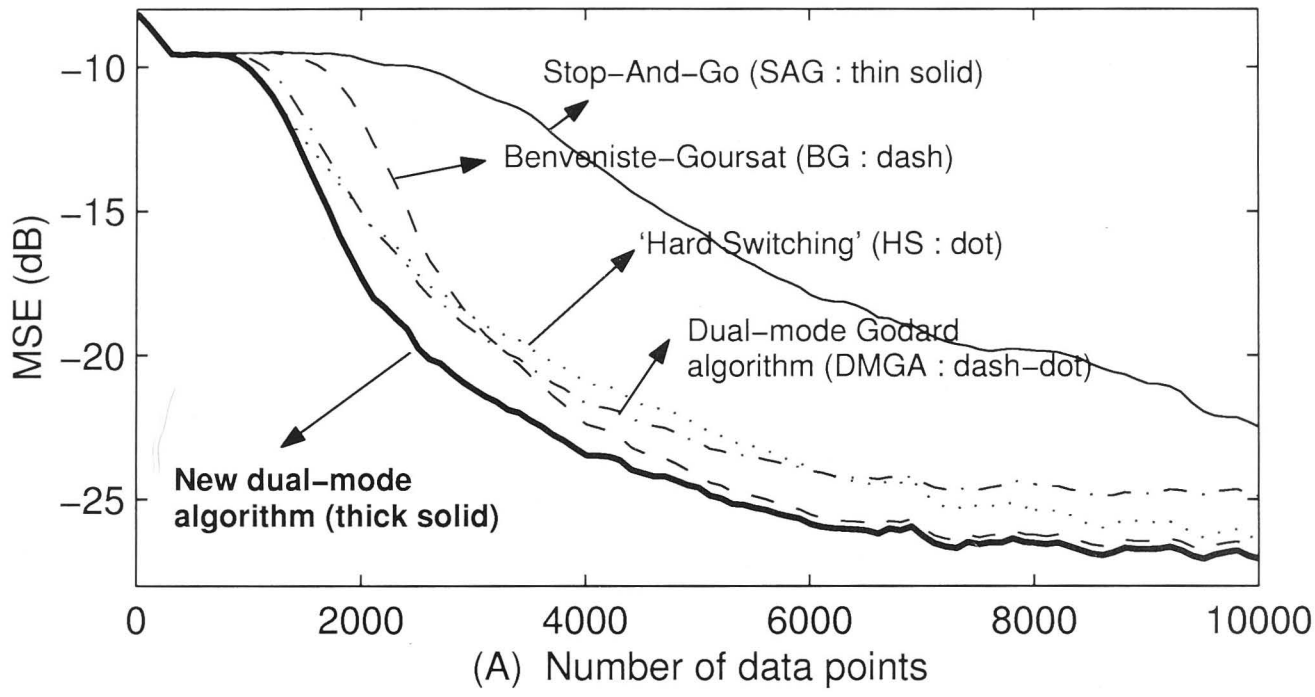

Figure 3.7: A comparison of several dual-mode algorithms equalizing channels (A) $\mathbf{h}'$, and (B) $\mathbf{h}''$, both at an SNR of 25dB for 4-PAM signalling.

3.6 Simulation Results

With stationary channels, the equalizer achievements can be characterized in terms of convergence speed and steady-state error. As our performance measure, we have used the DD MSE which can be estimated recursively via

$$\mathrm{MSE}_{\mathrm{DD}}(k+1) = \rho \mathrm{MSE}_{\mathrm{DD}}(k) + (1-\rho)[z(k) - \mathrm{Q}(z(k))]^2 \tag{3.41}$$

where $\rho = 0.99$ is the forgetting factor. Results have been obtained via Monte Carlo simulations using 200 independent runs on two nonminimum phase channels:

$$\mathbf{h}' = [0.04, -0.05, 0.07, -0.21, -0.5, 0.72, 0.36, 0, 0.21, 0.03, 0.07]$$
$$\mathbf{h}'' = [0.8264, -0.1653, 0.8512, 0.1636, 0.81].$$

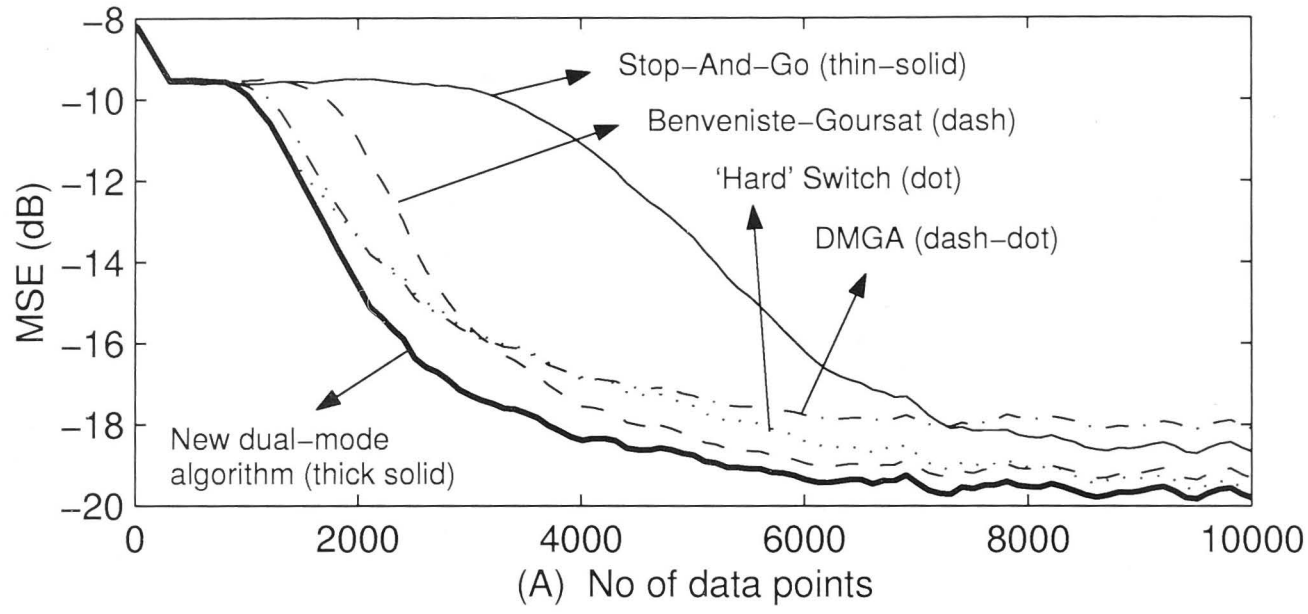

Figure 3.8: A comparison of several dual-mode algorithms equalizing channels $H_1(z)$ with an SNR of (A) 20 dB, and (B) 15 dB for 4-PAM signalling.

$\mathbf{h}'$ is a channel with a weaker coloring on the channel output [118], and $\mathbf{h}''$ is one with stronger coloring [75].

Simulations on both channels were carried out at an SNR of 25dB using 4-PAM signalling $\{\pm 1, \pm 3\}$. A total of 10^4 and 10^5 symbols were used in each simulation run when dealing with the channel $\mathbf{h}'$ and $\mathbf{h}''$, respectively. For the respective channels, we used a baud-rate equalizer with 20 taps and 40 taps initialized with a center tap strategy, employing a step size of 10^{-4} and 2.5×10^{-5}, respectively.

We compared the performance of the proposed dual-mode algorithm (new) with 4 other popular algorithms, namely the Benveniste-Goursat (BG) algorithm [15], the Stop-And-Go (SAG) algorithm [116], the dual-mode Godard algorithm (DMGA) [145] and the traditional "hard switching" (HS) algorithm. The HS algorithm will switch from the CMA to the LMS algorithm when $\mathrm{MSE}_{\mathrm{DD}}(k) < 0.25$. The choice of acquisition and tracking algorithms in all cases is the CMA 2-2 and the DD LMS algorithm, respectively, except for the DMGA [145] which employs the decision adjusted modulus algorithm (DAMA) [126] at steady state. Their

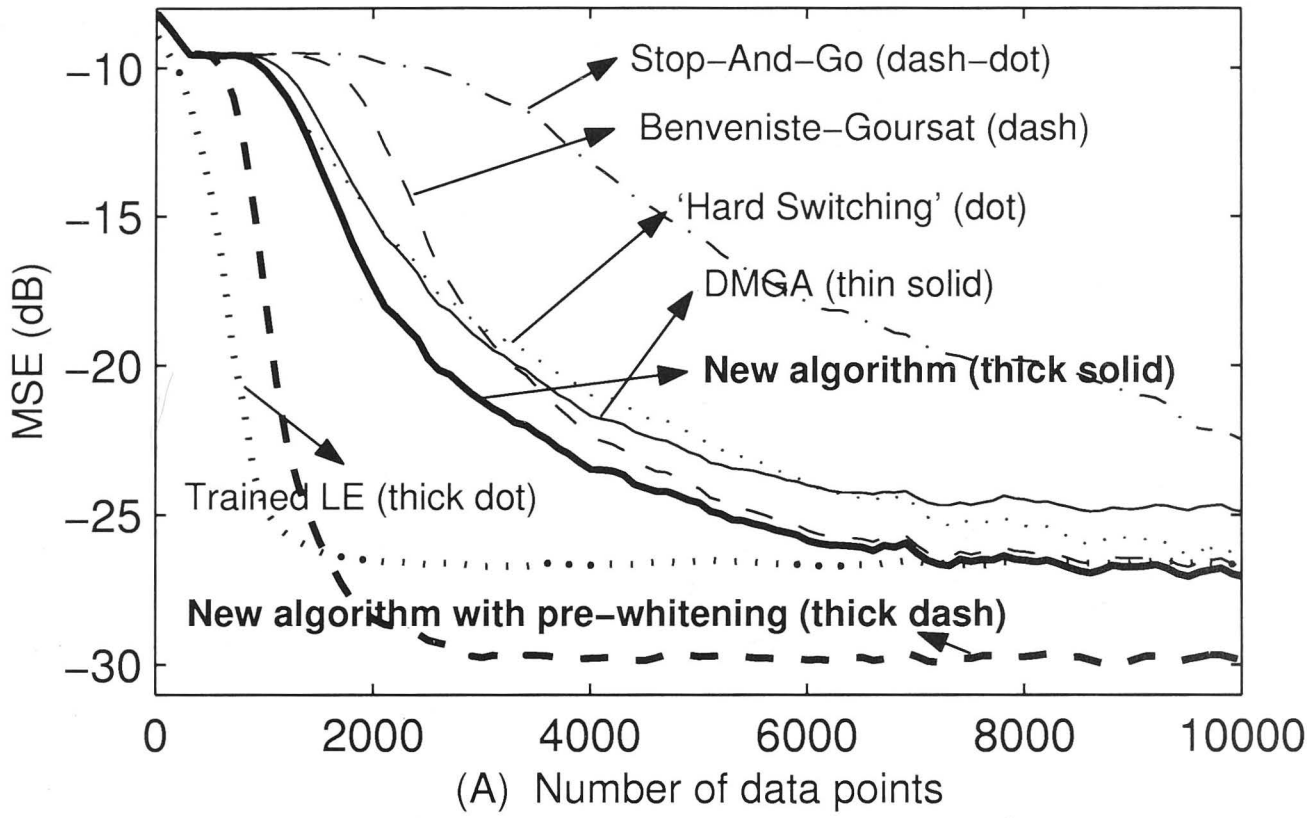

Figure 3.9: Similar to Fig. 3.7, this graph also includes the new equalizer that is prefixed with a recursive whitening filter where lower steady state MSE is achieved due to the additional whitening filter which increases the effective length of the equalizer, thereby being able to estimate the channel more accurately. The SNR is 25 dB, and a 4-PAM source is used.

respective error functions are given below:

$$\epsilon_{\mathrm{CMA}}(k) = z(k)(|z(k)|^2 - R_2) \tag{3.42}$$

$$\epsilon_{\mathrm{DD}}(k) = z(k) - \mathrm{Q}(z(k)) \tag{3.43}$$

$$\epsilon_{\mathrm{DAMA}}(k) = z(k)(|z(k)|^2 - \mathrm{Q}^2(z(k))) \tag{3.44}$$

where $\mathrm{Q}(\cdot)$ is the nearest neighbor quantizer.

For our simulations, the parameters of various algorithms of comparison are outlined below. In the notation of (3.22) we can express the error function in terms of the combination $\beta(k) = [\beta_1, \beta_2]$. The name of the associated algorithm is superscripted on $\beta(k)$. We have assigned $\beta^{\mathrm{BG}}(k) = [4, |\epsilon_{\mathrm{DD}}(k)|]$; $\beta^{\mathrm{SAG}}(k) = [40, 0]$ for $\mathbf{h}'$ and $\beta^{\mathrm{SAG}}(k) = [14, 0]$ for $\mathbf{h}''$ when $\mathrm{sgn}(\epsilon_{\mathrm{CMA}}(k)) = \mathrm{sgn}(\epsilon_{\mathrm{DD}}(k))$ and $\beta^{\mathrm{SAG}}(k) = [0, 0]$ when $\mathrm{sgn}(\epsilon_{\mathrm{CMA}}(k)) \neq \mathrm{sgn}(\epsilon_{\mathrm{DD}}(k))$; $\beta^{\mathrm{DMGA}}(k) = [1, 0]$, $\forall |z(k) - \mathrm{Q}(z(k))| < 0.2$, and $\beta^{\mathrm{DMGA}}(k) = [0, 1]$ otherwise; $\beta^{\mathrm{HS}}(k) = [0, 1]$ when $\mathrm{MSE}_{\mathrm{DD}}(k) > 0.25$, and $\beta^{\mathrm{HS}}(k) = [14, 0]$ otherwise. As for our dual-mode algorithm, we used $\alpha(k)$ from (3.27) and assigned $\gamma = 14$. The variance σ_v^2 is estimated from (3.32).

We simulated 200 runs for each dual-mode algorithm for both channels. The graphs of the DD MSE of the averaged runs are plotted in Fig. 3.7. Note that these graphs are obtained by averaging out only the MSE of the runs that have

Table 3.2: Summary of results in Fig. 3.7 including failure rates

Dual-mode Algorithm	$\mathbf{h}'$ Fail rate	$\mathbf{h}'$ Normalized time to reach -14dB	$\mathbf{h}'$ Normalized time to reach -18dB	$\mathbf{h}''$ Fail rate	$\mathbf{h}''$ Normalized time to reach -14dB	$\mathbf{h}''$ Normalized time to reach -18dB
New	0%	1	1	3.5%	1.11	1
BG	0%	1.45	1.35	16.5%	1.65	1.35
SAG	0%	2.70	2.88	55.0%	3.18	-
DMGA	0%	1.15	1.25	1.5%	1.06	1.07
HS	1.5%	1.13	1.31	53.5%	1	1.23
Time normalized by		1590	2110		28900	49700

been successful in convergence. The summary of the results that includes the failure rate of convergence is tabulated in Table 3.2. The failure rate is calculated by recording the number of runs where the MSE at the end of a particular run is higher than -13.15 dB $= 0.22$ then dividing by the total number of runs. The time to convergence has been normalized by the averaged number of symbols required by the fastest algorithm. From Table 3.2, the hard switching algorithm is the most unreliable as it yields high failure rates. Both our algorithm and the DMGA are the smoothest in terms of low failure rate, but the DMGA yields higher steady state MSE and is also carrier phase blind. The general conclusion is that the proposed new algorithm is superior than others in terms of convergence speed and steady state error based on the results in Fig. 3.7 and Table 3.2.

3.7 Conclusions

In this chapter, we proposed a new switch-mode algorithm. It reflects the reliability of the equalizer output as a function of the equalizer output itself and the effective noise variance. When employed, the technique exhibits superior performance in terms of convergence speed and steady state errors relative to several conventional switch-mode algorithms. This new technique also eliminates any requirement for the manual control of the parameters that govern the convergence speed and excess noise.

Chapter 4

Probabilistic-Based Switching Technique For Switch-Mode Algorithms

In this chapter, we propose a new switching technique that is intended to ensure a smooth switch-over between the acquisition and tracking algorithms in addition to achieving rapid convergence and low steady state errors. Like the reliability-based technique proposed in Chapter 3, this new technique also exploit both the equalizer output as well as its estimated distribution to achieve the above objectives. The most important property of this technique lies in the simplicity of this technique that accurately and promptly detects suitable conditions for the employment of the tracking algorithm. The technique incurs no additional costs but it achieves significant improvements over conventional switch-mode algorithms, for both hard-switching and soft-transition type. When coupled with the result of Chapter 6 which reveals that the multi-modulus algorithm (MMA) [105, 152] usually yields correctly oriented output constellation upon the convergence of the adaptive parameters, a tri-mode algorithm is proposed to perform joint blind equalization and phase recovery.

Conventionally, the switch-mode algorithm is perceived to be a combination of an acquisition algorithm and a tracking algorithm in such a way that they complement one another's weakness so that the performance of the algorithm is improved. However, in a more critical approach, we choose to treat switch-mode algorithms as merely a tracking algorithm whose acquisition is assisted by a temporary acquisition algorithm that is less desirable in the sense that it yields poorer steady state performance, but is absolutely necessary during acquisition. Thus, it becomes clear that failure to switch-over from the acquisition algorithm to the tracking algorithm

as soon as the data estimates become sufficiently reliable will consequently slow down the convergence of the equalizer. Hence, this motivates the design of better and more efficient techniques to detect the open-eye condition reliably.

In Section 4.1. of this chapter, we present a brief problem statement regarding the difficulty in performing switching. In Section 4.2, we illustrate graphically the error functions of the acquisition and tracking algorithms to help us appreciate the abruptness encountered due to the switching in switch-mode algorithms. In Section 4.3, we present our novel switching technique. In Section 4.4., we present a tri-mode algorithm that uses this switching technique to perform joint blind equalization and phase recovery effectively. Lastly, we present convincing simulation results to show the improvement of this technique over some previous criteria reported in the literature.

4.1 Problem Statement

Without a training sequence in blind equalizers, the detection of an open eye condition using for example the closed-eye measure (CLEM) in (2.7) is impossible. Thus, we must make a good guess regarding this condition based only on the available observed data. To make the existing task even more challenging, we will rely only on the most current data signal, for reduced complexity, to make this guess. Conventionally, it is common to estimate the mean-squared error (MSE) of the equalizer output, as in the "hard-switching" criterion, and periodically compare it to a pre-defined threshold value that usually corresponds to low error rates. These threshold values, unfortunately, need to be chosen in a very conservative manner to avoid switching too early that results in the ill-convergence of the tracking algorithm [92, 99]. The acquisition algorithm is thus employed much longer than desired, thereby sacrificing valuable acquisition time. Moreover, the MSE criterion may not be suitable because of the steady state MSE is comprised of the minimum achievable MSE plus an excess MSE term that is due to the fluctuation of the filter tap parameters about their optimal values [42, 58]. Large amount of excess MSE would often contribute to the total MSE especially for large non-constant modulus source constellations [65] such as higher order PAM and QAM constellations. Therefore, the MSE estimate that is high due to the excess MSE contributions may not drop below the pre-defined threshold value even though the channel eye is clearly open, thus preventing it from switching to the tracking algorithm. In addition to the above problems, most switching criteria are based on second order or higher order statistics of the data sequence, for example the MSE and the reliability

measure [82], which must be estimated. A large sample size is therefore required before an accurate estimate can be obtained. Again, the collection of large samples may take time and slow the estimation process and consequently, slow convergence.

4.2 Graphical Illustration of Switching Difficulty in Switch-Mode Algorithms

Since we have repeatedly mentioned about the switching difficulties encountered during the switch-over of the switch-mode algorithms, some graphical illustrations are provided in this section to assist us in appreciating this problem visually. Consider a transversal equalizer, $\mathbf{w}(k)$, whose regressor vector is denoted by $\mathbf{r}(k) = [r(k-N), \cdots, r(k), \cdots, r(k+N)]$. Then the stochastic gradient descent equation of the equalizer is

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \mu \frac{\partial J(k)}{\partial \mathbf{w}(k)} \tag{4.1a}$$

$$= \mathbf{w}(k) - \mu \epsilon(k) \mathbf{r}^*(k)^T. \tag{4.1b}$$

The error function of the switch-mode algorithm, $\epsilon(k)$, essentially combines the acquisition algorithm and the tracking algorithm. In what follows, we will plot the figures of $|\epsilon(k)|$ for some acquisition algorithms, namely, the CMA indexed by integers 1 and 2 (see (2.50)), and some tracking algorithms, namely, the DD algorithm and the decision adjusted modulus algorithm (DAMA) (2.70) [126]. Their cost and error functions are given below:

$$J_{\text{CMA1}}(k) = \text{E}\{(|z(k)| - \gamma_{\text{C}}^{(1)})^2\} \tag{4.2a}$$

$$J_{\text{CMA2}}(k) = \text{E}\{(|z(k)|^2 - \gamma_{\text{C}}^{(2)})^2\} \tag{4.2b}$$

$$J_{\text{DD}}(k) = \text{E}\{(z(k) - \text{Q}(z(k)))^2\} \tag{4.2c}$$

$$J_{\text{DAMA}}(k) = \text{E}\{|z(k)^2 - \text{Q}_R(z(k))^2|^2\} \tag{4.2d}$$

$$\epsilon_{\text{CMA1}}(k) = z(k) - \gamma_{\text{C}}^{(1)} \text{sgn}(z(k)) \tag{4.3a}$$

$$\epsilon_{\text{CMA2}}(k) = z(k)(|z(k)|^2 - \gamma_{\text{C}}^{(2)}) \tag{4.3b}$$

$$\epsilon_{\text{DD}}(k) = z(k) - \text{Q}(z(k)) \tag{4.3c}$$

$$\epsilon_{\text{DAMA}}(k) = z(k)\big(\big|z(k)^2 - \text{Q}_R(z(k))^2\big|^2\big) \tag{4.3d}$$

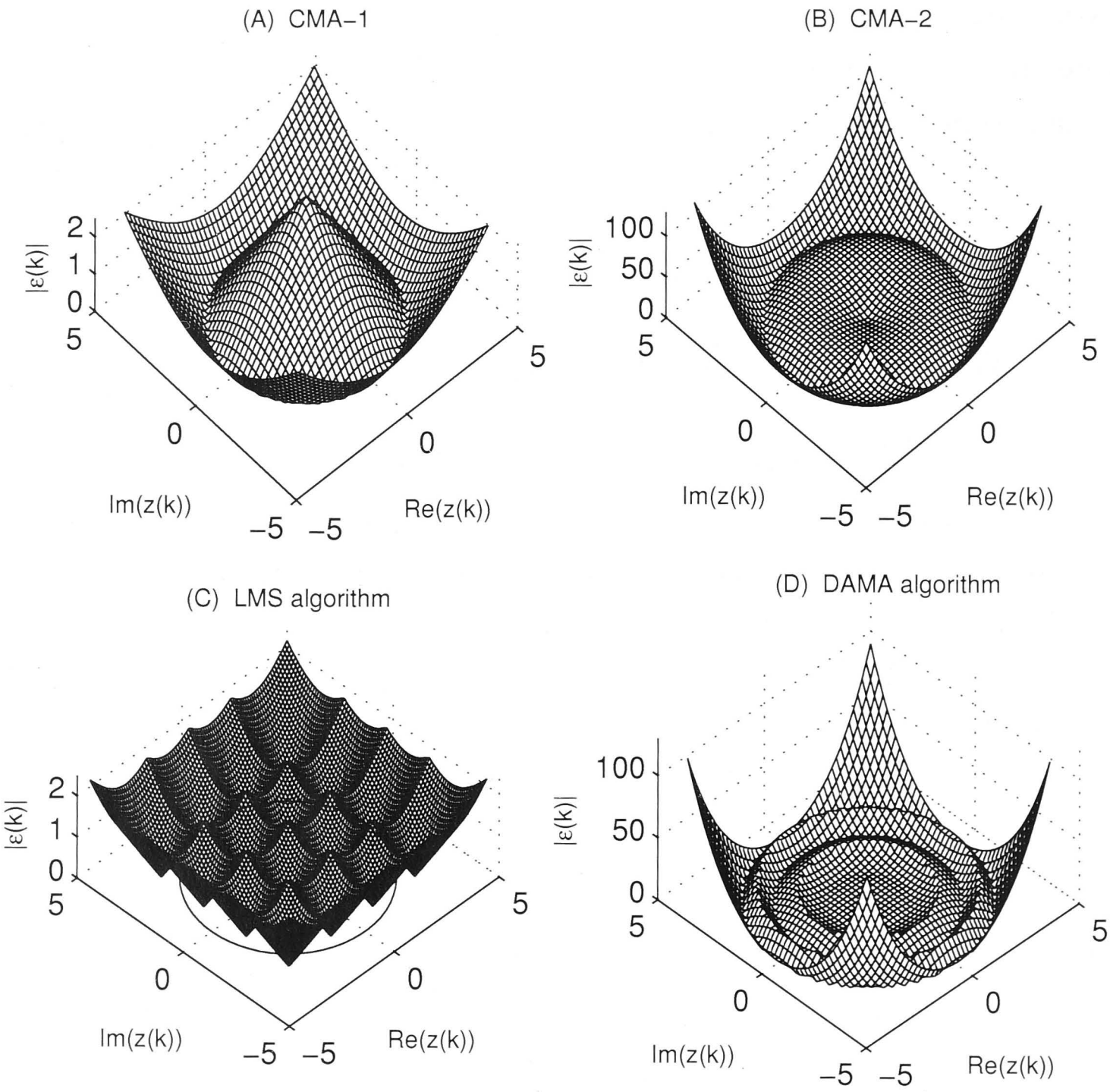

Figure 4.1: The plots of $|\epsilon(k)|$ for (A) CMA-1, (B) CMA-2, (C) DD algorithm, (D) DAMA algorithm, for the non-constant modulus 16-QAM source signals.

where $\gamma_{\mathrm{C}}^{(p)}$ is known as the dispersion constant that is parameterized by the positive integer p as in (2.51), $\mathrm{Q}(\cdot)$ is the nearest neighbor quantizer, $\mathrm{Q}_R = \arg\min_i \left\{ \big||z(k)| - r_i\big| \right\}$, where $\{r_i\}$ are the radii of circles that join the source constellation points. Further details on the DAMA can be found in Chapter 2.5.6 or [126].

Thus by plotting the figures of $|\epsilon(k)|$ against complex values of equalizer output $z(k)$, we can then appreciate the differences in the respective shapes of the acquisition and tracking algorithms that complicate switching. The shapes differ largely for data constellations which are constant modulus such as binary PAM, 4-QAM, 8-PSK and for those which are non-constant modulus such as 8-PAM and 16-QAM.

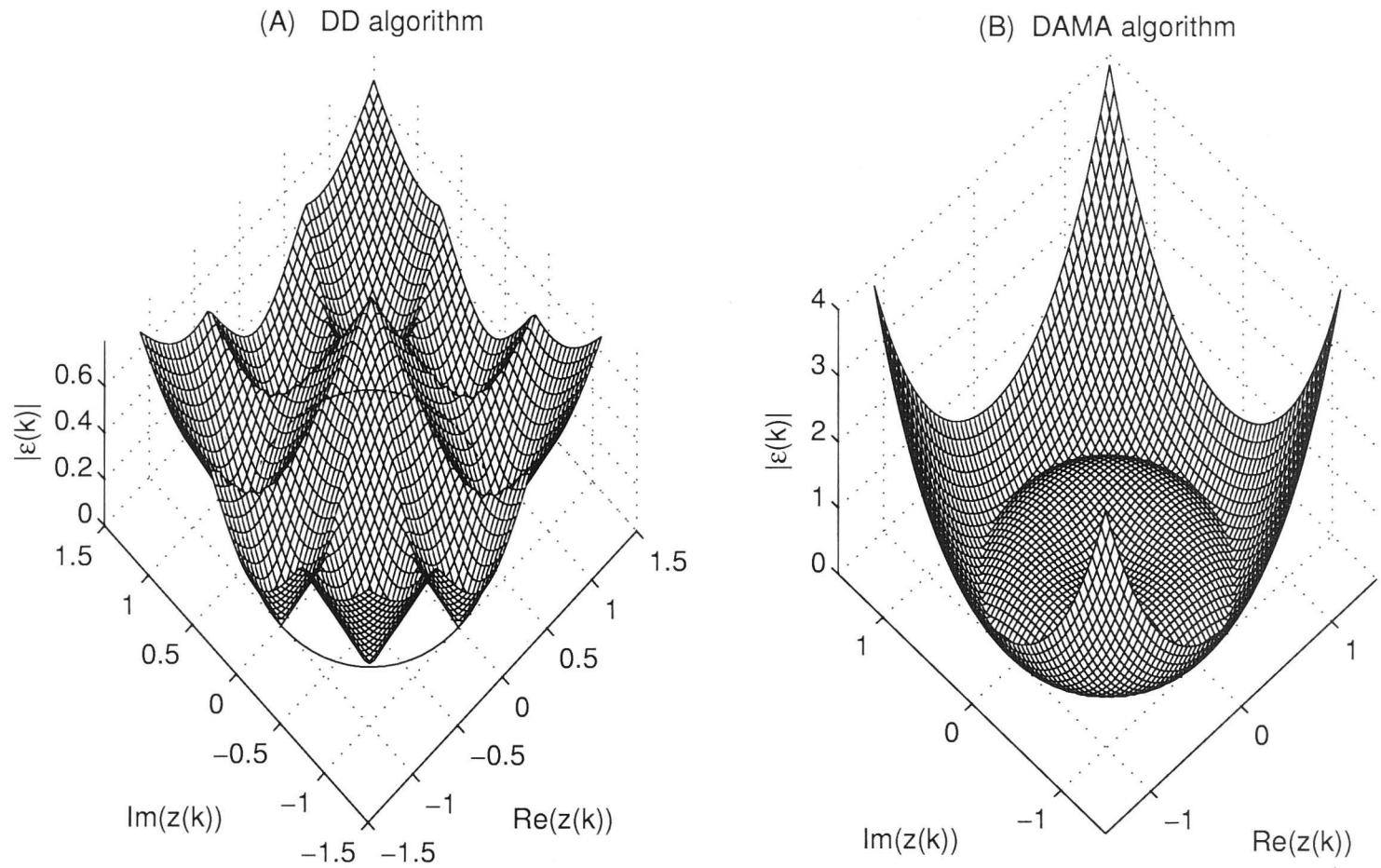

Figure 4.2: The plots of $|\epsilon(k)|$ for (A) DD algorithm, (B) DAMA algorithm, for the constant modulus 8-PSK source signals. The plots of $|\epsilon(k)|$ for CMA-1 and CMA-2 are identical to Fig. 4.1-A,B, except that $\gamma_{\mathrm{C}}^{(p)}$ has to be readjusted.

4.2.1 Non-Constant Modulus Constellation

We consider first the non-constant modulus constellation of 16-QAM. The absolute values of the error functions in (4.3) are plotted in Fig. 4.1. The minimum points as seem in the figures indicate zero update of the equalizer parameters. From Fig. 4.1-A,B, none of the sixteen (16) QAM data points lie on the minimum points. This is the reason why a finite excess MSE is still expected even when the CM equalizer has converged to its optimal settings [42,65]. This is because of the finite error term that constantly updates the equalizer parameters. On the contrary, the tracking algorithms such as the DD and the DAMA algorithms yield zero solutions when the equalizer has converged to its optimal settings since all 16 data points correspond to the minimum points of the plots in Fig. 4.1-C,D. It is clear from the figures that the error functions of the acquisition and the tracking algorithms are very different indeed, thereby resulting in a disruption in the convergence of the switch-mode equalizer when the switching occurs.

4.2.2 Constant Modulus Constellation

We consider now the constant modulus constellation of 8-PSK. From Fig. 4.2-A,B, it is clear that upon convergence, the eight (8) PSK data points would all

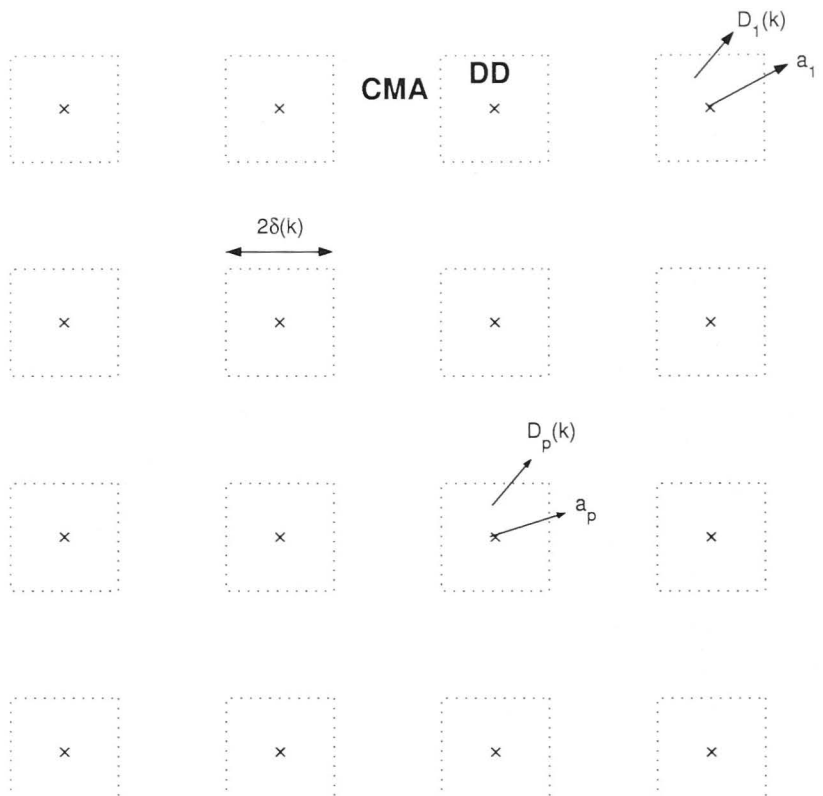

Figure 4.3: The new switch-mode algorithm featuring adaptation with the acquisition algorithm when $z(k)$ is outside the individual square regions surrounding each data point, and with the tracking algorithm otherwise.

lie on the surface that correspond to local minima. This is also the same for the acquisition algorithms, CMA-1 and CMA-2 (which have the same shapes as 4.2-A,B). The only difference between the acquisition and tracking algorithms is that upon convergence, the phase of the acquisition algorithm may differ (deviate) from the original phase. This result therefore implies that switching is easier if the data constellation is constant modulus than if it were non-constant modulus.

4.3 A Novel Probabilistic-Based Switching Technique

Using a concept that parallels the reliability measure derived from Bayes theorem in Chapter 3 and [82,83], we propose a new switching technique that is also dependent on *both* the equalizer output and its statistical distribution. This is because we have shown that according to Bayes theorem, the reliability of the equalizer output is a function of both the equalizer output and its distribution [82, 83]. Consequently, the new switch-mode algorithm may achieve more rapid convergence and lower steady state errors. On the other hand, the new technique in this chapter, unlike the reliability-based technique in Chapter 3 that is computationally intensive, is very simple but yet accurately reflects an open-eye condition. This switch-mode algorithm combines the acquisition and tracking algorithms in a convex manner

using a *binary* combination parameter[1].

We originally developed this smooth switching switch-mode technique for QAM systems. Its error function can be expressed as

$$\epsilon(k) = \alpha(k)\gamma\epsilon_{\mathrm{DD}}(k) + (1 - \alpha(k))\epsilon_{\mathrm{CMA}}(k) \tag{4.4}$$

or perhaps more conveniently as

$$\epsilon(k) = \begin{cases} \gamma\epsilon_{\mathrm{DD}}(k) & z(k) \in D(k) \\ \epsilon_{\mathrm{CMA}}(k) & z(k) \notin D(k) \end{cases} \tag{4.5}$$

where γ is appropriately chosen to compensate for the differences in the variance of the respective error functions as in (3.21) and details found in Appendix A, and the binary combination parameter is defined as

$$\alpha(k) = \begin{cases} 1 & z(k) \in D(k) \\ 0 & z(k) \notin D(k). \end{cases}$$

$D(k)$ is our reliable regions which are designed to correspond to regions of relatively low bit-error-rates so that the DD algorithm can be employed more reliably [92,99]. One relatively more reliable region in the QAM constellation space can be identified as regions enclosing each of the constellation point. Therefore, if we denote the individual region that encloses the p^{th} data point, a_p, as $D_p(k)$, then $D(k)$ is given by

$$D(k) = \cup D_p(k), \quad p = 1, 2, \cdots, M \tag{4.6}$$

where we have heuristically chosen $D_p(k)$ to be the square region, each of equal size for all p, in a similar fashion as [145]. The diagram illustrating this switch-mode algorithm is shown in Fig. 4.3. During the transient stages, the distortion introduced by the channel will cause the equalizer output to be scattered in a very large area around the transmitted data point. Thus, if $D(k)$ is sufficiently small during the initial acquisition mode, then the error function will be dominated by the acquisition algorithm. On the other hand, in the steady state, since the equalizer output will be close to the transmitted data point, then the error function will be dominated by the tracking algorithm. The more important contribution which we will now show is to vary the size of the region, $D(k)$, according the new probabilistic measure that is closely related to the distribution of the equalizer output. The

[1]It can sometimes be thought of as a binary approximation of the reliability measure.

increase of the size of $D(k)$ in steady state makes sense because in steady state, we want $D(k)$ to be a large area so that most equalizer output will be found in it and they can drive the equalizer with the DD algorithm instead of the acquisition algorithm. On the other hand, during the initial acquisition, we want $D(k)$ to be small so that most output will drive the equalizer with the acquisition algorithm.

Let the width of the each individual region, $D_p(k)$, be $2\delta(k)$ as in Fig. 4.3. Then we propose to update the width, which ultimately governs the size $D(k)$, according to *the probability of the equalizer output being found in the reliable region,* $D(k)$, i.e.,

$$\delta(k+1) = \Pr\{z(k) \in D(k)\} \tag{4.7a}$$

$$= \Pr\{\mathrm{Re}[z(k) - \mathrm{Q}(z(k))] \leq \delta(k)\} + \Pr\{\mathrm{Im}[z(k) - \mathrm{Q}(z(k))] \leq \delta(k)\}. \tag{4.7b}$$

Implementation-wise, the width of these individual regions, $\delta(k)$, can be updated according to the recursion

$$\delta(k+1) = \beta\delta(k) + (1-\beta)I(k) \tag{4.8}$$

where

$$I(k) = \begin{cases} 1 & \mathrm{Re}\big(\epsilon_{\mathrm{DD}}(k)\big) \leq \delta(k) \ \text{ and } \ \mathrm{Im}\big(\epsilon_{\mathrm{DD}}(k)\big) \leq \delta(k) \\ \delta(0) & \text{otherwise} \end{cases}$$

and β is a forgetting factor, say 0.99. Its initial value, $\delta(k)$, should be sufficiently small but non-zero. Consider first an initial closed-eye condition and $\delta(0)$ is small. Say under successful adaptation, the channel eye opens. Then $\Pr\{z(k) \in D(k)\}$ is increased regardless of the size of $D(k)$, as long as it is non-zero. This is because the highest concentration of equalizer output is found in the regions surrounding the data points. Consequently, the width of the individual region increases according to (4.8). The increase of $\delta(k)$ subsequently induces an increase in the size of the reliable region, $D(k)$. This will in turn increase $\Pr\{z(k) \in D(k)\}$ even more. This cycle repeats itself until an equilibrium is reached. For an open eye condition, $\delta(k) \to 1$ as $k \to \infty$. Conversely, for a closed eye condition, $\delta(k) \to \delta(0)$ as $k \to \infty$ in a similar manner.

Since the width $\delta(k)$ increases and decreases progressively, it provides a soft transition between the acquisition and the tracking algorithms. In addition, the simple technique incurs very little implementation cost but is very effective even

for large constellation sizes. Unlike higher-order criteria such as the DD MSE hard-switching technique, as well as the reliability-based technique [82, 83], this technique involves only first order estimation which are computationally simple and accurate with a small sample size.

4.4 Algorithm For Joint Blind Equalization And Phase Recovery

Extending the probabilistic switching technique developed in Section 4.3, we propose an algorithm that can perform joint blind equalization and phase recovery by exploiting reliability in the constellation space. We first describe the QAM equalizer model before delving into the development of the new algorithm.

4.4.1 System Model

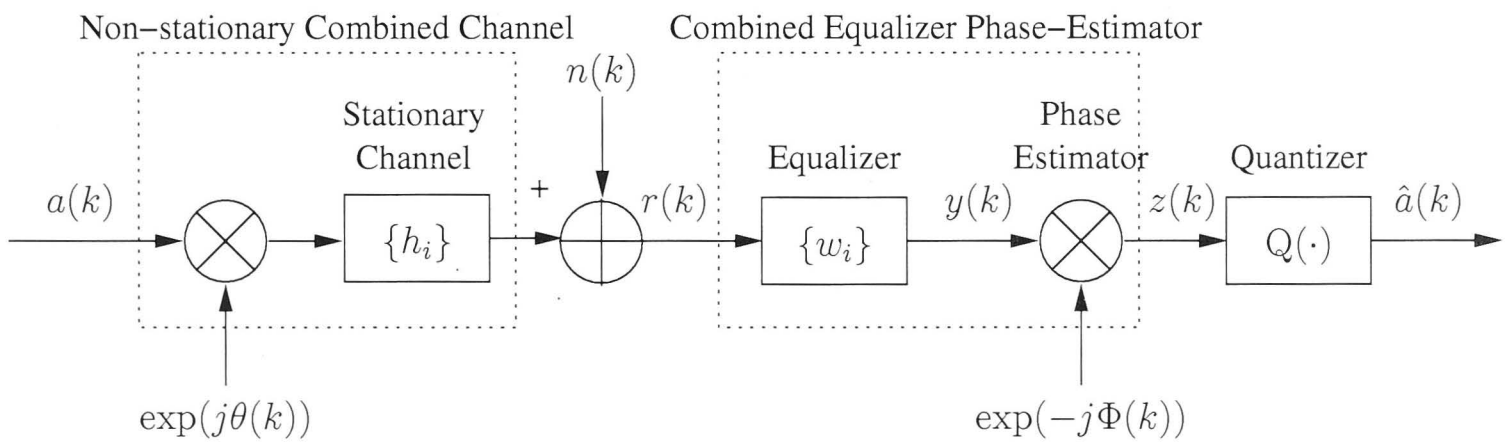

Figure 4.4: Baseband QAM system model for joint blind equalization and phase recovery.

The objective of the QAM digital baud-rate receiver that we will consider in this paper is to recover the data symbol up to a fixed time delay $\triangle_k$ and a possible phase shift $\bar{\theta}$ modulo 90°. The phase ambiguity of modulo 90° is allowed as it may be easily detected and corrected using differential encoding techniques. Thus the desired output of the combined equalizer-phase estimator should be in the form of

$$a(k - \triangle_k)e^{j\bar{\theta}}, \quad \text{where} \quad \bar{\theta} \in [0°, 90°, \cdots]. \tag{4.9}$$

where $a(k)$ is the transmitted symbol drawn from the M-QAM alphabet set $\mathbb{A} = \{a_1, a_2, \cdots, a_M\}$. The baseband channel output is corrupted by a linear distortive channel and additive white Gaussian noise (AWGN). In addition, we assume that demodulation phase errors may be present which may cause time-varying phase

rotations or simply an arbitrary phase offset in the output signal constellation. Therefore in a typical baseband receiver [49,63,139], its output prior to quantization can be expressed as (see Fig. 4.4)

$$z(k) = a(k)e^{j\Delta\theta(k)} + v(k) \tag{4.10}$$

where

$$\Delta\theta(k) = [\theta(k) - \Phi(k)] \tag{4.11}$$

is the instantaneous residual phase error term, $\theta(k)$ and $\Phi(k)$ are as depicted in Fig. 4.4, and

$$\theta(k) = 2\pi k f_\Delta + \theta_0 \tag{4.12}$$

is the demodulation phase error that consists of a normalized frequency offset, f_Δ, and an arbitrary phase offset, θ_0; and $v(k)$ is the residual-ISI-plus-channel-noise term. We will also make the simplifying assumption that $v(k)$ is Gaussian and circularly distributed with independent real and imaginary components each of variance σ_v^2 due to central limit theorem [57, Ch.2]. A typical receiver separately removes the ISI, $v(k)$, and compensates for the phase errors, $\Delta\theta(k)$, with an equalizer and a phase estimator, respectively. In our approach, we propose a new algorithm that accomplishes the joint tasks of ISI removal and phase estimation in a single filter. Our philosophy is as follows. Rather than treating the phase errors separately from $v(k)$ as in (4.10), we choose to express the equalizer output as

$$z(k) = a(k) + v'(k) \tag{4.13}$$

and then minimize $\mathrm{E}|v'(k)|^2$, where the $v'(k)$ is the ISI-plus-noise term of the non-stationary channel $e^{j\theta(k)} \cdot \mathbf{h}$, where $\mathbf{h}$ is the channel coefficient vector, $[h_1, h_2, \cdots]$.

4.4.2 Development of Novel Tri-Mode Algorithm

Having established the system model, we now propose an algorithm that performs joint blind equalization and phase recovery without having to rely on a separate phase estimator using the novel switching technique developed earlier in Section 4.3. In our approach, we have designed an algorithm that employs one of three blind algorithms depending on the location of the equalizer output in the constellation space. They are the DD algorithm, the MMA and the CMA, whose error functions

Table 4.1: Characteristics of blind algorithms

	CMA	MMA	DD algorithm
Robust Acquisition	Y	Y	N*
Low Excess Noise	N	N	Y
Phase-Awareness	N	Y	Y

(Y)es, (N)o.

* Requires the condition of an initial open eye.

are given below:

$$\epsilon_{\mathrm{DD}}(k) = z(k) - \mathrm{Q}(z(k)) \tag{4.14a}$$

$$\begin{aligned}\epsilon_{\mathrm{MMA}}(k) = {} & \mathrm{Re}\big(y(k)\big)\left[\mathrm{Re}\big(y(k)\big)^2 - \gamma_{\mathrm{M}}^2\right] \\ & + \sqrt{-1}\cdot \mathrm{Im}\big(y(k)\big)\left[\mathrm{Im}\big(y(k)\big)^2 - \gamma_{\mathrm{M}}^2\right]\end{aligned} \tag{4.14b}$$

$$\epsilon_{\mathrm{CMA}}(k) = y(k)\left(|y(k)|^2 - \gamma_{\mathrm{C}}^2\right) \tag{4.14c}$$

where $\gamma_{\mathrm{M}}^2 = \mathrm{E}\{\mathrm{Re}(a(k))^4\}/\mathrm{E}\{\mathrm{Re}(a(k))^2\}$ [152], $\gamma_{\mathrm{C}}^2 = \mathrm{E}\{|a(k)|^4\}/\mathrm{E}\{|a(k)|^2\}$ [49], and the names of the respective algorithms are subscripted on $\epsilon(k)$. These three algorithms display distinct characteristics which we will use, after identifying them, as the underlying reasons for the design of our algorithm. In particular, we are looking for robust acquisition abilities, low excess noise under convergence, and phase correcting capabilities. A summary of their properties is tabulated in Table 4.1.

Our proposed tri-mode algorithm, in terms of its error function, can be compactly expressed as

$$\epsilon(k) = \begin{cases} \gamma\epsilon_{\mathrm{DD}}(k) & z(k) \in D(k) \\ \epsilon_{\mathrm{MMA}}(k) & z(k) \in D_{\mathrm{out}} \\ \epsilon_{\mathrm{CMA}}(k) & z(k) \notin \cup\big(D(k), D_{\mathrm{out}}\big) \end{cases} \tag{4.15}$$

where γ is to compensate for the difference in the average values of $\epsilon_{\mathrm{DD}}(k)$ and $\epsilon_{\mathrm{MMA}}(k)$ as well as $\epsilon_{\mathrm{CMA}}(k)$; $D(k)$ and D_{out} are suitably defined regions of higher reliability in the constellation space, where the size of the former is time varying and the size of the latter is fixed. $D(k)$ should represent a region where $\epsilon_{\mathrm{DD}}(k)$ may be employed in a reliable manner where ill-convergence may be avoided with high probability. Reliability can be measured as a function of the variance of the residual ISI plus noise term [82] and the phase error. We choose to adopt $D(k)$ as proposed in Section 4.3 as described by equation 4.6. D_{out} is a fixed size region

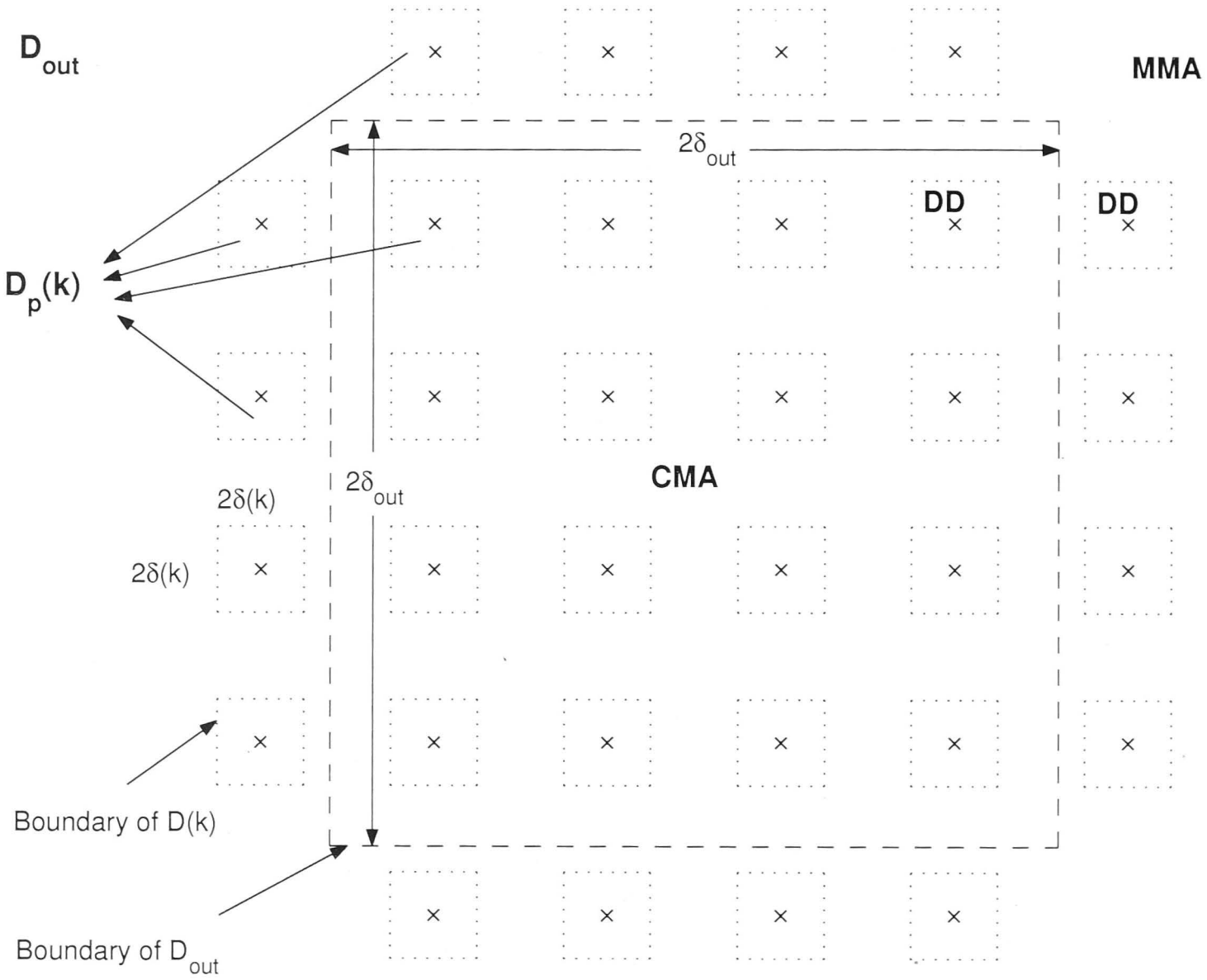

Figure 4.5: The new tri-mode blind equalization algorithm. The reliable regions $D(k) = \cup D_p(k), p = 1, 2, \cdots, 32$ and $D_{\text{out}}(k)$ where the DD algorithm and the MMA are respectively employed for the 32-QAM constellation are illustrated. The CMA is employed outside these regions.

outside a boundary set near the edge or corner symbols and it is exclusive of $D(k)$. Its main function is to separate the inner constellation from the outer constellation points so that the CMA is employed when smaller values of $z(k)$ are detected, otherwise the MMA is employed for larger values of $z(k)$. This design enables robust and rapid acquisition in the presence of phase errors which we now explain. According to (4.11-4.12), the phase error $\Delta\theta(k)$ consists of a constant phase offset θ_0 and a frequency component $2\pi k f_\Delta$. We make the following claim regarding phase recovery in the absence of frequency offsets, i.e., $f_\Delta = 0$.

In the absence of a frequency offset, phase recovery is guaranteed as long as $\Pr(z(k) \in D_{\text{out}}) > 0$, *i.e., the MMA is employed with a finite probability.*

This is true because the MMA must yield complete phase recovery when equalization is accomplished [81,152]. Its phase recovery properties are covered in Chapter 6. Therefore, the exact size of D_{out} is irrelevant as long as $\Pr(z(k) \in D_{\text{out}}) > 0$. Naturally the larger the size of D_{out}, the faster is the process of phase recovery.

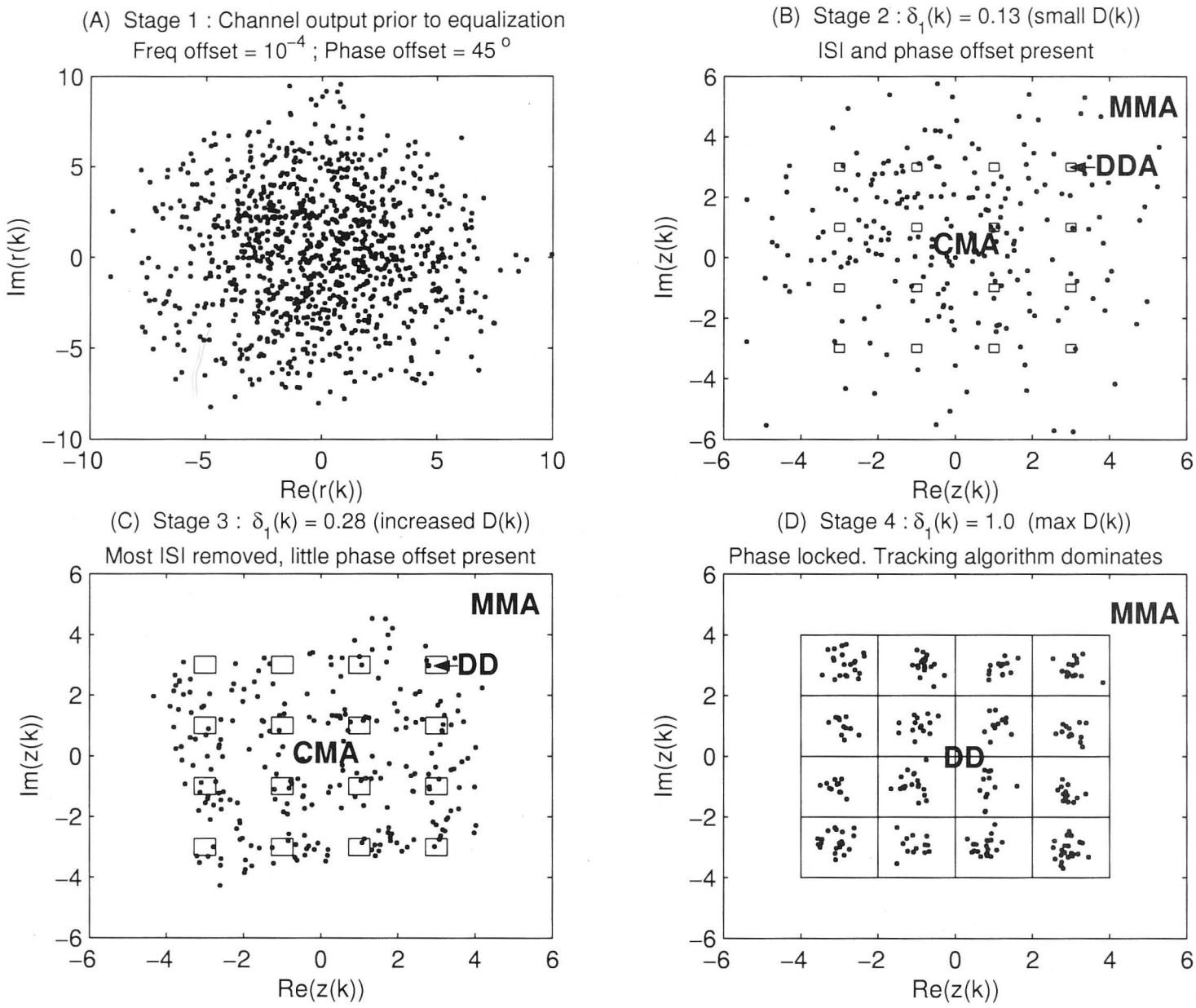

Figure 4.6: Figures illustrating the technique for joint equalization and phase recovery for 16-QAM signals in the presence of phase errors. The emphasis is on the variable-size reliable region, $D(k)$.

We propose that D_{out} be either

$$\left\{D_{\text{out}} : |z(k)| > \delta_{\text{out}}\right\} \tag{4.16}$$

which is a circle, or

$$\left\{D_{\text{out}} : \text{Re}(z(k)) > \delta_{\text{out}} \quad \text{and} \quad \text{Im}(z(k)) > \delta_{\text{out}}\right\} \tag{4.17}$$

which is a square. We have adopted the latter in our tri-mode algorithm as shown in Fig. 4.5. For fast phase recovery, δ_{out} should be as small as possible so that D_{out} becomes a large region that employs the MMA.

Consider now a (sizeable) frequency offset that corresponds to fast phase rotations where the MMA cannot recover. In such situations, the MMA will be combating fast phase rotations in addition to its usual task of ISI removal. This will result in large fluctuations in the equalizer parameters and hence increased MSE which we want to avoid. Thus, we design D_{out} to be the outer region where

the amplitude of $z(k)$ is large so that $\Pr(z(k) \in D_{\text{out}})$ is finite but small, so that the adaptation with the MMA is less frequent than with the CMA. Here we choose D_{out} to be the outer region because the equalizer output signals in this outer region are more reliable[2] [63, Sect. III]. Our design that performs phase discrimination only when $z(k)$ is large emulates several blind DD phase estimation techniques such as the "four-corner" technique [139] and the reduced-constellation phase-locked-loop (PLL) [63], with the exception that phase recovery is performed using a phase-aware algorithm instead of a separate phase estimation filter.

Discussion on characteristics of individual algorithms

Although employing the tracking algorithm is desirable in steady state because it yields less noisy solutions and possesses good tracking abilities, its convergence to undesirable local minima is highly likely if the initial parameter values cause significant number of incorrect decisions [92, 99]. The CMA and the MMA on the other hand, can perform reliable channel acquisition that results in significant ISI removal in a zero-forcing fashion [49, 65, 152]. However, their steady state solutions are significantly noisier than the DD solution, especially for large non-constant modulus modulation formats [65]. This is why the CMA or the MMA is usually employed during acquisition and later switched to the tracking algorithm when decision errors are less likely to occur. As for the task of simultaneous phase recovery, the MMA is chosen as the candidate algorithm that facilitates phase recovery in the acquisition mode, while the tracking algorithm caters for the tracking mode. The MMA is phase-aware because its cost function minimizes the dispersion of the output samples around straight moduli which fits into a reduced version of the original constellation in the statistical sense. The tracking algorithm is phase-aware as it minimizes $\mathrm{E}|v'(k)|^2$, where $v'(k)$ is the effective error term as a function of both residual ISI and phase error as expressed in (4.10). Table 4.1 shows a summary of the important characteristics of the above mentioned algorithms.

4.4.3 Performance Improvements via Pre-Whitening

Considerable performance improvements may be achieved if the channel output is first whitened prior to the equalizer. A whitening filter, or otherwise known as the minimum output energy (MOE) filter, can be implemented. Let $\{r_{\text{ch}}\}$ be the channel output sequence and $\{r\}$ be the equalizer input sequence. The MOE filter

[2]This is because they suffer less noise contributions from their adjacent data points.

minimizes the cost

$$J_{\text{MOE}} = \text{E}\{|r(k)|^2\} \tag{4.18}$$

which is its output energy. There are three areas where we may benefit from implementing the MOE filter prior to the equalizer.

1. Rapid acquisition: An MOE filter is able to compensate for the amplitude distortions of the channel. Thus, via pre-whitening, the proceeding equalizer will be left with the simpler task of compensation of residual phase distortions by the 'slower' CMA that computes higher moments of the received data. As demonstrated in [75], pre-whitening significantly increases convergence time compared to equalization without pre-whitening.

2. Low steady state MSE: Perhaps one of the nicest properties of the equalization scheme in [75] is that the equalizer acquires the channel efficiently as a linear equalizer, then switches to a decision feedback equalizer (DFE) to suppress steady state MSE when its output is sufficiently reliable. A predictive DFE scheme that is proposed in [5] can also be used in conjunction with the implementation of the MOE filter.

3. Larger step size for the DD algorithm for enhanced tracking: The step size of the DD algorithm is governed by the eigenvalue spread of the autocorrelation matrix of the received signals $r(k)$ [58]. It is bounded as follows:

$$0 < \mu < \frac{2}{\lambda_{\max}}$$

where $\lambda_{\max}$ is the largest eigenvalue of the autocorrelation matrix. Usually when the channel is unknown, a conservatively small value of μ that slows convergence is assigned in case of highly correlated channel output. Through whitening, $\lambda_{\max}$ may be minimized, hence allowing a larger step size to be employed for faster tracking purposes.

By incorporating the MOE filter, we summarize the algorithm and its associated step size assignments in Table 4.2.

4.5 Simulation Results

The objective of this section is to provide a performance comparison of our new tri-mode algorithm with probabilistic switching-technique and other conventional al-

Table 4.2: Progressive stages in new algorithm and associated step sizes employed in our simulations in the presence of phase errors

Stage	Step Sizes of MOE*	CMA	MMA	DD algorithm	Description
1	none				initialization
2	1×10^{-3}	2×10^{-5}	2×10^{-5}	2×10^{-5}	acquisition mode †
3	5×10^{-5}	5×10^{-5}	5×10^{-5}	5×10^{-5}	usual acquisition mode ††
4	2.5×10^{-5}	5×10^{-7}	5×10^{-7}	1.5×10^{-3}	DD mode †††

* MOE stands for the Minimum Output Energy algorithm

† MOE algorithm dominates in this mode.

†† Joint blind equalization and phase recovery is performed in this mode.

††† It can be switched to a DFE as in [5, 75] if required.

gorithms that have been proposed in the literature. The performance improvement of the tri-mode algorithm, as we will show later, varies according to the severity of the channel. Basically, the more severe the channel, the more distinctive is the tri-mode algorithm's performance over its competitors. The improvement is due to the way the tri-mode algorithm is designed and the smooth switching technique employed.

We separate the simulations into two major sections, where the first deals with a variety of channels assuming zero phase errors. The later section then deals with channels with phase errors. As the performance of the equalizer is usually characterized by the mean-squared error (MSE), we will estimate the decision directed MSE according to the following recursion:

$$\mathrm{MSE}_{\mathrm{DD}}(k+1) = \rho \mathrm{MSE}_{\mathrm{DD}}(k) + (1-\rho)\big|z(k) - \mathrm{Q}(z(k))\big|^2 \tag{4.19}$$

where $\rho = 0.99$ is our assigned forgetting factor. Throughout this section, we will use only three channels for all simulations. They consist of a lightly colored channel, $\mathbf{h}'$ [118], a medium colored channel, $\mathbf{h}''$ [118], and a heavily colored channel, $\mathbf{h}'''$ [75]. Their impulse responses are respectively given by

$$\mathbf{h}' = [0.04, -0.05, 0.07, -0.21, -0.5, 0.72, 0.36, 0, 0.21, 0.03, 0.07] \tag{4.20a}$$

$$\mathbf{h}'' = [2 - 0.4j, 1.5 + 1.8j, 1, 1.2 - 1.3j, 0.8 + 1.6j] \tag{4.20b}$$

$$\mathbf{h}''' = [0.8264, -0.1653, 0.8512, 0.1636, 0.81]. \tag{4.20c}$$

We will compare out tri-mode algorithm several other switch-mode algorithms,

where some employ the "hard-switching" criterion and others smooth switching techniques. The switch-mode algorithms that employ the "hard-switching" technique can be expressed as

$$\epsilon(k) = \begin{cases} \epsilon_{\text{acq}}(k) & \text{MSE}_{\text{DD}}(k) > \text{C}_{\text{MSE}} \\ \gamma\epsilon_{\text{DD}}(k) & \text{MSE}_{\text{DD}}(k) \leq \text{C}_{\text{MSE}}. \end{cases} \tag{4.21}$$

where C_{MSE} is a suitably chosen threshold value that would correspond to low error rates under the condition $\text{MSE}_{\text{DD}}(k) \leq \text{C}_{\text{MSE}}$. Two "hard-switching" algorithms under comparison are

1. The constant modulus algorithm (CMA) [49] and DD algorithm, where $\epsilon_{\text{acq}}(k) = \epsilon_{\text{CMA}}(k)$.

2. The multi-modulus algorithm (MMA) [152] and DD algorithm, where $\epsilon_{\text{acq}}(k) = \epsilon_{\text{MMA}}(k)$.

The switch-mode algorithms that employ smooth switching techniques are

3. Dual-mode Godard Algorithm (DMGA) [145] type where

$$\epsilon(k) = \begin{cases} \epsilon_{\text{MMA}}(k) & \text{Re}[z(k) - \text{Q}(z(k))] > \delta_{\text{DMGA}} \text{ or } \text{Im}[z(k) - \text{Q}(z(k))] > \delta_{\text{DMGA}} \\ \gamma\epsilon_{\text{DD}}(k) & \text{otherwise.} \end{cases} \tag{4.22}$$

and δ_{DMGA} is a suitably defined width between 0 and 1, normally small at around 0.2 [145].

4. Benveniste-Goursat [15] type where

$$\epsilon(k) = k_1\gamma\epsilon_{\text{DD}}(k) + k_2|\epsilon_{\text{DD}}(k)|\epsilon_{\text{CMA}}(k) \tag{4.23}$$

where k_1 and k_2 are a suitably chosen constants [15].

5. Tri-Mode Algorithm (4.15) where

$$\epsilon(k) = \begin{cases} \gamma\epsilon_{\text{DD}}(k) & z(k) \in D(k) \\ \epsilon_{\text{MMA}}(k) & z(k) \in D_{\text{out}} \\ \epsilon_{\text{CMA}}(k) & z(k) \notin \cup\big(D(k), D_{\text{out}}\big). \end{cases} \tag{4.24}$$

For our simulations, the parameters of the equalizer are as follows:

· Adaptation step size: $\mu = 5 \times 10^{-5}$, unless otherwise stated.

- $\gamma = 10$.
- For CMA-DD and MMA-DD, the MSE threshold value is assigned to be $C_{MSE} = \sqrt{2 \times 0.28^2} = 0.396$.
- For DMGA type algorithm, $\delta_{DMGA} = 0.25$, unless otherwise stated.
- For BG type algorithm, $k_1 = 5$ and $k_2 = 1$, unless otherwise stated.
- For the tri-mode algorithm, the initial width value is assigned as $\delta(0) = 0.25$ and the forget factor is $\beta = 0.95$, unless otherwise stated.
- The length of the equalizer for various channels are varied according to the length of the inverse response of the channel. For $\mathbf{h}'$, the equalizer has 31 taps. For $\mathbf{h}''$, the equalizer has 36 taps. For $\mathbf{h}'''$, the equalizer has 41 taps.

4.5.1 Results for Stationary Channel Without Phase Errors

We want to compare the performance of our tri-mode algorithm against several competitors as listed above under various conditions. We found a simple channel, $\mathbf{h}'$, and two difficult channels with deep spectral nulls, $\mathbf{h}''$ and $\mathbf{h}'''$. We also test the tri-mode algorithm against different modulation formats. For all simulations, we performed 50 independent Monte Carlo runs for each algorithm and plotted their average $MSE_{DD}(k)$ in Fig. 4.7 to Fig. 4.10.

16-QAM, h′

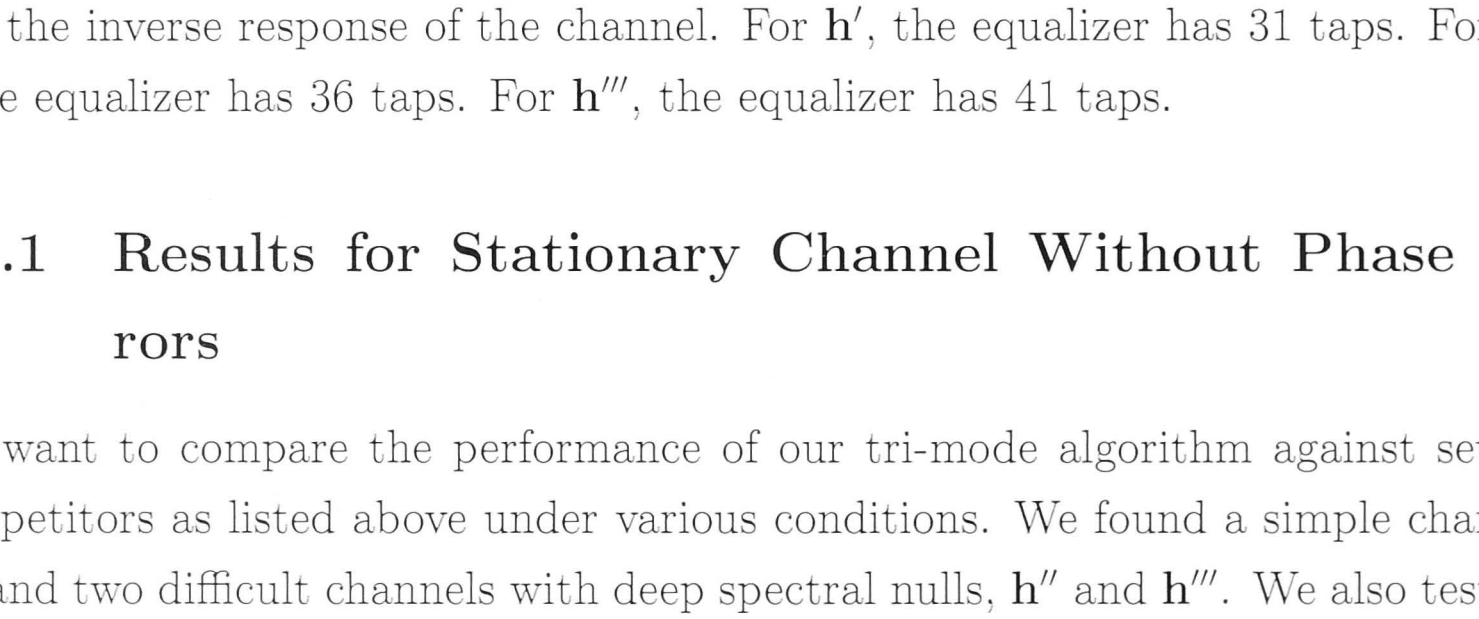

Figure 4.7: Comparing various algorithms using the channel $\mathbf{h}'$ under SNR levels of 15dB and 25dB only. The modulation format is 16-QAM. Equalizer has 31 taps.

For the first easy channel, $\mathbf{h}'$, we performed our simulations on the channel with an SNR levels of 15 dB and 25 dB for 16-QAM data using a total of 10,000 symbols for each of our runs. At lower SNR levels, e.g., 15 dB, smooth-switching algorithms outperform the hard switching algorithms in terms of steady state errors, whereby a gain of close to 1 dB is achieved over the MMA-DD algorithm at steady state. The speed of convergence for all the algorithms are almost the same, where the tri-mode algorithm has a marginal improvement over other competitors. The CMA-DD exhibit high steady state MSE because under adaptation, its output constellation has been regularly rotated because it is a phase-invariant algorithm, unlike the other algorithms.

Under higher SNR level, i.e., 25 dB, the tri-mode algorithm outperforms all other competitors in terms of convergence rate. In terms of steady state errors, it outperforms the BG type algorithm and marginally outperforms the DMGA-type algorithm. It achieves the same steady state performance as the MMA-DD algorithm. The CMA-DD algorithm exhibits higher steady state MSE and slower convergence due to the same reason as given above.

16-QAM, $\mathbf{h}''$

For the difficult channel, $\mathbf{h}''$, we performed our simulations on the channel with an SNR level of 15 dB for 16-QAM data using a total of 100,000 symbols for each of our runs. The step sizes for all algorithms are reduced by a factor of 50, i.e., $\mu = 10^{-6}$, except for the BG type algorithm where its step size is further reduced to $\mu = 2 \times 10^{-7}$ to ensure high rate of successful convergence. Here for the BG type algorithm, $k_1 = k_2 = 1$. The CMA-DD algorithm is omitted because it cannot converge at all unless the phase recovery is performed.

From Fig. 4.8, it is shown that our tri-mode algorithm converges the fastest on average among all four switch-mode algorithms, followed closely by the hard-switching method. The smooth switching algorithms converge much slower. The tri-mode algorithm, MMA-DD, and the DMGA-type algorithms achieve similar steady state errors in this special case. The BG-type algorithm is very slow in its convergence mainly due to a smaller step size that we assigned (5 times smaller). This is essential to ensure the convergence of the BG-algorithm with high probability of success. In other words, the BG-algorithm cannot be used reliably to open the channel-eye unless the step size is significantly reduced, thereby slows the convergence as shown in the figure.

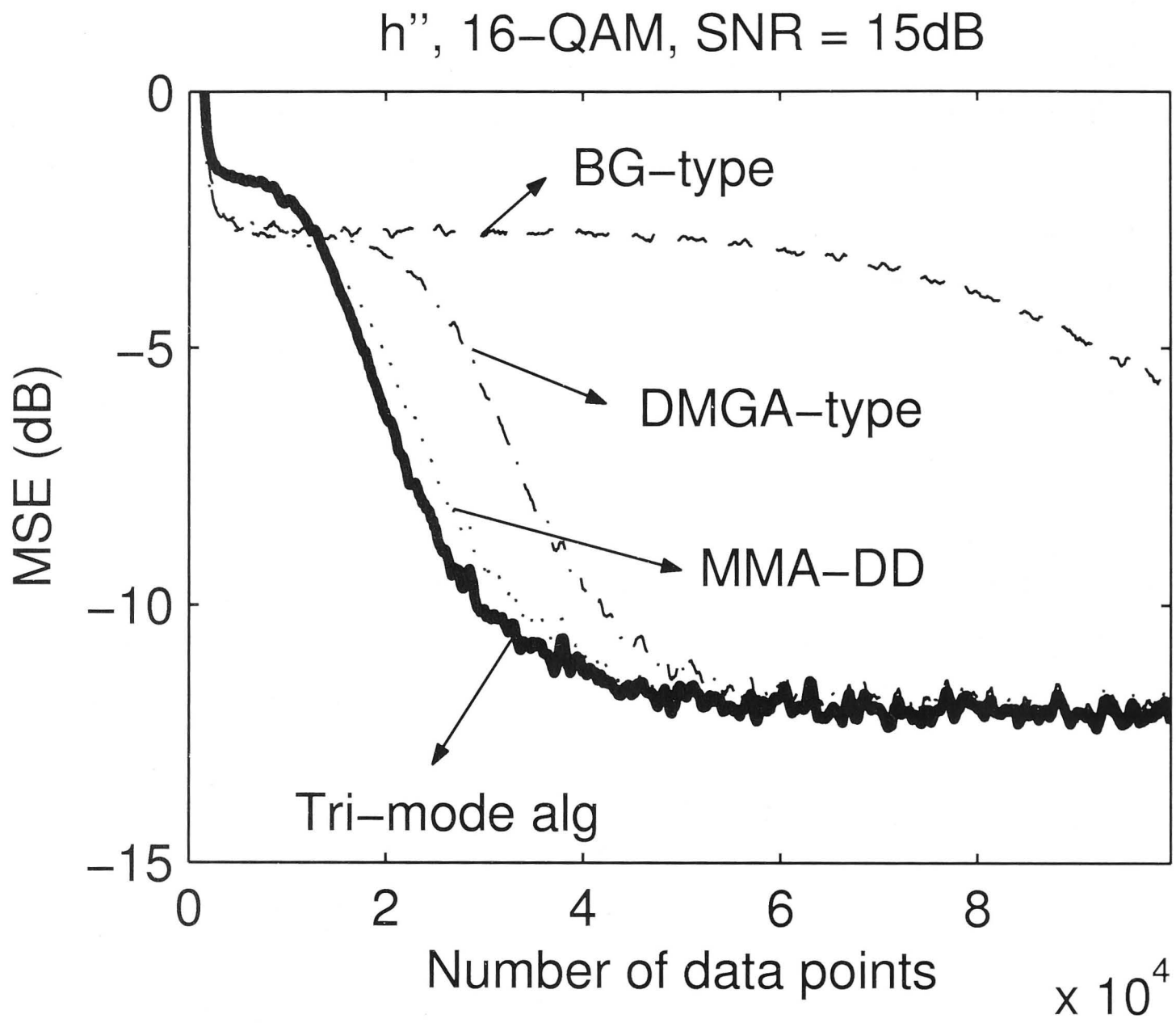

Figure 4.8: Comparing various algorithms using the channel $\mathbf{h}''$ under SNR level of 15dB. The modulation format is 16-QAM. Equalizer has 41 taps.

16-QAM, $\mathbf{h}'''$

For the difficult channel, $\mathbf{h}'''$, we performed our simulations on the channel with an SNR level of 17 dB for 16-QAM data using a total of 100,000 symbols for each of our runs. The step sizes for all algorithms are reduced by a factor of 5, i.e., $\mu = 10^{-5}$, except for the BG type algorithm where its step size is further reduced to $\mu = 2 \times 10^{-6}$ to ensure high rate of successful convergence. Here for the BG type algorithm, $k_1 = k_2 = 1$.

From Fig. 4.9, it is observed that our tri-mode algorithm once again achieves the fastest convergence. The more subtle advantage lies in its lower steady state MSE over all other algorithms. This is due to the new probabilistic technique that recognizes an open-eye condition under very noisy environments much earlier than did the hard-switching algorithm and the DMGA-type algorithm.

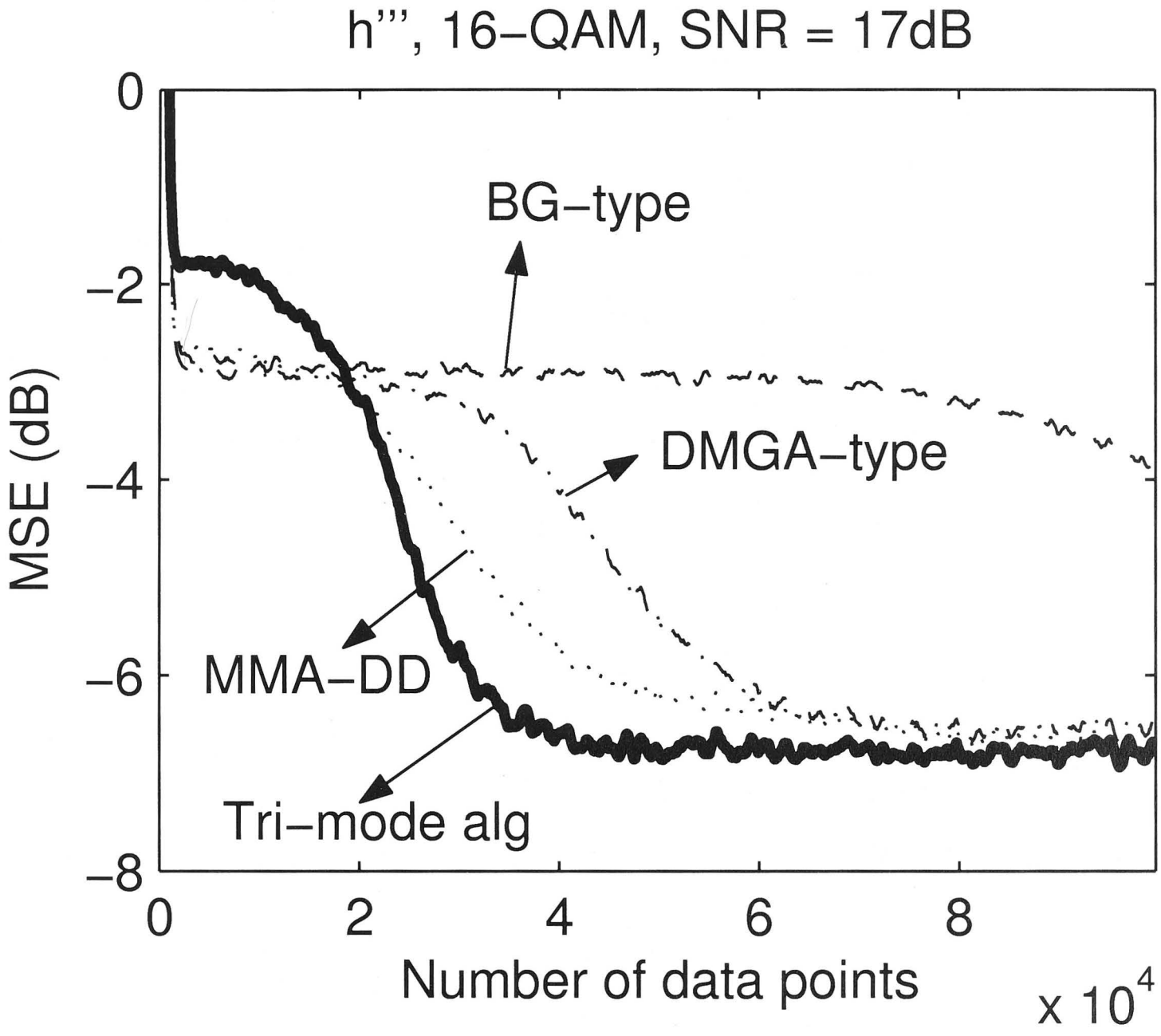

Figure 4.9: Comparing various algorithms using the channel $\mathbf{h}'''$ under SNR level of 17dB. The modulation format is 16-QAM. Equalizer has 41 taps.

32-QAM, $\mathbf{h}'$

For the easy channel, $\mathbf{h}'$, we performed our simulations on the channel with an SNR levels of 25 dB and 30 dB for 32-QAM data using a total of 50,000 symbols for each of our runs. The step sizes for all algorithms are reduced by a factor of 20, i.e., $\mu = 2.5 \times 10^{-6}$. As for γ, it is set to be at $\gamma = 50$ to compensate for the difference in the magnitudes of the error functions.

From Fig. 4.10, we draw several interesting conclusions. For (A) where SNR level is lower, it is observed that once again our tri-mode algorithm achieves the fastest convergence (only marginally slower than the BG-type algorithm) and the lowest steady state errors. The MMA-DD algorithm yields relatively higher MSE because it has failed to switch-over to the tracking algorithm because the DD MSE estimate never drops below the pre-defined MSE threshold. The BG-type and the DMGA-type switch-mode algorithms also display slightly marginally higher MSE because of they allow a significant percentage of the equalizer output to

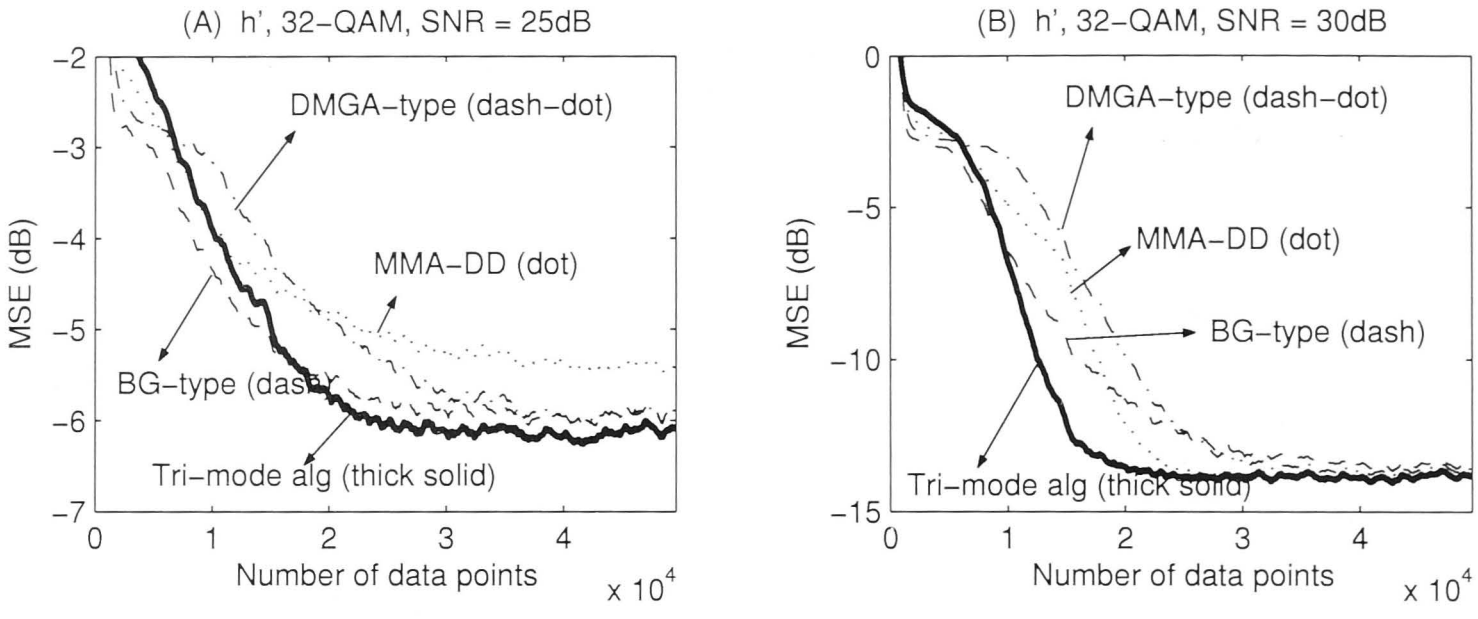

Figure 4.10: Comparing various algorithms using the channel $\mathbf{h}'$ under SNR levels of 25dB and 30dB. The modulation format is 32-QAM.

drive the equalizer parameters with the acquisition algorithm. For (B), i.e., the simulations performed at a higher SNR level, the tri-mode algorithm significantly outperforms all other algorithms. Only the MMA-DD algorithm achieves similar steady state MSE while the smooth-switching algorithms suffer slightly higher steady state MSE.

Only the tri-mode algorithm maintains its superiority in terms of convergence speed and low steady state errors for both relatively high and low SNR levels.

32-QAM, $\mathbf{h}'''$

For the difficult channel, $\mathbf{h}'''$, we performed our simulations on the channel with an SNR level of 30 dB for 32-QAM data using a total of 150,000 symbols for each of our runs. The step sizes for all algorithms are reduced by a factor of 20, i.e., $\mu = 2 \times 10^{-6}$, except for the BG-algorithm which has a step size of $\mu = 1.25 \times 10^{-6}$. As for γ, it is set to be at $\gamma = 50$ to compensate for the difference in the magnitudes of the error functions.

From Fig. 4.11, we notice that for this difficult channel, the convergence is still very rapid compared to the other algorithms. The DMGA-type algorithm, even though smooth, is slow and yields high steady state MSE as expected. The MMA-DD converges slower because of the DD MSE, which is difficult to accurately estimate, is used to govern the switching. Moreover, for larger constellation sizes, the excess MSE due to non-constant modulus input constellation may increase the DD MSE even though the channel eye is already open. As for the BG-type algorithm, it cannot employ a similar step size as the other algorithms otherwise it will not converge most of the time. That explains the slow rate of convergence

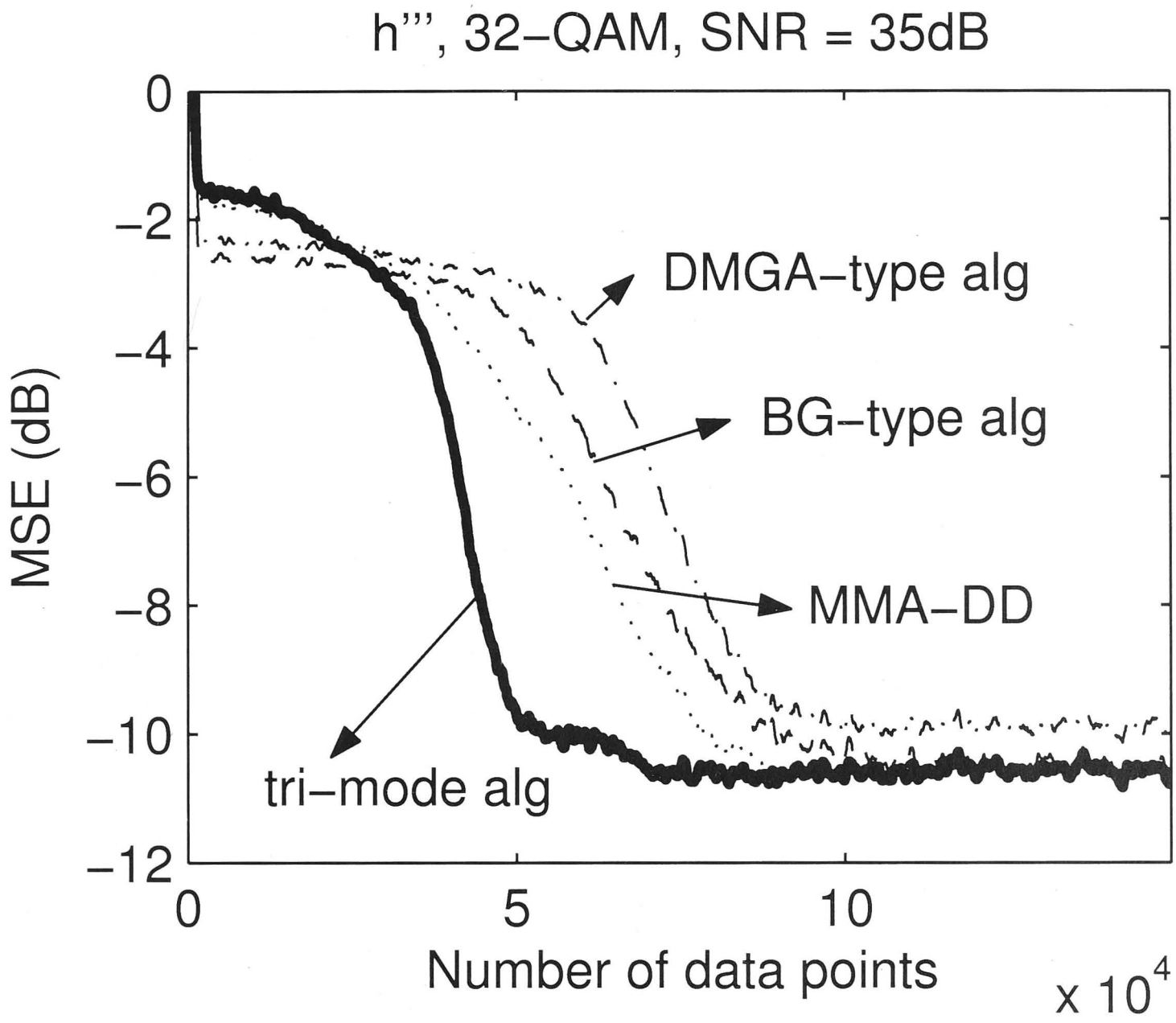

Figure 4.11: Comparing various algorithms using the channels $\mathbf{h}'''$ under SNR level of 30dB. The modulation format is 32-QAM.

which is due to a smaller step size assignment. Due to the smaller step size, it also achieves similar steady state errors as the tri-mode and MMA-DD algorithms.

4.5.2 Results for Joint Equalization and Phase Recovery

We compare the tri-mode algorithm with two phase-aware algorithms, namely the reduced constellation algorithm (RCA) and the multimodulus algorithm (MMA) [105, 152] in the presence of phase errors for the difficult channel, $\mathbf{h}'''$, only. Simulations were carried out for the noiseless channel and one with considerable noise using 16-QAM data. In addition, the channel is subjected to phase and frequency offsets in three different scenarios. In the first scenario there is no phase or frequency offsets. The second scenario has a 45° phase offset but no frequency offset. The third scenario has both 45° phase offset and $f_\Delta = 2 \times 10^{-4}$ of normalized frequency offset (4.10), i.e., a complete rotation of 2π in 5000 symbols. The MMA and the RCA are designed to switch to the DD algorithm once the DD MSE,

$\text{MSE}_{\text{DD}}(k)$, drops below 0.3960, as before in Section 4.5. Note that wehave not used an additional phase estimator for all the algorithms in order to assess their performances in the presence of phase and frequency offsets. In addition we also compared our algorithm with a trained linear equalizer.

The channel output sequence will be whitened by the MOE filter, $\mathbf{b}(k)$, prior to equalization as mentioned in Section 4.4.3, except for the case of the trained linear equalizer. We implemented an infinite impulse response (IIR) MOE filter which has four taps in the same manner as [75]. The taps of the MOE filter, $\mathbf{b}(k)$, are adapted as follows:

$$\mathbf{b}(k+1) = \mathbf{b}(k) - \mu_{\mathbf{b}} r(k) \mathbf{r}^*(k-1)$$

where $\mathbf{r}(k) = [r(k), r(k-1), r(k-2), r(k-3)]^T$ and $\mu_{\mathbf{b}}$ is the adaptation step size. The proceeding baud-rate equalizer, with the exception of the trained linear equalizer, has 24 taps and is initialized with a non-zero tap set at the 18^{th} position, i.e.,

$$\mathbf{w}(0) = [\underbrace{0 \cdots 0}_{1-17} \quad \underbrace{1}_{18} \quad \underbrace{0 \cdots 0}_{19-24}]^T.$$

More anti-causal taps than causal ones were assigned because the causal impulse response would be close to zero once the MOE filter has equalized the amplitude distortions. For the trained equalizer however, it has 40 taps and will be initialized with a center tap strategy. For our new algorithm, the parameters that control the sizes of $D(k)$ and $D_{\text{out}}(k)$ are $D(0) = 0.15$ and $\beta = 0.95$ according to (4.7), and $D_{\text{out}}(k) = 2.5$, $\forall k$. Its adaptation step sizes for different stages are shown in Table 4.2. As for the trained equalizer, we assigned a larger step size initially for fast convergence and then a smaller one to achieve a lower steady state MSE.

The results show fast convergence rate of the new algorithm in comparison to the CMA, MMA, and RCA schemes, in addition to automatic phase recovery and smooth transition between modes. In both noiseless and noisy cases without phase errors in Fig. 4.12-(A),(B), the tri-mode algorithm converges the fastest.

In Fig. 4.12 (C) and (D) the tri-mode algorithm corrects a 45^{o} phase offset and converges faster than the RCA but slower than the MMA. This is due to the lack of phase recovery in our algorithm as we only employ the MMA in a restricted region in $D_{\text{out}}(k)$.

In Fig. 4.12 (E) and (F), the RCA completely failed to acquire the frequency offset as depicted by the ripple-like MSE, where each notch corresponds to a mod-

ulo 90° phase offset, once in 1250 symbols. The tri-mode algorithm and the MMA perform better in the sense their MSE have descending staircase-like behaviors. This means at every modulo 90° phase offset, more simulation runs have successfully converged to the DD mode. The tri-mode algorithm reports more successful convergence compared to the hard-switching technique employed by the MMA-DD algorithm as depicted by the lower steady state MSE of the tri-mode algorithm.

4.6 Conclusions

The new tri-mode algorithm is the fastest converging algorithm among the hard-switching and smooth-switching algorithms according to our simulations. It is attributed to both the novel smooth-switching technique as well as the novel design of the tri-mode algorithm that assigns algorithms with different capabilities to different reliable regions in the constellation space. Apart from that, it also achieves the lowest steady state errors for a wide range of easy to difficult channels, different SNR levels, as well as constellation sizes. From the results obtained via simulations, we conclude that in general, smooth-switching algorithms outperform hard-switching algorithms under high noise conditions, but the opposite is true when the noise level is relatively low. The tri-mode algorithm however performs just as good under low and high noise conditions. The complexity of the algorithm is extremely low as the smooth-switching technique exploits first order estimates which are accurate and simple. In addition, it can also perform joint blind equalization and phase recovery in the presence of realistic demodulation phase errors without having to rely on a separate phase compensator.

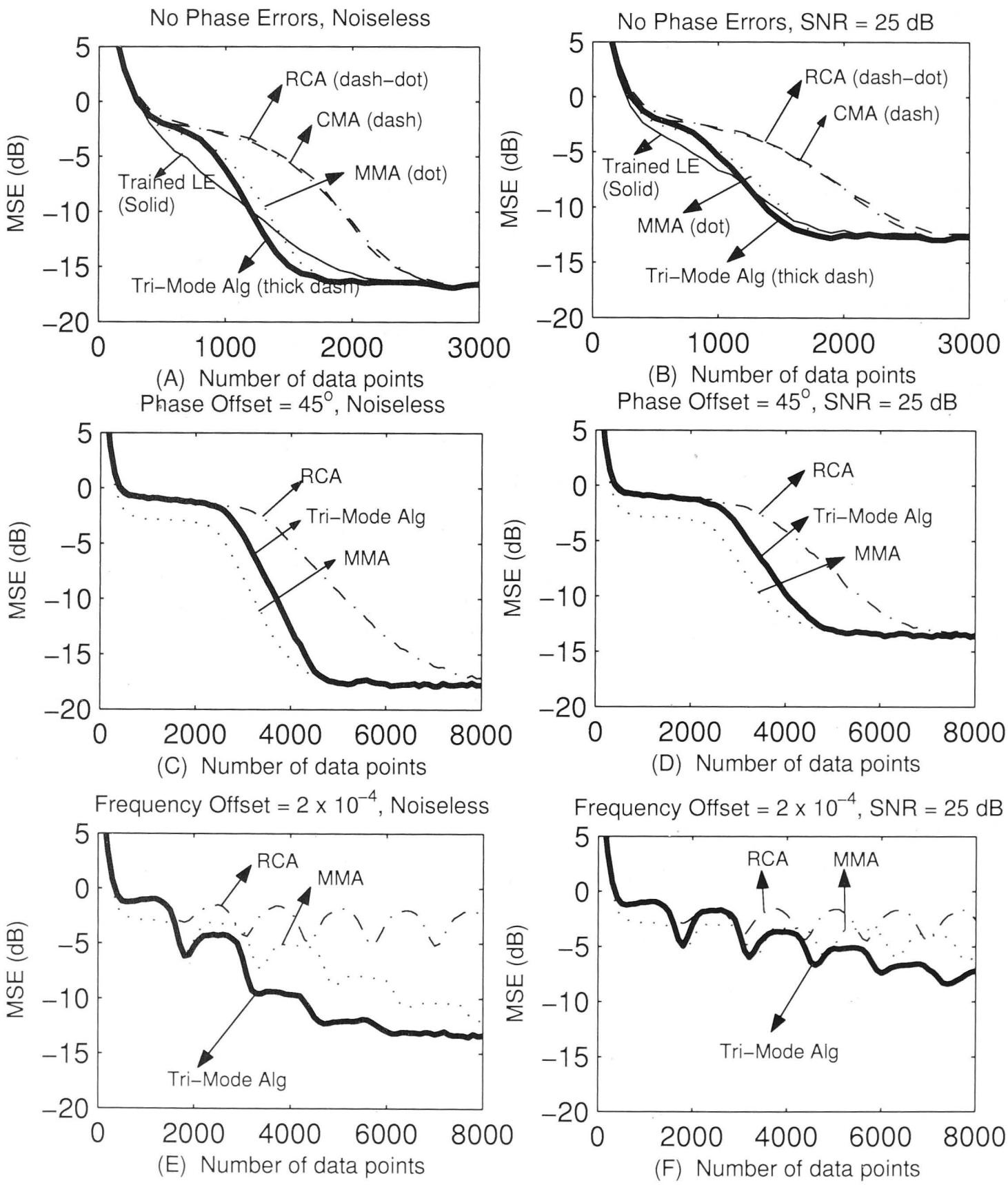

Figure 4.12: Comparing various algorithms in different equalizer setups.

Chapter 5

Fast Convergence Switch-Mode DFE Schemes

5.1 Introduction

Blind equalization compensates for channels distortions without relying on a training sequence. Relative to adaptation using training sequences, blind equalizers tend to exhibit slower convergence, higher steady state errors and possibly ill-convergence. In the domain of linear equalization the blind algorithm design and analysis problem has been exhaustively studied for several decades. In the domain of decision feedback equalization, it was only in 1998 that Labat, Macchi and Laot found a compelling realization of an effective blind decision feedback equalizer (DFE) [75]. In Sections 5.2 and 5.3 of this chapter, we build on their work to develop alternative designs which exhibit significant performance advantages.

In the innovative blind scheme presented in [75], the adaptation is broken up into two modes, an acquisition mode and a tracking mode. Not only is the adaptation algorithm switched between modes but also the actual equalization filtering structure is switched. In the acquisition mode a recursive linear structure is employed and later, once the eye diagram has opened sufficiently, the DFE is employed in the tracking mode. The algorithm in the acquisition mode uses a combination of a constrained convex minimum energy cost and the constant modulus algorithm (CMA), whilst in the tracking mode a decision directed (DD) algorithm is used. These choices are not arbitrary but are guided by the design objective to have the blind DFE converge as close as possible to the optimum MMSE solution.

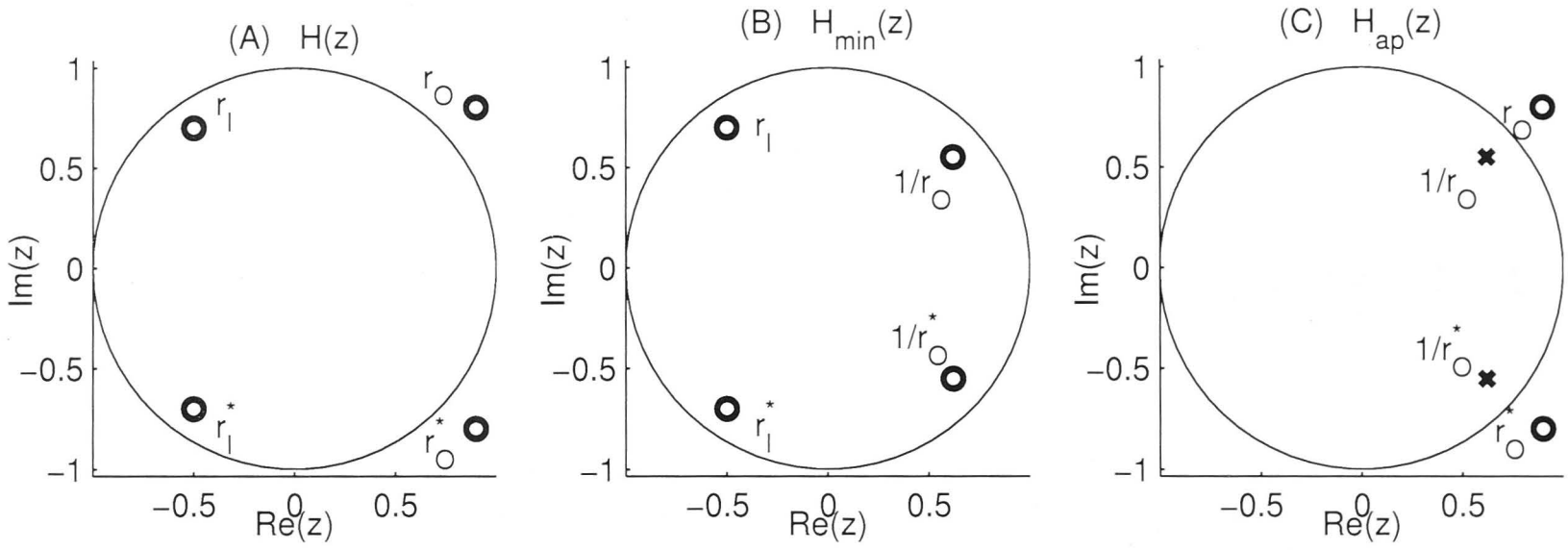

Figure 5.1: Decomposition of a transversal non-minimum phase channel via the channel decomposition property with $H(z) = \mathcal{G}H_{\min}(z)H_{\text{ap}}(z)$, assuming the complex gain, $\mathcal{G}$, is unity.

5.1.1 Channel Decomposition Property

The novel approach taken by Macchi, Labat and their colleagues in [22, 75] is to equalize a "decomposed" channel via the channel decomposition property to achieve rapid and effective equalization. According to the channel decomposition property [108], an arbitrary non-minimum phase channel $H(z)$ can be represented by the convolution of an equivalent minimum phase system, $H_{\min}(z)$, and an equivalent all-pass system, $H_{\text{ap}}(z)$, multiplied by a complex gain $\mathcal{G}$ in the following manner:

$$H(z) = \mathcal{G}H_{\min}(z)H_{\text{ap}}(z). \tag{5.1}$$

Consider the decomposition of a *transversal* non-minimum phase channel of the form

$$H(z) = \mathcal{G}\prod_{i=1}^{N_1}\left[1 - r_{\text{I},i}z^{-1}\right]\prod_{j=1}^{N_2}\left[r_{\text{O},j}^{-1} - z^{-1}\right] \tag{5.2}$$

where r_{I} and r_{O} are zeros of the system that are inside and outside the z-unit circle, (U), respectively. $H_{\min}(z)$ is the equivalent minimum phase system that contains r_{I} and r_{O}^{-*}, i.e., the zeros that are the conjugate reciprocals of r_{O}. $H_{\text{ap}}(z)$ is an all-pass system which comprises of zeros r_{O}, and poles r_{O}^{-*} to cancel the reflected zeros in $H_{\min}(z)$. Thus the decomposed channel consists of the following components

$$H_{\min}(z) = \prod_{i=1}^{N_1}\left[1 - r_{\text{I},i}z^{-1}\right]\prod_{j=1}^{N_2}\left[1 - (r_{\text{O},j}^{*}z)^{-1}\right] \tag{5.3a}$$

$$H_{\text{ap}}(z) = \mathcal{G} \prod_{j=1}^{N_2} \frac{r_{\text{O},j}^{-*} - z^{-1}}{1 - r_{\text{O},j}^{-*} z^{-1}} \tag{5.3b}$$

The decomposition process for an example with a pair of minimum phase zeros and a pair of maximum phase zeros is shown in Fig. 5.1. Its zeros and poles are respectively represented by circles, 'o', and crosses, 'x'. $H(z)$ is shown in Fig. 5.1-A, while its decomposed components, $H_{\text{min}}(z)$ and $H_{\text{ap}}(z)$, are shown in Fig. 5.1-B,C, respectively.

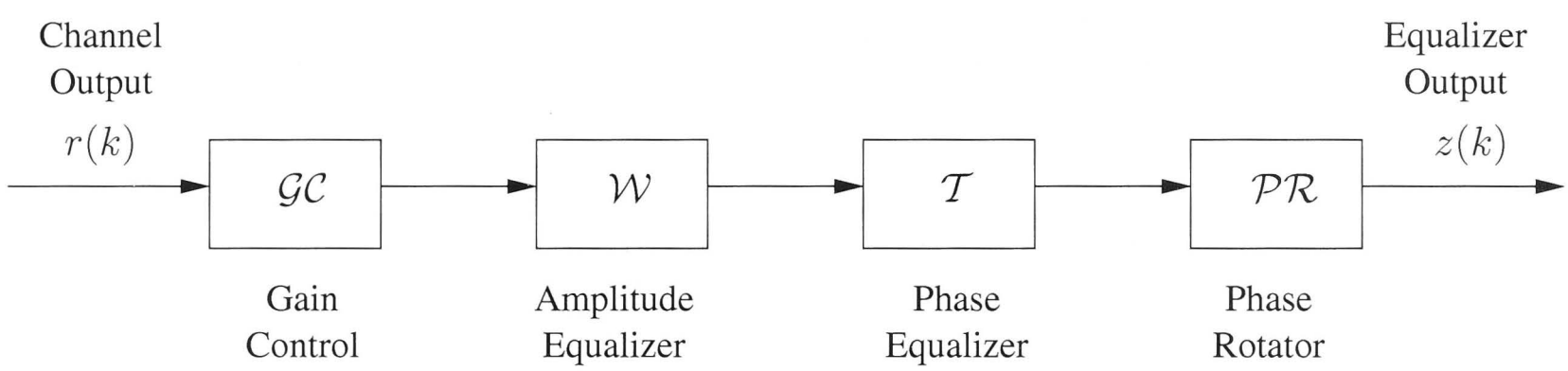

Figure 5.2: Novel blind linear equalizer setup [75] as a cascade of four devices used to acquire the "decomposed" channel.

The authors of [22, 73, 75] exploited this decomposition property to isolate $H_{\text{min}}(z)$ from $H_{\text{ap}}(z)$, where the equivalent minimum phase can be quickly equalized by a whitening filter that is based on second order statistics. It is also known as the amplitude equalizer in [22]. The output of the whitening filter which is now uncorrelated appears to have been produced by convolving the channel input with a new channel $\mathcal{G}H_{\text{ap}}(z)$ since the whitening filter has equalized $H_{\text{min}}(z)$. A transversal phase equalizer that employs the CMA will be used subsequently to equalize the remaining phase distortions due to $H_{\text{ap}}(z)$. The CMA, which is usually slower than the whitening algorithm, now has the simpler task of equalizing the all-pass channel without amplitude distortions [14,22]. Compensation for the complex gain, $\mathcal{G}$, is carried output by splitting the task into compensation of real and complex components, respectively, using a real gain control and complex phase rotator. It is suggested in [75] that the real gain control be placed upstream of the equalizer while the phase rotator downstream, as shown in Fig. 5.2 during the acquisition of the channel. This equalizer setup has been shown to speed up the convergence of the Godard equalizer significantly. In fact, at high signal-to-noise ratios (SNR), this cascaded equalizer structure has been shown to emulate the optimal linear minimum MSE equalizer and efficiently equalize the decomposed channel [75], where $\mathcal{W}$ will converge to $H_{\text{min}}^{-1}(z)$, $\mathcal{T}$ to $H_{\text{ap}}^{-1}(z)$, and $\mathcal{GC}$ and $\mathcal{PR}$ will jointly compensate for the complex gain, $\mathcal{G}$. Simulations in [75] show that its performance is equally impressive at low SNR levels, outperforming even the trained DFE in some cases.

Another attractive feature of the equalization scheme of [75] lies in the flexibility of the linear equalizer to switch to a DFE when the bit-error-rate is sufficiently low. This is accomplished by simply relocating $\mathcal{W}$ downstream, i.e., behind $\mathcal{T}$ and $\mathcal{PR}$, and incorporating a quantizer to facilitate decision feedback.

5.2 Design Objectives

For the equalizer scheme of sections 5.3 and 5.4

In our design we seek to improve on the design in [75], while retaining the simplicity of the equalization scheme, with the following principle goals in mind:

A) **Non-Recursive Linear Acquisition Filter** – In [75], the recursive form of the linear filter in the acquisition structure is used to permit direct transfer of the filter tap values to the DFE feedback filter in the tracking structure. Our design goal is to retain the ability of direct transfer of filter taps value but avoid some of the well-known problems with the adaptation of the recursive form (where the gradient can only be approximated, the dynamic range at the output may be large and the stability needs to be monitored). We show that this is indeed possible using a non-recursive linear filter provided we modify the DFE feedback filter appropriately.

B) **Reduced Switching Transients** – Switching structures and switching algorithms lead to transients in signals which disrupts convergence. The design goal here is to develop a strategy which provides for "smooth switching" and thereby significantly improved convergence. We achieve this on the structural side by employing a parallel adaptation strategy where the acquisition structure and the tracking structure are jointly adapted. Further on the algorithm side, rather than abruptly switch algorithms, we employ a technique to smoothly combine the acquisition and tracking algorithms into a single algorithm. In this way it is not necessary to make a distinction between acquisition and tracking and therefore the new scheme needs not be considered a dual-mode equalizer. This eliminates the requirement for the user to pre-specify certain parameter values that strictly depend on the channel to determine the exact and appropriate sampling period to perform the switch-over.

To accomplish our first goal, we develop new equalizer structures in the acquisition and tracking modes as described in Section 5.3. As the underlying concepts of our alternative scheme are identical to that of [75], simulation results would

subsequently show almost identical performance between both schemes. Our second goal, which is more important, features a novel strategy that enhances these fast-convergence DFE schemes (that include the scheme in [75] and our alternative scheme in Section 5.3) where they must undergo a switch in both their adaptation algorithms and filtering structures. This new switching strategy that results in a single-mode DFE by combining the respective equalizers in both modes in a parallel fashion is described in Section 5.4.

5.2.1 System Model

Consider a typical baseband symbol-rate blind equalizer whose main objective is to recover a corrupted version of the transmitted signal, $\{a(k)\}$, based only on the observable received signals, $\{r(k)\}$. Let the linear distortive channel and additive white Gaussian noise be denoted as $\{h_i\}$ and $n(k)$, respectively. Then the objective of equalization is to yield an output, $z(k)$, that well estimates the transmitted signal such that $z(k) \approx a(k - \triangle)$, where $\triangle$ is some time delay. We will treat the channel $\{h_i\}$, via the channel decomposition property [108], as a cascade of an equivalent minimum phase channel and an all-pass channel for the design of our linear equalizer in the starting mode in an identical fashion to [22, 75]. Thus, we separate the principle tasks of our equalizer into compensation for amplitude distortions only (due to the equivalent minimum phase channel), and compensation for the remaining phase distortions [73, 75].

The equalization scheme involves two distinct modes – the linear acquisition mode and the tracking DFE mode. In order to distinguish the filter parameters and signals of these two modes, we append the superscripts "(1)" and "(2)" to all filter parameters and signals in the "first" acquisition mode and "second" tracking mode, respectively.

5.3 Development of Alternative Fast-Convergence DFE

Following [75], a possible equalizer setup in the acquisition mode is the cascade of four filters, namely, a real gain control filter ($\mathcal{GC}$), a whitening filter ($\mathcal{W}$), a phase equalizer ($\mathcal{T}$), and a complex phase rotator ($\mathcal{PR}$). Thus in acquisition mode the equalizer is linear and decision feedback is not employed. This overall linear equalizer consisting of filters $\mathcal{GC}$, $\mathcal{W}$, $\mathcal{T}$, and $\mathcal{PR}$, is shown in Fig. 5.3.

In contrast to [75] for the reasons listed in the introduction, we propose to use a

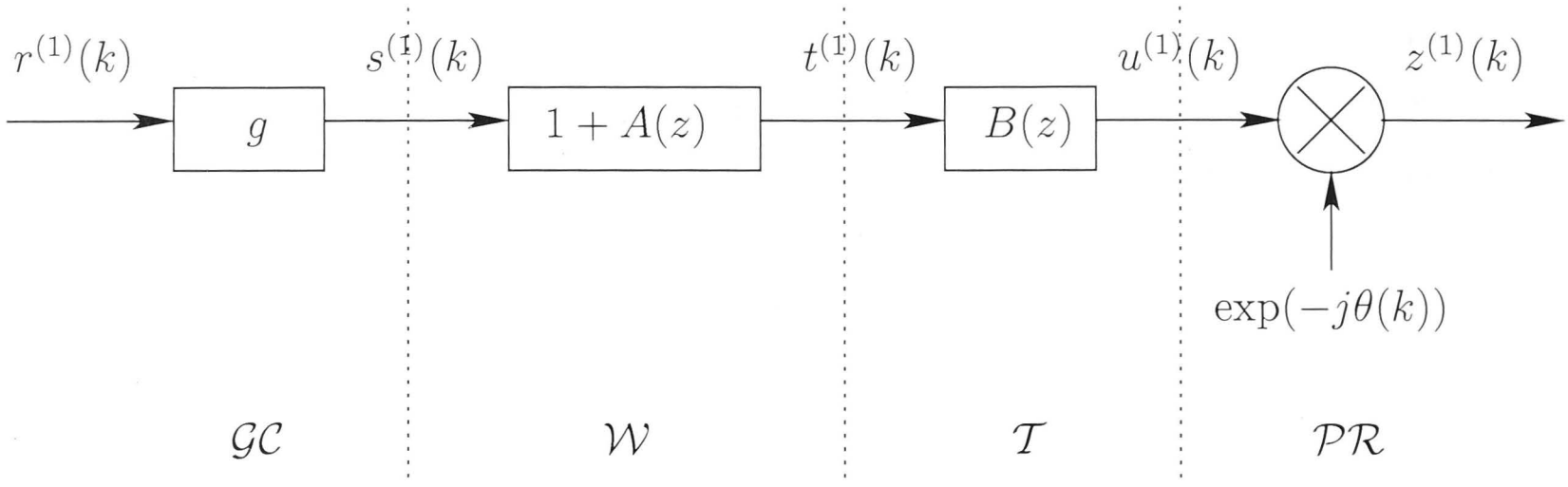

Figure 5.3: Linear equalizer of acquisition mode of the alternative DFE scheme to that of [75], featuring a non-recursive $\mathcal{W}$.

non-recursive whitening filter in the place of the originally proposed recursive filter. This requires a modification to the whitening algorithm. The whitening algorithm, also known as the minimum output energy (MOE) algorithm, minimizes the cost function $\mathrm{E}\{|t^{(1)}(k)|^2\}$, where $t^{(1)}(k)$ is the output of $\mathcal{W}$. In order to avoid the convergence of the equalizer to the trivial solution, it is necessary to impose a *linear* constraint upon the equalizer taps so that convexity of the cost function is not lost. Following [142], the simplest and recommended linear constraint fixes the leading tap, $a_0(k)$, to be unity for all k so that the transfer function of $\mathcal{W}$ becomes $1 + A(z)$ as shown in Fig. 5.3. Define $\mathbf{a}^{(1)}(k) = [a_1(k), a_2(k), \cdots, a_N(k)]^T$ as the weight vector of $\mathcal{W}$ of length N that *excludes* the leading tap, $a_0(k)$. Then under the usual stochastic gradient adaptation, $\mathbf{a}^{(1)}(k)$ is updated according to

$$\mathbf{a}^{(1)}(k+1) = \mathbf{a}^{(1)}(k) - \mu_{\mathbf{a}^{(1)}} \epsilon_{\mathrm{MOE}} \mathbf{s}^{(1)}(k-1) \tag{5.4}$$

where the MOE error takes the form

$$\epsilon_{\mathrm{MOE}} = t^{(1)}(k) \tag{5.5}$$

and $\mathbf{a}^{(1)}(0) = [0, \cdots, 0]^T$, $\mu_{\mathbf{a}^{(1)}}$ is the adaptation step size, $\mathbf{s}^{(1)}(k-1) = [s^{(1)}(k-1), s^{(1)}(k-2), \cdots, s^{(1)}(k-N)]^T$ and * is the complex conjugate operator. This algorithm attempts to minimize the cost $J_{\mathrm{MOE}}(\mathbf{a}^{(1)}) = \mathrm{E}\{|t^{(1)}(k)|^2\}$ subject to the first tap being fixed at unity, i.e., $a_0(k) = 1$.

As for $\mathcal{GC}$, $\mathcal{T}$ and $\mathcal{PR}$, our adaptation algorithms are no different from that of [75]. Briefly, $\mathcal{GC}$ is a real filter that fixes the average power level of the samples at $t^{(1)}(k)$ at a particular value, while $\mathcal{PR}$ is the complex phase rotator that compensates for any demodulation phase errors. As for the phase equalizer, $\mathcal{T}$, the CMA is the preferred algorithm to achieve the desired removal of the residual ISI.

The update equations of $\mathcal{GC}$, $\mathcal{T}$ and $\mathcal{PR}$ are as follows:

• For $\mathcal{GC}$:

$$G^{(1)}(k+1) = G^{(1)}(k) + \mu_{G^{(1)}}\left(1 - |u^{(1)}(k)|^2\right) \tag{5.6a}$$

$$g^{(1)}(k+1) = \sqrt{|G^{(1)}(k+1)|} \tag{5.6b}$$

where $G^{(1)}(0) = 1$, and $\mu_{G^{(1)}}$ is the positive step size that governs the rate of adaptation of $\mathcal{GC}$. The output of $\mathcal{GC}$ is $s^{(1)}(k) = g^{(1)}(k)r^{(1)}(k)$.

• For $\mathcal{T}$:

$$\mathbf{b}^{(1)}(k+1) = \mathbf{b}^{(1)}(k) - \mu_{\mathbf{b}^{(1)}}\epsilon_{\mathrm{CMA}}(k)\mathbf{t}^{(1)^*}(k) \tag{5.7}$$

where the CMA error takes the form

$$\epsilon_{\mathrm{CMA}}(k) = u^{(1)}(k)\left[|u^{(1)}(k)|^2 - \gamma_C^2\right] \tag{5.8}$$

and $\mathbf{b}^{(1)}(0) = [0, 0, \cdots, 0, 1, 0, \cdots, 0]^T$ and $\mu_{\mathbf{b}^{(1)}}$ is a small positive step size, $\mathbf{t}^{(1)}(k) = [t^{(1)}(k), t^{(1)}(k-1), \cdots, t^{(1)}(k-M+1)]^T$, where M is the equalizer length and $\gamma_C^2 \triangleq \mathrm{E}\{|a(k)|^4\}/\mathrm{E}\{|a(k)|^2\}$ is the so-called dispersion constant of the CMA, where $a(k)$ is the transmitted data symbol [49].

• For $\mathcal{PR}$:

$$\theta^{(1)}(k+1) = \theta^{(1)}(k) + \mu_{\theta^{(1)}}\epsilon_\Phi(k) \tag{5.9a}$$

$$\epsilon_\Phi(k) = \mathrm{Im}\left[\mathrm{Q}(z^{(1)}(k))z^{(1)*}(k)\right] \tag{5.9b}$$

$$\begin{aligned}\text{or}\quad \epsilon_\Phi(k) = {} & \mathrm{Im}\left[\mathrm{Q}(z^{(1)}(k))[\mathrm{Q}(z^{(1)}(k)) - z^{(1)*}(k)]\right] \\ & + \beta\sum_{j=1}^{k}\mathrm{Im}\left[\mathrm{Q}(z^{(1)}(j))[\mathrm{Q}(z^{(1)}(j)) - z^{(1)*}(j)]\right]\end{aligned} \tag{5.9c}$$

where $\theta^{(1)}(0) = 0$ and $\mu_{\theta^{(1)}}$ is a small positive step size, (5.9b) and (5.9c) correspond to the first [49] and second order [75] phase tracking loop, respectively, and β is an appropriate positive parameter.

In the tracking mode, the position of the whitening filter, $\mathcal{W}$, is interchanged with the positions of $\mathcal{T}$ and $\mathcal{PR}$, so that $\mathcal{W}$ is transformed to a non-linear filter and placed downstream. Unfortunately, unlike the recursive structure in [75], our *non-recursive* $\mathcal{W}$ is not immediately ready to facilitate decision feedback. Our subsequent objectives therefore are to firstly transform the non-recursive filter block, $\mathcal{W}$, into a new but equivalent filter block that is equipped with a feedback path;

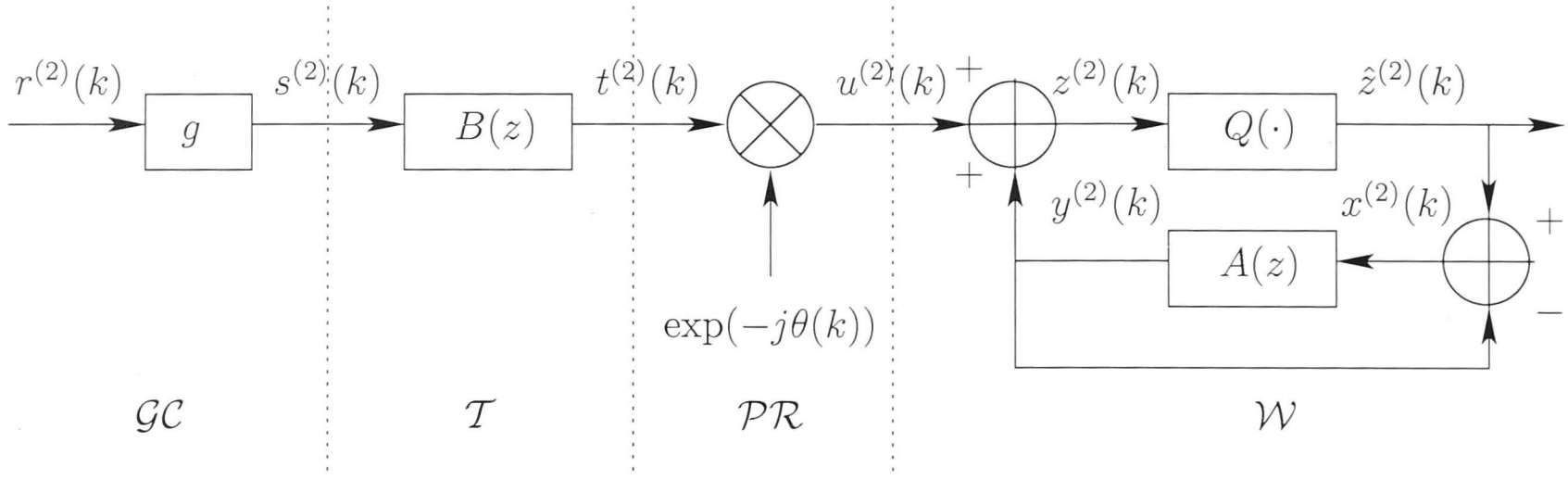

Figure 5.4: The non-linear equalizer (DFE) of the tracking mode.

and secondly, to retain the ability for the direct transfer of the parameters of $\mathcal{W}$ between the acquisition and the tracking modes. The objectives can be jointly achieved by performing the following block transformation:

$$1 + A(z) = \frac{1 + A(z)}{1 + A(z) - A(z)} = \frac{1}{1 - \frac{A(z)}{1+A(z)}}. \tag{5.10}$$

The term on the right hand side, which is equivalent to the non-recursive transfer function of $\mathcal{W}$ of the acquisition mode, is now equipped with a feedback path whose transfer function is $-\frac{A(z)}{1+A(z)}$. The transformed block $\mathcal{W}$ in terms of $A(z)$ only of the DFE is illustrated in Fig. 5.4.

The adaptation of the real gain control $\mathcal{GC}$ is no longer required in the tracking mode. As for $\mathcal{T}$, $\mathcal{W}$, and $\mathcal{PR}$, they are jointly adapted by minimizing the decision directed MSE criterion

$$J_{\mathrm{DD}}(\mathbf{a^{(2)}}, \mathbf{b^{(2)}}, \theta) = \mathrm{E}\left\{\left|z^{(2)}(k) - \mathrm{Q}\big(z^{(2)}(k)\big)\right|^2\right\} \tag{5.11}$$

where $\mathbf{a}^{(2)}$ and $\mathbf{b}^{(2)}$ are the weight vectors of $A(z)$ and $B(z)$, respectively, as in Fig. 5.4. They are updated as follows:

$$\mathbf{a}^{(2)}(k+1) = \mathbf{a}^{(2)}(k) - \mu_{\mathbf{a}^{(2)}} \epsilon_{\mathrm{DD}}(k) \mathbf{x}^{(2)^*}(k-1) \tag{5.12}$$

$$\mathbf{b}^{(2)}(k+1) = \mathbf{b}^{(2)}(k) - \mu_{\mathbf{b}^{(2)}} \epsilon_{\mathrm{DD}}(k) \mathbf{t}^{(2)^*}(k) \tag{5.13}$$

where $\epsilon_{\mathrm{DD}}(k) = z^{(2)}(k) - \mathrm{Q}(z^{(2)}(k))$ and

$$\begin{aligned}\mathbf{x}^{(2)}(k) = &[\hat{z}^{(2)}(k-1), \hat{z}^{(2)}(k-2), \cdots, \hat{z}^{(2)}(k-N)]^T \\ &- [y^{(2)}(k-1), y^{(2)}(k-2), \cdots, y^{(2)}(k-N)]^T\end{aligned} \tag{5.14a}$$

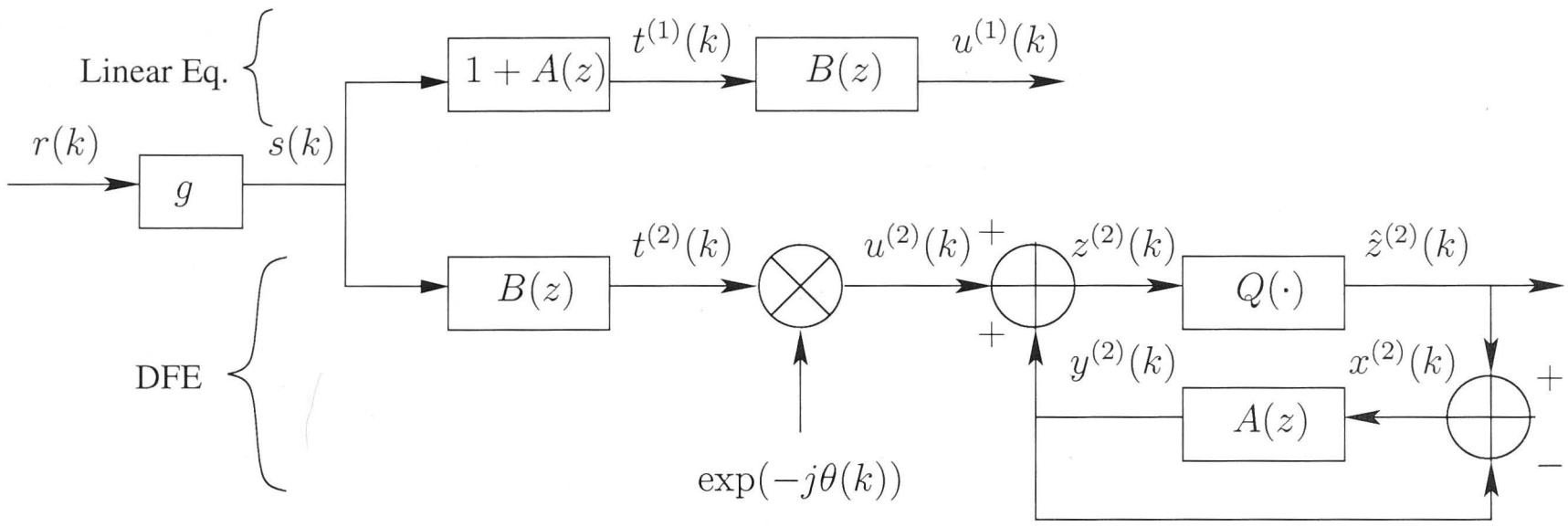

Figure 5.5: New DFE scheme under parallel adaptation. The taps are adapted and shared among filters from "top" and "bottom" paths of the new DFE.

$$\mathbf{t}^{(2)}(k) = [t^{(2)}(k), t^{(2)}(k-1), \cdots, t^{(2)}(k-M+1)]^T. \tag{5.14b}$$

The phase rotator is adapted as it would in the acquisition mode according to (5.9). Suppose switching occurred at $k = k_0$, then their initial values are [75]: $\mathbf{a}^{(2)}(k_0-1) = \mathbf{a}^{(1)}(k_0-1)$ with $\mathbf{x}^{(2)}(k_0-1) = \mathbf{0}$, $\mathbf{b}^{(2)}(k_0-1) = \mathbf{b}^{(1)}(k_0-1)$ and $\mathbf{t}^{(2)}(k_0-1) = \mathbf{0}$, and lastly $g(k) = g(k_0-1), \forall k > k_0 - 1$.

5.4 Parallel Adaptation Strategy for Dual Mode Equalization Schemes

Switching between the starting mode and the tracking mode involves both a rearrangement of the equalizer structure as well as a change between the acquisition and tracking algorithms. The switching often results in a slower rate of convergence because of the disruptions in the filtering structure and the algorithms employed. In addition, there is also a disruption in the states of the filters which adversely affect the algorithms and the output signals over several sample periods until they are flushed from the regressor vectors. Consequently, the output signals are more error-prone and the DFE, which is sensitive to incorrect decisions, may therefore exhibit pathological behavior [39, 71].

In the light of these problems, we propose a novel *parallel adaptation strategy* to ameliorate the transients by employing parallel adaptation of the linear equalizer of the acquisition mode and the DFE of the tracking mode such that only one set of filter parameters $\{A(z), B(z)\}$ is adapted and shared by the linear and decision feedback equalizers, as shown in Fig. 5.5. This means it is possible to always obtain the equalizer output from the output of the DFE, $z^{(2)}(k)$, as depicted in

Fig. 5.5. Thus, this strategy will transform the original dual-mode equalization scheme in [75] to a single-mode DFE scheme whose initial acquisition is *assisted* by a linear acquisition equalizer. The new update equations of the DFE filter weights, $\mathbf{a}(k)$ and $\mathbf{b}(k)$, under parallel adaptation are

$$\begin{aligned}\mathbf{a}(k+1) = \mathbf{a}(k) - \mu_{\mathbf{a}}\big\{&\alpha_1(k)\epsilon_{\mathrm{MOE}}\mathbf{s}^{(\mathbf{1})^*}(k) \\ &+ \alpha_2(k)\gamma_a\epsilon_{\mathrm{DD}}(k)\mathbf{x}^{(\mathbf{2})^*}(k)\big\} \qquad (5.15)\\ \mathbf{b}(k+1) = \mathbf{b}(k) - \mu_{\mathbf{b}}\big\{&\beta_1(k)\epsilon_{\mathrm{CMA}}(k)\mathbf{t}^{(\mathbf{1})^*}(k) \\ &+ \beta_2(k)\gamma_b\epsilon_{\mathrm{DD}}(k)\mathbf{t}^{(\mathbf{2})^*}(k)\big\} \qquad (5.16)\end{aligned}$$

where $\mathbf{a}(0) = [0, \cdots, 0]^T$, $\mathbf{b}(0) = [0, 0, \cdots, 0, 1, 0, \cdots, 0]^T$, $\epsilon_{\mathrm{DD}}(k) = z^{(2)}(k) - Q(z^{(2)}(k))$, $\alpha_1(k)$, $\alpha_2(k)$, $\beta_1(k)$, $\beta_2(k)$ are data dependent parameters in a manner that is described below, γ_a and γ_b are parameters assigned to compensate for the difference in the expected values of the respective error functions. $\mathbf{s}^{(1)}(k)$ and $\mathbf{t}^{(1)}(k)$ are regressor vectors of $1 + A(z)$ and $B(z)$ of length N and M, respectively, of the linear equalizer along the top path of the new DFE. A phase rotator in the linear equalizer is not required for this setup. The signals from the top path are obtained by employing the shared DFE taps such that

$$t^{(1)}(k) = s(k) + \sum_{j=1}^{N} a_j(k)s(k-j) \qquad (5.17a)$$

$$u^{(1)}(k) = \sum_{j=0}^{M} b_j(k)t^{(1)}(k-j). \qquad (5.17b)$$

where $a_j(k)$, $b_j(k)$ are the j^{th} taps of $A(z)$ and $B(z)$ of the new DFE, respectively. The gain control and the phase rotator are adapted as in (5.6) and (5.9). Therefore, due to the parallel adaptation strategy, the outputs of the "top" linear equalizer and the "bottom" DFE will be approximately equal, i.e.,

$$u^{(1)}(k)\exp(-j\theta(k)) \approx z^{(2)}(k). \qquad (5.18)$$

See Appendix-B for detailed explanation.

The choice of $\alpha_1(k)$, $\alpha_2(k)$, $\beta_1(k)$ and $\beta_2(k)$ is paramount to the success of our parallel adaptation strategy in terms of convergence speed and steady state errors, in addition to guaranteeing a smooth transition between "modes" which ultimately affects the convergence speed. In fact, certain soft switching techniques that were previously proposed for dual-mode algorithms are suitable for combining the ac-

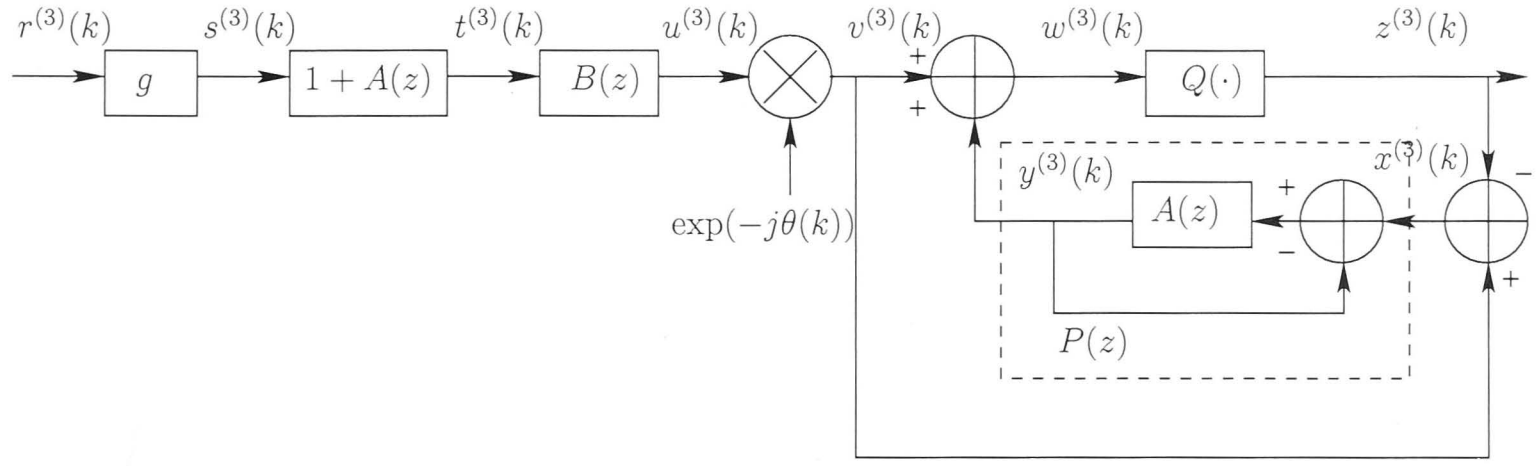

Figure 5.6: The new alternative fast-convergence predictive DFE.

quisition and tracking *increment* vectors[1] in (5.15) and (5.16) where intermediate values between 0 and 1 can be assigned to $\alpha_1(k)$, $\alpha_2(k)$, $\beta_1(k)$ and $\beta_2(k)$ according to a reliability measure of the equalizer output [82]. Alternative simpler techniques (albeit with poorer performance) that include the 'Stop-And-Go' algorithm [116], dual-mode type algorithm [145], as well as the simple Benveniste-Goursat algorithm [15] can also be employed. As a simple illustration, we will describe the Benveniste-Goursat soft-switching technique which assigns

$$\alpha_1(k) = \beta_1(k) = c_1|\epsilon_{\mathrm{DD}}(k)| \tag{5.19a}$$

$$\text{and} \quad \alpha_2(k) = \beta_2(k) = c_2 \tag{5.19b}$$

where c_1 and c_2 are user defined positive constants [15]. The new update equations of (5.15) and (5.16) under parallel adaptation strategy using the Benveniste-Goursat type parameters become

$$\begin{aligned}\mathbf{a}(k+1) = \mathbf{a}(k) - \mu_{\mathbf{a}}\big\{ & c_1|\epsilon_{\mathrm{DD}}(k)|\epsilon_{\mathrm{MOE}}\mathbf{s}^{(\mathbf{1})^*}(k) + \\ & c_2\gamma_a\epsilon_{\mathrm{DD}}(k)\mathbf{x}^{(\mathbf{2})^*}(k)\big\}\end{aligned} \tag{5.20}$$

$$\begin{aligned}\mathbf{b}(k+1) = \mathbf{b}(k) - \mu_{\mathbf{b}}\big\{ & c_1|\epsilon_{\mathrm{DD}}(k)|\epsilon_{\mathrm{CMA}}(k)\mathbf{t}^{(\mathbf{1})^*}(k) + \\ & c_2\gamma_b\epsilon_{\mathrm{DD}}(k)\mathbf{t}^{(\mathbf{2})^*}(k)\big\}.\end{aligned} \tag{5.21}$$

5.5 Alternative Fast-Convergence Predictive DFE Scheme

The predictive DFE (P-DFE) [10] is an alternative DFE structure that is equivalent to the conventional DFE under the condition that the forward filter has an infinite

[1]The increment vector of the stochastic gradient update equation is the regressor vector multiplied by the error function.

number of taps. A blind P-DFE that extended the equalization scheme in [75] has been proposed in [5]. The objectives of this P-DFE are to overcome the switching difficulties encountered by the switch-mode equalizer in [75] and to interface with trellis-based channel codings schemes [40,133,143]. They seek to overcome the same switching problem that our novel parallel adaptation strategy proposed in Section 5.4 seeks to overcome. However, in their novel equalizer setup, switching of filtering structures is not required at all because the taps of the whitening filter can be driven by the P-DFE filter downstream, and conversely, the P-DFE taps can be driven by the linear whitening filter upstream. The equivalence of the linear whitening filter (which is a linear forward predictor [58]) and the non-linear predictive DFE is shown in [5]. In the acquisition mode, the linear whitening filter will adapt its filter taps and transfer them directly to the P-DFE situated downstream. After convergence is achieved by this linear equalizer setup, the P-DFE will begin its adaptation and map its taps to the linear whitening filter upstream. In this way, a low MSE is achieved due to the P-DFE without the need to switch filter positions. As for the interface with trellis-based coding schemes, a vector deinterleaver is inserted in between the whitening filter and the transversal Godard equalizer to take advantage of the noise whitening effect [5].

In our alternative DFE scheme, we will use a non-recursive whitening filter in the place of the originally proposed recursive whitening filter. Then we design the feedback filter of the P-DFE as $P(z) = \frac{A(z)}{1+A(z)}$ as shown in Fig. 5.6 where it has been derived earlier in the same manner as our alternative *conventional* DFE scheme in Section 5.3 which is governed by Equation (5.10).

5.5.1 Switching Strategy

Switching for this P-DFE is not as difficult as it is for the conventional DFE scheme in Section 5.3. This is because the P-DFE only involves a switch in its algorithms while the positions of the filters need not be interchanged, which is the original intention of the authors of [5]. We adopt a similar concept as the parallel adaptation strategy described in Section 5.4. Let the taps of the whitening filter be $\mathbf{a}^{(3)}(k)$, which is also the filter taps of the P-DFE (see Fig. 5.6). Similarly, let the taps of the phase equalizer be $\mathbf{b}^{(3)}(k)$. Then the proposed update equations of $\mathbf{a}^{(3)}(k)$ and $\mathbf{b}^{(3)}(k)$ are as follows:

$$\mathbf{a}^{(3)}(k+1) = \mathbf{a}^{(3)}(k) - \mu_{a_p}\left\{\alpha_3 t^{(3)}(k)\mathbf{s}^{(3)^*}(k) + \alpha_4 \epsilon_{\mathrm{DD}}(k)\mathbf{x}^{(3)^*}(k)\right\} \tag{5.22}$$

$$\mathbf{b}^{(3)}(k+1) = \mathbf{b}^{(3)}(k) - \mu_{b_p}\left\{\beta_3 \epsilon_{\mathrm{CMA}}(k) + \beta_4 \epsilon_{\mathrm{DD}}(k)\right\}\mathbf{t}^{(3)^*}(k) \tag{5.23}$$

where $\mathbf{a^{(3)}}(0) = [0, \cdots, 0]^T$, $\mathbf{b^{(3)}}(0) = [0, 0, \cdots, 0, 1, 0, \cdots, 0]^T$, $\mathbf{x^{(3)}}(k)$ is the regressor vector for $P(z)$, $\mathbf{t^{(3)}}^*(k)$ the regressor vector for $B(z)$, and α_3, α_4, β_3, β_4 are user defined parameters. These parameters are assigned in the same manner as described for our alternative conventional DFE in Section 5.3. An example using Benveniste-Goursat combination technique [15] has been described earlier in Section 5.4.

5.6 Novel Dual Decision Feedback Equalizer

5.6.1 Problem Statement

The least-mean-square (LMS) algorithm has been widely embraced due to its simplicity, useability and convexity. It is also known for its satisfactory performance for tracking time-varying channel statistics [58]. When the LMS algorithm is employed by a DFE, it usually exhibits near-maximum likelihood (ML) steady state performance at a fraction of the ML implementation cost. In addition, the DFE is capable of equalizing channels with deep spectral nulls which are usually difficult with linear equalizers.

The DFE that is adapted using the stochastic gradient approach must employ a reasonably small adaptation step size to ensure the proper convergence of the LMS algorithm. When the channel order is high (and hence the equalizer length), and/or the impulse response of the channel is heavily colored, the step size parameter has to be further decreased. As the step size governs the speed of convergence and tracking, the tracking capabilities of the DFE may be limited once the equalizer length is increased. A bulk of research effort has focussed on blind equalization algorithms over the past few decades based on a fixed equalizer structures such as the linear equalizer [14, 49, 124] and the DFE [7, 10, 17, 133]. More recently, the approach by the research community is reversed whereby new equalizer structures and schemes based on conventional algorithms are proposed instead [5, 21, 75, 106]. In what follows, we propose a more robust DFE structure that employs the LMS algorithm.

5.6.2 Development

We propose a new DFE structure that incorporates both the conventional DFE and the predictive DFE in one equalizer for improved tracking. As this structure consists of two DFE's, we will refer to it as the dual DFE from hence forth [80]. It consists of three filter blocks in the equalizer path as shown by the solid lines

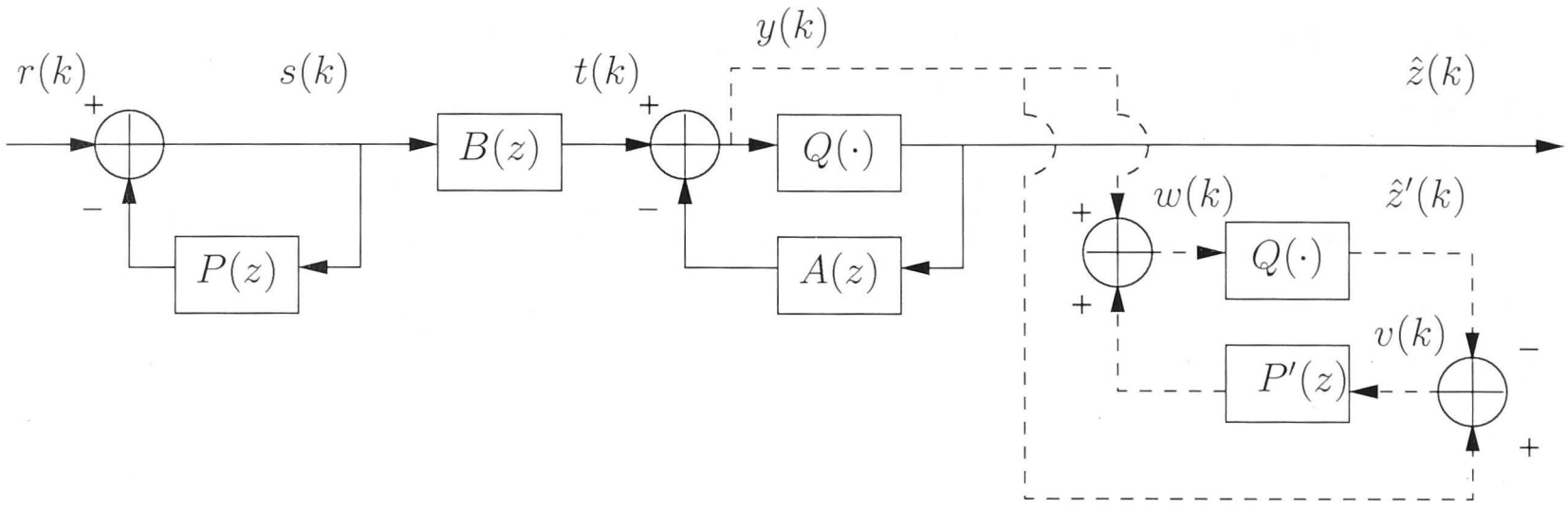

Figure 5.7: The dual DFE scheme which features a conventional DFE that is prefixed by a linear recursive filter whose taps are driven by a predictive DFE which is not in the equalizer path (illustrated by the dotted lines) but it minimizes the residual error at the soft output of the conventional DFE, $y(k)$.

connecting them in Fig. 5.7, i.e., the filters $P(z)$, $B(z)$ and $A(z)$. The combination of $B(z)$ and $A(z)$ constitutes the conventional DFE as previously shown in Chapter 2.2 in Fig. 2.6, where $B(z)$ is a non-causal transversal filter and $A(z)$ is a causal feedback filter. As for the newly appended $P(z)$ before the conventional DFE structure, it is a linear recursive filter is driven by a predictive DFE criterion that minimizes the residual error at the conventional DFE soft output. In fact the taps of $P(z)$ are copied directly from the predictive DFE, $P'(z)$, since the linear whitening filter upstream and the predictive DFE are forward predictors which are amplitude equalizers [58]. It can be shown that the prediction filter, $P'(z)$, is identical to $P(z)$ as follows [5]:

Assume first that $\hat{z}(k) = a(k)$ and $\hat{z}'(k) = a(k)$, i.e., the output signals of both the conventional DFE and the predictive DFE are correctly detected. Let the error sequences at the output of the conventional DFE and the predictive DFE be $\epsilon(k) = y(k) - \hat{z}(k)$ and $\epsilon'(k) = w(k) - \hat{z}'(k)$, respectively. Let $\mathcal{Z}(\cdot)$ denote the z-transformation. Then, the error sequence of the predictive DFE in the z-domain can be expressed as

$$W(z) = Y(z) + P'(z)[Y(z) - \hat{Z}'(z)] \tag{5.24}$$

$$\varepsilon'(z) = W(z) - \hat{Z}'(z) \tag{5.25}$$

$$= Y(z) + P'(z)[Y(z) - \hat{Z}'(z)] - \hat{Z}'(z) \tag{5.26}$$

$$= [1 + P'(z)][Y(z) - \hat{Z}'(z)] \tag{5.27}$$

$$= \varepsilon(z)[1 + P'(z)] \tag{5.28}$$

where $\varepsilon(z) = \mathcal{Z}(\epsilon(k))$ and $\varepsilon'(z) = \mathcal{Z}(\epsilon'(k))$. The power spectral density of the error

sequence at the output of the predictive DFE can be expressed as

$$\mathcal{Z}\big(\mathrm{E}[\epsilon'(k)\epsilon'^{*}(k-i)]\big) \tag{5.29a}$$

$$= \mathcal{Z}\big(\mathrm{E}[\epsilon(k)\epsilon^{*}(k-i)]\big)\big[1+P'(z)\big]\big[1+P'^{*}(z)\big] \tag{5.29b}$$

$$= \frac{\sigma_n^2\sigma_a^2}{\sigma_a^2 F(z)F^{*}(z^{-*})+\sigma_n^2}\big[1+P'(z)\big]\big[1+P'^{*}(z)\big] \tag{5.29c}$$

where $F(z)$ is the channel response, σ_n^2 and σ_a^2 are the variance of the AWGN and the data symbols, respectively. Further, the denominator of (5.29c) can be factorized into

$$\sigma_a^2 F(z)F^{*}(z^{-*})+\sigma_n^2 = S_0 G(z)G^{*}(z^{-*}) \tag{5.30}$$

where S_0 is a positive constant. Recall that the transfer function of the optimal linear MMSE equalizer can be expressed as [75, 118, 119]

$$C(z) = \frac{\sigma_a^2 F^{*}(z^{-*})}{\sigma_a^2 F(z)F^{*}(z^{-*})+\sigma_n^2} \tag{5.31}$$

and the transfer function of the linear recursive whitening filter that compensates for the minimum phase part of $C(z)$ is such that

$$\frac{1}{1+P(z)} = \frac{1}{G(z)} \tag{5.32a}$$

$$\therefore \quad P(z) = G(z) - 1. \tag{5.32b}$$

Thus the minimization of the MSE at the input to the decision device of the predictive DFE, i.e., $\mathrm{E}[\epsilon'(k)\epsilon'^{*}(k-i)]$, would lead to $P'(z) = G(z) - 1$ according to (5.29c). This expression is exactly identical to (5.32b), which is the transfer function of the linear whitening filter upstream.

5.6.3 Operation Details

The dual DFE's main objective is to enhance the tracking capability of the conventional DFE without compromising on the low steady state MSE of the conventional DFE. We wish to retain the advantage of implementing the conventional DFE whose steady state MSE is usually lower than that achievable by the predictive DFE in practice [119]. Thus, the predictive DFE should not 'compete' with the conventional DFE in acquiring the transfer function of the minimum phase channel expressed in (5.32), but rather complements the conventional DFE only

when rapid changes in the channel statistics are encountered.

Let the weight vectors of $P(z)$ and $P'(z)$ be $\mathbf{p}(k) = [p_1(k), p_2(k), \cdots, p_N(k)]^T$ and $\mathbf{p}'(k) = [p'_1(k), p'_2(k), \cdots, p'_N(k)]^T$. Let their associated regressor vectors be $\mathbf{s}(k) = [s_1(k), s_2(k), \cdots, s_N(k)]^T$ and $\mathbf{v}(k) = [v_1(k), v_2(k), \cdots, v_N(k)]^T$, respectively. The conventional DFE is allowed to converged first until a sufficiently low MSE, say MSE_{t1}, is detected at time $k = k_0$. The predictor coefficients are not updated before k_0, such that

$$\mathbf{p}(k) = \mathbf{0}, \mathbf{p}'(k) = \mathbf{0} \qquad \forall k < k_0 \tag{5.33a}$$

$$\mathbf{s}(k) = \mathbf{0}, \mathbf{v}(k) = \mathbf{0} \qquad \forall k < k_0. \tag{5.33b}$$

After time k_0, the regressor vectors are initialized with their input signals, but the adaptation of $P'(z)$ is not started yet as we are relying on the conventional DFE to achieve a low steady state MSE. However, when the MSE is detected to have risen above a pre-defined threshold, MSE_{t2}, at time $k > k_0$, either changes in the channel statistics have been encountered or the SNR level has dropped. Assume fixed SNR level, then the predictive DFE is started to assist in the tracking of the channel statistics. It is adapted as follows, for $k > k_0$, $\text{MSE} > \text{MSE}_{\text{t2}}$:

$$\mathbf{p}'(k+1) = \mathbf{p}'(k) - \mu_{\mathbf{p}'}[w(k) - x(k)]\mathbf{v}^*(k) \tag{5.34}$$

The weight vector of linear recursive filter upstream copies the adapted weights of (5.34). If tracking is successful and the channel statistics are vary slower, then the MSE would decrease given the same SNR levels. When the MSE drops below MSE_{t1}, then the adaptation of $P'(z)$ ceases and weight vector is slowly decreased to $\mathbf{0}$ by multiplying $\mathbf{p}'(k)$ with a decaying factor very close to unity.

Discussion

By incorporating a predictive DFE, the MSE can usually be reduced further at the output of the predictive DFE due to its noise whitening effect. However, this does not help the MMSE equalizer upstream to track the channel better since the MMSE equalizer is independent of the predictor coefficients [119]. The novelty of the dual DFE is in the mapping of the predictor taps into a linear recursive filter upstream of the MMSE equalizer, enabling a coupling between the taps in a special way so that the predictor coefficients may be reflected into the equalizer path (in solid lines in Fig. 5.7) to reduce the load on the conventional DFE. With this new configuration, a low steady state MSE due to the use of the conventional DFE in

stationary environments and enhanced tracking due to the predictive DFE driven criterion that minimizes the residual error at the conventional DFE's output are simultaneously accomplished.

Switch-mode fast-convergence blind dual DFE

For acquisition of the channel blindly, we choose to implement a fast convergence DFE scheme due to Labat *et al* [75]. The initial acquisition is performed using the cascaded equalizer structure that exploits the channel decomposition property as explained in Section 5.1 in this chapter. The equalizer is allowed to converge to its optimal settings as a conventional DFE. When the MSE estimate drops below a pre-defined threshold, then the dual DFE structure is subsequently implemented according to (5.33) and (5.34).

5.7 Simulation Results

This chapter deals with three broad classes of switch-mode DFE schemes, namely, the switch-mode conventional DFE scheme of Section 5.3, the switch-mode predictive DFE scheme of Section 5.5, as well as the switch-mode dual DFE scheme of Section 5.6. The simulation results are plotted for various switch-mode DFE schemes that will be addressed accordingly in the following subsections.

5.7.1 Switch-Mode Conventional DFE Scheme of Section 5.3 – Channel Equalization Results

The following simulations are generated to yield performance results regarding the new scheme in Section 5.3 with and without the enhancements using the parallel adaptation strategy as described in Section 5.4. Two stationary non minimum phase channels are selected for our simulations.The first is an auto-regressive moving average (ARMA) channel, $\mathbf{h}'$, which is more heavily colored, and the second is a moving average (transversal) channel, $\mathbf{h}''$, which is lightly colored [118]. The transfer functions (TF) of both channels are given by

$$\text{TF } (\mathbf{h}') = \frac{0.4 + z^{-1}}{1 - 0.3z^{-1} + 0.5z^{-2} - 0.2z^{-3}} \tag{5.35}$$

$$\begin{aligned}\text{TF } (\mathbf{h}'') = {} & 0.04 - 0.05z^{-1} + 0.07z^{-2} - 0.21z^{-3} - 0.5z^{-4} + \\ & 0.72z^{-5} + 0.36z^{-6} + 0.21z^{-8} + 0.03z^{-9} + 0.07z^{-10}.\end{aligned} \tag{5.36}$$

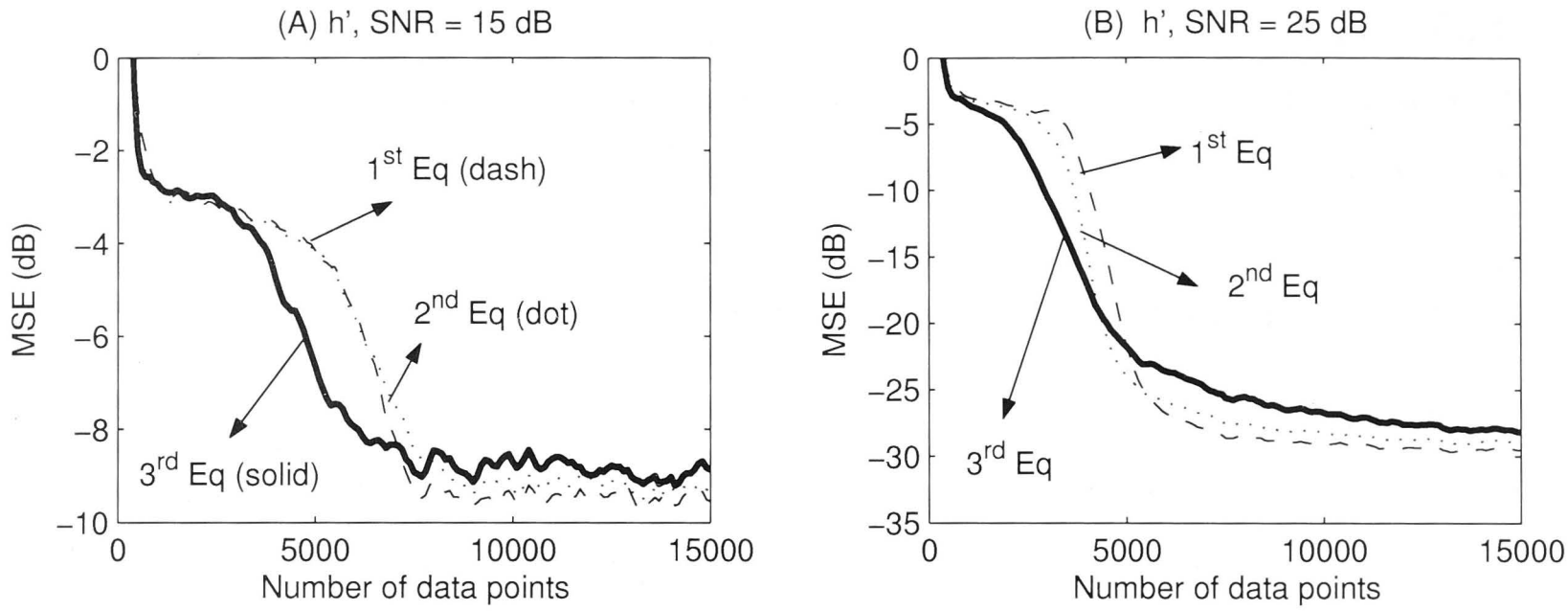

Figure 5.8: (A) and (B) are comparisons of three DFE schemes for the first channel, $\mathbf{h}'$, while (C) and (D) are for the second channel, $\mathbf{h}''$, using 16-QAM data signals. The first equalizer is the DFE of [75]. The second and third equalizers are the DFE described in Section 5.3 without and with parallel adaptation strategy, respectively. The parallel adaptation strategy employs the Benveniste-Goursat type combination parameters.

The transmit data format used is 16-QAM and assumed independent and identically distributed. The real and imaginary components of the transmit data are drawn from the $[-3, -1, 1, 3]$. To characterize the equalizer performance in terms of convergence speed and steady state errors, we employ the decision directed MSE which can be estimated using the following recursion

$$\mathrm{MSE}_{\mathrm{DD}}(k+1) = 0.99\ \mathrm{MSE}_{\mathrm{DD}}(k) + 0.01\ |\epsilon_{\mathrm{DD}}(k)|^2. \tag{5.37}$$

We tested with three types of adaptive equalizers. The first is the DFE in [75]. The second is the DFE of Section III without the use of parallel adaptation strategy. The third is similar to the second equalizer except that the parallel adaptation strategy with the Benveniste-Goursat type soft switching technique is used. The

system parameters of the three equalizers are as follows:

The length of $A(z)$ is 20, for the recursive whitening filter of [75] as well as our non-recursive one (see Fig. 5.3). The length of $B(z)$ is 21 for all three equalizers. For the first two DFE's (without parallel adaptation strategy), it switches from the linear acquisition mode to the tracking mode when $\mathrm{MSE}_{\mathrm{DD}}(k) < 0.5657$, i.e., -4.9dB. Once it is in the tracking mode, it may be switched back to the starting mode if $\mathrm{MSE}_{\mathrm{DD}}(k) > 0.7778$, i.e., -2.2dB. For the third equalizer there is no distinct switching point. Their adaptation equations are governed by the Benveniste-Goursat parameters $c_1 = c_2 = 1$ and $\gamma_a = 1, \gamma_b = 10$ for both channels (equations (5.20) and (5.21) are referred), with the exception c_1 for (5.20) for $\mathbf{h}''$ where we assigned $c_1 = 0.5$. The step sizes of the first two equalizers for $\mathbf{h}'$ in the acquisition mode are $\mu_{\mathbf{a}^{(1)}} = 5 \times 10^{-5}$ and $\mu_{\mathbf{b}^{(1)}} = 5 \times 10^{-6}$, while in the tracking mode they become $\mu_{\mathbf{a}^{(2)}} = \mu_{\mathbf{b}^{(2)}} = 10^{-4}$; for $\mathbf{h}''$, we assigned $\mu_{\mathbf{a}^{(1)}} = \mu_{\mathbf{a}^{(2)}} = \mu_{\mathbf{b}^{(2)}} = 2.5 \times 10^{-4}$ and $\mu_{\mathbf{b}^{(1)}} = 2.5 \times 10^{-5}$. As for the third equalizer, $\mu_{\mathbf{a}} = 5 \times 10^{-5}$ and $\mu_{\mathbf{b}} = 5 \times 10^{-6}$ for $\mathbf{h}'$; and $\mu_{\mathbf{a}} = 2.5 \times 10^{-4}$ and $\mu_{\mathbf{b}} = 2.5 \times 10^{-5}$ for $\mathbf{h}''$.

Fig. 5.8 shows the MSE plots of the three above equalizers averaged over 100 independent trials using 15,000 and 10,000 symbols for channels $\mathbf{h}'$ and $\mathbf{h}''$, respectively, at SNR levels of 15 dB and 25 dB. Two main conclusions are drawn. Firstly, the first and second equalizers exhibit almost identical performances in terms of convergence speed and steady state errors for both channels under both SNR levels. This shows that our simplified DFE as described in Section III has maintained the standard set by the original DFE in [75]. Secondly, the use of the parallel adaptation strategy using the Benveniste-Goursat type combination parameters improves performance of the third equalizer over the first two equalizers that are without the strategy. The success of the smooth switching strategy is implicit in the faster rate of convergence achieved by the third equalizer since the first two equalizers would usually switch between their acquisition and tracking modes several times before convergence is finally achieved. It is less obvious for easier channels, i.e., $\mathbf{h}''$, at high SNR levels (Fig. 5.8-D) since switching without parallel adaptation can still be "smooth" because the slicer errors that can cause ill-convergence of the DFE are fewer. On a different note, the third equalizer that employs the Benveniste-Goursat type parameters yield higher steady state MSE because of the finite contribution from the acquisition increment vectors, i.e., $c_1|\epsilon_{\mathrm{DD}}(k)|\epsilon_{\mathrm{MOE}}\mathbf{s}^{(1)^*}(k)$ and $c_1|\epsilon_{\mathrm{DD}}(k)|\epsilon_{\mathrm{CMA}}(k)\mathbf{t}^{(1)^*}(k)$ as in (5.20) and (5.21), even after convergence has been achieved. The problem with the high steady state MSE can be solved by employing the reliability measure technique [82] which is more computationally intensive.

5.7.2 Switch-Mode Conventional DFE Scheme of Section 5.3 – Channel Identification Results

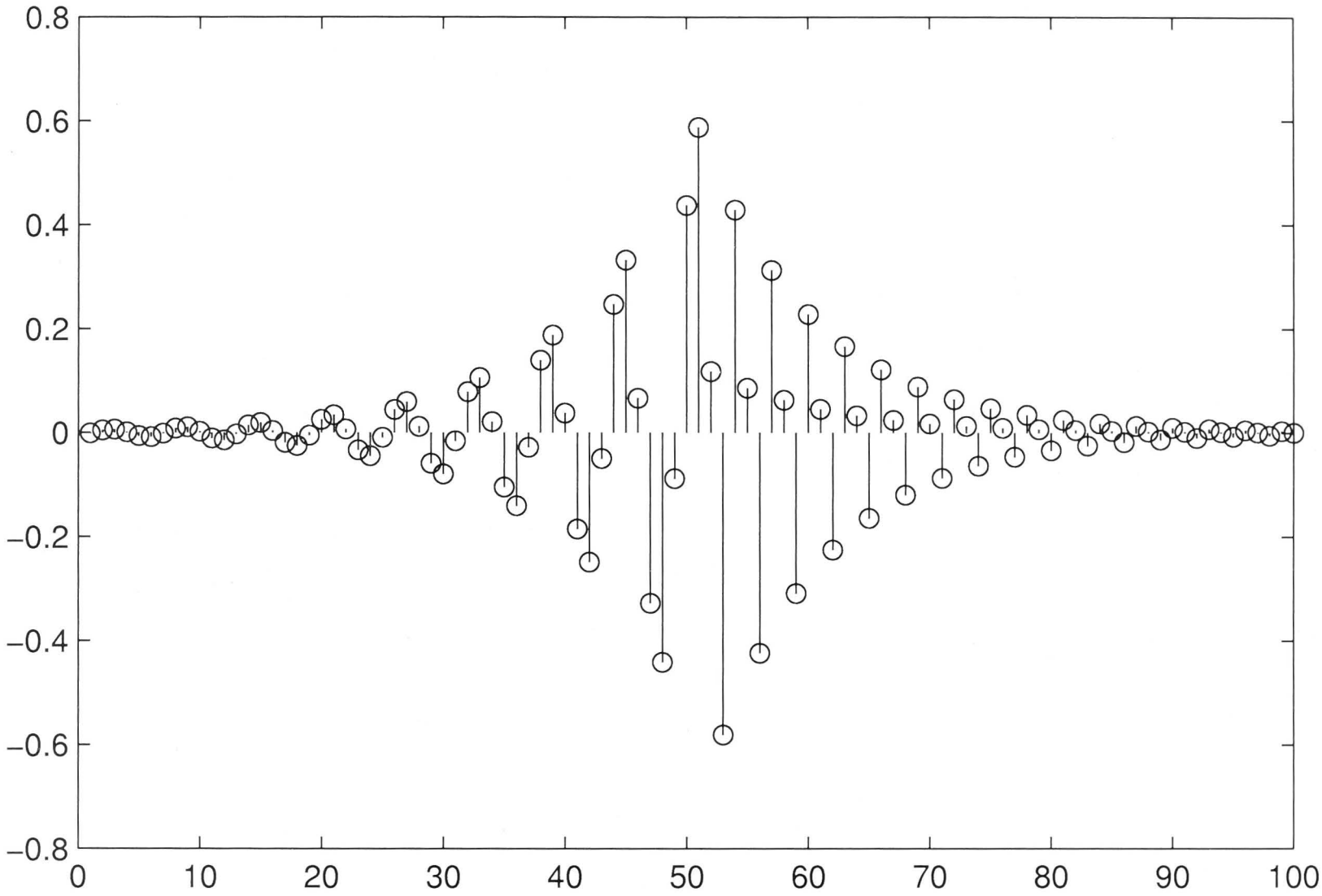

Figure 5.9: The 100 most dominant taps of the inverse response of channel $\mathbf{h}'''$.

For this subsection, we generated source data signals that are drawn from an 8-PAM alphabet set

$$\mathbb{A} = \{\pm 1, \pm 3, \pm 5, \pm 7\}.$$

We chose a 5 tap transversal channel whose transfer function is given by

$$\text{TF}\ (\mathbf{h}''') = 0.8264 - 0.1653z^{-1} + 0.8512z^{-2} + 0.1636z^{-3} + 0.81z^{-4}. \qquad (5.38)$$

It is a severe non-minimum phase channel that exhibits deep spectral nulls and has an initial kurtosis ratio of 0.3 (see [75]). The implication of the initial kurtosis ration that is less than 0.5, according to [79], is that convergence of the Godard equalizer would usually take much longer. As our equalizer is transversal, under convergence, its weight vector will be approximately the inverse of the channel's impulse response, which is shown in Fig. 5.9. Only 100 of its most dominant taps have been plotted.

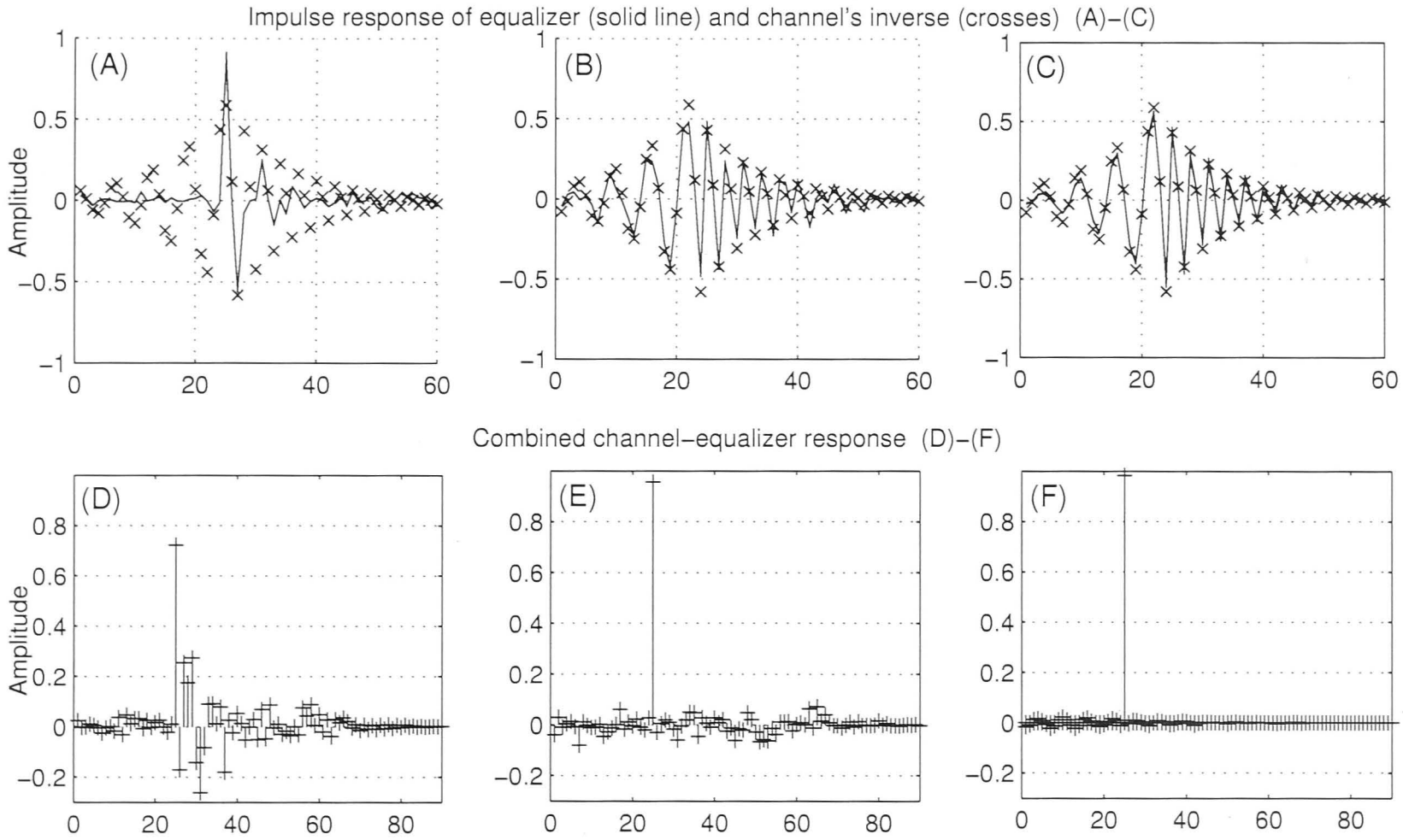

Figure 5.10: Channel identification results using new switch-mode DFE scheme proposed in Section 5.3. These results are obtained using (A) 300 symbols, (B) 800 symbols, (C) 2500 symbols. The equalizer is in its acquisition mode for both (A) and (B). As for (C), the equalizer has settled into its tracking (DD) mode.

The SNR level is set at 30 dB. Therefore, after some calculations, the standard deviation of the AWGN is obtained to be $\sigma = 0.15$. While the equalizer removes the ISI due to the channel distortions, it inevitably amplifies the noise. At the equalizer output, the filtered (amplified) noise has a standard deviation of 0.22, where amplification is by a factor of $\|1 + A(z)\|_2^2$, where $\|\cdot\|_2$ denotes the l_2 norm of a vector and assuming $A(z)$ has converged to its optimal values. This level of (filtered) noise makes equalization of this channel quite a challenging task, especially for non-constant modulus source signals like 8-PAM [42].

An equalizer can be designed for channel identification purposes by inverting the channel. In our case, we aim to find the inverse response of the channel using our non-recursive FIR acquisition equalizer. If $\mathrm{MSE}_{\mathrm{DD}}(k)$ drops below 0.27, it would then switch to the DFE as explained in Section 5.3. The inverse impulse response of the channel has more than 90 taps whose magnitudes exceed 10^{-3}, as well as many other smaller tails. A hundred of the most dominant taps of $H^{-1}(z)$ are plotted in Fig. 5.9.

Due to the large number of taps in the channel's inverse response, we assigned 40 taps for $A(z)$ and 50 taps for $B(z)$. We simulated 15 different runs where 300 and 800 baud-rate symbols were used. The equalizer's impulse response of each run

is plotted in Fig. 5.10-A,B, as a solid line and the desired impulse response of the channel inverse is plotted as circles on the same graphs. Even though the equalizer has not switched to its decision directed mode at this point, we are pleasantly surprised at the close proximity of the equalizer's taps to the channel's inverse after just 800 symbols despite the severity of the channel. Fig.5.10-B demonstrate the speedy acquisition of the linear acquisition equalizer.

5.7.3 Switch-Mode Conventional DFE Scheme Under Parallel Adaptation Strategy

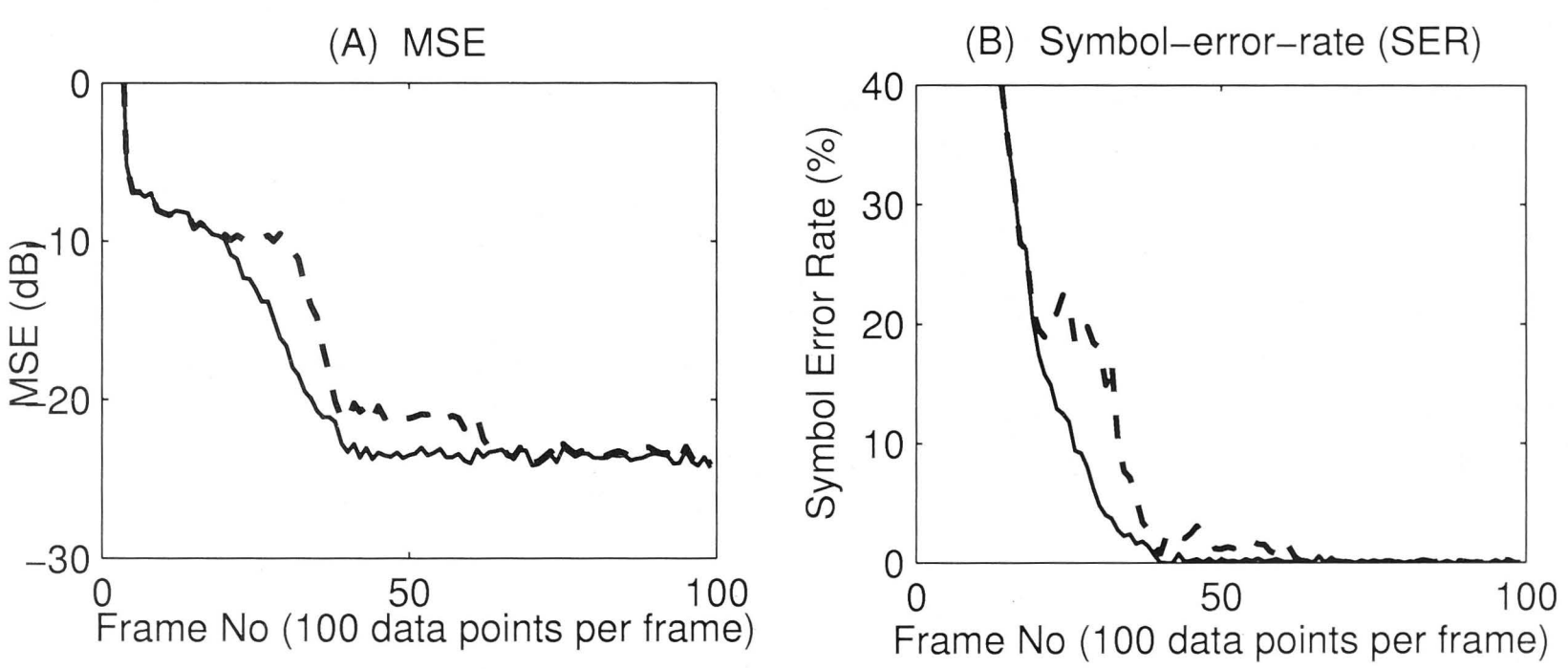

Figure 5.11: The average MSE and SER of 20 runs of fast convergence DFE scheme of Section 5.3, with (solid line) and without (dotted line) parallel adaptation strategy on 8-PAM signals at 30 dB for channel $\mathbf{h}'''$.

In Fig. 5.11, we plotted the averaged result from 20 separate runs to demonstrate the adverse effects of switching. When parallel adaptation is not used, the equalizer generally faces an abrupt change when switching occurs which disrupts the convergence of the equalizer parameters to its optimal decision directed settings. The average MSE plots of the equalizer output with and without parallel adaptation are drawn with the solid and dotted lines, respectively. Notice the 'bumps' encountered at the vicinity of switching when parallel adaptation is not employed. Consequently, the equalizer may switch back and forth between its acquisition and tracking mode until eventually convergence is achieved. As a result, the equalizer will generally take longer to converge. As the DFE is extremely sensitive to decision errors [39,71], the 'bump' in the symbol error rate (SER) at around 15% SER may cause the ill-convergence of the DFE.

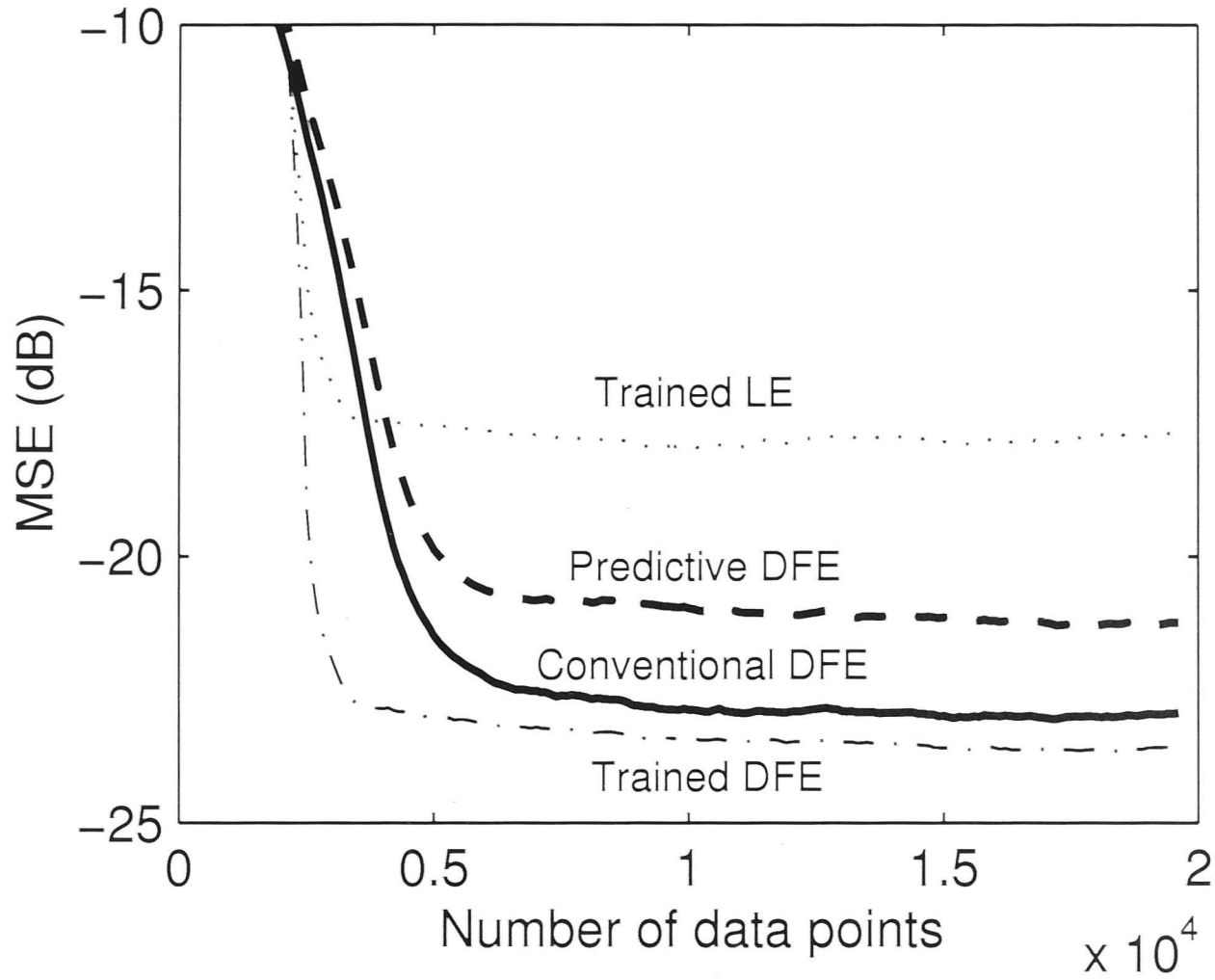

Figure 5.12: Results of two new alternative DFE's compared with trained equalizers.

The estimate MSE is obtained according to (5.37) and the symbol-error-rate (SER) is estimated via

$$\mathrm{SER}(k+1) = 0.99\ \mathrm{SER}(k) + 0.01\ I(k) \tag{5.39}$$

where

$$I(k) = \begin{cases} 1 & z(k) = a(k) \\ 0 & z(k) \neq a(k). \end{cases} \tag{5.40}$$

$I(k)$ is simply an indicator which yields unity when the equalizer output is correctly detected, i.e., $z(k) = a(k)$, and a zero if it is incorrectly detected, i.e., $z(k) \neq a(k)$.

5.7.4 New Switch-Mode DFE Schemes of Sections 5.3, 5.4 and 5.5 – Channel Equalization Results

In this section, we provide simulation results to draw evidence of fast convergence, low steady state errors, and high success rate of convergence that is unaffected by switching, of the new DFE. We will compare the performances of the blind conventional DFE of Section 5.3, the blind P-DFE of Section 5.5, the trained LE

and the trained DFE. The data format used is 4-QAM signalling. The SNR level is set at 15 dB. The channel to be equalized is $\mathbf{h}'''$ of (5.38). We used a total of 20,000 symbols for the above simulations.

The convergence rate of the blind equalizer is comparable to that of the trained equalizers (LE and DFE), and its steady state MSE is lower than the trained LE. Unlike the results in [75] that showed equal convergence speed for both their blind equalizer and the trained equalizers, our results show marginally slower convergence speed. This is largely due to the increased length of the non-recursive for this particular example. The length of the recursive whitening filter in [75] is 4 but ours is fixed at 21.

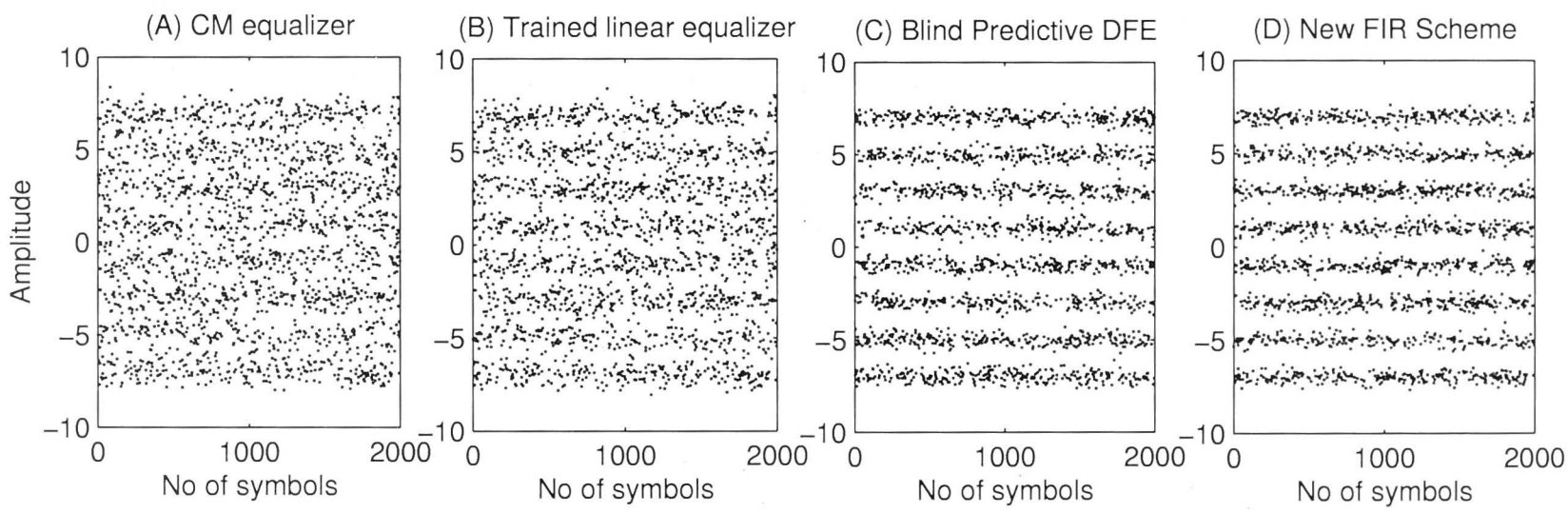

Figure 5.13: Plot of the equalizer output of various selected schemes extracted at the end of the data stream. (A) CMA equalizer. (B) Trained LMS equalizer. (C) Fast converging blind predictive DFE. (D) Our FIR scheme and the trained DFE, which yields similar MSE.

In Fig. 5.13, we plotted the last 2000 symbols of the data stream consisting of 20,000 symbols of the equalizer output for various selected equalizers from Fig. 5.13 in ascending order of their steady state MSE. The results is not surprising because a clear open eye is expected for our *conventional* DFE scheme, followed by a smaller eye by the predictive DFE scheme. This is because the predictive DFE's forward MMSE equalizer parameters are adapted *independently* of its DFE feedback parameters, whereas the parameters of both feedforward and feedback filters of the conventional DFE are jointly adapted [119]. The linear equalizers perform much worse, where the trained linear equalizer yields lower MSE compared to the CM equalizer because the data constellation is non-constant modulus [42].

5.7.5 Switch-Mode Dual DFE Scheme of Section 5.6 – Channel Equalization Results

The following simulations show the performance comparison of three schemes used to equalize the five-tap non-minimum phase channel that is used in [75], which is $\mathbf{h}'''$. The data format used here belongs to the real 8-PAM format. As the dual DFE equalizer scheme is one that only enhances tracking of a non-stationary channel after which convergence has been achieved, we let the channel be stationary for the first 15,000 symbols before introducing a time-varying zero. For the first simulation scenario whose MSE is as depicted in Fig. 5.14, a fifth zero denoted by $r_5(k)$ apart from the four original zeros of (5.38) is added. It starts from the origin at the $15,001^{\text{th}}$ symbol and it moves radially towards the coordinate $(1, 0)$ on the z-unit circle according to the following equation:

$$r_5(k) = \begin{cases} 0 & 0 \leq k \leq 15000 \\ \frac{0.06(k-15000)}{2000} & k > 15000 \end{cases} \tag{5.41}$$

As for the second simulation whose results are depicted in Fig. 5.14, a zero is also added on the $15,001^{\text{th}}$ symbol. It is defined by a similar linear motion plus a sine component in the following manner:

$$r_5(k) = \begin{cases} 0 & 0 \leq k \leq 15000 \\ \frac{0.05(k-15000)}{2000} + 0.1 \sin \frac{k\pi}{2000} & k > 15000 \end{cases} \tag{5.42}$$

Here, we compare the performances of three equalization scheme. The first is the linear equalizer of the acquisition mode of [75]. The second is the linear-equalizer-DFE scheme of [75], whereby it starts off as a linear equalizer (same as the first equalizer that we are testing on), and switches over to a DFE once the eye is sufficiently open. The switching occurs when $\text{MSE}_{\text{DD}}(k) < 0.27$. The third is the dual-DFE scheme that starts off as the DFE scheme of [75] (same os the second equalizer that we are testing on). It the dual-DFE is implemented once the MSE estimate further drops below a lower threshold, i.e., when $\text{MSE}_{\text{DD}}(k) < 0.23$.

Note that all three equalizers are extensions of the one before it. The first equalizer is merely the acquisition linear equalizer in [75]. Its recursive whitening filter, $A(z)$, has 25 taps. Its doubly finite Godard transversal filter, $B(z)$, has 50 taps. The second equalizer extends the first in the sense that acquisition is achieved using the similar linear equalizer, it switches to a conventional DFE once the estimate MSE drops below a predefined threshold value, i.e., $\text{MSE}_{\text{DD}}(k) < 0.27$. The third equalizer extends the second equalizer in the sense it is adapted

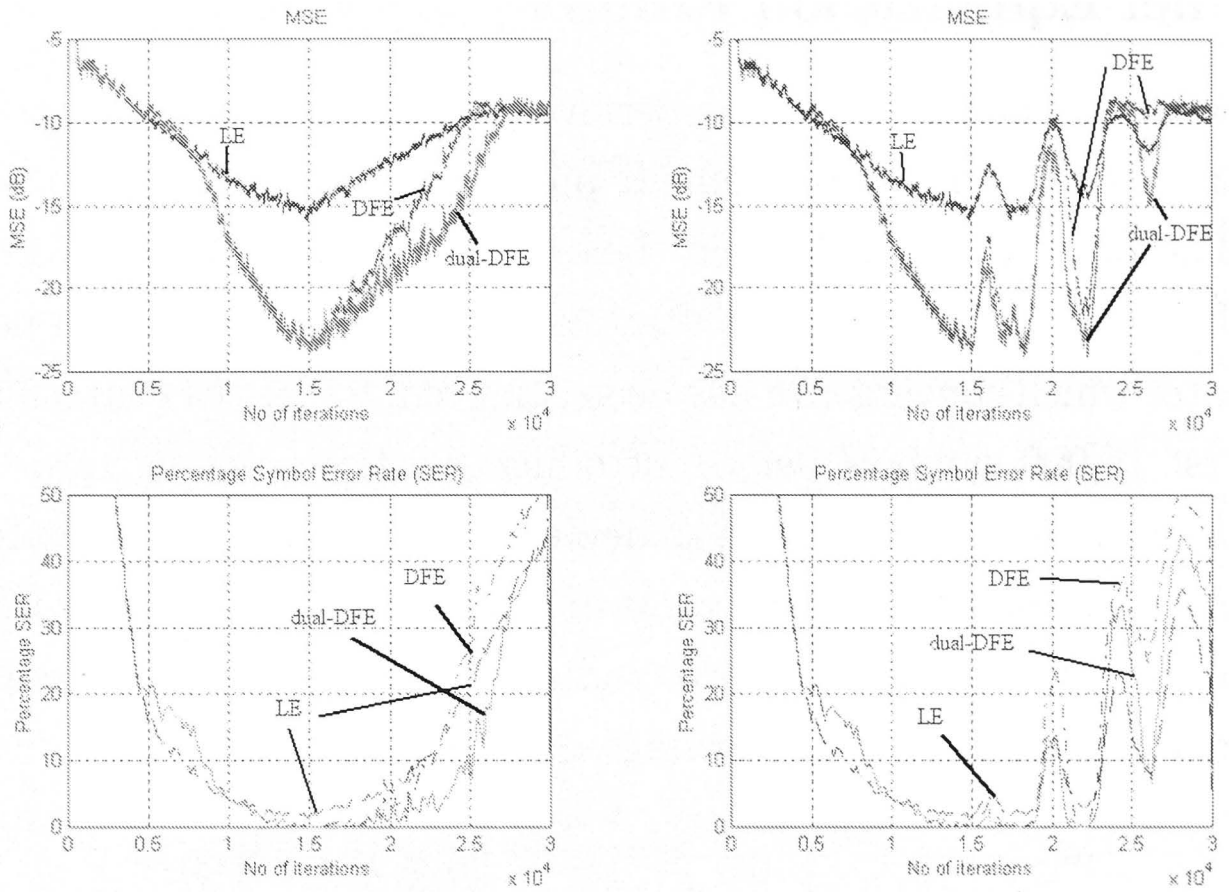

Figure 5.14: Comparing three equalizers: 1) the linear acquisition equalizer (LE) of [75], 2) the DFE scheme of [75], and 3) the dual-DFE that extends the DFE in [75].

exactly as the second equalizer until $\mathrm{MSE}_{\mathrm{DD}}(k) < 0.23$ after which the dual DFE is implemented. In the setup of the third equalizer, the additional predictor filters, $P(z)$ and $P'(z)$, have equal lengths of 25 taps.

The plots in the MSE figures are obtained using 20 different and independent runs at a signal to noise ratio (SNR) of 30 dB.

5.8 Conclusions

In this chapter, we have accomplished three distinct goals which describe the main contribution of this chapter. They can be itemized as follows:

1. We have successfully extended the switch-mode equalizer scheme of Labat *et al* [75] to incorporate a linear transversal whitening filter while retaining the simplicity of a direct transfer of its tap parameters to the DFE in the tracking mode. No observable loss in performance is encountered with this modification to the transversal mode as compared to the original recursive

scheme.

2. We addressed the smooth switching between operation modes of the equalization scheme of [75]. Even though it is advantageous to use different types of equalizer structures under different conditions related to the existing channel, it becomes difficult to ensure a smooth switching between these structures. We proposed a switch-mode technique called the parallel adaptation strategy [84] that ensures an automatic transition between a suboptimal linear acquisition structure and a non-linear tracking structure.

3. We proposed a new equalizer structure that incorporates two DFE's simultaneously to cope effectively with time-varying channel statistics as well as stationary ones.

Chapter 6

Rotational Analysis On Several Blind Equalization Algorithms

In this chapter, we study the rotational behaviors of the constant modulus algorithm (CMA), the reduced constellation algorithm (RCA) and the multi-modulus algorithm (MMA) for quadrature-amplitude-modulation (QAM) schemes, which are the behaviors of the equalizer output constellation under the adaptation of the above algorithms. In particular, we compare their susceptibilities to undesirable rotated solutions using the concept of torque where the output signal constellation is treated as a physical object and the update error term treated as an external force that induces rotation. We verify the phase-invariant property of the CMA and show that the RCA is prone to undesirable 45-degree solutions for cross-QAM constellations but much less likely for square-QAM constellations. We also found that the MMA cannot converge to such wrong solutions for any arbitrary constant phase offsets for a wide range of noise levels.

6.1 Introduction

Blind equalization algorithms can be employed to combat the intersymbol interference (ISI) when only the channel output and some statistical information regarding the data input are available. Certain blind algorithms are also capable of correcting the phase of the output constellation apart from their usual tasks of blind ISI removal. These *phase aware* algorithms include the reduced constellation algorithm (RCA) [15,50] and the multi-modulus algorithm (MMA) [105,152] that is developed more recently. Unfortunately, their convergence to undesirable rotated solutions whereby the output constellation is rotated by an arbitrary angle from its original orientation have been reported in the literature [105, 147, 152]. In this chapter, we

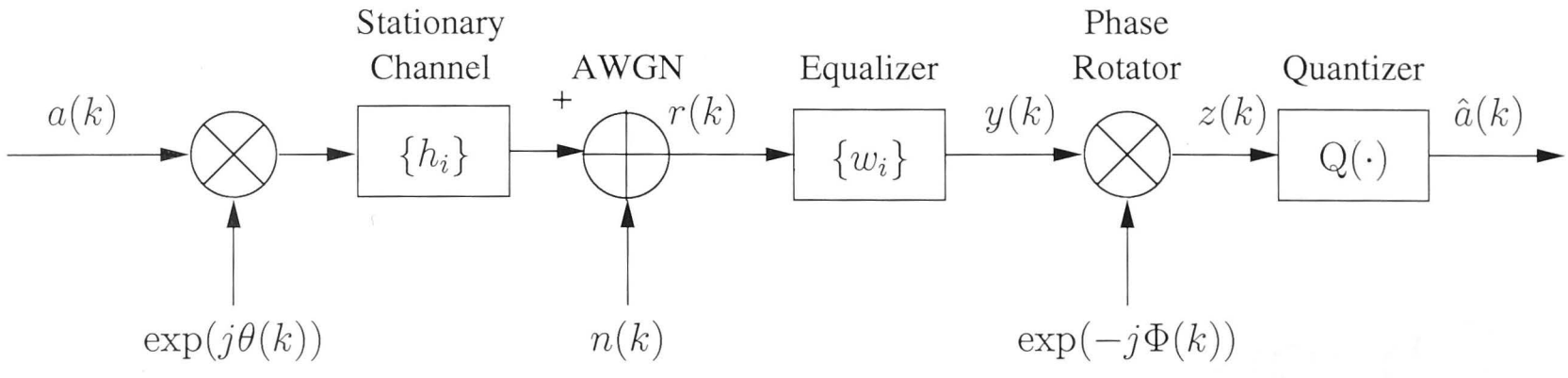

Figure 6.1: Baseband QAM system model.

will analyze the constant modulus algorithm (CMA) [49, 137], the RCA and the MMA, with respect to their immunities (or susceptibilities) to such undesirable solutions.

Blind equalization is used extensively in various emerging broadband access applications such as fiber-to-the-curb (FTTC) networks and x-digital-subscriber-line (xDSL). A phase aware blind algorithm that can perform the joint task of blind equalization and phase recovery satisfactorily is highly desirable [46, 147]. In carrierless amplitude and phase (CAP) modulation and quadrature amplitude modulation (QAM) receivers, for example, additional carrier phase tracking circuits are necessary due to the lack of an automatic constellation phase recovery property not inherent in the CMA. A solution is then to employ phase aware algorithms such as the RCA and the MMA so that the phase tracking circuits are no longer required. However their potential convergence to wrong solutions will seriously jeopardize the equalizer performance under practical environments. While the occurrences of such undesirable convergence are much fewer for the MMA compared to the RCA and the CMA [152], the reason for the superior behavior of the MMA has not been established. Moreover, even though the minimization of the costs of both the RCA and the MMA will result in the desirable 0^o solution [46], it is not clear it there exists other rotated solutions which may not yield minimum costs. We will study the existence of such ill-oriented solutions which helps in the design of an equalizer. In this chapter, we will use the concept of torque common in mechanics to address this rotational issue.

6.2 System Setup and Assumptions

Consider the channel-equalizer system depicted in Fig. 6.1. Let $\mathbf{h} \triangleq [h_0, h_1, \cdots, h_L]^T$ denote the coefficients of the stationary channel filter of length $L+1$. The channel is possibly non-minimum phase but unknown to the receiver. In addition, the data symbols may be rotated. Let the source data sequence be $\mathbf{a}(k) \triangleq$

$[a(k), a(k-1), \cdots, a(k-L)]^T$ with a time index k, drawn from M-QAM alphabet set, $M = 16, 32, 64, 128$, where $M = 16, 64$ are square constellations and $M = 32, 128$ are cross ones. At the receiving end, the input signal to the equalizer is

$$r(k) = \mathbf{h}^T \mathbf{a}(k) + n(k) \tag{6.1}$$

where $n(k)$ is the additive white Gaussian noise (AWGN) and ^{T} denotes transposition. Let $\mathbf{w}(k) \triangleq [w_{-N}(k), \cdots, w_0(k), \cdots, w_N(k)]^T$ be the $(2N+1)$ complex equalizer tap coefficients initialized with a center-tap strategy, which allows the causal development of approximate inverse filters for non-minimum phase systems. Then the equalizer output can be expressed as

$$z(k) = \sum_{n=-N}^{N} w_n(k) r(k-n) \tag{6.2a}$$

$$= a(k) e^{j\Delta\theta(k)} + v(k) \tag{6.2b}$$

where $\Delta\theta(k) = [\theta(k) - \Phi(k)]$ is the instantaneous residual phase error term, $\theta(k)$ and $\Phi(k)$ are depicted in Fig. 6.1, and $v(k)$ is the residual-ISI-plus-noise term of variance σ_v^2. In our analysis, we make the simplifying assumption that $v(k)$ is normally distributed so that we can assume that the data is transmitted over an AWGN channel due to the central limit theorem [57, ch. 2].

Under the usual stochastic gradient adaptation, the weight vector of the equalizer is updated as below:

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \mu_{\mathbf{w}} \frac{\partial J(\mathbf{w})}{\partial \mathbf{w}(k)} \tag{6.3a}$$

$$= \mathbf{w}(k) - \mu_{\mathbf{w}} \epsilon(k) \mathbf{r}_N^*(k) \tag{6.3b}$$

where $\mu_{\mathbf{w}}$ is the adaptation step size, $J(\mathbf{w})$ and $\epsilon(k)$ are the cost function and the error function of the blind adaptive algorithm, respectively, $\mathbf{r}_N(k) = [r(k), r(k-1), \cdots, r(k-N+1)]^T$ is the input regressor, and * denotes complex conjugation. In what follows, we will assess the cost functions $J(\mathbf{w})$, or equivalently, their error functions $\epsilon(k)$, with respect to their immunities (or susceptibilities) to undesirable rotational equilibria.

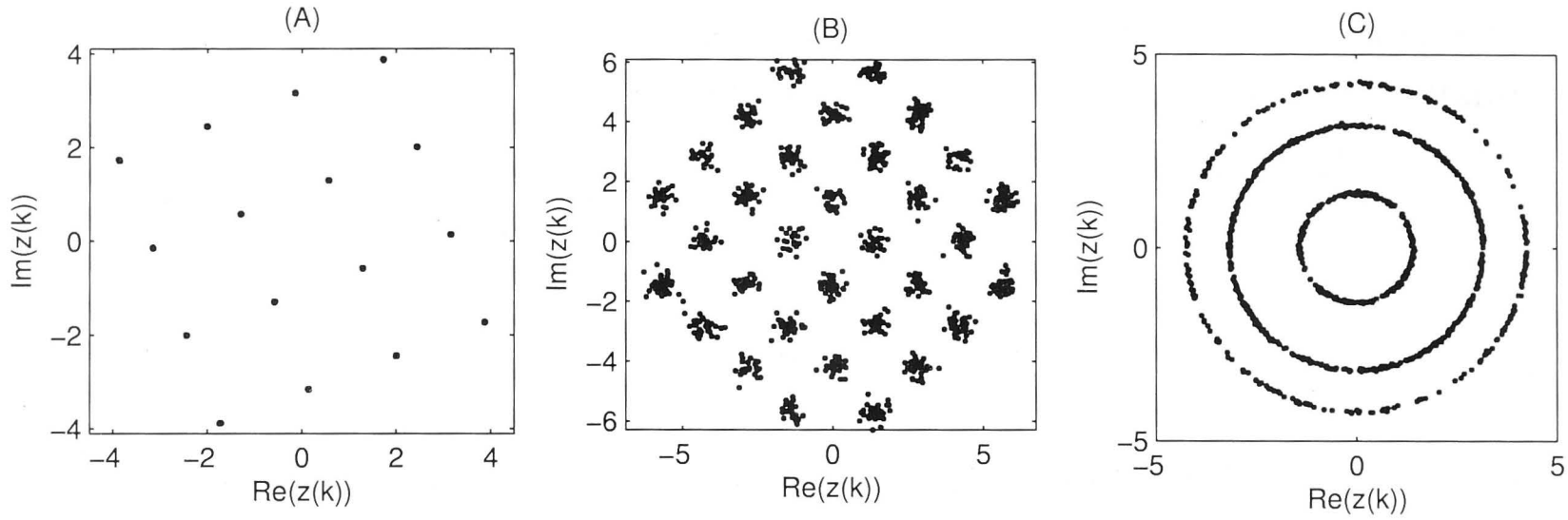

Figure 6.2: Examples of undesirable rotated solutions for (A) 16-QAM at 21°, obtained by RCA and CMA; (B) 32-QAM at 45°, obtained by RCA and CMA; (C) 16-QAM under a frequency offset, obtained by CMA.

6.3 Torque Analysis of Undesirable Rotated Solutions

6.3.1 Some Wrong Rotated Solutions

The cost functions of the CMA, the RCA and the MMA are respectively given by

$$J_{\mathrm{CMA}}(\mathbf{w}) = E\left\{\left(|z(k)|^2 - \gamma_{\mathrm{C}}^2\right)^2\right\} \tag{6.4a}$$

$$J_{\mathrm{RCA}}(\mathbf{w}) = E\left\{\left|z(k) - \gamma_{\mathrm{R}}\right|^2\right\} \tag{6.4b}$$

$$J_{\mathrm{MMA}}(\mathbf{w}) = E\left\{\left[\mathrm{Re}\big(z(k)\big)^2 - \gamma_{\mathrm{M}}^2\right]^2 + \left[\mathrm{Im}\big(z(k)\big)^2 - \gamma_{\mathrm{M}}^2\right]^2\right\} \tag{6.4c}$$

where $\gamma_{\mathrm{C}}^2 \triangleq \mathrm{E}\{|a(k)|^4\}/\mathrm{E}\{|a(k)|^2\}$ [49], $\gamma_{\mathrm{R}} \triangleq \mathrm{E}\{a(k)^2\}/\mathrm{E}\{|a(k)|\}$, $\mathrm{csgn}(\cdot) = \mathrm{sgn}(\mathrm{Re}(\cdot)) + \sqrt{-1}\cdot\mathrm{sgn}(\mathrm{Im}(\cdot))$ is the complex signum operator, and $\gamma_{\mathrm{M}}^2 \triangleq \mathrm{E}\{\mathrm{Re}(a(k))^4\}/\mathrm{E}\{\mathrm{Re}(a(k))^2\}$. Wrong rotated solutions should not be confused with the undesirable local minima in the channel-equalizer parameter space which will result in incomplete ISI removal [26, 31, 78]. Wrong rotated solutions, on the other hand, are solutions obtained even after the channel eye has been open except that its output signals are out of phase with its original QAM constellation. Note that modulo 90° is acceptable because such solutions may be detected and corrected with differential encoding techniques. Consequently with disorientated solutions, the MSE is not minimized. Thus such solutions present the notion of local minima in the *phase* parameter space, as opposed to local minima in the combined channel-equalizer parameter space.

Fig. 6.2 shows three wrong solutions that may be encountered during lab sim-

ulations, where Fig. 6.2-A, B show two possible wrong solutions at angles of 21° and 45°, respectively, when the CMA and the RCA are employed in the presence of a constant phase offset. Fig. 6.2-C shows a wrong solution of the CMA under non-stationary phase offsets. Even though the eye is clearly open for each of the three cases (three distinct circular lines in Fig. 6.2-C represent a well open eye condition), the MSE is not minimized due to the disorientation of the output constellation. It is also interesting to note that none of the wrong solutions has been reported when the MMA is employed. The torque analysis that proceeds will explain this phenomenon.

6.3.2 Preliminaries on Torque Concepts and Some Definitions

Unlike conventional approaches, we have resorted to using torque to analyze the rotational behavior of the output constellation under the adaptation of the tap weights. This is because the output constellation may be represented by an object, in the physical sense, that is being applied a net rotational force which causes rotation about its origin. Thus this translates into a problem that can be readily addressed by the concept of (mechanical) torque.

Torque is a measure of how much a force acting on an object causes that object to rotate. The object rotates about an axis called the *pivot point* or simply, the *origin*. The distance from the origin to the point where the force acts is called the *moment arm*. In our context, the origin of the object is in fact the origin of the QAM constellation. The object that is being rotated is the equalizer output vector

$$\overrightarrow{z(k)} = \big[\mathrm{Re}(z(k)), \mathrm{Im}(z(k))\big]. \tag{6.5}$$

Note that the moment arm is a vector. The acting force is the error function vector

$$\overrightarrow{\epsilon(k)} = \big[\mathrm{Re}(\epsilon(k)) - \mathrm{Re}(z(k)), \mathrm{Im}(\epsilon(k)) - \mathrm{Im}(z(k))\big]. \tag{6.6}$$

Thus, the effective torque applied on the output signal is defined as

$$\mathrm{T}(k) = \overrightarrow{z(k)} \times \overrightarrow{\epsilon(k)} = \overrightarrow{z(k)} \;\; \overrightarrow{\epsilon(k)} \;\; \sin\varphi \tag{6.7}$$

where $\times$ denotes the cross product of two vectors and φ is the angle between the two vectors. We now state the condition for rotational equilibrium.

Under rotational equilibrium, the sum of the torques acting on the object must

be equal to zero. In our context, there is only one force, $\overrightarrow{\epsilon(k)}$, acting on one moment arm, $\overrightarrow{z(k)}$, at any one instant. The moment arm and its associated force, however, are distributed over the constellation space. Thus, the condition for rotational equilibrium is for the expected torque value to be zero, i.e.,

$$\mathrm{E}\{\mathrm{T}(k)\} = 0. \tag{6.8}$$

6.3.3 Rotational Behavior of the CMA

Using the concept of torque, we will show that the CMA is a phase-blind algorithm, i.e., it is insensitive to the phase of its output. In other words, we will verify that all angles for the output signal constellation correspond to marginally stable equilibria. The error function of the CMA is

$$\epsilon_{\mathrm{CMA}}(k) = z(k)\big(|z(k)|^2 - \gamma_{\mathrm{C}}^2\big) \tag{6.9}$$

where the noisy phase rotated equalizer output (6.2b) given here again is

$$z(k) = a(k)e^{j\Delta\theta(k)} + v(k). \tag{6.10}$$

The force that acts on the moment arm is

$$\begin{aligned}
\overrightarrow{\epsilon(k)} &= \Big(\mathrm{Re}(z(k))\big(|z(k)|^2 - \gamma_{\mathrm{C}}^2\big) - \mathrm{Re}(z(k)), \\
&\qquad \mathrm{Im}(z(k))\big(|z(k)|^2 - \gamma_{\mathrm{C}}^2\big) - \mathrm{Im}(z(k))\Big) && \text{(6.11a)} \\
&= \big(|z(k)|^2 - \gamma_{\mathrm{C}}^2 - 1\big)\Big(\mathrm{Re}(z(k)), \mathrm{Im}(z(k))\Big) && \text{(6.11b)} \\
&= scalar \cdot \overrightarrow{z(k)}. && \text{(6.11c)}
\end{aligned}$$

The torque is

$$\begin{aligned}
\mathrm{T}(k) &= \overrightarrow{z(k)} \times \big(scalar \cdot \overrightarrow{z(k)}\big) && \text{(6.12a)} \\
&= scalar \cdot \big(\overrightarrow{z(k)} \;\; \overrightarrow{z(k)}\big) \sin 0 && \text{(6.12b)} \\
&= 0, \quad \forall z(k) && \text{(6.12c)}
\end{aligned}$$

because the cross product of two vectors in the same direction is zero (6.12a), which is also reflected by the term, $\sin 0$, in (6.12b). Since $z(k)$ is a function of the residual phase errors, $\Delta\theta(k)$, this implies that the rotational equilibrium is achieved for all

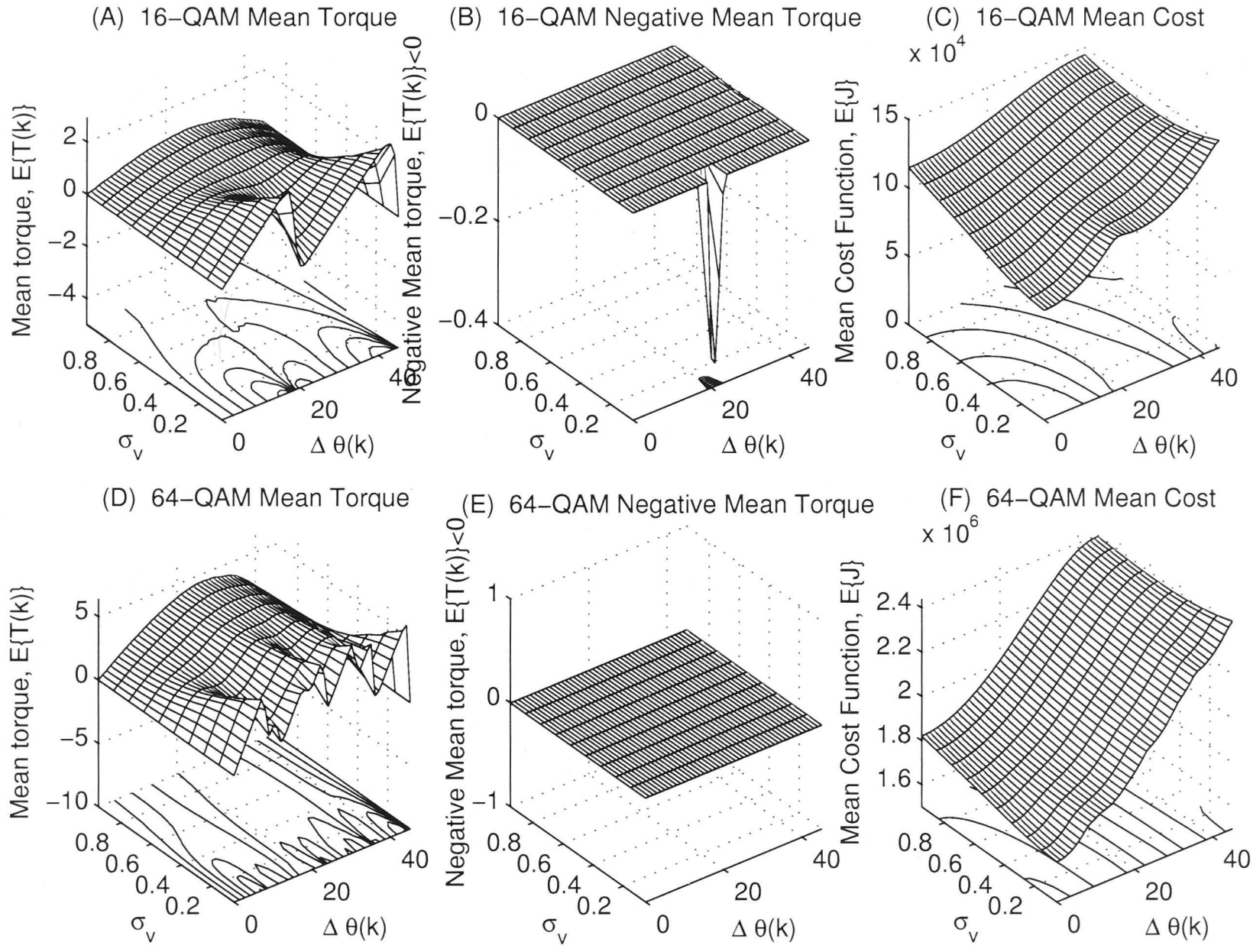

Figure 6.3: Rotational behavior of RCA for square-QAM constellations. The non-flat region in Fig. (B) indicates that there is an undesirable but rotationally stable equilibrium point which is at the vicinity of 21°. This is also reflected in Fig. (C) where a shallow local minimum is observed for $\sigma_v \approx 0$ and $\Delta\theta(k) \approx 21°$.

phase errors, i.e.,

$$\mathrm{E}\{\mathrm{T}(k)\} = 0, \quad \forall \Delta\theta(k). \tag{6.13}$$

In other words, there is no mean torque that acts on the output signal constellation for all values of $\Delta\theta(k)$ because the force that acts on $\overrightarrow{z(k)}$ will go through the origin, just like $\overrightarrow{z(k)}$ itself. This result supports the fact that the CMA is indeed phase-blind.

6.3.4 Rotational Behaviors of RCA and MMA

We will study the rotational behaviors of the RCA and the MMA by examining the mean torque acting on the output constellation for different phase errors, SNR levels, and constellation sizes. Rather than being vigorous in our analysis, we computed the cross products of large samples of equalizer output vectors, $\overrightarrow{z(k)}$, and

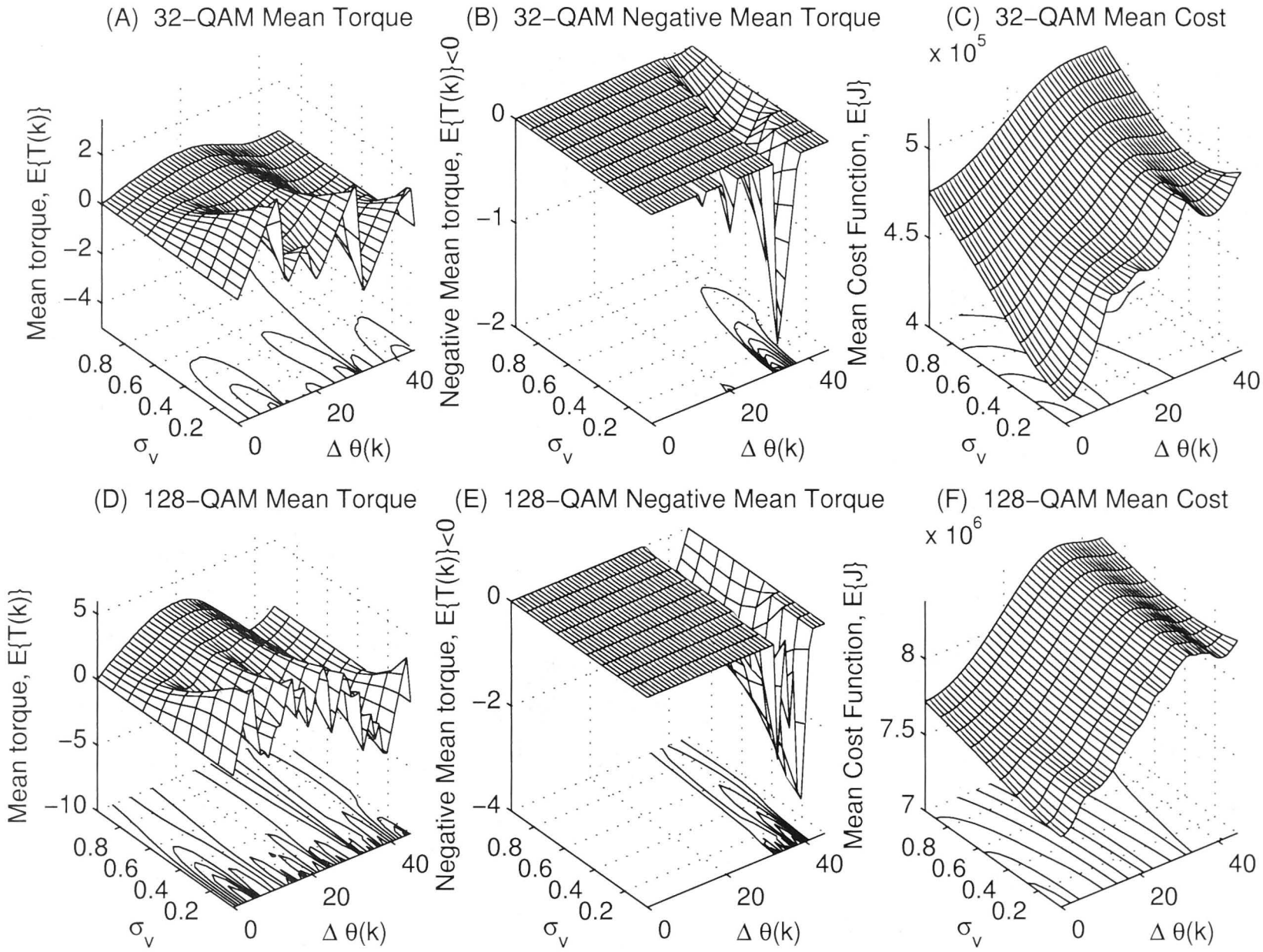

Figure 6.4: Rotational behavior of RCA for cross-QAM constellations. The large non-flat regions in Figures (B) and (E) at the vicinity of 45° indicate that the RCA is very susceptible to the '45°' wrong solutions. This is also reflected in the cost figures (C), (F) where a local minimum point clearly exists at 45°. There are two more stable points for 32-QAM as shown in Figure (B) at high SNR and around 13° and 21°.

their associated error function vectors, $\overrightarrow{\epsilon(k)}$, to obtain their mean torques. In addition, we also computed their mean costs, $\mathrm{E}\{J\}$, under the following environments:

i. A phase error of 0° to 45° at intervals of 0.5°.

ii. A standard deviation, σ_v, of the effective noise which is assumed to be Gaussian, from 0.05 to 1 at intervals of 0.05.

iii. Constellation sizes of $16, 32, 64, 128$, where $16, 64$ are square QAM constellations, while $32, 128$ are cross QAM constellations.

The empirical results using torque analysis regarding the rotational equilibria of the RCA and the MMA are shown in Fig. 6.3, 6.4, 6.5. In order to meaningfully interpret the figures, the following points must be understood first:

a) Due to the symmetry of the QAM data constellation about the x- and y-axes, the torque figures are identical for $0^\circ < \Delta\theta(k) < 45^\circ$, $90^\circ < \Delta\theta(k) < 135^\circ, \cdots$, while they are mirrored about the x-axis (i.e., the positive mean torque becomes the negative mean torque) for $45^\circ < \Delta\theta(k) < 90^\circ$, $135^\circ < \Delta\theta(k) < 180^\circ, \cdots$. Therefore, the results for the range $0^\circ \leq \Delta\theta(k) \leq 45^\circ$ can completely describe the rotational behavior of the algorithm.

b) By the definition of $\overrightarrow{z(k)}$ in (6.5) and $\overrightarrow{\epsilon(k)}$ in (6.6), a positive mean torque will result in a clockwise rotation in the constellation. Conversely, a negative mean torque will result in an anti-clockwise rotation. If you will refer to the figures of mean torque, a positive mean torque 'pushes' $\Delta\theta(k)$ towards the equilibrium point[1] on its *left*. A negative torque 'pushes' $\Delta\theta(k)$ towards the equilibrium point on its *right*. *Therefore, an undesirable equilibrium point between* 0° *and* 45° *exists if a negative mean torque exists over this range.* This is why we have also plotted the negative mean torque figures to ease the identification of such undesirable equilibria.

c) Not all equilibrium points are stable. For the range $0^\circ < \Delta\theta(k) < 45^\circ$, a stable equilibrium is one that must experience a negative mean torque on its left and a positive mean torque on its right. Take for example the 32-QAM for RCA and observe Figure 6.4-(B). The 45° solution is stable for $\sigma_v = 0.5$. However, at $\sigma_v = 0.1$, the stable angle corresponds to approximately 37°.

We now present the conclusions drawn from observing Fig. 6.3, 6.4, 6.5.

1) Square (16- and 64-) QAM constellations: Both algorithms exhibit little susceptibilities to wrong solutions, except for the RCA at very high SNR's and only for 16-QAM. There exists one stable equilibrium at approximately 21° and $\sigma_v = 0.1$.

2) Cross (32- and 128-) QAM constellations: The RCA, as opposed to the MMA, is much more likely to converge to the 45° solution for both 32- and 128-QAM sizes. Once trapped in these solutions, they are hard to 'escape'. This is because of the large area of negative mean torque that spans approximately 11° to the left of 45° and over all shown values of σ_v. Thus, once trapped in the 45° solution, a large amount of torque will 'push' the signal constellation back to the wrong solution if the constellation is rotated within 11° from 45°. Hence they are highly stable solutions. It is no wonder a high rate of undesirable convergence to 45° solutions has been observed in lab experiments [152].

[1] An equilibrium point corresponds to zero mean torque.

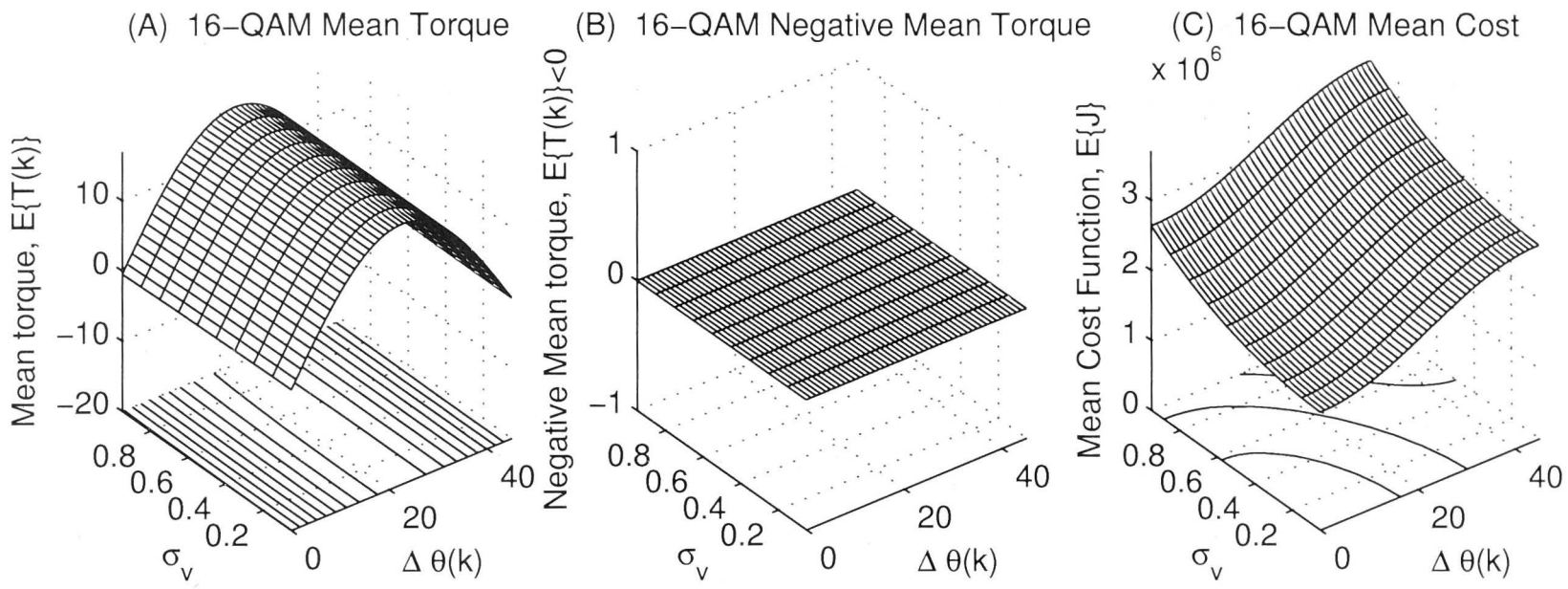

Figure 6.5: Rotational behavior of MMA for 16-QAM constellation. Identical-shaped figures are observed for 32-, 64- and 128-QAM, so they are not plotted. Figure (B) shows no signs of a negative mean torque and this is true also for 32-, 64- and 128-QAM. This phenomenon is also reflected in the cost function which does not have any undesirable local minimum.

3) The mean torque of the MMA that is always positive for all σ_v and $\Delta\theta(k)$ shows that a rotational force always pushes the output signal constellation back to its correct 0^o solution. Therefore, the MMA is not possible to converge to a wrong solution. In addition the mean torque has a constant shape even when the σ_v varies. This is a very interesting property which makes MMA a very robust blind phase recovery algorithm as it is not only immune to 45^o solutions, but its mean torque is also insensitive to fluctuations in the SNR.

Based on this torque analysis we therefore conclude that the MMA is far more superior than the RCA in terms of its convergence immunity to wrong solutions for QAM systems in both high and low noise environments.

6.4 Conclusions

The rotational behavior of gradient-type blind algorithms, namely, the CMA, the RCA and the MMA can be studied using the concept of torque. The CMA is shown to be phase invariant through the use of the torque analysis. As for the RCA, it is shown to suffer from potential wrong solutions of all sorts of angles especially for cross-QAM constellations. These wrong solutions are also highly stable for the RCA. The MMA on the other hand, is shown to be superior in terms of its immunity to undesirable rotated solutions in both high and low noise environments.

In this chapter, we have presented a new way to analyze the rotational behavior of blind algorithms which can be thought of as the local minima analysis in the

phase parameter space. Therefore, when it comes to selecting a blind algorithm for the purpose of blind equalization and phase recovery, the MMA is definitely the preferred choice over the RCA due to its superior phase recovery properties. From this analysis, we revealed that the MMA is the most suitable algorithm among the three algorithms to be used.

Chapter 7

Conclusions

In this chapter, we present a summary of the research findings of the thesis and the contributions found in each of the chapters. Also, guided by insights gained in conducting the research found in the thesis, we propose a program for further research.

7.1 Executive Summary

In this thesis, we have investigated, analyzed, and developed switch-mode equalization schemes with the global goals of fast convergence and smooth switching in mind. We have developed a new reliability concept based on Bayes theorem that is used to combine the acquisition and the tracking algorithms. This combination methodology is very flexible and can even be extended to combine algorithms based on different equalizer structures such as a linear equalizer and a decision feedback equalizer. We then focussed on developing simpler formulations of switching criterion inspired by our reliability combining concept. The new switch-mode algorithm that employs this switching criterion achieves very fast convergence and smooth switching at practically no additional cost. Finally, considered a novel analysis method for determining the phase/frequency locking ("rotational") behavior of blind algorithms and applied it to several blind algorithms to test their robustness and immunity to to wrongly rotated stable equilibria.

In Chapter 3, we developed a new reliability based switch-mode algorithm whereby the acquisition and the tracking algorithms are combined using a convex reliability parameter. This parameter has been derived using Bayes theorem and is found to be dependent on both the equalizer output and its estimated statistical distribution. Unlike conventional techniques found in the literature that employ heuristically chosen measures which are usually dependent on only the equalizer

output or other measures, our new technique (that depends on both equalizer output and its distribution) was shown to exhibit significant performance improvement over conventional techniques.

In Chapter 4, we developed a new switching criterion that is based on the probability of an equalizer output being found in square regions that enclose the constellation points. The novelty of this technique lies in the dynamically varying size of these square regions. The concept parallels the reliability measure developed in Chapter 3 whereby the combination of the acquisition and tracking algorithms is dependent not only on the equalizer output, but its estimated statistical distribution as well. This algorithm, unlike conventional algorithms including the reliability-based algorithm in Chapter 3, achieves very rapid convergence, low steady state error and automatic phase recovery. In addition it has the advantage, particularly in a practical setting, that it is very simple to implement with complexity comparable to conventional "hard-switching" techniques.

In Chapter 5, we developed a new DFE scheme that extends the work of of Labat, Macchi and Laot. The most important aspect of the new strategy is that it combines the linear equalizer and the DFE of this equalization scheme in the same way we combine the acquisition and tracking algorithms in Chapters 3 and 4. Once again, fast convergence and smooth switching between modes are achieved.

In Chapter 6, we proposed an alternative tool to analyze the phase/frequency locking ("rotational") behaviors of several blind algorithms. Specifically, we verified that the constant modulus algorithm (CMA) is rotationally invariant. The most important and somewhat surprising result is that the multi-modulus algorithm developed by Oh and Chin, and independently by Werner and Yang, is superior that the reduced constellation algorithm (RCA) developed by Benveniste *et al*, in terms of the algorithm's immunity to wrongly rotated stable equilibria for QAM signals.

7.2 Future Work

7.2.1 Direct Extensions

In terms of more direct extensions to the work there are a number of which we indicate a few. The new probabilistic switching criterion in Chapter 4 has been developed for QAM signals with the assumption that the effective noise[1] is a complex Gaussian noise term which is independent of the data signal. It can be extended to

[1]Effective noise is defined as the sum of the residual ISI and additive noise.

other data formats such as the real PAM format. It can also be extended to involve analysis that treats the effective noise as non-Gaussian and non-independent of the data signals, which is usually true in practise.

7.2.2 Generalizations

Dual mode (and possibly multi-mode) equalizers can be viewed as specific instances of a more general class of equalizers. In their most basic form this type of equalizer exists in two modes and there is either a hard switching or soft switching between the modes. The two modes differ in the filtering structure and the type of adaptive algorithm they use. This thesis has focussed on how to make this handover between the acquisition mode and the tracking mode as quick and smooth as possible.

The ability of an equalizer to switch structure and algorithm is in general a very powerful notion but the problem is not well studied. Generally a channel will be time-varying and a given equalizer can expect to have to operate over a variety of channels with varying degrees of difficulty. To be successful an equalizer needs to have a complexity comparable to the complexity of the channel dynamics. Hence when a channel is simple the equalizer should be simple and lowly parameterized. Conversely a complex channel would require an equalizer of many parameters.

The problem that needs to be addressed is how to have an equalizer change its structure automatically in response to the complexity of the channel and to do so in a blind way. Married with the changing structure is a requirement to change the associated algorithm. In this way the limited number of modes of today's equalizers (usually one or two modes) need to move towards a continuum of modes in a more general setting and change in a smooth way that tracks the channel complexity variations. Adaptation is no longer an issue of varying the parameter values in a fixed structure but also in addition varying the structure, parametrization and number parameters.

Appendix A

Computation of mean of absolute value of error functions, $\mathrm{E}\{|\epsilon(k)|\}$

The switch-mode algorithms that we deal with in Chapters 3, 4, 5 usually consists of two (or more) algorithms that will be switched from one to another at some point in time when the equalizer output is believed to be sufficiently reliable. The magnitudes of the error functions can be orders apart and thus they need to be scaled to yield same orders of error functions to ensure a smooth transition. In this Appendix, we will consider the minimum output energy (MOE) or the whitening algorithm, the constant modulus algorithm (CMA), and the *blind* least-mean-square (LMS) algorithm. We assume the input signal to the transversal equalizer is $r(k)$ and its output $z(k)$. The weight vector of the equalizer is denoted as $\mathbf{w}(k)$.

Firstly, we consider the MOE algorithm. Its cost function is

$$J_{\mathrm{MOE}}(k) \triangleq \mathrm{E}\{|z(k)|^2\} \tag{A.1}$$

and its error function, i.e., the derivative of the cost function and taking away the expectation operator, is therefore

$$\epsilon_{\mathrm{MOE}}(k) \triangleq z(k). \tag{A.2}$$

Thus, the mean of the absolute value of (A.2) is easily computed once the probability density function (p.d.f.) of $z(k)$ is known. The equalizer output can be expressed as the sum of the desired signal (originally transmitted data signal) and an *effective noise* term which consists of the residual ISI plus colored additive channel noise as in (2.4). In discrete time, it can be expressed as in (3.4) as follows:

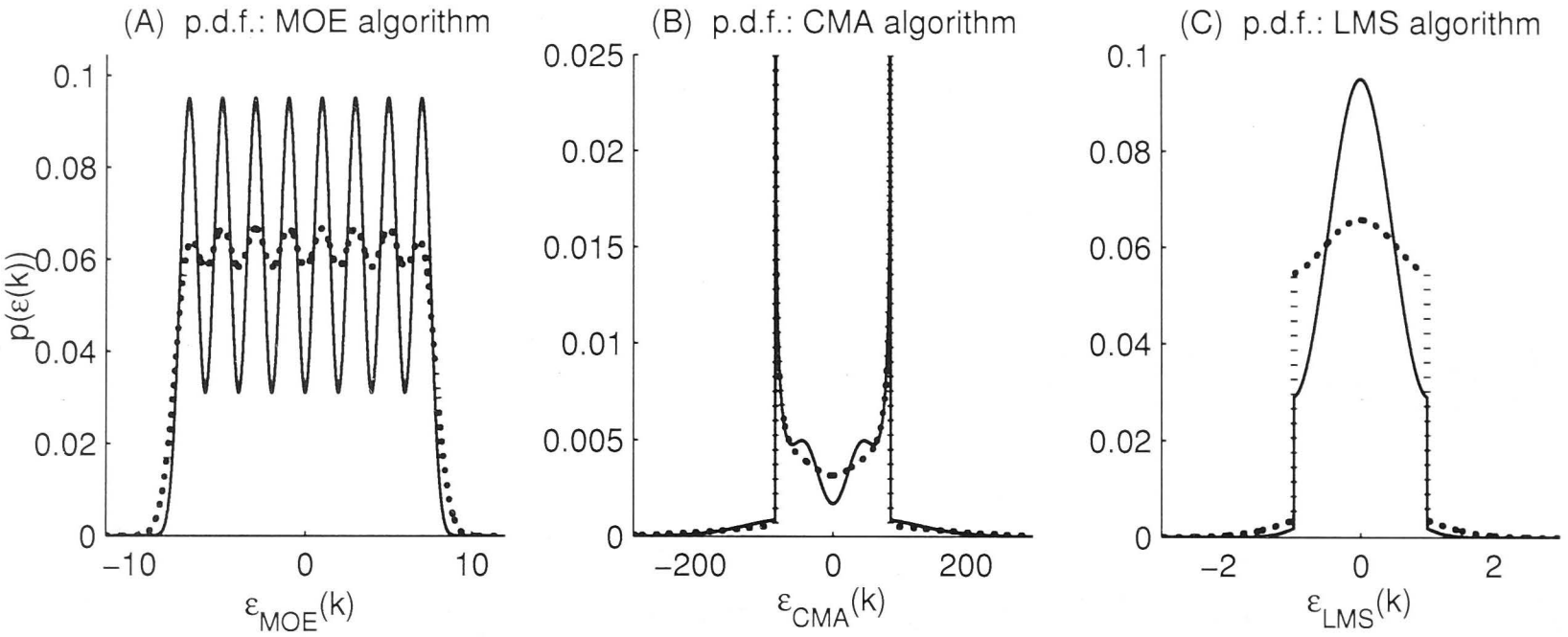

Figure A.1: The probability density functions (p.d.f.) of the error functions of the MOE, CMA, and the LMS algorithms. The solid lines represent the c.d.f. for $\sigma_v^2 = 0.28$ and the dotted lines for $\sigma_v^2 = 0.69$.

$$z(k) = \sum_{n=-N}^{N} c_n r(k-n) \tag{A.3a}$$

$$= s_0 a(k) + v(k) \tag{A.3b}$$

where

$$v(k) = \sum_{j\neq 0, j=-N}^{N+L} s_j a(k-j) + \sum_{n=-N}^{N} c_n \eta(k-n) \tag{A.4}$$

is the so-called *effective noise* with a variance of σ_v^2, and $\{s_j\}, j = -N, \cdots, N+L$ is the set of coefficients of the combined channel-equalizer filter. The p.d.f. of $z(k)$ therefore is the joint p.d.f. of the marginal p.d.f. of the transmit signal and the effective noise, given as

$$p(z(k)) = p(a(k), v(k)) = p(a(k)) \cdot p(v(k)) \tag{A.5}$$

where $p(\cdot)$ stands for the probability density function, and the second term to the third term is due to the assumption that $a(k)$ and $v(k)$ are statistically independent. With the simplifying assumption of Gaussian effective noise, we therefore obtain the following closed form expressions for the p.d.f. for both real data signals, which are one dimensional, and complex data signals, which are two dimensional, as follows:

$$\text{Real}: \quad p(z(k)) = \frac{1}{M}\sum_{j=1}^{M}\frac{1}{\sqrt{2\pi}\sigma_v}\exp\left[-\frac{(z(k)-d_j)^2}{2\sigma_v^2}\right] \tag{A.6a}$$

$$\text{Complex}: \quad p(z(k)) = \frac{1}{M}\sum_{j=1}^{M}\frac{1}{2\pi\sigma_v^2} \exp\left[-\frac{[\mathrm{Re}(z(k))-\mathrm{Re}(d_j)]^2+[\mathrm{Im}(z(k))-\mathrm{Im}(d_j)]^2}{2\sigma_v^2}\right] \tag{A.6b}$$

where d_j is the j^{th} alphabet in the data alphabet set. Therefore, we can plot the p.d.f. of the error function of the MOE algorithm, and from there evaluate the mean of its absolute value. This is shown in Fig. A.1-A where the real 8-PAM data is used assuming $\sigma_v^2 = 0.28$ and $\sigma_v^2 = 0.69$.

Next, we consider the CMA. Its cost function is

$$J_{\text{CMA}}(k) \triangleq \mathrm{E}\{(|z(k)|^2 - \gamma_{\text{C}}^2)^2\} \tag{A.7}$$

where γ_{C}^2 is the dispersion constant (2.51). Its error function is therefore

$$\epsilon_{\text{CMA}}(k) \triangleq z(k)(|z(k)|^2 - \gamma_{\text{C}}^2). \tag{A.8}$$

To obtain the p.d.f. of $\epsilon_{\text{CMA}}(k)$, it is possible to perform the p.d.f. transformation of $z(k)$, since we have closed form expression of $p(z(k))$ as in (A.6) according to [118, Ch. 1]. The p.d.f. transformation requires that

$$\epsilon_{\text{CMA}}(k) = z(k)(|z(k)|^2 - \gamma_{\text{C}}^2) \tag{A.9}$$

be solved, such that $\{z_1^*(k), z_2^*(k), \cdots, z_N^*(k)\}$ are the N *real* roots of the above equation of (A.9) which are functions of $\epsilon_{\text{CMA}}(k)$. It can be shown for real PAM data formats, $N = 3$, i.e., there are three real roots, when

$$|\epsilon_{\text{CMA}}(k)| \leq \pm\frac{\gamma_{\text{C}}^2}{3}\left(\sqrt{3\gamma_{\text{C}}^2} - \sqrt{\frac{\gamma_{\text{C}}^2}{3}}\right) \tag{A.10}$$

and $N = 1$ otherwise. This will be explained shortly afterwards. Thus, the p.d.f.

of $\epsilon_{\mathrm{CMA}}(k)$ may be expressed as [118]

$$p(\epsilon_{\mathrm{CMA}}(k)) = \sum_{j=1}^{N} \frac{p(z_j^*(k))}{|\epsilon'_{\mathrm{CMA}}(z_j^*(k))|} \tag{A.11}$$

where

$$\epsilon'_{\mathrm{CMA}}(z_j^*(k)) = 3[z_j^*(k)]^2 - \gamma_{\mathrm{C}}^2. \tag{A.12}$$

The p.d.f. plot of $\epsilon_{\mathrm{CMA}}(k)$ is shown in Fig. A.1-B for an 8-PAM data format with $\sigma_v^2 = 1$. It is clear from the figure that the p.d.f. is larger for $|\epsilon_{\mathrm{CMA}}(k)| \leq 86.63$ due to (A.10), where $\gamma_{\mathrm{C}}^2 = 37$. In fact, at $|\epsilon_{\mathrm{CMA}}(k)| = 86.63$, the p.d.f. tends to be much larger because at these points, the denominator of (A.11), $|\epsilon'_{\mathrm{CMA}}(z_j^*(k))|$, tends to zero and the p.d.f. thus tends to infinity.

Equation (A.10) is obtained by finding the values of $\epsilon_{\mathrm{CMA}}(k)$ that correspond to $3z^2(k) - \gamma_{\mathrm{C}}^2 = 0$. The $z(k)$ value is found to be $\pm\sqrt{\gamma_{\mathrm{C}}^2/3}$. By substituting this value into $\epsilon_{\mathrm{CMA}}(k)$ of (A.9), we thus arrive at equation (A.10).

Lastly, we consider the blind LMS algorithm. It is rather difficult to obtain closed form expression for the p.d.f. of the LMS algorithm. Therefore, we generated large amount of data samples and plotted its p.d.f. as shown in Fig. A.1-C for the same 8-PAM data format and $\sigma_v^2 = 0.28$ and $\sigma_v^2 = 0.69$.

Conclusions drawn from observation of the figures in Fig. A.1 include :

1. The magnitudes of the respective error functions are indeed orders apart for 8-PAM. In fact, their differences increase as the constellation size increases, say for 32-PAM compared to 8-PAM.

2. The "shapes" of the p.d.f. of $\epsilon_{\mathrm{MOE}}(k)$ and $\epsilon_{\mathrm{LMS}}(k)$ are rather similar, suggesting that a switch between these two algorithms would be smooth under appropriate scaling of the error function. The "shapes" of the p.d.f. of $\epsilon_{\mathrm{CMA}}(k)$ and $\epsilon_{\mathrm{LMS}}(k)$ are rather different, suggesting that even under appropriate scaling of $\epsilon_{\mathrm{LMS}}(k)$, the switching between these algorithms may not be smooth.

The means of the absolute value of the error functions can be computed by evaluating the cumulative distribution function (c.d.f.) of the $\epsilon(k)$ [118, Ch. 1], defined as

$$D(x) = \Pr(\epsilon(k) \leq x) \qquad -\infty < x < \infty \tag{A.13}$$

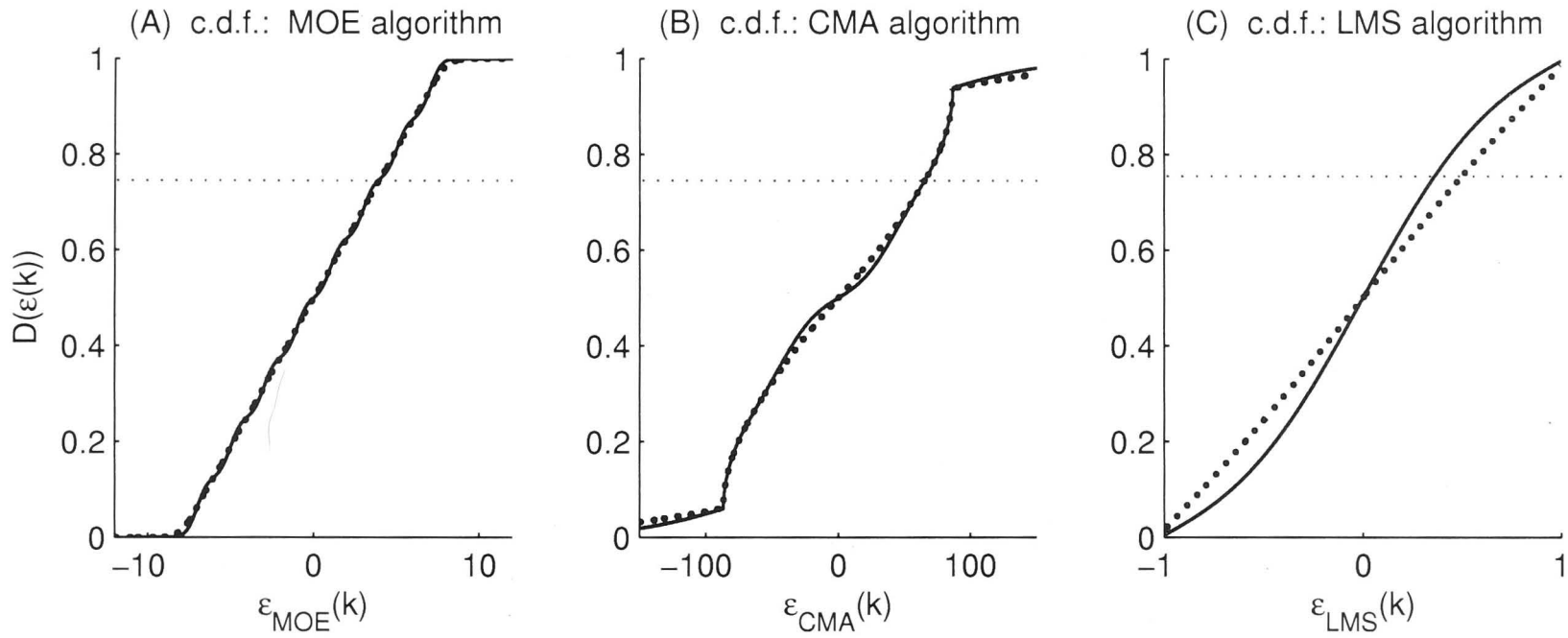

Figure A.2: The cumulative distribution functions (c.d.f.) of the error functions of the MOE, CMA, and the LMS algorithms. The solid lines represent the c.d.f. for $\sigma_v^2 = 0.28$ and the dotted lines for $\sigma_v^2 = 0.69$. The horizontal dotted lines show $D(\epsilon(k)) = 0.75$.

where x is any real number in the interval $(-\infty, \infty)$. Note that the mean of the error function is obtained when $D(\epsilon(k)) = 0.5$, but the mean of the *absolute value* of the error function is obtained when $D(\epsilon(k)) = 0.75$. This is shown in Fig. A.2.

Appendix B

Validity of (5.18) due to parallel adaptation

In this appendix, we show that via parallel adaptation, the output signals of the "top" and "bottom" paths of the new DFE in Section 5.4 are almost similar in value, but differs in a manner that is proportional to the adaptation step sizes of the respective filters.

Refer to Fig. 5.5 for the notations of the signals and filters. Consider first the linear equalizer of the "top" path. The output signal of the non-recursive whitening filter is

$$t^{(1)}(k) = \mathbf{A}(k)\mathbf{S}^T(k) \tag{B.1a}$$

$$= [1, a_1(k), a_2(k), \cdots][s(k), s(k-1), \cdots]^T. \tag{B.1b}$$

The output signal of the transversal (Godard) equalizer is

$$u^{(1)}(k) = \mathbf{B}(k)\mathbf{T^{(1)}}^T(k) \tag{B.2a}$$

$$= [b_1(k), b_2(k), \cdots][t^{(1)}(k), t^{(1)}(k-1), \cdots]^T \tag{B.2b}$$

$$= [b_1(k)\mathbf{A}(k)\mathbf{S}^T(k), b_2(k)\mathbf{A}(k-1)\mathbf{S}^T(k-1), \cdots,$$

$$b_{N+M}(k)\mathbf{A}(k-N-M+1)\mathbf{S}^T(k-N-M+1)]. \tag{B.2c}$$

The output signal $u^{(1)}(k)$ depends on $(N+M)$ previous input signals as shown in (B.5c).

Consider now the DFE in the "lower" path of the new DFE under parallel adaptation. In order to arrive at meaningful results, we "linearize" the DFE in the "lower" path by ignoring the quantizer for the following derivation. In that case, the transformed $\mathcal{W}$ of the DFE will have exactly the same transfer function

as the $\mathcal{W}$ of the linear equalizer, i.e., $1 + A(z)$. The output signal of the transversal (Godard) equalizer is

$$t^{(2)}(k) = \mathbf{B}(k)\mathbf{S}^T(k) \tag{B.3a}$$

$$= [b_1(k), b_2(k), \cdots][s(k), s(k-1), \cdots]^T. \tag{B.3b}$$

After the phase rotator, the signal and its vector becomes

$$u^{(2)}(k) = t^{(2)}(k)e^{-j\theta(k)} \tag{B.4a}$$

$$\mathbf{U^{(2)}}(k) = \mathbf{T^{(2)}}(k)e^{-j\theta(k)}. \tag{B.4b}$$

The equalizer output (ignoring the effect of the quantizer due to linearization), eventually, becomes

$$z^{(2)}(k) = \mathbf{A}(k)\mathbf{U^{(2)}}^T(k) \tag{B.5a}$$

$$= [1, a_1(k), a_2(k), \cdots][u^{(2)}(k), u^{(2)}(k-1), \cdots]^T \tag{B.5b}$$

$$= [1 \cdot \mathbf{B}(k)\mathbf{S}^T(k)e^{-j\theta(k)}, a_1(k)\mathbf{B}(k-1)\mathbf{S}^T(k-1)e^{-j\theta(k-1)}, \cdots,$$
$$a_{N+M-1}(k)\mathbf{B}(k-N-M+1)\mathbf{S}^T(k-N-M+1)e^{-j\theta(k-N-M+1)}]. \tag{B.5c}$$

Recall that Equation (5.18) states that

$$u^{(1)}(k)e^{-j\theta(k)} \approx z^{(2)}(k). \tag{B.6}$$

From (B.2) and (B.5), we get

$$u^{(1)}(k)e^{-j\theta(k)} = [b_1(k)\mathbf{A}(k)\mathbf{S}^T(k), b_2(k)\mathbf{A}(k-1)\mathbf{S}^T(k-1), \cdots,$$
$$b_{N+M}(k)\mathbf{A}(k-N-M+1)\mathbf{S}^T(k-N-M+1)]e^{j\theta(k)} \tag{B.7a}$$

$$z^{(2)}(k) = [1 \cdot \mathbf{B}(k)\mathbf{S}^T(k)e^{-j\theta(k)}, a_1(k)\mathbf{B}(k-1)\mathbf{S}^T(k-1)e^{-j\theta(k-1)}, \cdots,$$
$$a_{N+M-1}(k)\mathbf{B}(k-N-M+1)\mathbf{S}^T(k-N-M+1)e^{-j\theta(k-N-M+1)}]. \tag{B.7b}$$

After a close examination of (B.7a) and (B.7b), it is clear that under the condition of *non-adaptive* equalization, both terms are exactly equal. The difference in magnitude of these two terms increases as the step sizes of the adapting filters are increased. If $\Pr\{Q(z^{(2)}(k)) = a(k)\}$ is high enough, i.e., the output signal is sufficiently reliable, then incorporating the non-linear quantizer would still yield output signals that are almost similar from both the "top" and "bottom" paths.

Bibliography

[1] K. Abed-Meraim, P. Duhamel, D. Gesbert, P. Loubaton, S. Mayrargue, E. Moulines, and D. Slock, "Prediction error methods for time-domain blind identification of multichannel FIR filters," *Proc ICASSP*, vol. 3, pp. 1968–1971, 1995.

[2] B. Agee, "The least-squares CMA: A new technique for rapid correction of Constant Modulus signals," *Proc. ICASSP 86 Conference, Tokyo, Japan.*, pp. 953–956, 1986.

[3] B. G. Agee, "The property restoral approach to blind signal extraction," *Ph.D. Dissertation, University of California, Davis, CA*, 1989.

[4] N. Al-Dhahir and J. M. Cioffi, "MSE decision-feedback equalizers: finite-length results," *IEEE Trans. Information Theory*, vol. 41, no. 4, Jul 1995.

[5] G. Ananthaswamy and D. L. Goeckel, "A fast acquiring blind predictive DFE," *IEEE Trans. Commun.*, vol. 50, no. 10, pp. 1557–1560, Oct 2002.

[6] Arsenault and Roberts, *Reliability and Maintainability of Electronic Systems*, 1980.

[7] M. E. Austin, "Decision-feedback equalization for digital communication over dispersive channels," *MIT Lincoln Laboratory, Lexington, Mass., Tech. Report*, no. 437, Aug 1967.

[8] R. A. Axford, L. B. Milstein, and J. R. Zeidler, "A dual-mode algorithm for blind equalization of QAM signals: CADAMA," *Proceedings of the 29th Asilomar Conference on Signals, Systems and Computers*, 1995.

[9] B. F. Barrett and D. G. Lampard, "An expansion for some second-order probability distributions and its applications to noise problems," *IRE - Trans. Inform. Theory*, pp. –, Mar 1955.

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

[11] S. Bellini, "Bussgang techniques for blind equalization," *GLOBECOM*, pp. 1634–1640, 1986.

[12] ——, "Blind equalization and deconvolution," *SPIE Proc*, vol. 1565, pp. 88–101, Jul 1991.

[13] T. R. Benedict and T. T. Soong, "The joint estimation of signal and noise from the sum envelope," *IEEE Trans. Information Theory*, vol. IT-13, pp. 447–454, Jul 1967.

[14] A. Benveniste, M. Goursat, and G. Ruget, "Robust identification of a non-minimum phase system: Blind adjustment of a linear equalizer in data communication," *IEEE Trans. on Automatic Control*, vol. 25, no. 3, pp. 385–399, 1980.

[15] A. Benveniste and M. Goursat, "Blind equalizers," *IEEE Trans. Commun.*, vol. COM-32, no. 8, pp. 871–883, Aug 1984.

[16] D. R. Brillinger, *Time Series: Data Analysis and Theory.* Holt, Rinehart, and Winston, New York, 1974.

[17] R. A. Casas, Z. Ding, R. A. Kennedy, C. R. Johnson, Jr., and R. Malamut, "Blind adaptation of decision feedback equalizers based on the constant modulus algorithm," *Conference Record of the Twenty-Ninth Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 698–702, Oct 1995.

[18] Y. Chen, C. L. Nikias, and J. G. Proakis, "Blind equalization with criterion with memory nonlinearity," *Optical Engeineering*, vol. 31, no. 6, pp. 1200–1210, Jun 1992.

[19] M. Chiani, "Introducing erasures in decision-feedback equalization to reduce error propagation," *IEEE Trans. Commun.*, vol. 45, no. 7, pp. 757–760, Jul 1997.

[20] C. A. F. da Rocha and O. Macchi, "A novel self-learning adaptive recursive equalizer with unique optimum for QAM," *Proc. ICASSP-94, Adelaide, Australia*, pp. III–481–484, Apr 1994.

[21] C. A. F. da Rocha, O. Macchi, and J. M. T. Romano, "An adaptive nonlinear IIR filter for self-learning equalization," *Proc. IEEE International Telecommunications Symp., Rio de Janeiro, Brazil*, pp. 6–10, 1994.

[22] C. A. F. da Rocha, J. M. T. Romano, and O. Macchi, "Self-learning deconvolution using a cascade of magnitude and phase equalizers," *Proceedings of the 38th Midwest Symposium on Circuits and Systems*, vol. 1, pp. 255–258, Aug. 13-16 1995.

[23] N. Deshpande, "Fast recovery equalization techniques for DTV signals," *IEEE Trans. on Broadcasting*, vol. 43, no. 4, pp. 370–377, Dec 1997.

[24] Z. Ding, "Characterisitcs of band-limited channels unidentifiable from second-order cyclostationary statistics," *IEEE Signal Processing Letters*, vol. 3, pp. 150–152, May 1996.

[25] ——, "On convergence analysis of fractionally spaced adaptive blind equalizers," *IEEE Trans. Acoust., Speech, Signal Proceesing*, vol. 45, no. 3, pp. 650–657, Mar 1997.

[26] Z. Ding, B. D. O. Anderson, R. A. Kennedy, and C. R. Johnson, Jr., "Local convergence of the Sato blind equalizer and generalizations under practical constraints," *IEEE Trans. Information Theory*, vol. 39, pp. 129–144, Jan 1993.

[27] Z. Ding and C. R. Johnson, Jr., "Existing gap between theory and application of blind equalization," *SPIE Proc, Adaptive Signal Processing*, vol. 1565, pp. 154–165, 1991.

[28] Z. Ding, C. R. Johnson, Jr., and R. A. Kennedy, "Nonglobal convergence of blind recursive identifiers based on gradient descent of continuous cost functions," *Decision and Control, 1990., Proceedings of the 29th IEEE Conference on*, vol. 1, pp. 225–230, Dec 1990.

[29] ——, "On the admissibility of blind adaptive equalizers," *Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on*, vol. 3, pp. 1707–1710, Apr 1990.

[30] ——, "On the admissibility of blind adaptive equalizers," *Proc. ICASSP*, pp. 1707–1710, 1990.

[31] ——, "Ill-convergence of Godard blind equalizers in data communication systems," *IEEE Trans. Commun.*, vol. 39, no. 9, pp. 1313–1327, Sep 1991.

[32] ——, "Local convergence of 'globally convergent' blind adaptive equalization algorithms," *Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference on*, pp. 1533–1536, Apr 1991.

[33] ——, "On the (non)existence of undesirable equilibria of Godard blind equalizers," *IEEE Trans. Signal Processing*, vol. 40, pp. 2425–2432, Oct 1992.

[34] Z. Ding and R. A. Kennedy, "A new adaptive algorithm for joint blind equalization and carrier recovery," *1991 Conference Record of the Twenty-Fifth Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 699–703, Nov 4-6 1991.

[35] Z. Ding, R. A. Kennedy, C. R. Johnson, Jr., and B. D. O. Anderson, "Ill-convegence of Godard blind equalizers in data communication systems," *IEEE Trans. Commun.*, vol. 39, no. 9, pp. 1313–1326, Sep 1991.

[36] S. C. Douglas, S. ichi Amari, and S.-Y. Kung, "On gradient adaptation with unit-norm constraints," *IEEE Trans. Signal Processing*, vol. 48, no. 6, pp. 1843–1847, Jun 2000.

[37] A. Duel-Hallen and C. D. Heegard, "Delayed decision feedback sequence estimation," *IEEE Trans. Commun.*, vol. 37, pp. 428–436, May 1989.

[38] P. Duhamel, "Blind equalization," *Proc. of the International Conference on Acoustics, Speech and Signal Processing*, 1995.

[39] D. L. Duttweiler, J. E. Mazo, and D. G. Messerschmitt, "An upper bound on the error probability in decision-feedback equalization," *IEEE Trans. Information Theory*, vol. IT-20, pp. 490–497, Jul 1974.

[40] M. V. Eyuboglu, "Detection of coded modulation signals on linear, severely distorted channels using decision-feedback noise prediction with interveaving," *IEEE Trans. Commun.*, vol. 36, no. 4, pp. 401–409, Apr 1988.

[41] D. D. Falconer, "Jointly adaptive equalization and carrier recovery in two-dimensional digital communication systems," *Bell Syste. Tech. J.*, vol. 55, pp. 316–334, 1976.

[42] I. Fijalkow, C. Manlove, and C. R. Johnson, Jr., "Adaptive fractionally spaced blind CMA equalization: Excess MSE," *IEEE Trans. Signal Processing*, vol. 46, pp. 227–231, Jan 1998.

[43] J. B. D. Filho, G. Favier, and J. M. T. Romano, "Neural networks for blind equalization," *Global Telecommunications Conference, 1996*, vol. 1, pp. 196–200, Nov 18-22 1996.

[44] G. D. Forney, "The Viterbi algorithm," *Proc. of IEEE*, vol. 61, no. 3, pp. 268–278, Mar 1973.

[45] W. A. Gardner, "A new method of channel identification," *IEEE Trans. Commun.*, vol. 39, pp. 813–817, Jun 1991.

[46] L. M. Garth, J. Yang, and J.-J. Werner, "Blind equalization algorithms for dual-mode CAP-QAM reception," *IEEE Trans. Commun.*, vol. 49, no. 3, pp. 455–466, Mar 2001.

[47] M. Ghosh and C. L. Weber, "Maximum-likelihood blind equalization," *SPIE Proc*, vol. 31, pp. 1224–1228, Jun 1992.

[48] G. B. Giannakis and J. M. Mendel, "Identification of nonminimum phase systems using higher-order statistics," *IEEE Trans. Acoust., Speech, Signal Proceesing*, vol. 37, pp. 360–377, Mar 1989.

[49] D. N. Godard, "Self-recovering equalization and carrier tracking in two dimensional data communication system," *IEEE Trans. Commun.*, vol. 28, no. 11, pp. 1867–1875, Nov 1980.

[50] D. N. Godard and P. E. Thirion, "Method device for training an adaptive equalizer by means of an unknown data signal in a QAM transmission system," *U.S. Patent 4 227 152*, Oct 7 1980.

[51] R. Godfrey and F. Rocca, "Zero memory nonlinear deconvolution," *Geophys. Prospecting*, vol. 29, pp. 189–228, 1981.

[52] A. J. Goldsmith and S. G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1218–1230, Oct 1997.

[53] P. Ha, "Dual-mode filtered regressor constant modulus algorithm with automatically controlled parameter," *ICT'99, Cheju, Korea. International Conference on Telecommunications*, pp. 393–396, Jun 1999.

[54] D. Hatzinakos, "Stop-and-go sign algorithms for blind equalization," *Proc. SPIE, Adaptive Signal Processing*, vol. 1565, pp. 118–129, 1991.

[55] ——, "Blind equalization using stop-and-go adaptation rules," *Optical Engineering*, vol. 31, no. 6, pp. 1181–1188, Jun 1992.

[56] D. Hatzinakos and C. L. Nikias, "Blind equalization using a tricepstrum-based algorithm," *IEEE Trans. Commun.*, vol. 39, pp. 669–682, May 1991.

[57] S. Haykin, *Blind Deconvolution.* Prentice Hall Information and System Sciences Series, 1994.

[58] ——, *Adaptive Filter Theory*, 3rd ed. Prentice Hall Information and System Sciences Series, 1996.

[59] K. Hilal and P. Duhamel, "A convergence study of the Constant Modulus Algorithm leading to a normalized CMA and a normalized block-CMA," *EUSIPCO, European Sig. Proc. Conf., (Brussels)*, pp. 135–138, Aug 24-27 1992.

[60] ——, "A blind equalizer allowing soft transition between the CMA and the DD algorithm for PSK modulated signals," *Proc. ICC, Geneva, Switzerland*, pp. 1144–1148, May 1993.

[61] I. D. Holland, H. J. Zepernick, and M. Caldera, "Analysis of an adaptive QAM scheme with non-zero delay for Rayleigh fading channels," *Proc. IEEE Veh. Tech. Conf., Orlando, USA*, Oct 2003.

[62] Institute of Electrical and Electronics Engineers, *IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries.* New York, NY, 1990.

[63] N. K. Jablon, "Joint blind equalization, carrier recovery, and timing recovery for high-order QAM signal constellations," *IEEE Trans. Signal Processing*, vol. 40, pp. 1383–1398, Jun 1992.

[64] C. R. Johnson, Jr., S. Dasgupta, and W. A. Sethares, "Averagin analysis of local stability of a real constant modulus algorithm adaptive filter," *IEEE Trans. Acoust., Speech, Signal Proceesing*, vol. 36, pp. 900–910, Jun 1988.

[65] C. R. Johnson, Jr. *et al.*, "Blind equalization using the constant modulus criterion: A review," *Proc. IEEE*, vol. 86, no. 10, pp. 1927–1950, Oct 1998.

[66] R. A. Kennedy, B. D. O. Anderson, and R. R. Bitmead, "Blind adaptation of decision feedback equalizers: gross convergence properties," *International Journal of Adaptive Control and Signal Processing*, vol. 7, pp. 497–523, 1993.

[67] R. A. Kennedy, B. D. O. Anderson, Z. Ding, and C. R. Johnson, Jr., "On the failure of proposed blind recursive identification schemes based on the Sato algorithm and its generalizations," *Inter. Sym. on Signal Processing and its Applications, Gold Coast, Australia.*, pp. 267–270, 1990.

[68] R. A. Kennedy and Z. Ding, "Blind adaptive equalizers for QAM communication systems based on convex cost functions," *Optical Engieering (SPIE)*, vol. 31, pp. 1189–1199, 1992.

[69] ——, "Quantizer design and optimization in decision feedback equalization," *Signal Processing and Its Applications, 1996. ISSPA 96., Fourth International Symp. on*, vol. 2, pp. 614–617, 25-30 Aug 1996.

[70] R. A. Kennedy and B. D. O. Anderson, "Recovery times of decision feedback equalizers on noiseless channels," *IEEE Trans. Commun.*, vol. COM-35, no. 10, pp. 1012–1021, Oct 1987.

[71] R. A. Kennedy, B. D. O. Anderson, and R. R. Bitmead, "Tight bounds on the error probabilities of decision feedback equalizers," *IEEE Trans. Commun.*, vol. COM-35, no. 10, pp. 1022–1028, Oct 1987.

[72] ——, "Channels leading to rapid error recovery for decision feedback equalizers," *IEEE Trans. Commun.*, vol. 37, no. 11, pp. 1146–1155, Nov 1989.

[73] J. Labat, O. Macchi, C. Laot, and N. Lesquin, "Is training of adaptive equalizers still useful?" *Proc. Globecom '96, London, U.K.*, vol. 2, pp. 968–972, Nov. 18-22 1996.

[74] J. Labat, C. Laot, and O. Macchi, "Dispositif degalization adaptatif pour systemes de communications numeriques," *French Patent 9510832*, Sep 1995.

[75] J. Labat, O. Macchi, and C. Laot, "Adaptive decision feedback equalization: Can you skip the training period?" *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 921–930, Jul 1998.

[76] M. G. Larimore and J. R. Treichler, "Convergence behavior of the Constant Modulus Algorithm," *ICASSP (Boston)*, pp. 13–16, Apr 1983.

[77] F. K. Li, D. Held, J. Curlander, and C. Wu, "Doppler parameter estimation for spaceborne synthetic aperture radars," *IEEE Trans. on Geoscience and Remote Sensing*, vol. GE-23, pp. 47–51, Jan 1985.

[78] Y. Li, J. K. Liu, and Z. Ding, "Length- and cost-dependent local minima of unconstrained blind channel equalizers," *IEEE Trans. Signal Processing*, vol. 44, pp. 2726–2735, Nov 1996.

[79] Y. Li and Z. Ding, "Convergence analysis of finite length blind adaptive equalizers," *IEEE Trans. Signal Processing*, vol. 43, pp. 2120–2129, Sep 1995.

[80] W. G. Lim, "Tri mode equalizer (trime) for time varying channels," *Patent Disclosure*, Jun 2001.

[81] ——, "Torque based rotational study of several blind equalization algorithms for QAM signals," *Australian Communications Theory Workshop, Newcastle, Australia*, Feb 2004.

[82] W. G. Lim, T. D. Abhayapala, and R. A. Kennedy, "Reliability based soft transition between blind startup and decision directed algorithms," *submitted to IEEE Trans. Commun. for publication*, Oct 2003.

[83] ——, "Reliability based soft transition technique for dual-mode blind equalizers," *submitted to IEEE International Conference on Communications, Paris, 2004, for publication*, Sept 2003.

[84] W. G. Lim, R. A. Kennedy, and T. D. Abhayapala, "Parallel adaptation strategy for robust and rapid convergence to the decision directed mode in a blind adaptive decision feedback equalizer," *Patent*, 2002.

[85] ——, "Fast-convergence decision feedback equalizer under parallel adaptation," *submitted to IEEE Trans. Commun. for publication*, Dec 2003.

[86] H. Liu and G. Xu, "Closed-form blind symbol estimation in digital communications," *IEEE Trans. Acoust., Speech, Signal Proceesing*, vol. 43, no. 11, pp. 2714–2723, Nov 1995.

[87] Y. Lou, "Channel estimation standard and adaptive blind equalization," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 182–186, Feb/Mar/Apr 1995.

[88] R. W. Lucky, "Automatic equalization for digital communication," *Bell System Tehnical Journal*, vol. 44, pp. 547–588, Apr 1965.

[89] ——, "Techniques for adaptive equalization of digital communciation systems," *Bell System Tehnical Journal*, vol. 45, Feb 1966.

[90] ——, "A survey of the communication theory literature, 1968-1973," *IEEE Trans. Information Theory*, vol. IT-9, pp. 725–739, Nov 1973.

[91] O. Macchi, *Adaptive processing. The LMS Approach with Applications in Transmission.* New York: Wiley, 1995.

[92] O. Macchi and E. Eweda, "Convergence analysis of self-adaptive equalizers," *IEEE Trans. Information Theory*, vol. IT-30, no. 2, pp. 161–176, Mar 1984.

[93] O. Macchi and Y. Gu, "Self-adaptive equalization with a mixed backward and forward predictor," *Proc. Intern. Symp. on Electronic Devices, Circuits and Systems (ISELDECS), Kharagpur, Inde*, pp. 437–440, 1987.

[94] O. Macchi and M. Jaidane-Saidane, "Stability of adaptive recursive filters," *International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 1503–1505, Apr. 11-14 1988.

[95] O. Macchi, C. A. F. D. Rocha, and J. M. T. Romano, "Eqalization adaptative autodidacte par retroprediction et prediction," *14th GRETSI Symp., Juan-Les-Pins, France*, pp. 491–494, 1993.

[96] R. Matzner, "An SNR estimation algorithm for complex baseband signals using higher order statistics," *Facta Universitatis, Series: Electronics and Energetics*, no. 6, pp. 41–52, 1993.

[97] R. Matzner and F. Englberger, "An SNR estimation algorithm using fourth-order moments," *Int. Symp. Inf. Theory, (Trondheim)*, 1994.

[98] R. Matzner and K. Letsch, "SNR estimation and blind equalization (deconvolution) using the kurtosis," *IEEE/IMS Workshop Inform. Theory and Statistics, Alexandria (VA)*, Oct 1994.

[99] J. E. Mazo, "Analysis of decision-directed equalizer convergence," *Bell. System Technical Journal*, vol. 59, pp. 1857–1876, Dec 1980.

[100] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications," *Proc. IEEE*, vol. 79, pp. 278–305, 1991.

[101] K. Metzger, "On the probability density of intersymbol interference," *IEEE Trans. Commun.*, vol. COM-35, no. 4, pp. 396–402, Apr 1987.

[102] ——, "Author's reply to "a correction to Kurt Metzger's algorithm in the paper 'on the probability density of intersymbol interference' "," *IEEE Trans. Commun.*, vol. 45, no. 7, p. 756, Jul 1997.

[103] P. Monsen, "Feedback equalization for fading dispersive channels," *IEEE Trans. Information Theory*, vol. IT-17, pp. 56–64, Jan 1971.

[104] C. L. Nikias, "Blind deconvolution using higher order statistics," *Int. Sig. Proc. Workshop on HOS, (Chamrousse, France)*, pp. 155–160, July 10-12 1991.

[105] K. N. Oh and Y. O. Chin, "Modified constant modulus algorithm: Blind equalization and carrier phase recovery algorithm," *Proc. 1995 IEEE International Conference on Communications, Seattle, Washington*, pp. 498–502, Jun 1995.

[106] ——, "New blind equalization techniques based on constant modulus algorithm," *Proc. 1995 IEEE Global Telecommunications Conf., Singapore*, pp. 865–869, Nov 14-16 1995.

[107] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing.* Prentice-Hall, Englewood Cliffs, NJ, 1975.

[108] ——, *Discrete-Time Signal Processing*, international ed. Prentice Hall Signal Processing Series, 1989.

[109] J. Palicot, "A weighted decision feedback equalizer with limited error propagation," *ICC 2000, New Orleans, US*, Jun 2000.

[110] R. Pan and C. L. Nikias, "The complex cepstrum of higher order cumulants and nonminimum phase identification," *IEEE Trans. Acoust., Speech, Signal Proceesing*, vol. ASSP-36, pp. 186–205, 1988.

[111] C. B. Papadias, "Methods for blind equalization and identificatin of linear channels," *PhD Thesis*, 1995.

[112] C. B. Papadias and A. Paulraj, "Decision-feedback equalization and identification of linear channels using blind algorithms of the Bussgang type," *Proc. 29th Asilomar Conf. on Signals, Systems and Computers, Pacific Grove, CA*, 1995.

[113] C. B. Papadias and D. T. M. Slock, "Normalized sliding window Constant Modulus Algorithm," *Proc. Quatorzieme Collogque sur le Traitement du Signal et des Images, Juan-les-Pins, France*, pp. 507–510, Sept 13-16 1993.

[114] ——, "Normalized sliding window constant modulus and decision-directed algorithms: A link between blind equalization and classical adaptive filtering," *IEEE Trans. Signal Processing*, vol. 45, pp. 231–235, Jan 1997.

[115] D. R. Pauluzzi and N. C. Beaulieu, "A comparison of SNR estimation techniques for the AWGN channel," *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 1681–1691, Oct 2000.

[116] G. Picchi and G. Prati, "Blind equalization and carrier recovery using a "Stop-and-Go" decision-directed algorithm," *IEEE Trans. Commun.*, vol. COM-35, no. 9, pp. 877–887, Sep 1987.

[117] J. G. Proakis and C. L. Nikias, "Blind equalization," *SPIE Proc*, vol. 1565, pp. 76–87, Jul 1991.

[118] J. G. Proakis, *Digital Communications*, 2nd ed. McGraw-Hill International Editions, 1989.

[119] S. U. H. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol. 73, no. 9, pp. 1349–1387, Sep 1985.

[120] M. J. Ready and R. P. Gooch, "Blind equalization based on radius directed adaptation," *Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on*, vol. 3, pp. 1699–1702, Apr 1990.

[121] 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, no. 11, pp. 2028–2041, Nov 2001.

[122] F. Rocca, B. Godfrey, and F. Muir, "Bussgang processes," *sepwww.stanford.edu/oldreports/sep16/16_23_abs.html.*

[123] F. J. Ross and D. P. Taylor, "An enhancement to blind equalization algorithms," *IEEE Trans. Commun.*, vol. 39, no. 5, pp. 636–639, May 1991.

[124] Y. Sato, "A method of self-recovering equalization for multi-level amplitude-modulations systems," *IEEE Trans. Commun.*, pp. 679–682, Jun 1975.

[125] N. Seshadri, "Joint data and channel estimation using blind trellis search techniques," *IEEE Trans. Commun.*, vol. 23, pp. 1000–1011, Feb/Mar/Apr 1994.

[126] W. A. Sethares, G. A. Rey, and C. R. Johnson, Jr, "Approaches to blind equalization of signals with multiple modulus," *ICASSP-89*, vol. 2, pp. 972–975, 1989.

[127] O. Shalvi and E. Weinstein, "Super-exponential methods for blind deconvolution," *IEEE Trans. Information Theory*, vol. 39, pp. 504–519, Mar 1993.

[128] ——, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. Information Theory*, vol. 36, no. 2, Mar 1990.

[129] C. E. Shannon, "A mathematical theory of communication." *The Bell System Technical J. 27*, vol. 27, pp. 623–656, Jul and Oct 1948.

[130] J. J. Shynk, C. K. Chan, and R. P. Gooch, "Comparative performance study of several blind equalization algorithms," *SPIE 91, San Diego, CA (Optical Engineering)*, 1991.

[131] D. T. M. Slock, "On the convergence behavior of the LMS and the normalized LMS algorithms," *IEEE Trans. Signal Processing*, vol. 41, no. 9, pp. 2811–2825, 1993.

[132] T. Stockham, T. Cannon, and R. Ingebretsen, "Blind deconvolution through digital signal processing," *Proc. of IEEE*, vol. 63, pp. 678–692, Apr 1975.

[133] L. Tong and D. Liu, "Blind predictive decision-feedback equalization via the constant modulus algorithm," *ICASSP-97, IEEE International Conference on Accoustics, Speech, and Signal Processing*, vol. 5, pp. 3901–3904, Apr 21-24 1997.

[134] L. Tong, G. Xu, B. Hassibi, and T. Kailath, "Blind channel identification based on second-order statistics: A frequency-domain approach," *IEEE Trans. Information Theory*, vol. 41, pp. 329–334, Jan 1995.

[135] L. Tong, G. Xu, and T. Kailath, "A new approach to blind identification and equalization of multipath channels," *Conference Record of the Twenty-Fifth Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 856–860, Nov. 4-6 1991.

[136] ——, "Blind identification and equalization basd on second-order statistics: A time domain approach," *IEEE Trans. Information Theory*, vol. 40, pp. 340–349, Mar 1994.

[137] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Signal Processing*, vol. 31, pp. 349–372, Apr 1983.

[138] J. R. Treichler and M. G. Larimore, "New processing techniques based on the constant modulus adaptive algorithm," *IEEE Trans. on Acoust. Speech and Signal Processing*, vol. 33, pp. 420–431, Apr 1985.

[139] J. R. Treichler, M. G. Larimore, and J. C. Harp, "Practical blind demodulators for high-order QAM signals," *Proc. IEEE*, vol. 86, no. 10, pp. 1907–1926, Oct 1998.

[140] J. K. Tugnait, "Blind equalization and estimation of digital communication FIR channels using cumulant matching," *IEEE Trans. Commun.*, vol. 43, pp. 1240–1245, Feb/Mar/Apr 1995.

[141] J. K. Tugnait, L. Tong, and Z. Ding, "Single-user channel estimation and equalization," *IEEE Signal Processing Magazine*, vol. 17, pp. 16–28, May 2000.

[142] S. Verdu, B. D. O. Anderson, and R. A. Kennedy, "Blind equalization without gain identification," *IEEE Trans. Information Theory*, vol. 39, no. 1, pp. 292–297, Jan 1993.

[143] T. Wang and C.-L. Wang, "On adaptive decision-feedback equalization of intersymobl interference channels in coded modulation systems," *IEEE Trans. Commun.*, vol. 44, no. 11, pp. 1404–1408, Nov 1996.

[144] V. Weerackody, S. A. Kassam, and K. R. Laker, "Sign algorithms for blind equalizaiton and their convergence analysis," *Circuits Systems Signal Processing*, vol. 10, no. 4, pp. 393–431, 1991.

[145] V. Weerackody and S. A. Kassam, "Dual-mode type algorithms for blind equalization," *IEEE Trans. Commun.*, vol. 42, no. 1, pp. 22–28, Jun 1994.

[146] J.-J. Werner and J. Yang, "Multimodulus blind equalization using piece-wise linear contours," *U.S. Patent 5 793 807*, Aug. 11 1998.

[147] J.-J. Werner, J. Yang, D. D. Harman, and G. A. Dumont, "Blind equalization for broadband access," *IEEE Communications Magazine*, pp. 87–93, Apr 1999.

[148] K. Wesolowski, "Adaptive blind equalizers with automatically controlled parameters," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 170–172, Feb/Mar/Apr 1995.

[149] R. A. Wiggins, "Minimum entropy deconvolution," *Geoexploration*, vol. 16, pp. 21–35, Feb 1978.

[150] K. Yamakazi, R. A. Kennedy, and Z. Ding, "Globally convergent blind equalization algorithms for complex data systems," *Proc. of the International Conference on Acoustics, Speech and Signal Processing, San Francisco*, vol. IV, pp. 553–556, 1992.

[151] J. Yang, J. J. Werner, and G. A. Dumont, "The multimodulus blind equalization algorithm," *Proc. Thirteenth International Conference on Digital Signal Processing, Santorini, Greece*, Jul 1997.

[152] J. Yang, J.-J. Werner, and G. A. Dumont, "The multimodulus blind equalization and its generalized algorithms," *IEEE Journal on selected areas in communications*, vol. 20, no. 5, pp. 997–1015, Jun 2002.

[153] R. Yatsuboshi, N. Sata, and K. Aoki, "A convergence of automatic equalizer by maximum level error control," *Nat. Conv. Rec., IECE Japan*, no. 2192, 1974.

[154] J. L. Yen, *Image reconstruction in synthesis radiotelescope arrays*, array signal processing, s. haykin, ed. ed. Prentice-Hall, Englewood Cliffs, NJ, 1985.

[155] H. H. Zeng and L. Tong, "Blind channel estimation using the second-order statistics: Algorithms," *IEEE Trans. Signal Processing*, vol. 45, pp. 1785–1794, Jul 1997.

[156] H. H. Zeng, L. Tong, and C. R. Johnson, Jr., "Relationships between the constant modulus and weiner receivers," *IEEE Trans. Information Theory*, vol. 44, pp. 1523–1538, Jul 1998.