

**MIXED NORM EQUALIZATION WITH APPLICATIONS  
IN TELEVISION MULTIPATH CANCELLATION**

By

**JINSHI HUANG**

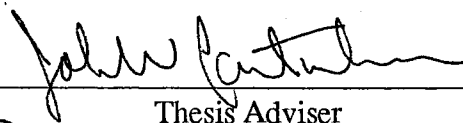
**Bachelor of Science  
Wuhan Technical University  
of Surveying and Mapping  
Wuhan, China  
1985**

**Master of Science  
Oklahoma State University  
Stillwater, Oklahoma  
1988**

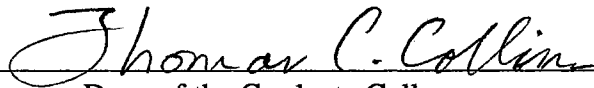
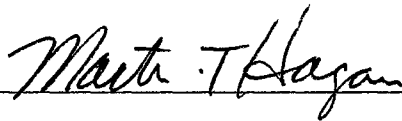
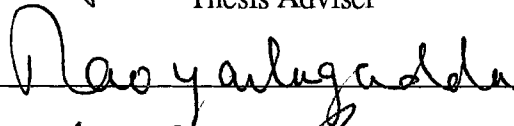
**Submitted to the Faculty of the  
Graduate College of the  
Oklahoma State University  
in partial fulfillment of  
the requirements for  
the Degree of  
DOCTOR OF PHILOSOPHY  
December, 1995**

MIXED NORM EQUALIZATION WITH APPLICATIONS  
IN TELEVISION MULTIPATH CANCELLATION

Thesis Approved:



Thesis Adviser



Dean of the Graduate College

## ACKNOWLEDGMENTS

I wish to express my sincere appreciation to the advisory committee: Dr. Rao Yarlagadda, Dr. Jack Cartinhour, Dr. Martin Hagan and Dr. Michael Branson. Dr. Yarlagadda initiated this program. His well-known contributions to the area of non-Gaussian signal processing had significant influence on the formulation of the concept of mixed norm equalization in this paper. Dr. Cartinhour provided the specific guidance and suggestions and handled all the necessary work to keep the program running. His encouragement and friendship are invaluable. Dr. Hagan introduced the author to the area of stochastic signal processing. His academic attitude provides a role model for the author. I would like to thank Dr. Branson for serving in the committee for all these years.

I would like to express my sincere gratitude to my father, who passed away in June, 1995. He had constantly emphasized the importance of education for my life since my early years, and strongly insisted that I receive higher education. I deeply regret that he could not see me finish this program. My appreciation extends to all the family members: my mother, brother and sisters for their concern and support in all these years.

I would like to give my special appreciation to my wife, Luping, for her help in preparing this paper. Her strong support at times of difficulty, her love, patience and understanding throughout this whole process are essential for me to finish this program.

I would also like to thank the Department of Electrical Engineering for allowing me to resume this program after I left the school.

## TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION .....	1
II. THE MIXED NORM .....	9
1. The $L_p$ Norm .....	9
2. Bayesian Estimation and the Mixed Norm .....	12
3. Summary .....	14
III. MINIMUM NORMS EQUALIZATION .....	15
1. Communication System Model .....	15
2. Intersymbol Interference and Linear Receivers .....	17
(1) Discrete-time Model .....	17
(2) Linear Zero-forcing Equalizer .....	19
a. Infinite Length .....	19
b. Finite Length .....	21
(3) Linear MSE Equalizer .....	23
a. Infinite Length .....	23
b. Finite Length .....	26
(4) Mixed Norm Linear Equalizer .....	30
a. Infinite Length .....	30
b. Finite Length .....	33
3. Feedback Equalizer .....	34
(1) Why Do We Need a Feedback Equalizer ? .....	34
(2) Zero-forcing Feedback Equalizer .....	38
a. Infinite Length .....	38
b. Finite Length .....	45
(3) MSE Feedback Equalizer .....	47
a. Infinite Length .....	47
b. Finite Length .....	50
(4) Mixed Norm Feedback Equalizer .....	52
a. Infinite Length .....	52
b. Finite Length .....	53
4. Summary .....	54

IV. ADAPTIVE AND AUTOMATIC EQUALIZATION BY MINIMIZING NORMS .....	58
1. Automatic Linear Equalizer .....	59
(1) Minimizing the $L_1$ Norm and the ZF Algorithm .....	59
a. Minimizing the $L_1$ Norm .....	59
b. Minimizing Peak Distortion and Lucky's ZF Algorithm .....	61
c. Stochastic ZF Algorithm .....	64
(2) Minimizing the $L_2$ Norm and LMS Algorithm .....	66
(3) Minimizing the Mixed Norm .....	68
2. Automatic Feedback Equalizer .....	69
(1) Updating the Feedback Portion .....	69
(2) Automatic Zero-forcing Feedback Equalizer .....	73
(3) Automatic MSE Feedback Equalizer .....	74
(4) Automatic Mixed Norm Feedback Equalizer .....	76
3. Summary .....	77
V. AUTOMATIC TELEVISION MULTIPATH CANCELLATION .....	81
1. Introduction .....	81
2. The Filter Structures for Ghost Cancellation .....	83
3. The Ghost Cancellation Reference (GCR) Signal .....	87
4. The Ghost Cancellation System .....	92
5. Ghost Cancellation using a Linear Equalizer .....	99
6. Ghost Cancellation using a Feedback Equalizer .....	102
VI. CONCLUSIONS .....	189
REFERENCES .....	194
APPENDIXES .....	207
APPENDIX A--THE $L_1$ AND $L_2$ NORMS AND THEIR PROPERTIES .....	207
APPENDIX B--THE MSE EQUALIZER AND WIENER FILTER .....	218
APPENDIX C--CLOSED FORM EXPRESSIONS OF THE MSE FOR FEEDBACK EQUALIZERS WITH INFINITE LENGTH .....	225
APPENDIX D--COMPARISON OF THE MSE BETWEEN LINEAR AND FEEDBACK EQUALIZERS .....	228
APPENDIX E--AN ALTERNATIVE MODEL OF THE GHOSTING PROBLEM .....	229

## LIST OF FIGURES

Figure	Page
2-1. $p$ -Gaussian pdf for $p=1,2,4,10,50$ .....	11
3-1. Communication System .....	15
3-2. Baseband Model of the Communication System .....	17
3-3. Discrete-time Model with Equalizer .....	18
3-4. Communication System with Linear Equalizer .....	23
3-5. Frequency Domain Equivalence of Figure 3-4 .....	24
3-6. Simplified Version of Figure 3-5 .....	24
3-7. Decision Feedback Equalizer .....	36
3-8. Feedback Equalizer with Reference Signal.....	36
3-8a. Typical coefficients distribution of the forward filter $h(n)$ .....	36
3-8b. Typical coefficients distribution of the feedback filter $g(n)$ .....	37
3-9. ZF Feedback Equalizer Model .....	39
3-10. Equivalent ZF Feedback Equalizer Model .....	40
3-11. Rearrangement of Noise Path for ZF-FE .....	40
3-12. MSE Feedback Equalizer Model .....	47
3-13. Model to Calculate the Error for MSE-FE .....	48
3-14. Equivalent Model of Figure 3-13 .....	48

4-1. Two-dimensional Vector Updates in ZF and LMS .....	68
4-2. Model for the Output Error .....	70
5-1a. A Noncausal Feedback Equalizer for Ghost Canceler .....	85
5-1b. A Noncausal Linear Equalizer for Ghost Canceler .....	85
5-2a. Actual Implementation of the Feedback Equalizer .....	86
5-2b. Actual Implementation of the Linear Equalizer .....	86
5-3. Japan Standard GCR Signal .....	90
5-4. Differential Japan Standard GCR Signal .....	90
5-5. US Standard GCR Signal .....	91
5-6. Spectrum of the US Standard GCR Signal .....	91
5-6a. Autocorrelation of the US Standard GCR Signal .....	92
5-7. The Ghost Cancellation System .....	93
5-8. Received GCR Signal for Case 1 (linear equalizer) .....	106
5-8a. Spectrum of the Received GCR Signal for Case 1 (linear equalizer) .....	106
5-9. Estimated channel impulse response $f(n)$ for Case 1 (linear) .....	107
5-10. Coefficients of Linear ZF Equalizer for Case 1 .....	108
5-11. Frequency Response of Linear ZF Equalizer for Case 1 .....	108
5-12. Output Signal of the Linear ZF Equalizer for Case 1 .....	109
5-13. Output Signal Spectrum of the Linear ZF Equalizer for Case 1 .....	109
5-14. Coefficients of Linear ZF ( $L_1$ ) Equalizer for Case 1 .....	110
5-15. Frequency Response of Linear ZF ( $L_1$ ) Equalizer for Case 1 .....	110
5-16. Output Signal of the Linear ZF ( $L_1$ ) Equalizer for Case 1 .....	111

5-17. Output Signal Spectrum of the Linear ZF ( $L_1$ ) Equalizer for Case 1 .....	111
5-18. Coefficients of Linear MSE Equalizer for Case 1 .....	112
5-19. Frequency Response of Linear MSE Equalizer for Case 1 .....	112
5-20. Output Signal of the Linear MSE Equalizer for Case 1 .....	113
5-21. Output Signal Spectrum of the Linear MSE Equalizer for Case 1 .....	113
5-22. Coefficients of Linear Mixed Norm Equalizer for Case 1 .....	114
5-23. Frequency Response of Linear Mixed Norm Equalizer for Case 1 .....	114
5-24. Output Signal of the Linear Mixed Norm Equalizer for Case 1 .....	115
5-25. Output Signal Spectrum of the Linear Mixed Norm Equalizer for Case 1 ....	115
5-26. Peak Distortions of Linear Equalizers for Case 1 .....	116
5-27. Detailed Peak Distortions of Linear Equalizers for Case 1 .....	117
5-28. MSE of Linear Equalizers for Case 1 .....	118
5-29. Detailed MSE of Linear Equalizers for Case 1 .....	119
5-30. MSE Values versus Mixed Norm Parameter $\lambda$ for Case 1 (linear) .....	120
5-31. Received GCR Signal for Case 2 (linear equalizer) .....	121
5-31a. Spectrum of the Received GCR Signal for Case 2 (linear equalizer) .....	121
5-32. Estimated channel impulse response $f(n)$ for Case 2 (linear) .....	122
5-33. Coefficients of Linear ZF Equalizer for Case 2 .....	123
5-34. Frequency Response of Linear ZF Equalizer for Case 2 .....	123
5-35. Output Signal of the Linear ZF Equalizer for Case 2 .....	124
5-36. Output Signal Spectrum of the Linear ZF Equalizer for Case 2 .....	124
5-37. Coefficients of Linear ZF ( $L_1$ ) Equalizer for Case 2 .....	125



5-38. Frequency Response of Linear ZF ( $L_1$ ) Equalizer for Case 2 .....	125
5-39. Output Signal of the Linear ZF ( $L_1$ ) Equalizer for Case 2 .....	126
5-40. Output Signal Spectrum of the Linear ZF ( $L_1$ ) Equalizer for Case 2 .....	126
5-41. Coefficients of Linear MSE Equalizer for Case 2 .....	127
5-42. Frequency Response of Linear MSE Equalizer for Case 2 .....	127
5-43. Output Signal of the Linear MSE Equalizer for Case 2 .....	128
5-44. Output Signal Spectrum of the Linear MSE Equalizer for Case 2 .....	128
5-45. Coefficients of Linear Mixed Norm Equalizer for Case 2 .....	129
5-46. Frequency Response of Linear Mixed Norm Equalizer for Case 2 .....	129
5-47. Output Signal of the Linear Mixed Norm Equalizer for Case 2 .....	130
5-48. Output Signal Spectrum of the Linear Mixed Norm Equalizer for Case 2 .....	130
5-49. Peak Distortions of Linear Equalizers for Case 2 .....	131
5-50. Detailed Peak Distortions of Linear Equalizers for Case 2 .....	132
5-51. MSE of Linear Equalizers for Case 2 .....	133
5-52. Detailed MSE of Linear Equalizers for Case 2 .....	134
5-53. MSE Values versus Mixed Norm Parameter $\lambda$ for Case 2 (linear) .....	135
5-54. Received GCR Signal for Case 3 (linear equalizer) .....	136
5-54a. Spectrum of the Received GCR Signal for Case 3 (linear equalizer) .....	136
5-55. Estimated channel impulse response $f(n)$ for Case 3 (linear) .....	137
5-56. Coefficients of Linear ZF Equalizer for Case 3 .....	138
5-57. Frequency Response of Linear ZF Equalizer for Case 3 .....	138
5-58. Output Signal of the Linear ZF Equalizer for Case 3 .....	139

5-59. Output Signal Spectrum of the Linear ZF Equalizer for Case 3 .....	139
5-60. Coefficients of Linear ZF ( $L_1$ ) Equalizer for Case 3 .....	140
5-61. Frequency Response of Linear ZF ( $L_1$ ) Equalizer for Case 3 .....	140
5-62. Output Signal of the Linear ZF ( $L_1$ ) Equalizer for Case 3 .....	141
5-63. Output Signal Spectrum of the Linear ZF ( $L_1$ ) Equalizer for Case 3 .....	141
5-64. Coefficients of Linear MSE Equalizer for Case 2 .....	142
5-65. Frequency Response of Linear MSE Equalizer for Case 2 .....	142
5-66. Output Signal of the Linear MSE Equalizer for Case 3 .....	143
5-67. Output Signal Spectrum of the Linear MSE Equalizer for Case 3 .....	143
5-68. Coefficients of Linear Mixed Norm Equalizer for Case 2 .....	144
5-69. Frequency Response of Linear Mixed Norm Equalizer for Case 2 .....	144
5-70. Output Signal of the Linear Mixed Norm Equalizer for Case 3 .....	145
5-71. Output Signal Spectrum of the Linear Mixed Norm Equalizer for Case 3 .....	145
5-72. Peak Distortions of Linear Equalizers for Case 3 .....	146
5-73. Detailed Peak Distortions of Linear Equalizers for Case 3 .....	147
5-74. MSE of Linear Equalizers for Case 3 .....	148
5-75. Detailed MSE of Linear Equalizers for Case 3 .....	149
5-76. MSE Values versus Mixed Norm Parameter $\lambda$ for Case 3 (linear) .....	150
5-77. Received GCR Signal for Case 1 (feedback equalizer) .....	151
5-77a. Spectrum of the Received GCR Signal for Case 1 (feedback equalizer) .....	151
5-78. Estimated channel impulse response $f(n)$ for Case 1 (feedback) .....	152
5-79. Coefficients of Forward Filter $h(n)$	

of ZF Feedback Equalizer for Case 1 .....	153
5-80. Coefficients of Feedback Filter $g(n)$	
of ZF Feedback Equalizer for Case 1 .....	153
5-81. Output Signal of the Forward filter of the	
ZF Feedback Equalizer for Case 1 .....	154
5-82. Spectrum of the Forward filter Output of the	
ZF Feedback Equalizer for Case 1 .....	154
5-83. Output Signal of the ZF Feedback Equalizer for Case 1 .....	155
5-84. Spectrum of the Output Signal of the ZF Feedback Equalizer for Case 1 .....	155
5-85. Coefficients of Forward Filter $h(n)$	
of MSE Feedback Equalizer for Case 1 .....	156
5-86. Coefficients of Feedback Filter $g(n)$	
of MSE Feedback Equalizer for Case 1 .....	156
5-87. Output Signal of the Forward filter of the	
MSE Feedback Equalizer for Case 1 .....	157
5-88. Spectrum of the Forward filter Output of the	
MSE Feedback Equalizer for Case 1 .....	157
5-89. Output Signal of the MSE Feedback Equalizer for Case 1 .....	158
5-90. Spectrum of the Output Signal of the MSE Feedback Equalizer for Case 1 ..	158
5-91. MSE of the Feedback Equalizers for Case 1 .....	159
5-92. Detailed MSE of the Feedback Equalizers for Case 1 .....	160
5-93. MSE Values versus Mixed Norm Parameter $\lambda$ for Case 1 (feedback) .....	161

5-94. Received GCR Signal for Case 2 (feedback equalizer) .....	162
5-94a. Spectrum of the Received GCR Signal for Case 2 (feedback equalizer) ....	162
5-95. Estimated channel impulse response $f(n)$ for Case 2 (feedback) .....	163
5-96. Coefficients of Forward Filter $h(n)$ of ZF Feedback Equalizer for Case 2 .....	164
5-97. Coefficients of Feedback Filter $g(n)$ of ZF Feedback Equalizer for Case 2 .....	164
5-98. Output Signal of the Forward filter of the ZF Feedback Equalizer for Case 2 .....	165
5-99. Spectrum of the Forward filter Output of the ZF Feedback Equalizer for Case 2 .....	165
5-100. Output Signal of the ZF Feedback Equalizer for Case 2 .....	166
5-101. Spectrum of the Output Signal of the ZF Feedback Equalizer for Case 2 ....	166
5-102. Coefficients of Forward Filter $h(n)$ of MSE Feedback Equalizer for Case 2 .....	167
5-103. Coefficients of Feedback Filter $g(n)$ of MSE Feedback Equalizer for Case 2 .....	167
5-104. Output Signal of the Forward filter of the MSE Feedback Equalizer for Case 2 .....	168
5-105. Spectrum of the Forward filter Output of the	

MSE Feedback Equalizer for Case 2 .....	168
5-106. Output Signal of the MSE Feedback Equalizer for Case 2 .....	169
5-107. Spectrum of the Output Signal of the MSE Feedback Equalizer for Case 2 .....	169
5-108. MSE of the Feedback Equalizers for Case 2 .....	170
5-109. Detailed MSE of the Feedback Equalizers for Case 2 .....	171
5-110. MSE Values versus Mixed Norm Parameter $\lambda$ for Case 2 (feedback) .....	172
5-111. Received GCR Signal for Case 3 (feedback equalizer) .....	173
5-111a. Spectrum of the Received GCR Signal for Case 3 (feedback equalizer) ..	173
5-112. Estimated channel impulse response $f(n)$ for Case 3 (feedback) .....	174
5-113. Coefficients of Forward Filter $h(n)$ of ZF Feedback Equalizer for Case 2 .....	175
5-114. Coefficients of Feedback Filter $g(n)$ of ZF Feedback Equalizer for Case 2 .....	175
5-115. Output Signal of the Forward filter of the ZF Feedback Equalizer for Case 3 .....	176
5-116. Spectrum of the Forward filter Output of the ZF Feedback Equalizer for Case 3 .....	176
5-117. Output Signal of the ZF Feedback Equalizer for Case 3 .....	177
5-118. Spectrum of the Output Signal of the ZF Feedback Equalizer for Case 3 .....	177
5-119. Coefficients of Forward Filter $h(n)$	

	of MSE Feedback Equalizer for Case 3 .....	178
5-120.	Coefficients of Feedback Filter $g(n)$ of MSE Feedback Equalizer for Case 3 .....	178
5-121.	Output Signal of the Forward filter of the MSE Feedback Equalizer for Case 3 .....	179
5-122.	Spectrum of the Forward filter Output of the MSE Feedback Equalizer for Case 3 .....	179
5-123.	Output Signal of the MSE Feedback Equalizer for Case 3 .....	180
5-124.	Spectrum of the Output Signal of the MSE Feedback Equalizer for Case 3 .....	180
5-125.	MSE of the Feedback Equalizers for Case 3 .....	181
5-126.	Detailed MSE of the Feedback Equalizers for Case 3 .....	182
5-127.	MSE Values versus Mixed Norm Parameter $\lambda$ for Case 3 (feedback) .....	183
5-128.	Received GCR Signal for Case 4 .....	184
5-128a.	Spectrum of the Received GCR Signal for Case 4 .....	184
5-129.	Estimated channel impulse response $f(n)$ for Case 4 .....	185
5-130.	Coefficients of Linear MSE Equalizer for Case 4 .....	186
5-131.	Frequency Response of Linear MSE Equalizer for Case 4 .....	186
5-132.	Coefficients of Forward Filter $h(n)$ of MSE Feedback Equalizer for Case 4 .....	187
5-133.	Coefficients of Feedback Filter $g(n)$	

of MSE Feedback Equalizer for Case 4 .....	187
5-134. MSE values of the linear and feedback MSE equalizers for Case 4 .....	188
A-1. Two Dimensional Convex Hull .....	204
A-2. $d_1(c)$ versus $c$ in Example A-1 .....	209
A-3. $d_1(c)$ and $d_2(c)$ versus $c$ in Example A-2 .....	210

## NOMENCLATURE

<b>A</b>	covariance matrix
<b>a</b>	$M \times 1$ vector
<b>a*</b>	$M \times 1$ vector, mean of <b>a</b>
<b>AM</b>	amplitude modulation
<b>AR</b>	autoregressive
<b><math>B_x</math></b>	bandwidth of signal
<b><math>B_f</math></b>	bandwidth of channel
<b>b</b>	$M \times 1$ vector
<b>c</b>	$M \times 1$ vector, unknown variables
<b><math>d_p(\cdot)</math></b>	$L_p$ norm
<b>DSB</b>	double-side band
<b>DSP</b>	digital signal processing
<b><math>\mathbb{E}</math></b>	linear norm space
<b><math>\mathcal{E}[\cdot]</math></b>	expectation operator
<b><math>E(z)</math></b>	Z transform of $e(n)$
<b>e</b>	error vector
<b>e</b>	error variable, element of <b>e</b>
<b><math>F(z)</math></b>	frequency response of channel
<b>f</b>	coefficient vector of $F(z)$
<b>f</b>	element of <b>f</b>
<b>F</b>	matrix made of <b>f</b>
<b>FE</b>	feedback equalizer
<b>FIFO</b>	first-in-first-out buffer
<b>FM</b>	frequency modulation
<b>FSK</b>	frequency-shift keying
<b><math>G(z)</math></b>	frequency response of time-delayed ghosting system; frequency response of predictor frequency response of feedback filter
<b>GCR</b>	ghost cancellation reference
<b>g</b>	coefficient vector of $G(z)$
<b>g</b>	element of <b>g</b>
<b><math>H(z)</math></b>	frequency response of FIR equalizer
<b>h</b>	coefficient vector of $H(z)$
<b>h</b>	element of <b>h</b>
<b>I</b>	identity matrix
<b>i</b>	time index
<b>j</b>	time index
<b>J</b>	direction vector of directional derivative



$K(z)$	frequency response of time-advanced ghosting system
$\mathbf{k}$	coefficient vector of $H(z)$
$k_i$	element of $\mathbf{k}$
$k$	time index
$l$	prediction error
$L_p$	$L_p$ norm
LE	linear equalizer
$M$	number of unknowns; order of forward filter of feedback equalizer
$m$	time index
MN	mixed norm
MSE	mean square error
$N$	number of measurements; order of feedback filter order of IIR filter and time-delayed ghosting system
$n$	time index
$N_0$	noise spectral density
NTSC	National Television Standard Committee
$p$	order of norms, $1 \leq p \leq \infty$
$p_i$	$i$ -th pole
$P$	order of time-advanced ghosting system
PAL	a television system used mainly in Europe and some Asian countries
PAM	pulse amplitude modulation
PSK	phase-shift keying
$Q(z)$	frequency response of combined channel
$\mathbf{q}$	coefficient vector of $Q(z)$
$q$	element of $\mathbf{q}$
QAM	quadrature amplitude modulation
$R(\cdot)$	autocorrelation; cross-correlation
$\mathbf{r}$	residual vector
$r$	element of $\mathbf{r}$
$\mathbf{S}$	convex hull
$S$	$L_1$ solution set
$S_i$	ghost cancellation reference (GCR) signal sequence
$S(\omega)$	power spectral density
SSB	single-side band
$T$	sample period
$T(z)$	frequency response of IIR filter
$\mathbf{t}$	coefficient vector of $T(z)$
$t_i$	element of $\mathbf{t}$
$\mathbf{u}$	output vector of channel without noise
$u$	element of $\mathbf{u}$
$\mathcal{V}$	convex space
$\mathbf{v}$	output vector of entire ghost canceler
$v$	element of $\mathbf{v}$

VSB	vestigial-side band
$\mathbf{w}$	received signal vector
$w$	element of $\mathbf{w}$
$\mathbf{x}$	transmitted signal vector
	reference signal vector
$x$	element of $\mathbf{x}$
$\mathbf{X}$	matrix made of $\mathbf{x}$
$\mathbf{x}_i^T$	$i$ -th row of $\mathbf{X}$
$\mathbf{y}$	signal vector of equalizer output
$y$	element of $\mathbf{y}$
$z$	Z transform variable
$\mathcal{Z}$	Z transform operator
ZF	zero-forcing
$\alpha$	cross-correlation vector
$\alpha$	constant
$\beta$	constant
$\gamma$	variable for noise modelling
$\gamma_n$	Fourier series
$\omega$	angle frequency
$\Gamma(\cdot)$	gamma function
$\Psi$	gradient vector
$\psi$	element of $\Psi$
$\lambda$	mixed norm parameter
$\sigma_n^2$	noise variance
$\sigma_x^2$	signal variance
$\Phi(\omega)$	noise spectral density
$\phi$	intermediate variable for noise modelling
$\Theta(\omega)$	power spectral density
$\phi_n$	Fourier series
$\Delta$	update step size
$\delta(\cdot)$	unit impulse response
$\delta$	vector made of unit impulse
$\varepsilon^2$	mean square error
$\tau$	parameter, $\tau \in [0,1]$
$\xi$	prior distribution parameter
$\eta$	noise variable
$\rho$	coefficient of Wiener filter
$ \cdot $	absolute value
$\ \cdot\ _p$	$p$ -norm
$(\cdot)^*$	conjugate operator
$(\cdot)^T$	transpose operator

## CHAPTER I

### INTRODUCTION

Communications through wire line or radio have existed for over a hundred years. Today, communication is part of our daily life: from telephones to the Internet; from cellular phones to the global positioning system, etc. In a broader sense, broadcasting is also a kind of communication: it is the communication between the stations and the viewers or listeners. Nyquist's telegraph theory [1] probably lays down the first foundation for the framework of modern communication theory, where certain basic problems of communication are formulated. Shannon's information theory [2] is a key milestone on the long journey towards mastering of the communication problems.

One of the basic problems in communication is how to **efficiently** utilize the **limited** transmission media capacity (e.g. bandwidth) for carrying **useful** information. "Efficiently" means more information will be transmitted with the same transmission media. "Useful" generally means less interference and disturbance. For example, in data communication, this problem is translated as how to transmit data faster with less bit error rate or block error rate.

In electrical communication, the "efficiency" problem is usually solved by properly **modulating** the signal and/or encoding the information at the transmitter. The "usefulness" is usually enforced by properly processing at the receiver. One of the key elements of this processing is equalization. In a broad sense, an **equalizer** is a device

whose goal is to restore the transmitted signal by processing the received signal. Usually, the received signal is a **distorted** version of the transmitted signal. These distortions are mainly caused by the transmission media. For example, in an analog telephone line, distortion results when the channel frequency response deviates from the ideal of constant amplitude and linear phase [3, 4]; in radio communication, distortion is mainly due to multipath propagation [5, 6,7,8,9,10], which may be viewed as transmission through a group of channels with various relative amplitudes and delays.

The result of the distortion is the time dispersion of the signal [11,12,13,14,15], generally called **intersymbol interference (ISI)**, because the information is usually carried by elements of the signal called **symbols**. In a broader sense of communication, such as analog television broadcasting, there is no such thing as a “symbol”. We still use the term “intersymbol interference” to refer the time dispersion of the signal introduced by the channel distortion.

In addition to channel distortion, the received signal is generally corrupted by noise. This noise can be additive and/or multiplicative. In this paper, only additive noise is considered.

The straightforward way of equalization is to attempt to eliminate the intersymbol interference completely, which is exactly what the zero-forcing equalizer does [4,16]. The zero-forcing (ZF) criterion results in an equalizer which effectively inverts the channel frequency response. Unfortunately, this inversion may severely enhance the noise at the frequencies where the nulls of the channel frequency response exist [12,13,14]. In addition, if the channel transfer function is not minimum phase, the theoretical inverse

channel transfer function would be noncausal (assuming the solution is not allowed to be unstable). As an alternative to the ZF criterion, the mean square error (MSE) criterion is used [15,17]. The MSE criterion produces an equalizer which minimizes the mean square of the error between the equalizer output and a given reference signal. It was later realized that for an equalizer with a tapped delay line structure and an infinite number of taps, the MSE equalizer is indeed a superset of the ZF equalizer [11,18]: it strikes a balance between reducing the intersymbol interference (inverting the channel) and minimizing the noise enhancement. When there is no noise corrupting the received signal, the MSE equalizer will reduce to a ZF equalizer. On the other hand, when the received signal is corrupted by noise, the MSE equalizer will leave some intersymbol interference at the equalizer output so that the noise is not severely amplified at the frequencies where the channel frequency response manifests nulls. The result is that the linear MSE equalizer with infinite length will produce less mean square error value than the ZF counterpart. For communication systems where the signal waveform itself is the ultimate information (such as analog audio and video signals), the MSE is a meaningful measure of merit.

In most practical applications, the channel distortion is statistical. In some cases, some prior knowledge about the channel distribution, typically the type of probability density function (e.g. Laplacian distribution) of the channel impulse response, is available based on previous experience. For example, in multipath propagation, the channel impulse response typically consists of a series of impulses which correspond to the various reflections produced in the process of transmission [19,20,21,22]. Histogram of the “weights” of these reflections in certain area can be used to suggest a suitable probability

density function. Based on Bayesian theory, the maximum *a posteriori* (MAP) estimator results in a criterion which is a summation of two norms (called a **mixed norm (MN)**), if  $p$ -Gaussian distributions [23] are assumed for both channel impulse response and the noise at the equalizer input. The concept of a mixed norm is not new in other areas, such as deconvolution [24,25]. But to the best of the author's knowledge, this is the first time that the mixed norm concept is introduced into the equalization problem. For the equalizer structure of a tapped delay line with an infinite number of taps, the mixed norm equalizer has a performance in between the ZF equalizer and the MSE equalizer.

In addition to the choice of criteria, extensive research has been done on the structure of the equalizer [26,27]. In addition to the tapped delay line, the **feedback equalizer** is another important equalizer structure [28,29,30,31,32]. As mentioned earlier, the linear (tapped delay line) equalizer is faced with a dilemma: leaving some intersymbol interference or enhancing the noise. On the other hand, the feedback equalizer which is generally the cascade of a linear filter (called **forward filter**) with a **feedback filter**, has the benefit of compensating for severe channel distortion without enhancing the noise, provided that the delay line of the feedback portion is fed with noise-free signal (which is either from a known reference signal or from the correct decisions at the equalizer output) [11,33]. The choice of the previous three criteria (ZF, MSE, MN) generally will not affect the design of the feedback portion of the equalizer. It is the performance of the forward filter portion that affects the overall performance of a feedback equalizer [34].

In this research, the ISI problem under consideration is the "ghosting" problem in television reception, which is caused by the existence of multiple signal propagation paths

between the transmitter and the receiver. At the receiver, the “main” signal (which is defined as the **strongest** signal and is not necessarily the first signal received) is corrupted by the addition to it of time shifted, attenuated copies, plus noise. In actual implementation, the “raw” signal received is correlated with the local reference signal. The peak of this cross-correlation corresponds to the position of the main signal. The “raw” signal is then shifted by the distance between this peak position and the nominal main signal position, before it is fed to the filter. By doing so, the main signal position is always fixed and known to the filter. The copies which precede the main signal in time are called the **precursor ISI**. The copies that arrive after the main signal are called the **postcursor ISI**. If we adopt the convention that channel introduces no pure delay, then the precursor part of the ISI is caused by the anticausal part of the channel impulse response. This concept is also commonly used in the communication theory [11, 13]. In spite of the fact that all physically realizable systems have to be causal, we use the noncausality concept to facilitate the theoretical analysis. We will discuss the physical implementation in Chapter V.

The feedback filter is usually causal for realizability and stability [34]. This configuration makes the feedback filter able to cancel only the postcursor ISI. Therefore, the forward filter is solely responsible for the cancellation of the precursor ISI. In addition, since the feedback filter deals with the output of the forward filter, the performance of the forward filter will have direct impact on the feedback filter, hence the overall performance of the feedback equalizer. On the other hand, as mentioned earlier, the choice of the

criterion will affect the performance of the forward (linear) filter. Therefore, the choice of the criterion will affect the overall performance of the feedback equalizer.

However, due to the structural difference, the choice of the criterion will have a different degree of impact on the feedback equalizer compared to the linear equalizer. For example, the feedback equalizer based on the zero-forcing criterion is generally less sensitive to the channel spectral nulls than the corresponding linear equalizer [13].

So far, our analysis is based on the ideal equalizer with an infinite number of taps. In practical applications, the equalizer has to be of finite length. Generally speaking, for the ideal equalizer with an infinite length, the frequency domain analysis provides more insight. On the other hand, for the practical equalizer with finite number of taps, the time domain method is more efficient. We follow this convention in this paper with a few exceptions. The practical restriction of finite length also makes the theoretical analysis more involved. For example, for the three criteria mentioned earlier, it is very difficult to find closed form expressions for the MSE of the feedback equalizer with finite length. Therefore, evaluations for the performance of the equalizer in practice rely heavily on numerical solutions. Nevertheless, the theoretical analysis of the ideal equalizer provides a guideline for the practical design, which proves to be indispensable.

Another important practical restriction is that the channel impulse response, or equivalently the channel frequency response, is not available. Moreover, it is usually changing with the environment and possibly with time [35,36]. Therefore, an equalizer should be able to “learn” the channel characteristics by itself and adjust itself to reflect the changes of the channel characteristics. This is usually done by minimizing an objective



function, i.e., the criterion mentioned earlier. For the same equalizer structure, the way in which the equalizer is updated, hence the performance of the equalizer, depends on the criterion chosen. Although some theoretical results can be obtained as a guideline, it was necessary to resort to numerical methods to evaluate the performance of the adaptive or automatic equalizer design.

This paper is organized as follows. In Chapter II, the theoretical background about “norms”, the common abstract representation of our three criteria (ZF, MSE, MN), is reviewed. In particular, the concept of mixed norm is formulated in the framework of Bayesian theory. In Chapter III, the discrete-time communication system model is described, and the optimum linear and feedback equalizers based on the ZF and MSE criteria are reviewed in the context of norm theory, with the assumption of known channel characteristics. These results are extended to include the mixed norm criterion. The expressions for the MSE of the equalizer with infinite length are obtained for various criteria to show the inter-relationship between the reduction of the intersymbol interference and the enhancement of the noise. Further insight is made possible by the comparison of the linear equalizer with the feedback equalizer, and the inclusion of the linear prediction model of the feedback equalizer.

In Chapter IV, adaptive or automatic equalizers with unknown channel characteristics are described. Lucky’s ZF algorithm is re-cast into the  $L_1$  norm framework, resulting in the so-called stochastic ZF algorithm. Widrow’s LMS algorithm is revisited. Finally, an automatic equalizer minimizing the mixed norm is proposed for both the linear and feedback structures.

In Chapter V, numerical evaluation of the various equalizers discussed herein is described. In addition, applications of these equalizers are described, with particular emphasis on the mixed norm feedback equalizer to the television multipath (ghost) cancellation problem. The significance of the mixed norm parameter  $\lambda$  and its relationship to the performance of the equalizer is discovered for the first time. Numerous experiments are conducted with various ghosting scenarios.

Chapter VI concludes this paper by summarizing the main results of the research leading to this paper.

## CHAPTER II

### THE MIXED NORM

#### 1. The $L_p$ Norm

For a general regression problem, assume there are  $N$  points in the  $(M+1)$ -dimensional Euclidean space  $(\mathbf{x}_i, y_i) \in R^{M+1}$ , we can set up a system of linear equations

$$\mathbf{X} \mathbf{c} = \mathbf{y} \quad (2.1)$$

where  $\mathbf{X}$  is an  $N \times M$  matrix,  $\mathbf{c}$  is an  $M \times 1$  vector,  $\mathbf{y}$  is an  $N \times 1$  vector. If  $N > M$ , (2.1) is an overdetermined system of equation; if  $N < M$ , (2.1) is an underdetermined system of equations. ( In this paper, we only consider the case where  $N \geq M \geq 1$ ). For both cases, the solution to (2.1) is not defined unless an additional constraint is imposed. One class of solutions is to minimize the norm of the residual vector. Define the residual vector  $\mathbf{r}(\mathbf{c}) = \mathbf{y} - \mathbf{X} \mathbf{c}$ . The  $L_p$  problem is to find a vector  $\mathbf{c}$  such that the  $p$ -norm  $\|\mathbf{r}(\mathbf{c})\|_p$  is minimized, where

$$\|\mathbf{r}(\mathbf{c})\|_p = \left( \sum_{i=1}^N |r_i(\mathbf{c})|^p \right)^{1/p} \quad (2.2)$$

and

$$r_i(\mathbf{c}) = y_i - \mathbf{x}_i^T \cdot \mathbf{c} \quad (2.3)$$

where  $\mathbf{x}_i^T$  is the  $i^{\text{th}}$  row of  $\mathbf{X}$  and  $1 \leq p \leq \infty$ . For  $1 < p \leq \infty$  and non-negative argument the function  $(\cdot)^{1/p}$  is a monotonically increasing function. Therefore, minimizing  $\|\mathbf{r}(\mathbf{c})\|_p$  is equivalent to minimizing

$$d_p(\mathbf{c}) = \sum_{i=1}^N |r_i(\mathbf{c})|^p \quad (2.4)$$

The  $L_p$  criterion can be further elaborated by introducing the p-Gaussian distribution [23]. Let  $r_i$  ( $i=1, 2, \dots, N$ ) be independent and identically p-Gaussian distributed, i.e., the joint probability density function of  $\mathbf{r}_i$  is

$$f(\mathbf{r}) = \alpha \exp\left\{ -\frac{\beta}{\sigma^p} \sum_{i=1}^N |r_i - \mu|^p \right\} \quad (2.5)$$

where

$$\alpha = \frac{p}{2\Gamma(1/p) \left[ \frac{\Gamma(1/p)}{\Gamma(3/p)} \right]^{1/2} \cdot \sigma} \quad (2.6)$$

$$\beta = \left[ \frac{\Gamma(3/p)}{\Gamma(1/p)} \right]^{p/2}$$

where  $\Gamma(\cdot)$  is the gamma function,  $\mu$  is the mean, and  $\sigma$  is the standard deviation. Figure 2-1 shows the one-dimensional probability density function of  $r_i$  for various values of  $p$ . For  $p=1, 2$  and  $p = \infty$ , it corresponds to the Laplacian, Gaussian and uniform distribution, respectively. Maximizing the likelihood function of  $r_i$ ,  $\ln\{f(\mathbf{r})\}$ , is equivalent to minimizing  $d_p(\mathbf{c})$ . That is to say, when the  $r_i$  are independent and identically p-Gaussian distributed, the  $L_p$  solution is the maximum likelihood estimate of  $\mathbf{c}$ .

The choice of the value of  $p$  in practice depends on the characteristics of  $r_i$ . When  $p=1$ , the  $L_p$  problem falls back to the  $L_1$  problem, which characterizes an  $r_i$  with small values most of the time and huge values occasionally. Similarly, when  $p=2$ , the  $L_p$  problem reduces to the  $L_2$  problem, which characterizes an  $r_i$  of “normal” behavior. When

$p = \infty$ , the  $L_p$  problem corresponds to an “minimax” problem, which characterizes an  $r_i$  of uniform distributions.

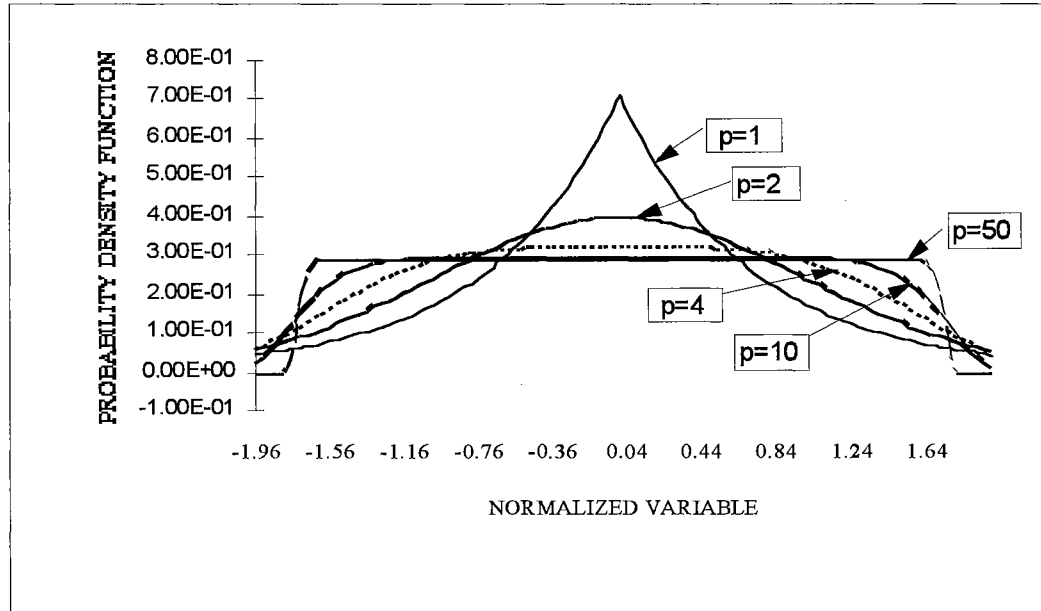


Figure 2-1. p-Gaussian pdf for p=1, 2, 4, 10, 50  
 ( p=1: ——— ; p=2: - - - - ; p=4: ····· ; p=10: - · - · - ; p=50: - - - - - )

Appendix A shows many useful properties of the  $L_1$  and  $L_2$  norms. Generally speaking, the  $L_1$  norm is more resistant to the effect of “outliers” in the measurements. It is more robust with respect to the signal with long-tailed distribution. But the solution to the  $L_1$  problem may not be unique unless an additional constraint is imposed. On the other hand, the  $L_2$  solution is unique, which allows the algorithms like the steepest descent algorithm to be employed to find the  $L_2$  solution.

## 2. Bayesian Estimation and the Mixed Norm

In the previous section, we pointed out that minimizing the  $L_p$  norm of the residual vector is in fact the maximum likelihood estimation of the unknown vector  $\mathbf{c}$ , with the assumption that the elements of the residual vector are independent and identically  $p$ -Gaussian distributed. We have made no assumption about the distribution of the unknown vector  $\mathbf{c}$  itself. In some applications, past experience will provide some prior knowledge about the vector  $\mathbf{c}$ . For example, in the seismic deconvolution problem [37, 38, 25], experience shows that the impulse response (which is the unknown of deconvolution) appears like a series of impulses. Hence the prior assumption of the Laplacian distribution of the unknown seems appropriate.

We now re-write the Gaussian distribution of the measurement vector  $\mathbf{y}$ , with explicit dependency on the unknown vector  $\mathbf{c}$  [24]:

$$f(\mathbf{y}|\mathbf{c}, \sigma) = \left(\frac{1}{2\pi\sigma^2}\right)^{N/2} \exp\left(-\frac{1}{2\sigma^2}\|\mathbf{y} - \mathbf{X}\mathbf{c}\|_2^2\right) \quad (2.7)$$

Assume  $\mathbf{c}$  has a prior distribution  $f(\mathbf{c}|\xi)$  with  $\xi$  unknown. According to Bayes's rule, the posteriori distribution for  $\mathbf{c}$  is

$$f(\mathbf{c}|\mathbf{y}, \xi, \sigma) = \frac{f(\mathbf{y}|\mathbf{c}, \sigma)f(\mathbf{c}|\xi)}{f(\mathbf{y}|\xi, \sigma)} \quad (2.8)$$

$f(\mathbf{y}|\xi, \sigma)$  provides the likelihood function for  $\xi$  and  $\sigma$ . This function has more practical implications, and we leave this topic to Chapter V where we are going to discuss the choice of  $\xi$  (equivalently, the choice of  $\lambda=1-\xi$  as defined in the following) based on a practical application.

If we choose a Laplacian prior distribution, then [24]

$$f(\mathbf{c}|\xi) = \left(\frac{1}{2}\right)^{M/2} \xi^M \exp(-\sqrt{2}\xi\|\mathbf{c} - \mathbf{c}^*\|_1) \quad (2.9)$$

where  $\mathbf{c}^*$  is the mean of  $\mathbf{c}$  and  $\xi$  is the parameter. It is evident from (2.9) that the larger the  $\xi$  is, the smaller the variance of  $f(\mathbf{c}|\xi)$ .

The maximum *a posteriori* (MAP) estimate of  $\mathbf{c}$  will minimize

$$d'(\mathbf{c}) = \|\mathbf{y} - \mathbf{X}\mathbf{c}\|_2^2 + \xi\sigma\sqrt{8}\|\mathbf{c} - \mathbf{c}^*\|_1 \quad (2.10)$$

Equation (2.10) is what we call the **mixed norm**. It combines the  $L_2$  norm of the residual vector and our belief on the conjecture that  $\mathbf{c}$  is Laplacian distributed with mean of  $\mathbf{c}^*$ . The degree of belief is characterized by the parameter  $\xi$ .

Let the middle element of  $\mathbf{c}^*$  be 1 and the rest be 0, i.e.,

$$\mathbf{c}^* = \delta_u = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (2.11)$$

This choice of  $\mathbf{c}^*$  assumes that the mean of  $\mathbf{c}$  is an unit impulse, which makes sense in some applications like the equalization, where the expected value of the combined channel impulse response is a pure delay. If we further let  $\sigma$  be absorbed by  $\xi$  and define

$$\lambda = 1 - \xi \quad (2.12)$$

and modify (2.10) by normalizing the sum, we then obtain the following form of mixed norm with more practical value:

$$d(\mathbf{c}) = \lambda\|\mathbf{y} - \mathbf{X}\mathbf{c}\|_2^2 + (1 - \lambda)\|\mathbf{c} - \delta_u\|_1 \quad (2.13)$$

Since the first term of  $d(\mathbf{c})$  in (2.13) is a strictly convex function of  $\mathbf{c}$ , and the second term is a convex function of  $\mathbf{c}$  (see Appendix A), we have that  $d(\mathbf{c})$  is a strictly convex function of  $\mathbf{c}$ . Therefore, a steepest descent algorithm can be used to find the  $\mathbf{c}$  that minimizes  $d(\mathbf{c})$ .

Similarly, if we chose a Gaussian prior distribution of  $\mathbf{c}$ , the corresponding mixed norm will be

$$d(\mathbf{c}) = \lambda \|\mathbf{y} - \mathbf{X}\mathbf{c}\|_2^2 + (1 - \lambda) \|\mathbf{c} - \delta_u\|_2^2 \quad (2.14)$$

### 3. Summary

In this chapter, we introduced the basic concepts of the  $L_p$  norm, leaving the detailed descriptions of the  $L_1$  and  $L_2$  norms to Appendix A. The normal definition of the  $L_p$  norm of the residual vector does not assume any prior knowledge about the unknown itself. The maximum *a posteriori* estimation was then introduced, which results in the mixed norm. The theoretical mixed norm was then modified for the convenience of practical usage.



## CHAPTER III

### MINIMUM NORMS EQUALIZATION

#### 1. Communication System Model

A communication system generally includes a transmitter, transmission medium and a receiver (Figure 3-1). Depending on the specific application, the transmission medium can have physically different forms: twisted pair telephone cables for voice communication; wireless channel for line-of-sight terrestrial radio and satellite transmission; subscriber loops for integrated services digital networks (ISDN), etc.

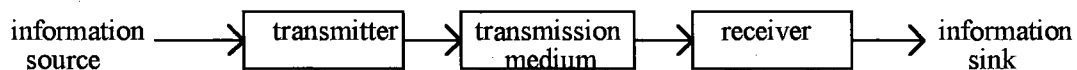


Figure 3-1. Communication system

The **information** can be expressed in various forms, e.g., voice, image, data, etc.. In an electrical communication system, the information is carried by the **signal**. In order to more reliably transmit the signal and more efficiently utilize the transmission medium, the signal is usually **modulated** before it is transmitted. At the receiver, the received signal is

**demodulated** and information is restored to the original form. In the analog world, typical modulation schemes include amplitude modulation (AM) and frequency modulation (FM). With respect to AM, depending on the bandwidth usage, there are three modulation schemes: double-side band (DSB), single-side band (SSB) and vestigial-side band (VSB). For example, in the NTSC (National Television Standard Committee) television system, the image signal is VSB-AM modulated, and the audio signal is FM modulated.

In the past few decades, digital communication has become more and more popular. It has many advantages over the analog counterpart. For example, state-of-the-art digital signal processing (DSP) methods can be used to significantly improve the communication quality, and to utilize the transmission medium even more efficiently. Typical digital modulation schemes include pulse amplitude modulation (PAM), frequency-shift keying (FSK), phase-shift keying (PSK) and quadrature amplitude modulation (QAM).

The modulation and demodulation schemes may or may not be linear. For example, the QAM scheme is linear, but the FM scheme is nonlinear. This paper only concerns the linear modulation schemes. With the linear modulation, the distortion on the modulated signal can be modeled as an equivalent baseband effect. With reasonable approximation for practical interest, most transmission media are also linear. Therefore, the communication system in Figure 3-1 can be approximated by the baseband model shown in Figure 3-2. A signal  $x(t)$  is transmitted through a linear time-invariant (or slowly time-varying relative to the signal rate) channel  $f(t)$ . At the receiver, the signal  $w(t)$  is received. All the noise sources are modeled by an equivalent white Gaussian noise  $\eta(t)$ .

The channel model  $f(t)$  includes the effects of the transmitter filter, the transmission media and the receiver filter. By linear system theory, we have

$$w(t) = x(t) * f(t) + \eta(t) \quad (3.1)$$

where “\*” stands for convolution. If there is no time-dispersion in the channel, i.e.,  $f(t) = \delta(t)$ , where  $\delta(t)$  is the unit impulse function, and there is no noise, i.e.,  $\eta(t) = 0$ , then

$$w(t) = x(t) \quad (3.2)$$

i.e., the transmitted signal can be completely recovered if the channel does not introduce any distortion and there is no noise in the channel.

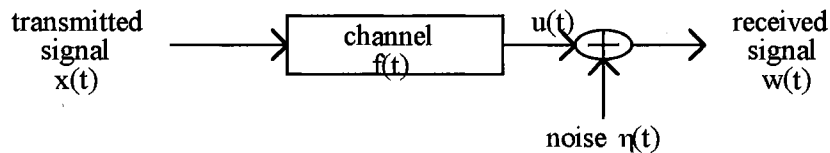


Figure 3-2. Baseband model of the communication system

## 2. Intersymbol Interference and Linear Receivers

### (1) Discrete-time Model

Since most practical channels are not distortion-free, compensation is needed at the receiver side. The scheme used to compensate for the channel distortion is the equalizer. In this section, we focus on the equalizer with linear structure. Modern equalizers are generally implemented by a digital signal processor. Therefore the received signal has to be sampled before it is equalized. Suppose the bandwidth of  $x(t)$  and  $f(t)$  are

$B_x$  and  $B_f$ , respectively. If  $u(t)$  is sampled at a sample frequency  $f_s \geq 2 \cdot \max\{B_x, B_f\}$ , then we have the equivalent discrete-time model shown in Figure 3-3, where  $h(n)$  is the discrete-time equalizer, which is implemented by a tapped delay line. In other words,  $h(n)$  is the impulse response of a FIR system.

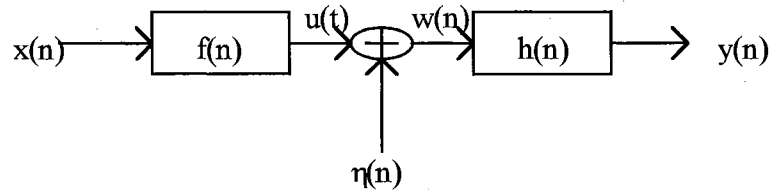


Figure 3-3. Discrete-time model with equalizer

In most communication system today, the timing and phase information needed for proper synchronization are extracted from the received signal itself without a secondary channel. In addition, the equalizer has to be trained with a signal such that fast convergence is possible. This requires the transmitted signal be broad band. In digital communication system, this is implemented by a device called “scrambler” which randomizes the transmitted data before it is modulated. At the receiver, a device called “descrambler” is used to recover the data. For the analog television system where usually no scrambler is used (except in the encrypted cable TV signal which is irrelevant to the topic in this paper), a broad band signal called the Ghost Cancellation Reference (GCR) signal is inserted periodically. This signal is used at the receiver to trained the equalizer.

Therefore, for practical purpose, we can assume a broadband transmitted

signal  $x(n)$  which is wide sense stationary and uncorrelated at different sampling instances and has the power  $\sigma_x^2$ , i.e.

$$\mathcal{E}[x(n) x(n+m)] = \sigma_x^2 \cdot \delta(m) \quad (3.3)$$

where  $\mathcal{E}(\cdot)$  is the expectation operator. Then the problem arises: among all receivers  $h(n)$  with linear structure, which one is the “best”? The answer is that we have to first define what is the meaning of the “best”. In most communication systems, minimizing the probability of error is the goal. Since the probability of error is generally a non-linear function of  $h(n)$ , and the nonlinearity depends on the modulation scheme, usually some simplified measures such as the mean square error (MSE) and intersymbol interference (ISI) are used. In other applications, such as the television multipath cancellation problem we are going to discuss later, reconstructing the transmitted waveform itself is the goal. In this case the MSE and ISI have more practical meaning.

## (2) Linear Zero-forcing Equalizer

**a. Infinite Length.** Assume the equalizer  $h(n)$  has an infinite number of taps. Then the cascade of the discrete-time channel model  $f(n)$  and the equalizer  $h(n)$  can be represented by an equivalent combined channel  $q(n)$ , where

$$q(n) = \sum_{j=-\infty}^{\infty} h(j) f(n-j) \quad (3.4)$$

the output of the equalizer can be expressed as

$$\begin{aligned}
y(n) &= \sum_{j=-\infty}^{\infty} x(j)q(n-j) + \sum_{j=-\infty}^{\infty} \eta(j)h(n-j) \\
&= q(0)x(n) + \sum_{j=-\infty, j \neq n}^{\infty} x(j)q(n-j) + \sum_{j=-\infty}^{\infty} \eta(j)h(n-j)
\end{aligned} \tag{3.5}$$

The first term of (3.5) represents a scaled version of the desired signal  $x(n)$ ; the second term of (3.5) represents the intersymbol interference at the output of the equalizer; and the third term of (3.5) is the noise at the output of the equalizer.

The worst case intersymbol interference is measured by the **peak distortion**, defined as [39]

$$D = \sum_{n=-\infty, n \neq 0}^{\infty} |q(n)| \tag{3.6}$$

Therefore,  $D$  is a function of equalizer taps  $h(n)$ . Given the infinite number of equalizer taps, it is possible to choose a set of equalizer taps such that  $D$  is zero. This is called the **zero-forcing criteria**.

In order to have  $D=0$ , we need  $q(n)=0$  except at  $n=0$ , i.e.,

$$q(n) = \sum_{j=-\infty}^{\infty} h(j)f(n-j) = \delta(n) \tag{3.7}$$

Taking the  $z$  transform of (3.7) leads to

$$H(z)F(z)=1$$

or

$$H(z) = \frac{1}{F(z)} \tag{3.8}$$

where  $H(z)$  and  $F(z)$  are the  $Z$  transforms of  $h(n)$  and  $f(n)$ , respectively. Equation (3.8) indicates that the linear equalizer that minimizes the intersymbol interference is simply the inverse filter of the channel. This equalizer is called zero-forcing linear equalizer,

abbreviated as ZF-LE. Clearly, if the channel is not minimum phase, the resulting  $H(z)$  will be noncausal, since  $H(z)$  is a FIR filter here by definition.

We define the **MSE achieved by an equalizer** as the mean square error of the equalizer output with respect to the desired output, i.e.,

$$\varepsilon^2(\text{equalizer}) = \mathcal{E}\{[x(n) - y(n)]^2\} \quad (3.9)$$

where the “equalizer” is  $\varepsilon^2(\text{equalizer})$  will be replaced by the particular equalizer under study. We will use  $\varepsilon^2(\ )$  as a measure of merit to compare the various equalizers.

Assume the noise  $\eta(n)$  at the equalizer input is white Gaussian noise with spectral density of  $N_o$ , then the spectral density of the noise at the output of the equalizer will be

$$\frac{N_o}{|F(e^{j2\pi fT})|^2} \quad 0 \leq f \leq \frac{1}{T}$$

where  $T$  is the sampling period, with the above definition the MSE achieved by such equalizer is

$$\varepsilon^2(\text{ZF-LE}) = T \int_0^{1/T} \frac{N_o}{|F(e^{j2\pi fT})|^2} df \quad (3.10)$$

From (3.10), it can be seen that the ZF equalizer will enhance the noise over those frequency regions where the channel frequency response has nulls.

**b. Finite Length.** For practical applications, the equalizer has to be of finite length. For an equalizer with  $(2M+1)$  taps, Equation (3.7) is re-written as

$$\sum_{j=-M}^M h(j)f(n-j) = \delta(n) \quad (3.11)$$

Although in practical applications,  $h(n)$  needs to represent a causal system, here, as mentioned earlier, we adopt the perspective that the channel impulse response is noncausal, and our equalizer impulse response can also be noncausal for theoretical convenience. Equation (3.11) can be written in matrix form, evaluated at  $n=-M, \dots, +M$ :

$$\mathbf{F} \mathbf{h} = \boldsymbol{\delta} \quad (3.12)$$

where

$$\mathbf{F} = \begin{bmatrix} f(0) & f(-1) & \dots & f(-2M) \\ f(1) & f(0) & \dots & f(-2M+1) \\ \dots & \dots & \dots & \dots \\ f(2M) & f(2M-1) & \dots & f(0) \end{bmatrix} \quad \mathbf{h} = \begin{bmatrix} h(-M) \\ \vdots \\ h(0) \\ \vdots \\ h(M) \end{bmatrix} \quad \boldsymbol{\delta} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad (3.13)$$

The ZF equalizer coefficient vector is given by

$$\mathbf{h}_{(ZF-LE)} = \mathbf{F}^{-1} \boldsymbol{\delta} \quad (3.14)$$

Therefore, if the impulse response of the channel,  $f(j)$  is known and the matrix  $\mathbf{F}$  is not singular, the ISI within the range covered by the equalizer,  $q(n)$  ( $-M \leq n \leq M, n \neq 0$ ), can be canceled exactly. However, in most applications, the channel impulse response is not known; it has to be estimated from the received signal. An automatic equalizer or adaptive equalizer will be discussed in Chapter IV.



### (3) Linear MSE Equalizer

a. **Infinite Length.** As mentioned earlier, the ZF equalizer can eliminate the ISI completely but at the expense of possible severe enhancement of noise. The minimum mean square error (MSE) equalizer, which minimizes the mean square error between the output of the equalizer and the desired signal, does not suffer from this drawback. Assume the transmitted signal  $x(n)$  is white with power of  $\sigma_x^2$ , as shown in Equation (3.3). The MSE is, as defined in (3.9),

$$\mathcal{E}[e^2(n)] = \mathcal{E}[(x(n) - y(n))^2] \quad (3.15)$$

where  $e(n)$  is the error at the equalizer output, shown in Figure 3-4.

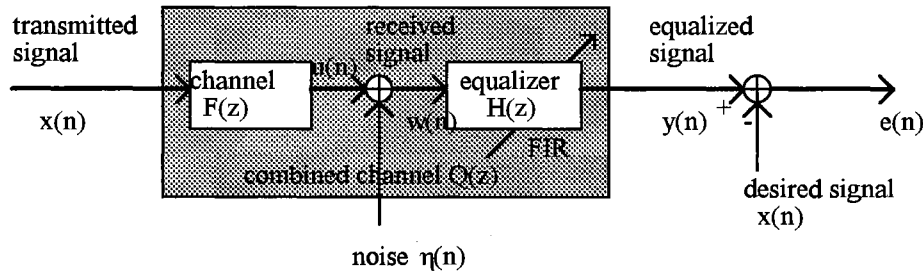


Figure 3-4. Communication system with linear equalizer

By Parseval's theorem, the MSE can also be expressed as the integral of the power spectral density of the error, i.e.,

$$\mathcal{E}[e^2(n)] = T \int_0^{1/T} |E(e^{j2\pi fT})|^2 df \quad (3.16)$$

where  $E(\cdot)$  denotes the spectrum of  $e(n)$ . To further evaluate Equation (3.16), we start with the frequency domain equivalence of Figure 3-4, shown in Figure 3-5.

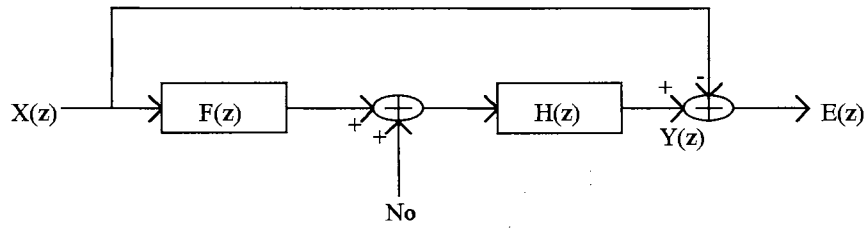


Figure 3-5. Frequency domain equivalence of Figure 3-4

Figure 3-5 can be further simplified as Figure 3-6.

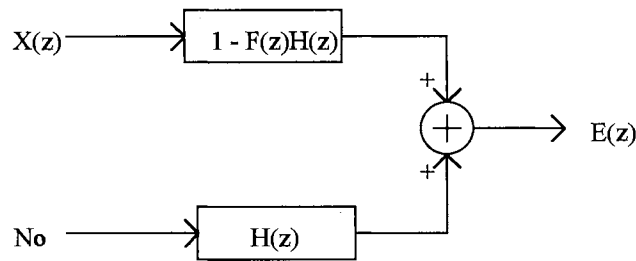


Figure 3-6. Simplified version of Figure 3-5

We have assumed that the sequence  $x(n)$  is uncorrelated and we further assume that  $x(n)$  and the noise  $n(n)$  are uncorrelated. Therefore the power spectral density of  $e(n)$  is

$$\left| E(e^{j2\pi fT}) \right|^2 = \sigma_x^2 \left| 1 - F(e^{j2\pi fT}) H(e^{j2\pi fT}) \right|^2 + N_o \left| H(e^{j2\pi fT}) \right|^2 \quad (3.17)$$

The MSE is expressed as

$$\mathcal{E}[e^2(n)] = T \int_0^{1/T} \sigma_x^2 \left| 1 - F(e^{j2\pi fT}) H(e^{j2\pi fT}) \right|^2 + N_o \left| H(e^{j2\pi fT}) \right|^2 df \quad (3.18)$$

From (3.18), it is evident that the MSE consists of two parts: the first term in the integrand is the contribution from ISI ; the second term is the contribution from the noise. The equalizer that minimizes the MSE will minimize the sum of both contributions. Since there is no constraint on the complexity of the equalizer except that it has to be linear , we can minimize the MSE of Equation (3.18) by minimizing the power spectral density of Equation (3.17) at each frequency. We do this by rearranging the terms in Equation (3.17):

$$\begin{aligned}
\left| E(e^{j2\pi fT}) \right|^2 &= \sigma_x^2 \{1 - 2 \operatorname{Re}(FH) + |F|^2 |H|^2\} + N_o |H|^2 \\
&= (\sigma_x^2 |F|^2 + N_o) |H|^2 - 2 \sigma_x^2 \cdot \operatorname{Re}(F \cdot H) + \sigma_x^2 \\
&= (\sigma_x^2 |F|^2 + N_o) \left[ |H|^2 - \frac{2 \sigma_x^2 \cdot \operatorname{Re}(F \cdot H)}{\sigma_x^2 |F|^2 + N_o} + \left( \frac{\sigma_x^2 \cdot |F|}{\sigma_x^2 |F|^2 + N_o} \right)^2 \right] \\
&\quad - \frac{|\sigma_x^2 \cdot F|^2}{\sigma_x^2 |F|^2 + N_o} + \sigma_x^2 \\
&= (\sigma_x^2 |F|^2 + N_o) \left| H - \frac{F^*}{|F|^2 + \frac{N_o}{\sigma_x^2}} \right|^2 + \frac{N_o}{|F|^2 + \frac{N_o}{\sigma_x^2}}
\end{aligned} \tag{3.19}$$

where  $\operatorname{Re}(\cdot)$  means the real part of a complex number. The second term of (3.19) is independent of  $H$ . Therefore  $\left| E(e^{j2\pi fT}) \right|^2$  can be minimized by minimizing the first term of (3.19) , which can be done by setting:

$$H = \frac{F^*}{|F|^2 + \frac{N_o}{\sigma_x^2}} \tag{3.20}$$

This is indeed a type of noncausal Wiener filter. The relationship between the MSE equalizer and the Wiener filter is further elaborated in Appendix B. The MSE corresponding to the equalizer of (3.20) is

$$\mathcal{E}^2(MSE - LE) = T \int_0^{1/T} \frac{N_o}{|F|^2 + \frac{N_o}{\sigma_x^2}} df \quad (3.21)$$

where ‘‘MSE-LE’’ stands for mean-square-error linear equalizer. Since  $N_o \geq 0$  and  $\sigma_x^2 \geq 0$ , we always have

$$\frac{N_o}{|F|^2 + \frac{N_o}{\sigma_x^2}} \leq \frac{N_o}{|F|^2}, \quad 0 \leq f \leq \frac{1}{T} \quad (3.22)$$

Therefore, from (3.10) and (3.22), we have

$$\mathcal{E}^2(MSE - LE) \leq \mathcal{E}^2(ZF - LE) \quad (3.23)$$

In other words, for a linear equalizer with an infinite number of taps, the MSE equalizer always yields less MSE value than the ZF equalizer.

We come back to examine the implication of Equation (3.20). We can see that the equalizer  $H$  strikes a balance between inverting the channel frequency response and minimizing the noise enhancement. For those frequency regions where signal to noise ratio is large, i.e.,

$$|F|^2 \gg \frac{N_o}{\sigma_x^2}$$

the equalizer tries to **invert** the channel frequency response:

$$H \approx \frac{F^*}{|F|^2} = \frac{1}{F}$$

For those frequency regions where signal-to-noise ratio is small, i.e.,

$$|F|^2 \ll \frac{N_o}{\sigma_x^2}$$

the equalizer tries to **match** the channel:

$$H \approx \frac{\sigma_x^2}{N_o} \cdot F^*$$

This makes a lot of sense, since the matched filter will maximize the output signal-to-noise ratio if the input signal is corrupted only by additive white Gaussian noise.

**b. Finite Length.** For an equalizer with  $(2M+1)$  taps, the output of the equalizer

$$\begin{aligned} y(n) &= \sum_{j=-M}^M h(j)w(n-j) \\ &= \mathbf{h}^T \cdot \mathbf{w} \end{aligned} \quad (3.24)$$

where

$$\begin{aligned} \mathbf{w} &= [w(n+M) \dots w(n) \dots w(n-M)]^T \\ &= \mathbf{u} + \boldsymbol{\eta} = \mathbf{F}^T \mathbf{x} + \boldsymbol{\eta} \end{aligned} \quad (3.24a)$$

is the input to the equalizer, i.e. the received signal with noise,  $\mathbf{F}$  is defined in (3.13),  $\mathbf{u} = \mathbf{F}^T \mathbf{x}$  is the noise-free output of the channel,  $\mathbf{x}$  is the vector consisting of the transmitted signal samples and  $\boldsymbol{\eta}$  is a vector consisting of noise samples. The MSE is

$$\mathcal{E}[(x(n) - y(n))^2] = \mathcal{E}[x^2(n) - 2x(n)\mathbf{h}^T \mathbf{w} + \mathbf{h}^T \mathbf{w} \mathbf{w}^T \mathbf{h}] \quad (3.25)$$

The MSE is minimized by setting the gradient of the MSE with respect to  $\mathbf{h}$  to 0, which leads to

$$\mathbf{h}(\text{MSE} - \text{LE}) = \mathbf{A}^{-1} \boldsymbol{\alpha} \quad (3.26)$$

where

$$\begin{aligned}
\mathbf{A} &= \mathcal{E}[\mathbf{w}\mathbf{w}^T] \\
&= \mathcal{E}[(\mathbf{F}^T \mathbf{x} + \boldsymbol{\eta}) \cdot (\mathbf{F}^T \mathbf{x} + \boldsymbol{\eta})^T] = \sigma_x^2 \cdot \mathbf{I} \cdot \mathbf{F}^T \mathbf{F} + N_o \mathbf{I}
\end{aligned} \tag{3.27}$$

is the covariance matrix of the input signal, where  $\mathbf{I}$  is the  $(2M+1)$  by  $(2M+1)$  identity matrix, and

$$\begin{aligned}
\alpha &= \mathcal{E}[\mathbf{w}x(n)] \\
&= \mathcal{E}[(\mathbf{F}^T \mathbf{x} + \boldsymbol{\eta}) \cdot x(n)] \\
&= \begin{bmatrix} f(M) \\ \vdots \\ f(0) \\ \vdots \\ f(-M) \end{bmatrix} \sigma_x^2 \\
&= \mathbf{F}^T \boldsymbol{\delta} \cdot \sigma_x^2
\end{aligned} \tag{3.28}$$

is the cross-correlation vector. Assuming that  $x(n)$  has zero mean and variance  $\sigma_x^2$ , from

Equation (3.28), the corresponding MSE is given by

$$\begin{aligned}
\varepsilon^2_{(MSE - LE)} &= \mathcal{E}[(x(n) - \mathbf{h}^T_{(MSE - LE)} \mathbf{w})^2] \\
&= \sigma_x^2 - \alpha^T \mathbf{A}^{-1} \mathcal{E}[\mathbf{w}\mathbf{w}^T] \mathbf{A}^{-1} \alpha \\
&= \sigma_x^2 - \alpha^T \mathbf{A}^{-1} \alpha \\
&= \sigma_x^2 - \alpha^T \mathbf{h}_{(MSE - LE)}
\end{aligned} \tag{3.29}$$

Where we have used the fact that  $\mathbf{A}$  is symmetric, which is obvious from (3.27).

Now, we can establish the relationship between the MSE of the linear ZF equalizer and that of the linear MSE equalizer. From (3.25), (3.26) and (3.27), we have

$$\begin{aligned}
\varepsilon^2_{(ZF - LE)} &= \sigma_x^2 - 2\mathbf{h}^T_{(ZF - LE)} \alpha + \mathbf{h}^T_{(ZF - LE)} \mathbf{A}^{-1} \mathbf{h}_{(ZF - LE)} \\
&= \sigma_x^2 - 2\mathbf{h}^T_{(ZF - LE)} \mathbf{A} \mathbf{h}_{(MSE - LE)} + \mathbf{h}^T_{(ZF - LE)} \mathbf{A}^{-1} \mathbf{h}_{(ZF - LE)}
\end{aligned} \tag{3.29a}$$

From (3.29) and (3.29a), we obtain

$$\begin{aligned}
& \mathcal{E}^2(ZF - LE) - \mathcal{E}^2(MSE - LE) \\
&= -2\mathbf{h}^T(ZF - LE)\mathbf{A}\mathbf{h}(MSE - LE) + \mathbf{h}^T(ZF - LE)\mathbf{A}^{-1}\mathbf{h}(ZF - LE) \\
&+ \mathcal{Q}^T\mathbf{h}(MSE - LE) \\
&= -2\mathbf{h}^T(ZF - LE)\mathbf{A}\mathbf{h}(MSE - LE) + \mathbf{h}^T(ZF - LE)\mathbf{A}^{-1}\mathbf{h}(ZF - LE) \\
&+ \mathbf{h}^T(ZF - LE)\mathbf{A}^T\mathbf{h}(MSE - LE)
\end{aligned} \tag{3.29b}$$

Because  $\mathbf{A}$  is symmetric by the definition in (3.27), we have

$$\begin{aligned}
\mathcal{E}^2(ZF - LE) &= \mathcal{E}^2(MSE - LE) - 2\mathbf{h}^T(ZF - LE)\mathbf{A}\mathbf{h}(MSE - LE) + \mathbf{h}^T(ZF - LE)\mathbf{A}^{-1}\mathbf{h}(ZF - LE) \\
&+ \mathbf{h}^T(ZF - LE)\mathbf{A}\mathbf{h}(MSE - LE) \\
&= \mathcal{E}^2(MSE - LE) + (\mathbf{h}(ZF - LE) - \mathbf{h}(MSE - LE))^T \mathbf{A} (\mathbf{h}(ZF - LE) - \mathbf{h}(MSE - LE))
\end{aligned} \tag{3.30}$$

where “ZF-LE” stands for zero-forcing linear equalizer.

It will be interesting to compare the ZF solution with the MSE solution. First, we observe that

$$\mathbf{u} = \mathbf{F}^T \mathbf{x} \tag{3.31}$$

and

$$\mathbf{w} = \mathbf{u} + \boldsymbol{\eta} \tag{3.32}$$

where  $\boldsymbol{\eta}$  is the vector consisting of noise sample. Then from Equation (3.27), we have

$$\begin{aligned}
\mathbf{A} &= \mathcal{E}[(\mathbf{u} + \boldsymbol{\eta})(\mathbf{u} + \boldsymbol{\eta})^T] \\
&= \mathcal{E}[(\mathbf{F}^T \mathbf{x} + \boldsymbol{\eta})(\mathbf{F}^T \mathbf{x} + \boldsymbol{\eta})^T] \\
&= \mathbf{F}^T \cdot \mathcal{E}(\mathbf{x}\mathbf{x}^T)\mathbf{F} + N_o \cdot \mathbf{I} \\
&= \sigma_x^2 \mathbf{F}^T \mathbf{F} + N_o \cdot \mathbf{I}
\end{aligned} \tag{3.33}$$

Similarly, from Equation (28), we have

$$\begin{aligned}
\alpha &= \mathcal{E}[(\mathbf{u} + \eta)x(n)] \\
&= \mathcal{E}[(\mathbf{F}^T \mathbf{x} + \eta)x(n)] \\
&= \sigma_x^2 \mathbf{F}^T \delta
\end{aligned} \tag{3.34}$$

From Equation (3.26), we have

$$\mathbf{h}_{(MSE - LE)} = (\sigma_x^2 \mathbf{F}^T \mathbf{F} + N_0 \cdot \mathbf{I})^{-1} \sigma_x^2 \mathbf{F}^T \delta \tag{3.35}$$

Therefore, if there is no noise at all, i.e.,  $N_0 = 0$ , then

$$\mathbf{h}_{(MSE - LE)} = \mathbf{F}^{-1} \delta \tag{3.35a}$$

which is the same as the ZF equalizer.

#### (4) Mixed Norm Linear Equalizer

**a. Infinite Length.** While the ZF equalizer tries to eliminate the ISI completely without considering the effect of noise, the MSE equalizer treats the ISI and the noise equally. The mixed norm equalizer can highlight the relative significance of the two by adjusting a parameter.

As mentioned in Chapter II, section 4, the mixed norm consists of two parts. The first part is the contribution from the error at the equalizer output; the second part is the prior knowledge about the combined channel  $q(n)$ . Since the peak distortion defined in (3.6) is independent of the signal power, and the MSE term is related to the signal power, the peak distortion term is multiplied by  $\sigma_x^2$  before it is added to the MSE term for proper scaling. Nonetheless, this new measure is still indicative of how well the peak distortion is



minimized. With this modification, the mixed norm problem under consideration is to minimize

$$\lambda \cdot \mathcal{E}[e^2(n)] + (1 - \lambda) \cdot \mathcal{E}[\{[q(n) - \delta(n)]x(n)\}^2]$$

Note that the second term also implicitly assumes the choice of the mean value of the unknown  $q(n)$  to be  $\delta_u$ , as in (2.11).

In practice, before the equalizer, an automatic gain control (AGC) device is usually used to adjust the received signal to the nominal level. Therefore, for simplicity, we can assume that  $q(0)=1$ . For the case of a Gaussian distribution, according to Parseval's theorem, the peak distortion, defined as an  $L_2$  norm here, can be expressed as the integral of the power spectral density,

$$\sum_{n \neq 0} q^2(n) = T \int_0^{1/T} |\mathcal{Z}\{q(n) - \delta(n)\}|_{z=e^{j2\pi f}}^2 df \quad (3.36)$$

where  $\mathcal{Z}\{.\}$  is the Z transform operator, and

$$\begin{aligned} & \mathcal{Z}\{q(n) - \delta(n)\} \\ &= \mathcal{Z}\{f(n) * h(n)\} - \mathcal{Z}\{\delta(n)\} \\ &= F(z) \cdot H(z) - 1 \end{aligned} \quad (3.37)$$

The power spectral density function corresponding to the linear MSE equalizer was given by equation (3.17). The power spectral density function corresponding to the linear mixed norm equalizer is the weighted sum of (3.17) and the contribution of the peak distortion:

$$\begin{aligned}
& \lambda \left[ \sigma_x^2 |1 - F \cdot H|^2 + N_o |H|^2 \right] + (1 - \lambda) |F \cdot H - 1|^2 \cdot \sigma_x^2 \\
&= \lambda \left[ \sigma_x^2 (1 - 2 \cdot \text{Re}(F \cdot H) + |F|^2 |H|^2) + N_o |H|^2 \right] + (1 - \lambda) |F \cdot H - 1|^2 \cdot \sigma_x^2 \\
&= (\sigma_x^2 |F|^2 + \lambda \cdot N_o) |H|^2 - 2 \sigma_x^2 \cdot \text{Re}(F \cdot H) + \sigma_x^2 \\
&= (\sigma_x^2 |F|^2 + \lambda \cdot N_o) \left\{ |H|^2 - \frac{2 \sigma_x^2 \cdot \text{Re}(F \cdot H)}{(\sigma_x^2 |F|^2 + \lambda \cdot N_o)} + \left| \frac{\sigma_x^2 F}{(\sigma_x^2 |F|^2 + \lambda \cdot N_o)} \right|^2 \right\} - \frac{|\sigma_x^2 F|^2}{(\sigma_x^2 |F|^2 + \lambda \cdot N_o)} + \sigma_x^2 \\
&= (\sigma_x^2 |F|^2 + \lambda \cdot N_o) \left| H - \frac{F^*}{|F|^2 + \lambda \frac{N_o}{\sigma_x^2}} \right|^2 + \frac{\lambda \cdot N_o}{|F|^2 + \lambda \frac{N_o}{\sigma_x^2}}
\end{aligned} \tag{3.38}$$

The equalizer that minimizes this spectral density at each frequency is

$$H = \frac{F^*}{|F|^2 + \lambda \frac{N_o}{\sigma_x^2}} \tag{3.39}$$

and the corresponding MSE is

$$\mathcal{E}^2(MN - LE) = T \int_0^{1/T} \frac{\lambda N_o}{|F|^2 + \lambda \frac{N_o}{\sigma_x^2}} df \tag{3.40}$$

Since  $0 \leq \lambda \leq 1$ , from (3.10), (3.21) and (3.40) we have

$$\mathcal{E}^2(MSE - LE) \leq \mathcal{E}^2(MN - LE) \leq \mathcal{E}^2(ZF - LE) \tag{3.40a}$$

where ‘‘MN-LE’’ stands for mixed-norm linear equalizer.

From Equation (3.39), we observe that the mixed norm (MN) equalizer, like the MSE equalizer, also strikes a balance between inverting the channel to reduce the ISI and matching the channel to reduce the noise at the output. The difference is that a parameter,  $\lambda$ , is introduced in the MN equalizer to adjust the relative significance of the ISI and the

noise. One may ask: since the MSE equalizer already can optimize the overall MSE, why do we need the MN equalizer which yields a higher MSE? The answer is that the MSE equalizer yields a lower MSE only in the case of infinite number of equalizer taps. In practice, since we have finite number of equalizer taps, the MSE equalizer may not always yield less MSE than the MN equalizer, especially when the feedback mechanism is introduced, as will be evident later.

**b. Finite Length.** The finite version of the mixed norm is given by

$$\begin{aligned}
& \mathcal{E}\left\{\lambda[x(n)-y(n)]^2+(1-\lambda)[f(n)*h(n)-\delta(n)]^2x^2(n)\right\} \\
&= \lambda \cdot \mathcal{E}\left\{[x(n)-\mathbf{h}^T \mathbf{w}]^2\right\}+(1-\lambda) \sigma_x^2 \cdot \mathcal{E}\left\{[\mathbf{Fh}-\delta]^T[\mathbf{Fh}-\delta]\right\} \\
&= \lambda \cdot \mathcal{E}\left\{x^2(n)-2x(n) \mathbf{h}^T \mathbf{w}+\mathbf{h}^T \mathbf{w} \mathbf{w}^T \mathbf{h}\right\}+(1-\lambda) \sigma_x^2\left\{\mathbf{h}^T \mathbf{F}^T \mathbf{Fh}-2 \mathbf{Fh}+1\right\}
\end{aligned} \tag{3.41}$$

Taking the gradient of (3.41) with respect to  $\mathbf{h}$  and setting it to zero leads to

$$-2 \lambda \cdot \mathcal{E}\{x(n) \mathbf{w}\}+2 \lambda \cdot \mathcal{E}\{\mathbf{w} \mathbf{w}^T\} \mathbf{h}+2(1-\lambda) \cdot \sigma_x^2 \cdot \mathbf{F}^T(\mathbf{Fh}-\delta)=0 \tag{3.42}$$

Recall from equations (3.33) and (3.34), we have

$$\mathcal{E}\{x(n) \mathbf{w}\}=\sigma_x^2 \mathbf{F}^T \delta \tag{3.42a}$$

and

$$\mathcal{E}\{\mathbf{w} \mathbf{w}^T\}=\sigma_x^2 \mathbf{F}^T \mathbf{F}+N_o \cdot \mathbf{I} \tag{3.42b}$$

Equations (3.42), (3.42a) and (3.42b) together lead to

$$\left(\sigma_x^2 \mathbf{F}^T \mathbf{F}+\lambda \cdot N_o \cdot \mathbf{I}\right) \mathbf{h}-\mathbf{F}^T \delta \cdot \sigma_x^2=0 \tag{3.43}$$

or

$$\mathbf{h}_{(MN-LE)}=\left(\mathbf{F}^T \mathbf{F}+\lambda \cdot \frac{N_o}{\sigma_x^2} \cdot \mathbf{I}\right)^{-1} \mathbf{F}^T \delta \tag{3.44}$$

On the other hand, from (3.26), (3.42a) and (3.42b), we have

$$\mathbf{h}_{(MSE - LE)} = (\mathbf{F}^T \mathbf{F} + \frac{N_0}{\sigma_x^2} \cdot \mathbf{I})^{-1} \mathbf{F}^T \delta \quad (3.44a)$$

### 3. Feedback Equalizer

#### (1) Why Do We Need A Feedback Equalizer ?

In theory, for a channel with a frequency response of

$$F(e^{j\omega T}) = |F(e^{j\omega T})| e^{j\angle F(e^{j\omega T})}$$

a linear equalizer with an infinite number of taps and frequency response function  $H(e^{j\omega T})$  could be designed to remove all phase distortion from the received signal without increasing the noise power at the equalizer output. This could be accomplished by making  $H(e^{j\omega T})$  an all pass filter, such that  $|H(e^{j\omega T})| = 1$  and  $\angle H(e^{j\omega T}) = -\angle F(e^{j\omega T})$ . However, removing the amplitude distortion would require

$$|H| = \frac{1}{|F|} \quad (\text{ZF-LE})$$

or

$$|H| = \frac{|F|}{|F|^2 + N_0 / \sigma_x^2} \quad (\text{MSE-LE})$$

or

$$|H| = \frac{|F|}{|F|^2 + \lambda \cdot N_0 / \sigma_x^2} \quad (\text{MN-LE})$$

which could greatly enhance the noise power at the equalizer output. The problem is inherent in the linear structure. The inclusion of a feedback portion has the potential to compensate for severe amplitude distortion without enhancing the noise. The data in the feedback filter is either from the previous decision as in the case of the decision feedback equalizer (DFE), or from the reference signal itself as in the applications we are going to discuss in chapter V. In the case where a reference signal is available, the data in the feedback delay line is noise free. In the case of the DFE, for a reasonable low error probability in the decision, the decision output is basically noise free, although a wrong decision can cause error propagation [13]. Therefore, generally speaking, the feedback portion does not provide noise enhancement.

Another motivation for using the feedback filter arises from the practical restriction of a finite number of equalizer taps. In equalizing a multipath signal, the multipath is generally modeled by an FIR system. Therefore, to effectively compensate for the multipath distortion the equalizer is generally a feedback filter[6,7,8,9,10], although this feedback filter may be unstable if the multipath FIR system has zeros outside the unit circle. If the equalizer is restricted to be FIR, the number of taps of the equalizer will be much more than the one with feedback structure.

Figure 3-7 shows the structure of the decision feedback equalizer where the input of the feedback delay line comes from the decision output. Figure 3-8 shows another form of feedback equalizer where the input of the feedback delay line comes from a known reference signal. In both cases, the feedback portion  $g(n)$  strives to cancel the post cursor (after the reference position) ISI at the output of the combined channel, which consists of

the channel  $f(n)$  and the linear equalizer  $h(n)$ . The forward filter  $h(n)$  is responsible for canceling only the precursor ISI, and therefore is anticausal. Figures 3-8a and 3-8b illustrate the taps of the forward filter  $h(n)$  and feedback filter  $g(n)$ , respectively.

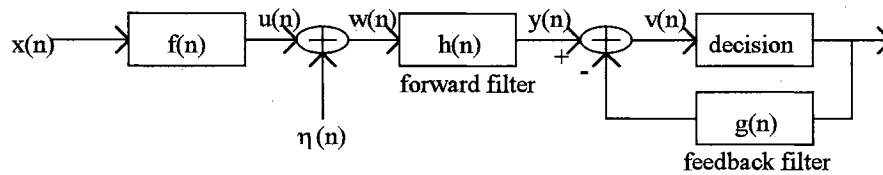


Figure 3-7. Decision feedback equalizer

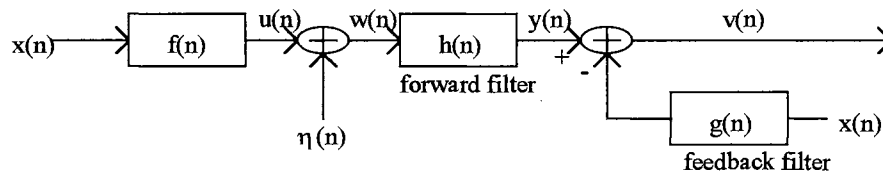


Figure 3-8. Feedback equalizer with reference signal

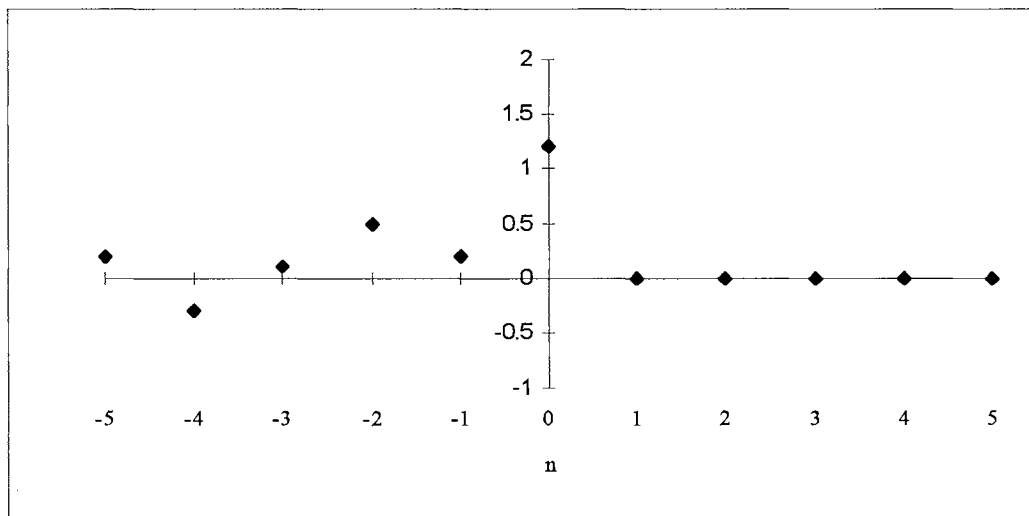


Figure 3-8a. Typical coefficients distribution of the forward filter  $h(n)$

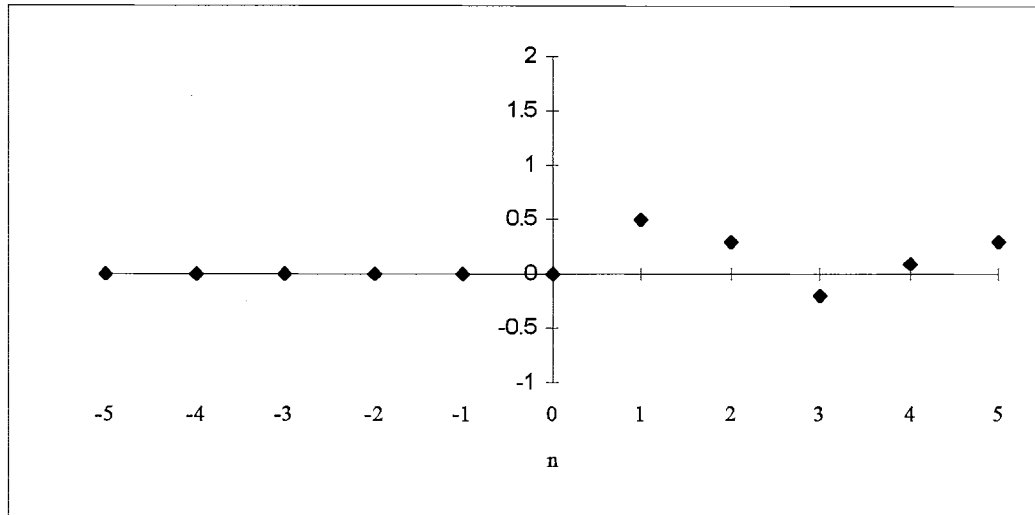


Figure 3-8b. Typical coefficients distribution of the feedback filter  $g(n)$

In the following, we will focus on the feedback equalizer with a reference signal (Figure 3-8), since it is the waveform that needs to be restored in the application we are going to discuss later.

As mentioned earlier, the reference position is always known to the equalizer (by correlating the “row” signal received to the local reference signal). In the feedback equalizer, the forward filter  $h(n)$  is responsible for eliminating the precursor ISI (samples of the channel impulse response before the reference position); the feedback portion is responsible for canceling the postcursor ISI (samples of the channel impulse response after the reference position). Since removing only the precursor ISI using the forward filter does not enhance noise as much as removing all ISI, and the feedback portion plays no part in the effect of noise, it is expected that the feedback equalizer will have less overall noise enhancement than the linear equalizer.

The feedback portion  $g(n)$  must be strictly causal. Therefore  $g(n)=0$  for  $n \leq 0$ . If we allow the feedback part to have an infinite number of taps, we can completely eliminate the postcursor ISI by choosing suitable set of coefficients, as shown in the following .

Referring to Figure 3-8, the output of the entire feedback equalizer is

$$\begin{aligned} v(n) &= \sum_{j=-\infty}^{\infty} q(j)x(n-j) + \sum_{j=-\infty}^{\infty} \eta(j)h(n-j) - \sum_{j=1}^{\infty} g(j)x(n-j) \\ &= \sum_{j=-\infty}^{-1} q(j)x(n-j) + \sum_{j=1}^{\infty} (q(j) - g(j))x(n-j) + \sum_{j=-\infty}^{\infty} \eta(j)h(n-j) \end{aligned} \quad (3.46)$$

and the MSE at the equalizer output is

$$\begin{aligned} &\mathcal{E}\{[v(n) - x(n)]^2\} \\ &= \mathcal{E}\left\{\left[\sum_{j=-\infty}^{-1} q(j)x(n-j) + \sum_{j=1}^{\infty} (q(j) - g(j))x(n-j) + \sum_{j=-\infty}^{\infty} \eta(j)h(n-j) - x(n)\right]^2\right\} \end{aligned} \quad (3.47)$$

Given that the sequence  $x(n)$  is uncorrelated, and  $x(n)$  and  $\eta(n)$  are uncorrelated with each other,

$$\begin{aligned} &\mathcal{E}\{[v(n) - x(n)]^2\} \\ &= \sigma_x^2 \sum_{j=-\infty}^{-1} q^2(j) + \sigma_x^2 (q(0) - 1)^2 + \sigma_x^2 \sum_{j=1}^{\infty} (q(j) - g(j))^2 + N_o \sum_{j=-\infty}^{\infty} h^2(j) \end{aligned} \quad (3.48)$$

From Equation (3.48), once the forward filter is chosen, the overall MSE can be minimized by choosing the feedback filter

$$g(j) = q(j) \quad \text{for } j=1, 2, 3, \dots \quad (3.49)$$

where  $q(j)$  is the combined channel impulse response. Equation (3.49) shows that for a given forward filter, the feedback filter that minimizes the overall MSE actually cancels the postcursor ISI at the output of the forward filter completely.



It remains to find a suitable forward filter. As with the linear equalizer, there are several criteria regarding how to choose the forward filter, and we name the corresponding feedback equalizers after these criteria.

## (2) Zero-forcing Feedback Equalizer

**a. Infinite Length.** We adopt a method proposed by Messerschmitt [34], since it provides more insight into the physical meaning of the feedback equalizer. We start with introducing the predictor  $(1+G(z))$ . The forward filter of the zero-forcing feedback equalizer (ZF-FE) is modeled as a cascade of the zero-forcing linear equalizer (ZF-LE) and a predictor, shown in Figure 3-9. The ZF-LE will invert the channel, and the predictor only introduces post cursor ISI since  $G(z)$  is causal. The ISI introduced by the predictor will be completely canceled by the feedback portion  $G(z)$ . Note that the model of the forward filter is just for illustration purpose. It is not actually how the forward filter is implemented.

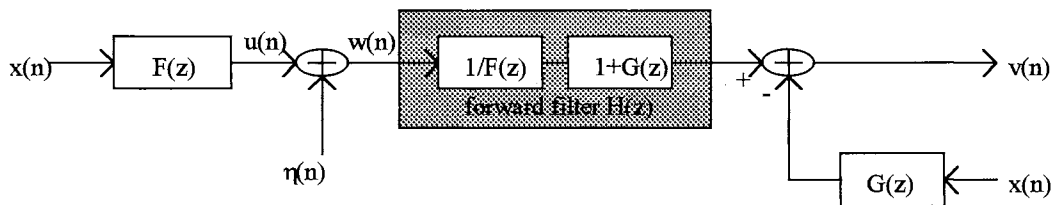


Figure 3-9. ZF feedback equalizer model

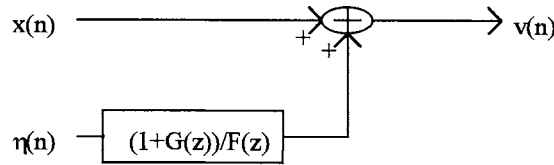


Figure 3-10. Equivalent ZF feedback equalizer model

An equivalent model of Figure 3-9 is shown in Figure 3-10, where the signal and noise are treated separately. Since the ISI is completely canceled (precursor ISI is canceled by the forward filter, and postcursor ISI is canceled by the feedback filter), there is no ISI at the output of the entire equalizer. Therefore, the choice of the predictor  $(1+G(z))$  has no effect on the signal itself, but it does have impact on the noise at the output. Then  $G(z)$  can be chosen to minimize the noise variance at the output. Therefore, the ZF-FE has more degrees of freedom in minimizing the noise than the ZF-LE whose sole purpose is to invert the channel regardless of how much noise is in the system. This qualitatively explains why the ZF-FE enhances less noise than the ZF-LE. We will compare the MSE values more quantitatively later.

In order to find the predictor  $(1+G(z))$  which will minimize the noise variance, we rearrange the noise path in Figure 3-10, shown in Figure 3-11:

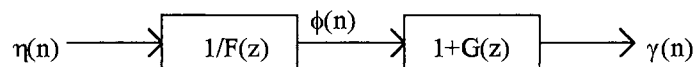


Figure 3-11. Rearrangement of noise path for ZF-FE

where the white noise sequence  $\eta(n)$  is colored by the ZF-LE (which inverts the channel) to yield  $\phi(n)$ . The predictor  $(1+(G(z)))$  strives to fit an autoregressive (AR) model to the sequence  $\phi(n)$  by taking advantage of the correlation between the successive samples of  $\phi(n)$  to minimize the prediction error  $\mathcal{E}[\gamma^2(n)]$ . From linear prediction theory [40], we know that for a predictor that minimizes  $\mathcal{E}[\gamma^2(n)]$ ,

$$\mathcal{E}[\gamma(n+m)\phi(n)] = R_{\gamma\phi}(m) = 0 \quad \text{for } m > 0 \quad (3.50)$$

i.e., the prediction error is uncorrelated with the past inputs. Since  $\gamma(n)$  is a linear combination of  $\phi(n), \phi(n-1), \dots$ , it follows that

$$\mathcal{E}[\gamma(n+m)\gamma(n)] = R_{\gamma}(m) = 0 \quad \text{for } m > 0 \quad (3.51)$$

Since the autocorrelation function is symmetric, then

$$R_{\gamma}(m) = 0 \quad m \neq 0$$

or

$$R_{\gamma}(m) = \sigma_{\gamma}^2 \cdot \delta(m) \quad (3.52)$$

i.e., the predictor that minimizes MSE of prediction error yields a white noise sequence.

On the other hand, since  $G(z)$  is causal,  $g(m)$  ( $m > 0$ ) and  $g(0)=1$  together form the impulse response of the predictor. Thus the autocorrelation of the  $\gamma(n)$  sequence is

$$\begin{aligned} R_{\gamma}(m) &= \mathcal{E}\{\gamma(n+m)\gamma(n)\} \\ &= \mathcal{E}\left\{\sum_j \phi(j)g(n+m-j) \cdot \sum_i \phi(i)g(n-i)\right\} \\ &= \mathcal{E}\left\{\sum_j \sum_i g(n+m-j)g(n-i)\phi(j)\phi(i)\right\} \end{aligned}$$

Let  $i=j+k$  and  $t=n+m-j$ , then

$$\begin{aligned}
R_\gamma(m) &= \mathcal{E}\left\{\sum_j \sum_k g(n+m-j)g(n-j-k)\phi(j)\phi(j+k)\right\} \\
&= \sum_t \sum_k g(t)g(t-m-k)R_\phi(k) \\
&= g(m)*g(-m)*R_\phi(m)
\end{aligned} \tag{3.53}$$

Combining (3.52) and (3.53), we have

$$g(m)*g(-m)*R_\phi(m) = \sigma_\gamma^2 \cdot \delta(m) \tag{3.53a}$$

From the definition of the Z transform, it is easy to show that, for a real sequence  $g(m)$ , the Z transform of the sequence  $g(-m)$  is

$$\begin{aligned}
\mathcal{Z}\{g(-m)\} &= \sum_{m=-\infty}^{\infty} g(-m)z^{-m} \\
&= \sum_{n=-\infty}^{\infty} g(n)z^n \\
&= 1+G(z^{-1})
\end{aligned} \tag{3.54}$$

Note that  $(1+G(z))$  is the Z transform of the sequence  $g(m)$  with  $g(0)=1$  and  $g(m)=0$  for  $m < 0$ . We know from the properties of the autocorrelation that the Z transform of the autocorrelation function  $R_\phi(m)$  is the power spectrum of the sequence  $\phi(n)$ , i.e.

$$\begin{aligned}
\mathcal{Z}\{R_\phi(m)\} &= |\Phi(z)|^2 \\
&= \frac{N_o}{|F(z)|^2}
\end{aligned} \tag{3.55}$$

Taking the Z transform of (3.53a) leads to

$$[1+G(z)][1+G(z^{-1})] = \frac{\sigma_\gamma^2(z^F - FE)}{N_o} \cdot |F(z)|^2 \tag{3.56}$$

Given  $|F(z)|^2$ , finding  $G(z)$  in equation (3.56) is known as the spectral factorization problem. Since  $(1+G(z))$  is causal,  $(1+G(z^{-1}))$  is anti-causal. Therefore, Equation (3.56)

says that the predictor that minimizes the noise variance is the causal part of the factorization of channel power spectrum  $|F(z)|^2$ .

In the following, we shall find this predictor for the special case where  $|F(z)|^2$  is rational.

Since  $N_o/|F(z)|^2$  is the Z transform of  $R_\phi(m)$ , and by definition  $R_\phi(m)=R_\phi(-m)$ , then

$$\begin{aligned}
 \frac{N_o}{|F(z)|^2} &= \sum_{m=-\infty}^{\infty} R_\phi(m)z^{-m} \\
 &= \sum_{m=-\infty}^{\infty} R_\phi(-m)z^{-m} \\
 &= \sum_{n=-\infty}^{\infty} R_\phi(n)z^n \\
 &= \frac{N_o}{|F(z^{-1})|^2}
 \end{aligned} \tag{3.57}$$

Equation (3.57) implies that if  $z_o$  is a pole (zero) of  $|F(z)|^2$ , then  $z_o^{-1}$  is also a pole (zero) of it. Let  $a_1, \dots, a_m$  be the  $m$  simple zeros and  $b_1, \dots, b_n$  be the  $n$  simple poles of  $|F(z)|^2$  inside the unit circle. Then  $1/a_1, \dots, 1/a_m$  will be the  $m$  zeros and  $1/b_1, \dots, 1/b_n$  the  $n$  poles of  $|F(z)|^2$  outside the unit circle. Since  $|F(z)|^2$  is rational, it can be expressed as

$$|F(z)|^2 = K \cdot \frac{\prod_{i=1}^m (1-a_i z^{-1})(1-a_i z)}{\prod_{i=1}^n (1-b_i z^{-1})(1-b_i z)} \tag{3.58}$$

where  $K$  is a constant.

We now want to factorize the right hand side of Equation (3.58) into  $(1+G(z))$  and  $(1+G(z^{-1}))$ . Since  $G(z)$  is causal, we want its poles inside the unit circle. We could assign

the zeros arbitrarily as long as one zero is assigned to  $(1+G(z))$  and the reciprocal of it is assigned to  $(1+G(z^{-1}))$ . We choose such that the inverse filter  $1/(1+G(z))$  has its poles inside the unit circle and hence is stable. This results in a predictor which is minimum phase, i.e., both its poles and zeros are inside the unit circle:

$$(1+G(z)) = \frac{\prod_{i=1}^m (1-a_i z^{-1})}{\prod_{i=1}^n (1-b_i z^{-1})} \quad (3.59)$$

with this factorization, from Equation (3.56) the MSE of the predictor error, which is also the MSE of the entire ZF-FE since there is no ISI at the output, is

$$\mathcal{E}^2(ZF - FE, rational) = \sigma_\gamma^2 = \frac{N_o}{K} \quad (3.60)$$

One interesting observation from Equation (3.60) is that even though the channel frequency response like the one of Equation (3.58) may have zeros (nulls), the final MSE of the ZF-FE output can be finite. This is in clear contrast to the ZF-LE which may result in infinite MSE (Equation (3.10)) by trying to invert a channel frequency response with nulls.

From Figure 3-9 and Equation (3.56), we know that the forward filter of the ZF-FE is

$$\begin{aligned} H(z) &= \frac{1}{F(z)} \cdot (1+G(z)) \\ &= \frac{\sigma_\gamma^2}{N_o} \cdot \frac{|F(z)|^2}{1+G(z^{-1})} \cdot \frac{1}{F(z)} \\ &= \frac{\sigma_\gamma^2}{N_o} \cdot \frac{F^*(z)}{1+G(z^{-1})} \end{aligned} \quad (3.61)$$

Therefore, for a rational channel frequency response, the procedure to find the ZF-FE is :

(1) factorize  $|F(z)|^2$ , which yields the feedback portion  $G(z)$  and the factor  $[1+G(z^{-1})]$ ; (2) calculate the forward filter by using Equation (3.61). Note that the forward filter obtained from (3.61) is a noncausal system, which is expected from Figure 3-8a.

For a general case where  $F(z)$  is not rational, the solution can be found by the geometry method introduced in [34]. It is shown in Appendix C that for a general channel frequency response  $F(z)$ , the ZF-FE with infinite length yields the output MSE:

$$\varepsilon^2 (ZF - FE) = N_o \cdot \exp\left\{T \int_{-T/2}^{T/2} \ln[1/|F(e^{j2\pi f})|^2] df\right\} \quad (3.62)$$

It is further shown that

$$\varepsilon^2 (ZF - FE) \leq \varepsilon^2 (ZF - LE) \quad (3.63)$$

Equation (63) quantitatively shows what we have explained earlier, i.e., the ZF-FE always has less noise enhancement than the ZF-LE.

**b. Finite Length.** For a finite length forward equalizer with length of  $(M+1)$ , we can only guarantee the ISI in the equalizer span  $(-M \leq n \leq 0)$  to be zero, i.e.,

$$\begin{aligned} q(n) &= \sum_{j=-M}^0 h(j)f(n-j) = 1, & \text{for } n = 0 \\ &= 0, & \text{for } -M \leq n < 0 \end{aligned} \quad (3.64)$$

That is to say, with respect to the reference position,  $h(n)$  is strictly anticausal, because  $h(n)$  is designed to cancel only the precursor ISI. Equation (3.64) is written in matrix form as

$$\mathbf{F}_u \mathbf{h} = \delta_u \quad (3.65)$$

Where

$$\mathbf{F}_u = \begin{bmatrix} f(0) & f(-1) & \dots & f(-M) \\ f(1) & f(0) & \dots & f(-M+1) \\ \dots & \dots & \dots & \dots \\ f(M) & f(M-1) & \dots & f(0) \end{bmatrix} \quad \mathbf{h} = \begin{bmatrix} h(-M) \\ h(-M+1) \\ \vdots \\ h(0) \end{bmatrix} \quad \delta_u = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \quad (3.66)$$

If  $\mathbf{F}_u$  has full rank, i.e., the inverse of  $\mathbf{F}_u$  exists, then the forward filter of ZF-FE is

$$\mathbf{h}(\text{ZF-FE}) = \mathbf{F}_u^{-1} \cdot \delta_u \quad (3.67)$$

Equations (3.65) to (3.67) are the anticausal versions of equations (3.12) to (3.14). On the other hand, from Equation (3.49), we have for the feedback filter with length  $N$

$$g(n) = q(n) = \sum_{j=-M}^0 h(j) f(n-j) \quad \text{for } 0 < n \leq N \quad (3.68)$$

or in matrix form

$$\mathbf{F}_b \mathbf{h} = \mathbf{g} \quad (3.69)$$

where

$$\mathbf{F}_b = \begin{bmatrix} f(1+M) & f(M) & \dots & f(1) \\ f(2+M) & f(1+M) & \dots & f(2) \\ \dots & \dots & \dots & \dots \\ f(N+M) & f(N+M-1) & \dots & f(N) \end{bmatrix} \quad \mathbf{h} = \begin{bmatrix} h(-M) \\ h(-M+1) \\ \vdots \\ h(0) \end{bmatrix} \quad \mathbf{g} = \begin{bmatrix} g(1) \\ g(2) \\ \vdots \\ g(N) \end{bmatrix} \quad (3.70)$$

Then the feedback filter of the ZF-FE is

$$\begin{aligned} \mathbf{g}(\text{ZF-FE}) &= \mathbf{F}_b \cdot \mathbf{h}(\text{ZF-FE}) \\ &= \mathbf{F}_b \cdot \mathbf{F}_u^{-1} \cdot \delta_u \end{aligned} \quad (3.71)$$



Due to the finite length, both the forward and feedback filters do not cancel the ISI completely. Therefore, there is some ISI at the output of the entire ZF-FE. There is no simple closed form expression for the MSE in this case. Numerical solutions are generally used to evaluate the performance. Several factors may have an effect on the ultimate performance, including the particular channel characteristics being equalized, number of taps for forward and feedback filters and the noise variance, etc.. Some of these will be shown in Chapter V for the case of television multipath cancellation.

### (3) MSE Feedback Equalizer

**a. Infinite Length.** The MSE feedback equalizer (MSE-FE) minimizes the overall MSE value at the output of the entire equalizer. Unlike the ZF-FE, there is no constraint on the forward filter other than that it has to be linear.

The method we use to find the forward filter and feedback filter is similar to the one used in the ZF-FE case. We first model the forward filter as the cascade of a linear equalizer and a linear predictor, as shown in Figure 3-12,

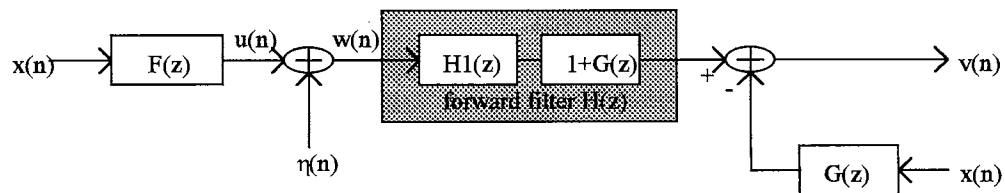


Figure 3-12. MSE feedback equalizer model

where  $H_1(z)$  is a linear equalizer. The error at the equalizer output can be found by using Figure 3-13., which can be further simplified to Figure 3-14.

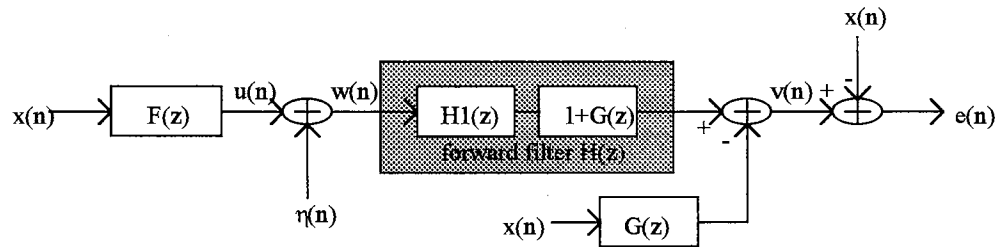


Figure 3-13. Model to calculate the error for MSE-FE

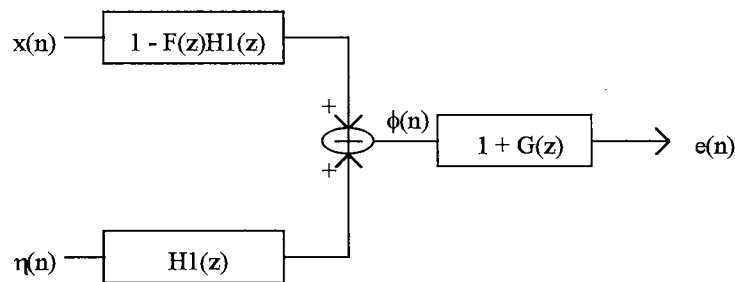


Figure 3-14. Equivalent model of Figure 3-13

To minimize  $\mathcal{E}[e^2(n)]$ , we are going to minimize the power spectral density of  $e(n)$ ,  $|E(z)|^2$ , at each frequency. From Figure 3-14, we can see that  $|E(z)|^2$  can be minimized in two steps: (1) choose  $H_1(z)$  so that  $|\Phi(z)|^2$  is minimized; (2) choose  $G(z)$  so that  $|E(z)|^2$  is minimized given  $\phi(n)$  as input to the predictor. Comparing Figure 3-14 with Figure 3-6, we can see that to find  $H_1(z)$  is actually to find the optimum MSE-LE.

From Equation (3.20), we have

$$H_1(z) = \frac{F^*(z)}{|F(z)|^2 + N_o / \sigma_x^2} \quad (3.72)$$

and the spectral density of  $\phi(n)$  is

$$\begin{aligned}
|\Phi(z)|^2 &= \sigma_x^2(1-FH_1)(1-F^*H_1^*) + N_o \cdot |H_1|^2 \\
&= \sigma_x^2[1-FH_1 - F^*H_1^* + |F|^2|H_1|^2] + N_o|H_1|^2 \\
&= \sigma_x^2[|F|^2 + \frac{N_o}{\sigma_x^2}]|H_1|^2 + \sigma_x^2[1-FH_1 - F^*H_1^*] \\
&= \sigma_x^2 \cdot \frac{|F|^2}{|F|^2 + \frac{N_o}{\sigma_x^2}} + \sigma_x^2 \left[ 1 - \frac{2|F|^2}{|F|^2 + \frac{N_o}{\sigma_x^2}} \right] \\
&= \frac{N_o}{|F|^2 + \frac{N_o}{\sigma_x^2}}
\end{aligned} \tag{3.73}$$

The optimum predictor can be found by spectral factorization in a similar fashion as in the ZF-FE case. Similar to Equation (3.56), we have

$$[1+G(z)][1+G(z^{-1})] = \frac{\sigma_e^2(MSE-FE)}{N_o} \left( |F|^2 + \frac{N_o}{\sigma_x^2} \right) \tag{3.74}$$

As in the case of ZF-FE, the prediction error, i.e., the error of the entire MSE-FE, is white. But, in contrast to the ZF-FE, the predictor error is not Gaussian due to the residual ISI [11,13]. The factorization can be done in the same any as in ZF-FE, except that this time we want to factorize  $(|F|^2 + N_o/\sigma_x^2)$  instead of  $|F|^2$ . Of Course, the resulting forward filter  $H(z)$  is also anticausal.

The MSE for the general channel frequency response can also be found in a similar way (see Appendix C):

$$\mathcal{E}^2(MSE-FE) = N_o \exp \left\{ T \int_{-T/2}^{T/2} \ln [1 / (|F(e^{j2\pi f})|^2 + N_o/\sigma_x^2)] df \right\} \tag{3.75}$$

It is also shown that

$$\mathcal{E}^2(MSE-FE) \leq \mathcal{E}^2(MSE-LE) \tag{3.76}$$

as expected. From Equation (3.62) and Equation (3.75), it can be seen that

$$\mathcal{E}^2(MSE - FE) \leq \mathcal{E}^2(ZF - FE) \quad (3.77)$$

i.e., the MSE-FE with **infinite length** always yields less or equal noise enhancement than the ZF-FE. Note that this conclusion is true only for the equalizers with infinite length. For the equalizers with finite length, there is no general conclusion, and we have to resort to numerical solutions. As will be seen in Chapter V, for some particular channel characteristics and equalizers with finite length, the ZF-FE may outperform the MSE-FE in obtaining smaller overall MSE.

An alternative way to derive the MSE feedback equalizer is to use the Wiener-Hopf theory, as shown in Appendix B [41].

**b. Finite Length.** From Figure 3-8, the output of the entire equalizer is

$$\begin{aligned} v(n) &= \sum_{j=-M}^0 h(j)w(n-j) - \sum_{j=1}^N g(j)x(n-j) \\ &= \mathbf{h}^T \cdot \mathbf{w}_M - \mathbf{g}^T \cdot \mathbf{x}_N \end{aligned} \quad (3.78)$$

where

$$\begin{aligned} \mathbf{w}_M &= [w(n+M), \dots, w(n)]^T \\ &= \mathbf{F}_u^T \mathbf{x}_M + \boldsymbol{\eta} \end{aligned} \quad (3.79)$$

where

$$\begin{aligned} \mathbf{x}_N &= [x(n-1), \dots, x(n-N)]^T \\ \mathbf{x}_M &= [x(n), \dots, x(n+M)]^T \end{aligned}$$

and  $\mathbf{F}_u$  is defined in (3.66) and  $\boldsymbol{\eta}$  is the vector consisting of noise samples. The MSE is

$$\begin{aligned} \mathcal{E}\{[x(n) - v(n)]^2\} &= \mathcal{E}\{[x(n) - \mathbf{h}^T \cdot \mathbf{w}_M + \mathbf{g}^T \cdot \mathbf{x}_N]^2\} \\ &= \mathcal{E}\{[x(n) - \mathbf{h}^T \cdot \mathbf{w}_M]^2 - 2[x(n) - \mathbf{h}^T \cdot \mathbf{w}_M] \cdot \mathbf{g}^T \cdot \mathbf{x}_N + \mathbf{g}^T \cdot \mathbf{x}_N \cdot \mathbf{x}_N^T \cdot \mathbf{g}\} \end{aligned} \quad (3.80)$$

The MSE is minimized by setting the gradient with respect to  $\mathbf{g}$  to  $\mathbf{0}$ . Considering Equation (3.79) and taking into account the fact that  $\mathbf{x}(n)$  is an uncorrelated sequence and  $\mathbf{x}(n)$  and  $\eta(n)$  are uncorrelated, this leads to [14]

$$\mathbf{g} = \mathbf{F}_p \cdot \mathbf{h} \quad (3.81)$$

Where

$$\mathbf{F}_p = \begin{bmatrix} f(0) & f(-1) & \dots & f(-M) \\ f(1) & f(0) & \dots & f(1-M) \\ \dots & \dots & \dots & \dots \\ f(N-1) & f(N-2) & \dots & f(N-M-1) \end{bmatrix} \quad (3.82)$$

Considering Equation (3.81) and setting the gradient of MSE with respect to  $\mathbf{h}$  to zero leads to

$$\mathbf{h}_{(MSE - FE)} = \mathbf{A}_p^{-1} \boldsymbol{\alpha}_p \quad (3.83)$$

Where

$$\begin{aligned} \mathbf{A}_p &= \mathcal{E}[\mathbf{w}_M \cdot \mathbf{w}_M^T] \\ &= \mathcal{E}[(\mathbf{F}_u^T \mathbf{x}_M + \boldsymbol{\eta})(\mathbf{F}_u^T \mathbf{x}_M + \boldsymbol{\eta})^T] \\ &= \mathcal{E}[\mathbf{F}_u^T \mathbf{x}_M \mathbf{x}_M^T \mathbf{F}_u + \boldsymbol{\eta} \cdot \boldsymbol{\eta}^T + \mathbf{F}_u^T \mathbf{x}_M \boldsymbol{\eta}^T + \boldsymbol{\eta} \mathbf{x}_M^T \mathbf{F}_u] \\ &= \sigma_x^2 \cdot \mathbf{F}_u^T \cdot \mathbf{F}_u + N_o \cdot \mathbf{I} \end{aligned} \quad (3.84)$$

and

$$\begin{aligned}
\alpha_p &= \mathcal{E} [\mathbf{w} \mathbf{x}(n)] \\
&= \mathcal{E} [(\mathbf{F}_u \mathbf{x}_M + \eta) \mathbf{x}(n)] \\
&= \sigma_x^2 \begin{bmatrix} f(0) \\ \vdots \\ f(-M) \end{bmatrix}
\end{aligned} \tag{3.85}$$

From Equations (3.81) and (3.83), the optimum feedback filter that minimizes the MSE is

$$\mathbf{g}_{(MSE - FE)} = \mathbf{F}_p \cdot \mathbf{A}_p^{-1} \cdot \alpha_p \tag{3.86}$$

#### (4) Mixed Norm Feedback Equalizer

**a. Infinite Length.** We use a similar method as the one used for MSE-FE.

Referring to Figure 3-14, we first want to find the linear equalizer  $H_1(z)$  that minimizes the mixed norm of  $\phi(n)$ . This is exactly the same as finding the optimum MN-LE. From Equation (3.38), we have

$$H_1(z) = \frac{F^*(z)}{|F(z)|^2 + \lambda \frac{N_o}{\sigma_x^2}} \tag{3.87}$$

We then want to find the optimum linear predictor that minimizes the MSE of the prediction error. Note that for the prediction error, there is no such thing as a mixed norm.

Following a similar procedure as the one used in ZF-FE, this optimum predictor can be found by factorization:

$$[1 + G(z)][1 + G(z^{-1})] = \frac{\sigma_x^2 (MN - FE)}{N_o} (|F|^2 + \lambda \frac{N_o}{\sigma_x^2}) \tag{3.88}$$

and the MSE can be expressed as (see Appendix D)

$$\varepsilon^2(MSE - FE) = N_o \exp\left\{T \int_{-1/2T}^{1/2T} \ln[1/(|F|^2 + \lambda N_o / \sigma_x^2)] df\right\} \quad (3.89)$$

From Equations (3.62), (3.75) and (3.89), we have

$$\varepsilon^2(MSE - FE) \leq \varepsilon^2(MN - LE) \leq \varepsilon^2(ZF - FE) \quad (3.90)$$

**b. Finite Length.** The finite version of the mixed norm for the FE case is given by

$$\mathcal{E}\{\lambda[x(n) - v(n)]^2 + (1 - \lambda)[q(n) - \delta(n)]^2 x^2(n)\} \quad (3.91)$$

Where

$$q(n) = \sum_{j=-M}^0 h(j) f(n-j) \quad (3.92)$$

Equation (3.91) can be written in matrix form as

$$\begin{aligned} & \lambda \cdot \mathcal{E}[x^2(n) - 2x(n)\mathbf{h}^T \cdot \mathbf{w} + \mathbf{h}^T \mathbf{w} \mathbf{w}^T \mathbf{h}] \\ & + (1 - \lambda)(\mathbf{F}_u \mathbf{h} - \delta_u)^T (\mathbf{F}_u \mathbf{h} - \delta_u) \sigma_x^2 \end{aligned} \quad (3.93)$$

Where

$$\mathbf{w} = \mathbf{F}_u^T \mathbf{x}_M + \boldsymbol{\eta} \quad (3.94)$$

Following the same arguments as in MN-LE and MSE-FE, it can be shown that the

optimum forward filter is

$$\mathbf{h}(MN - FE) = \mathbf{A}_{MN}^{-1} \boldsymbol{\alpha}_{MN} \quad (3.95)$$

Where

$$\begin{aligned} \mathbf{A}_{MN} &= \mathcal{E}[\lambda \mathbf{w} \mathbf{w}^T + (1 - \lambda) \mathbf{F}_u^T \mathbf{F}_u \sigma_x^2] \\ &= \sigma_x^2 \mathbf{F}_u^T \mathbf{F}_u + \lambda N_o \mathbf{I} \end{aligned} \quad (3.96)$$

and

$$\begin{aligned}
\alpha_{MN} &= \mathcal{E} [\mathbf{w} \mathbf{x}(n)] \\
&= \mathcal{E} [(\mathbf{F}_u^T \mathbf{x}_M + \eta) \mathbf{x}(n)] \\
&= \sigma_x^2 \begin{bmatrix} \mathbf{f}(0) \\ \vdots \\ \mathbf{f}(-M) \end{bmatrix}
\end{aligned} \tag{3.97}$$

as in Equation (3.85).

The feedback filter of MN-FE is

$$\mathbf{g}(MN - FE) = \mathbf{F}_p \mathbf{A}_{MN}^{-1} \alpha_{MN} \tag{3.98}$$

where  $\mathbf{F}_p$  is defined in equation (3.82).

Again, there is no simple closed form expression for the MSE of the MN-FE. The performance evaluation relies heavily on a numerical solution. It is expected that the final MSE value depends on the number of taps for both forward and feedback filters, the channel characteristics, the noise variance and the value of the parameter  $\lambda$ . This subject will be explored later in Chapter V.

#### 4. Summary

In this chapter, the communication system model was introduced. It was further simplified and cast into the digital signal processing framework. The concept of intersymbol interference (ISI) was then reviewed and the problem of eliminating it, i.e. equalization, was formulated. Finding the equalizer coefficients was shown to be equivalent to minimizing norms. Several norm-based criteria were introduced. The zero-



forcing (ZF) criterion aims at eliminating the ISI regardless of the noise environment. The equalizer it leads to has a potential of enhancing noise for a channel with spectral nulls. The mean square error (MSE) criterion strikes a balance between reducing ISI and minimizing noise enhancement. The mixed norm (MN) criterion assumes some *a priori* knowledge about the underlying channel impulse response. (In fact, in this chapter, all three (ZF, MSE and MN) equalizers assume a known channel impulse response. But the MN equalizer has a distinct way to use this knowledge, which will become more apparent in Chapter V where no explicit channel impulse response is available). The relative significance of this *a priori* knowledge and the MSE is quantified by a parameter  $\lambda$ .  $\lambda$  will be chosen according to practical applications.

Starting with linear structure, equalizers with infinite and finite number of taps were derived for the three criteria, and the resulting MSE values were compared. The feedback equalizer was then studied. It was noted that the feedback equalizer can compensate for severe amplitude distortion without significantly enhancing noise. Further insight was obtained by the introduction of the linear predictor in the model of the feedback equalizer. The feedback portion of the equalizer can be found by spectral factorization and the forward filter can be calculated as a function of the channel impulse response.

Closed form expressions for the MSE of feedback equalizers with infinite length were derived using different criteria. It was shown quantitatively that the feedback equalizers yield less overall MSE than the linear counterparts.

The following paragraph summarizes the main results covered in this chapter for the linear and feedback equalizers [42]:

**A. Linear equalizers:**

\* Block diagram: Figure 3-4

\* Definition for  $\mathbf{F}$ ,  $\mathbf{h}$ ,  $\delta$ : Equation (3.13)

(i) ZF-LE (zero-forcing linear equalizer):

$$\mathbf{h}_{(ZF-LE)} = \mathbf{F}^{-1} \delta$$

(ii) MSE-LE (mean-square-error linear equalizer):

$$\mathbf{h}_{(MSE-LE)} = (\sigma_x^2 \mathbf{F}^T \mathbf{F} + N_o \cdot \mathbf{I})^{-1} \sigma_x^2 \mathbf{F}^T \delta$$

(iii) MN-LE (mixed-norm linear equalizer):

$$\mathbf{h}_{(MN-LE)} = (\mathbf{F}^T \mathbf{F} + \lambda \cdot \frac{N_o}{\sigma_x^2} \cdot \mathbf{I})^{-1} \mathbf{F}^T \delta$$

**B. Feedback Equalizers:**

\* Block diagram: Figure 3-9

\* Definition for  $\mathbf{F}_u$ ,  $\mathbf{F}_b$ ,  $\mathbf{F}_p$ ,  $\alpha_p$ ,  $\alpha_{MN}$ ,  $\mathbf{h}$ ,  $\mathbf{g}$ ,  $\delta_u$ :

$$\text{Equations (3.66), (3.70), (3.82), (3.85), (3.97)}$$

\* Note:  $\mathbf{h}$  is anticausal and  $\mathbf{g}$  is causal

(i) ZF-FE (zero-forcing feedback equalizer):

$$\mathbf{h}_{(ZF-FE)} = \mathbf{F}_u^{-1} \cdot \delta_u$$

$$\mathbf{g}_{(ZF-FE)} = \mathbf{F}_b \cdot \mathbf{h}_{(ZF-FE)}$$

(ii) MSE-FE (mean-square-error feedback equalizer):

$$\mathbf{h}_{(MSE-FE)} = [\sigma_x^2 \cdot \mathbf{F}_u^T \cdot \mathbf{F}_u + N_o \cdot \mathbf{I}]^{-1} \alpha_p$$

$$\mathbf{g} = \mathbf{F}_p \cdot \mathbf{h}$$

(iii) MN-FE (mixed norm feedback equalizer):

$$\mathbf{h}_{(MN-FE)} = [\sigma_x^2 \cdot \mathbf{F}_u^T \cdot \mathbf{F}_u + \lambda \cdot N_o \cdot \mathbf{I}]^{-1} \alpha_p$$

$$\mathbf{g} = \mathbf{F}_p \cdot \mathbf{h}$$

## CHAPTER IV

### ADAPTIVE AND AUTOMATIC EQUALIZATION BY MINIMIZING NORMS

In chapter III, we found the optimum equalizer taps by minimizing norms: the  $L_1$  norm of the combined impulse response for ZF; the  $L_2$  norm of the error signal for MSE; and the mixed norm for MN. We did the same thing for two different equalizer structures: the linear equalizer (LE) and the feedback equalizer (FE). In all of the above cases, we assumed that the channel impulse response (or equivalently the channel frequency response) is available. In most practical situations, this information is unknown. For example, in the television multipath cancellation problem, the channel characteristics depend on such factors as the locations of the TV receiver and the broadcasting station (or its transmitting antenna); the geographical environment in the area (mountains, building, etc.), etc.. All these factors are not available when the equalizer is designed. Therefore, the equalizer has to be able to identify the channel characteristics by itself.

There are two distinct techniques in identifying the channel characteristics. One is generally called the “automatic” equalizer which uses a known signal as a reference. For example, in most communication systems, the equalizer is trained with a training signal (which is known to the receiver) before the application signal is transmitted. The reference signal can also be interleaved with the application signal. For example, in the TV multipath cancellation problem, the ghost cancellation reference (GCR) signal is transmitted during the vertical blanking period in every field. The second technique is

called the “adaptive” equalizer which adjusts the equalizer taps based on some statistics of the signal without an explicit reference signal. The “blind equalizer” belongs to this category, which strives to restore the transmitted signal only by observation of the received signal over a nonminimum phase channel [43, 44, 45, 46 ]. Typical criteria used in the blind equalizer include entropy, higher order cumulants or bispectrum.

In this chapter, we are going to focus on the automatic equalizer, i.e., a known reference signal is available. The chapter starts with the automatic linear equalizer obtained by minimizing norms. The result is then extended to the feedback equalizer. General comments on the performance of these equalizers are given as conclusions of the chapter.

## **1. Automatic Linear Equalizer**

### **(1) Minimizing the $L_1$ Norm and the ZF Algorithm**

**a. Minimizing the  $L_1$  Norm.** Effort to solve the  $L_1$  problem began with Edgeworth's work in 1880's. Since then a lot of algorithms have been developed, among which there are three major categories, namely, linear programming based algorithms, the iteratively reweighted least squares (IRLS) method, and the residual steepest descent (RSD) method.

There are a number of different formulations of the  $L_1$  problem in linear programming, among which the Bartel-Roberts algorithm [47], the Bartel-Conn-Sinclair algorithm [48], and the Bloomfield-Steiger algorithm [49] are three most representative

ones. All these algorithms are of the extreme fits type [50]. Later Ruzinsky and Olsen [51] proposed an  $L_1$  optimization algorithm using a variant of the Karmakar's linear programming technique, in which the worst case computational complexity is the polynomial of the number of unknowns. By contrast, the worst case computational complexity of the extreme fits type is the exponential of the number of unknowns.

The IRLS algorithm solves the  $L_1$  problem by iteratively computing a weighted least squares solution with weights optimized in the  $L_1$  sense. Schlossmaker [52] studied the IRLS and compared it with the linear programming-based methods then available and showed that the IRLS is much more efficient than the linear programming-based counterparts. But later Fair and Peck indicated that the IRLS can be numerically unstable and converge to the wrong value [53].

The IRLS was re-studied by Yarlagadda et al. [37]. They used the fast Fourier transform to implement the IRLS algorithm and illustrated that large overdetermined systems of equations have been solved successfully and the solutions converged in every case.

The third category of algorithms for the  $L_1$  problem is the residual steepest descent (RSD) method [54,37]. Consider the set of linear equations of (2.1), the RSD algorithm finds the solutions to the equations using the following iteration:

$$\mathbf{c}(k+1) = \mathbf{c}(k) - \Delta_k \cdot (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \Psi(k) \quad (4.1)$$

where

$$\Psi^T = [\psi_1(k), \psi_2(k), \dots, \psi_M(k)] \quad (4.2)$$

and

$$\psi_i(k) = \text{sgn}\{\mathbf{X}\mathbf{c}(k) - \mathbf{y}\}_i \quad (4.3)$$

The simplest case is to let  $\Delta_k$  be some constant. But this does not guarantee fast convergence or may result in convergence to the wrong value. The optimum  $\Delta_k$  can be chosen by minimizing the  $L_1$  norm [37]

$$E(k+1) = \left| \mathbf{r}(k) - \Delta_k \cdot \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \Psi(k) \right| \quad (4.4)$$

where

$$\mathbf{r}(k) = -\mathbf{y} + \mathbf{X} \mathbf{c}(k) \quad (4.5)$$

is the residual vector. The minimization process can be carried out by the IRLS algorithm mentioned earlier. Since this is a one variable ( $\Delta_k$ ) optimization, the computation involved is much less than the IRLS algorithm for solving the  $\mathbf{c}$  vector itself.

**b. Minimizing Peak Distortion and Lucky's ZF Algorithm.** The  $L_1$  solution that minimizes the peak distortion defined in (3.6) may not be unique. If we defined the truncated peak distortion

$$D_M = \sum_{n=-M, n \neq 0}^M |q(n)| \quad (4.6)$$

where  $q(n)$  is the combined channel impulse response, then finding the ZF-LE equalizer with  $(2M+1)$  taps that minimizes  $D_M$  is unique, as long as the matrix  $\mathbf{F}$  defined in (3.13) is of full-rank. It has been shown in [39] that if  $h(0)$  is used to satisfy the constraint  $q(0)=1$ , then  $D_M$  is a convex function of the equalizer taps  $h(n)$  ( $-M \leq n \leq M, n \neq 0$ ). Therefore, steepest descent technique can be used to estimate the set of  $h(n)$  that minimizes  $D_M$ .

Before we derive the algorithm, we define a measure called **initial distortion**:

$$D_o = \sum_{n=-\infty, n \neq 0}^{\infty} |f(n)| \quad (4.7)$$

where  $f(n)$  is the channel impulse response.  $D_o$  indicates how severe the ISI is before equalization. In binary baseband transmission,  $D_o < 1$  means the “eye” of the eye pattern is open prior to equalization.

In the following derivation, the transmitted signal  $x(n)$  is assumed to be a periodic pulse sequence with period much longer than the pulse duration. Therefore, the input signal to the equalizer closely approximates the impulse response of the channel  $f(n)$ , and the signal at the equalizer output closely approximates the impulse response of the combined channel  $q(n)$ .

The steepest descent algorithm updates the coefficient vector of the equalizer in the opposite direction of the gradient of  $D_M$ :

$$\nabla D_M = \sum_{j=-M, j \neq 0}^M \frac{\partial D_M}{\partial h(j)} \cdot \mathbf{a}_j \quad (4.8)$$

Where  $\mathbf{a}_j$  is a unit vector in the direction of the  $h(j)$  coordinate, and

$$\begin{aligned} \frac{\partial D_M}{\partial h(j)} &= \sum_{n=-M, n \neq 0}^M \frac{\partial q(n)}{\partial h(j)} \cdot \text{sgn}(q(n)) \\ &= \sum_{n=-M, n \neq 0}^M f(n-j) \cdot \text{sgn}(q(n)) \end{aligned} \quad (4.9)$$

Assume that  $f(n)$  (for  $n \neq 0$ ) are small compared to  $f(0)$ , then

$$\frac{\partial D_M}{\partial h(j)} \approx f(0) \cdot \text{sgn}(q(j)) \quad (4.10)$$

The equalizer taps can be updated as follows:

$$h(j, \text{new}) = h(j, \text{old}) - \Delta \cdot \text{sgn}(q(j, \text{old})) \quad (4.11)$$



Where  $\Delta$  is the step size. Equation (4.11) says that the equalizer that minimizes the truncated peak distortion,  $D_M$ , can be updated, based on the sign of the combined channel impulse response. When the transmitted signal,  $x(n)$ , is a sequence of impulses as assumed here, the combined channel impulse response can be approximated by the equalizer output.

The convergence of this algorithm is guaranteed only if the initial distortion  $D_0$  is less than one, and the step size  $\Delta$  is sufficiently small. This is shown as follows [39]:

$$\begin{aligned}
q(n, new) &= \sum_{j=-M, j \neq 0}^M h(j, new) f(n-j) + f(n) \\
&= \sum_{j=-M, j \neq 0}^M [h(j, old) - \Delta \cdot \text{sgn}(q(j, old))] f(n-j) + f(n) \\
&= q(n, old) - \Delta \sum_{j=-M, j \neq 0}^M \text{sgn}[q(j, old)] f(n-j) \\
&= q(n, old) - \Delta \cdot \text{sgn}[q(n, old)] f(0) \\
&\quad - \Delta \sum_{j=-M, j \neq 0, n}^M \text{sgn}[q(j, old)] f(n-j)
\end{aligned} \tag{4.12}$$

Therefore,

$$\begin{aligned}
|q(n, new)| &\leq \left| [ |q(n, old)| - \Delta \cdot f(n) ] \text{sgn}[q(n, old)] \right| \\
&\quad + \Delta \sum_{j=-M, j \neq 0, n}^M |f(n-j)|
\end{aligned} \tag{4.13}$$

But

$$\sum_{j=-M, j \neq 0, n}^M |f(n-j)| \leq \sum_{j=-\infty, j \neq 0}^{\infty} |f(n-j)| = f(0) \cdot D_0 < f(0) \tag{4.14}$$

where we have assumed that  $D_0 < 1$  and  $f(0) > 0$ .

So

$$|q(n, new)| < \left| |q(n, old)| - \Delta \cdot f(0) \right| + \Delta \cdot f(0) \tag{4.15}$$

If  $|q(n, old)| > \Delta \cdot f(0)$ , then

$$|q(n, new)| < |q(n, old)| \quad (4.16)$$

i.e., the ISI is consistently decreasing.

If  $|q(n, old)| < \Delta \cdot f(0)$ , then

$$|q(n, new)| < 2\Delta \cdot f(0) \quad (4.17)$$

i.e., the ISI will eventually be within the bound of  $2\Delta f(0)$ . As  $\Delta$  approaches zero, the truncated peak distortion  $D_M$  will approach zero.

There are two problems with this algorithm. First,  $D_o < 1$  may not always be guaranteed in practical applications ( although this is a sufficient but not necessary condition for the algorithm to work). Secondly, even if  $D_o < 1$ , this algorithm only guarantees to minimize the ISI within the span of the equalizer. It may create some new ISI outside the span of the equalizer.

**c. Stochastic ZF Algorithm.** In Lucky's ZF algorithm, the transmitted signal is assumed to be a sequence of impulses. In most practical applications, this is generally not the case, because this kind of signal does not perform well in a noisy environment. For the same peak power level (which is generally the measure of power in the transmission power amplifier) and noise environment, this kind of signal will have less signal to noise ratio than such signals as the pseudo random sequence (e.g. M-sequence) and the GCR signal. Therefore, signals with higher average power are used instead. These signals are usually zero-mean and uncorrelated, i.e.,

$$R_x(m) = \mathcal{E}[x(n)x(n-m)] = \sigma_x^2 \cdot \delta(m) \quad (4.18)$$

For this kind of signal , Lucky's algorithm is no longer valid, since the equalizer output can no longer approximate the combined channel impulse response. It has to be estimated.

From Figure 3-3, the error at the output of the equalizer is

$$\begin{aligned} e(n) &= y(n) - x(n) \\ &= \sum_j x(j)q(n-j) + \sum_j \eta(j)h(n-j) - x(n) \end{aligned} \quad (4.19)$$

Note that the transmitted signal,  $x(n)$ , is also the desired output of the equalizer.

The cross correlation between  $e(n)$  and  $x(n)$  is

$$\begin{aligned} R_{ex}(m) &= \mathcal{E}\{e(n)x(n-m)\} \\ &= \mathcal{E}\{[\sum_k x(k)q(n-k) - x(n)] \cdot x(n-m)\} \\ &= \sum_k q(n-k) \cdot R_x(n-m-k) - R_x(m) \\ &= \sum_k q(n-k) \cdot \sigma_x^2 \cdot \delta(n-m-k) + \sigma_x^2 \cdot \delta(m) \\ &= \begin{cases} \sigma_x^2 \cdot (q(0) - 1), & \text{if } m = 0 \\ \sigma_x^2 \cdot q(m), & \text{if } m \neq 0 \end{cases} \end{aligned} \quad (4.20)$$

Therefore,  $R_{ex}(m)$  can be used to estimate the combined channel impulse response, because  $R_{ex}(m)=0$  for all  $m$  would imply that  $q(m)=\delta(m)$ , as desired for zero-forcing. In practice, the expectation in Equation (4.20) is approximated by the instantaneous value:

$$R_{ex}(m) \approx e(n)x(n-m) \quad (4.21)$$

Using this estimation, Equation (4.11) can be rewritten as

$$h(j, new) = h(j, old) - \Delta \cdot \text{sgn}[e(n)x(n-j)] \quad (4.22)$$

Equation (4.22) is called the **stochastic version of Lucky's ZF algorithm**.

If  $\sum_{n=-M, n \neq 0}^M q^2(n)$ , instead of  $D_M$  defined in (4.6), is to be minimized, the equalizer can

be updated by

$$h(j, new) = h(j, old) - \Delta \cdot e(n)x(n-j) \quad (4.23)$$

Equation (4.23) is also commonly referred to as a **stochastic ZF algorithm** in the literature.

## (2) Minimizing the $L_2$ Norm and the LMS Algorithm

In Chapter III, section 2.3.b, we have shown (see Equation (3.25) ) that the MSE at the linear equalizer output is

$$\begin{aligned} & \mathcal{E}\{[x(n) - y(n)]^2\} \\ &= \mathcal{E}\{x^2(n) - 2x(n)\mathbf{h}^T \mathbf{w}(n) + \mathbf{h}^T \mathbf{w}(n)\mathbf{w}^T(n)\mathbf{h}\} \end{aligned} \quad (4.24)$$

Where  $\mathbf{w}(n)=[w(n+M) \dots w(n) \dots w(n-M)]^T$ . This is a convex function of  $\mathbf{h}$ . Therefore, steepest descent algorithm can be used to find the coefficient vector  $\mathbf{h}$ :

$$\mathbf{h}(n+1) = \mathbf{h}(n) - \Delta \cdot \frac{\partial MSE}{\partial \mathbf{h}(n)} \quad (4.25)$$

where the gradient of MSE with respect to  $\mathbf{h}$  can be found from Equation (4.24).

$$\begin{aligned} \frac{\partial MSE}{\partial \mathbf{h}(n)} &= 2\mathcal{E}[\mathbf{w}(n)\mathbf{w}^T(n)]\mathbf{h} - 2\mathcal{E}[\mathbf{w}(n)x(n)] \\ &= 2\mathbf{A}\mathbf{h} - 2\alpha \end{aligned} \quad (4.26)$$

where  $\mathbf{A}$  and  $\alpha$  are defined in Equations (3.27) and (3.28), respectively.

From (4.25) and (4.26), we have

$$\mathbf{h}(n+1) = \mathbf{h}(n) - 2\Delta[\mathbf{A}\mathbf{h}(n) - \alpha] \quad (4.27)$$

It has been shown [11] that as long as

$$0 < \Delta < 1 / \zeta_{\max} \quad (4.28)$$

where  $\zeta_{\max}$  is the maximum eigenvalue of the matrix  $\mathbf{A}$ , the coefficient calculated by (4.27) will always converge to  $\mathbf{h}(\text{MSE-LE})$  defined in (3.26).

In practice, the ensemble averages  $\mathbf{A}$  and  $\alpha$  are approximated by the unbiased estimates, which are the instantaneous values:

$$\begin{aligned} \mathbf{A} &= \mathcal{E}[\mathbf{w}(n)\mathbf{w}^T(n)] \\ &\approx \mathbf{w}(n)\mathbf{w}^T(n) \\ \alpha &= \mathcal{E}[\mathbf{w}(n)x(n)] \\ &\approx \mathbf{w}(n)x(n) \end{aligned} \quad (4.29)$$

Then

$$\begin{aligned} \mathbf{A}\mathbf{h}(n) - \alpha &\approx \mathbf{w}(n)\mathbf{w}^T(n)\mathbf{h}(n) - \mathbf{w}(n)x(n) \\ &= \mathbf{w}(n)[y(n) - x(n)] \\ &= \mathbf{w}(n) \cdot e(n) \end{aligned} \quad (4.30)$$

From Equations (4.27) and (4.30), we have

$$\mathbf{h}(n+1) = \mathbf{h}(n) - 2\Delta e(n)\mathbf{w}(n) \quad (4.31)$$

Equation (4.31) is the celebrated least-mean squares (LMS) algorithm.

Comparing with the stochastic ZF algorithm of (4.32), rewritten in vector form,

$$\mathbf{h}(n+1) = \mathbf{h}(n) - 2\Delta \cdot e(n)\mathbf{x}(n) \quad (4.32)$$

where  $\mathbf{x}(n) = [x(n+M) \dots x(n) \dots x(n-M)]^T$  is the reference signal vector (which is also the transmitted signal vector), we can see that the LMS algorithm updates the equalizer

coefficients vector in the direction of the received signal vector, whereas the ZF algorithm updates it in the direction of the reference signal vector. Figure 4-1 shows this difference.

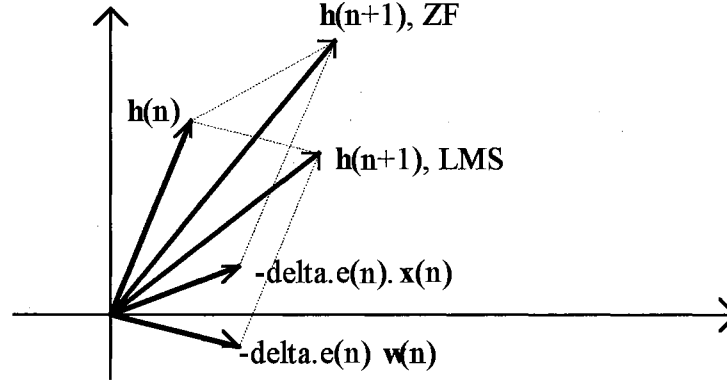


Figure 4-1. Two-dimensional vector updates in ZF and LMS

### (3) Minimizing the Mixed Norm

The mixed norm defined in (3.41) is rewritten as

$$\begin{aligned}
 & \mathcal{E}\{\lambda[x(n) - y(n)]^2 + (1 + \lambda)[q(n) - \delta(n)]^2 x^2(n)\} \\
 &= \lambda \mathcal{E}\{x^2(n) - 2x(n)\mathbf{h}^T \mathbf{w}(n) + \mathbf{h}^T \mathbf{w}(n)\mathbf{w}^T(n)\mathbf{h}\} \\
 &+ (1 - \lambda) \mathcal{E}\{[\mathbf{h}^T \mathbf{f} - \delta(n)][\mathbf{h}^T \mathbf{f} - \delta(n)]^T x^2(n)\}
 \end{aligned} \tag{4.33}$$

In most practical applications, we can assume  $f(0) \gg f(n)$  ( $n \neq 0$ ). For example, in the ghost cancellation problem, this means the main signal is much stronger than the ghosts.

Taking the gradient of the above mixed norm with respect to  $\mathbf{h}$  yields

$$\frac{\partial MSE}{\partial \mathbf{h}} = 2\lambda(\mathbf{A}\mathbf{h} - \boldsymbol{\alpha}) + 2(1 - \lambda)f(0)(\mathbf{q} - \delta)\sigma_x^2 \tag{4.34}$$

Where

$$\mathbf{q} = \begin{bmatrix} q(-M) \\ \vdots \\ q(0) \\ \vdots \\ q(M) \end{bmatrix} \quad \delta = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad (4.35)$$

From Equation (4.20), we know that

$$q(n) - \delta(n) = R_{ex}(n) / \sigma_x^2 \quad (4.36)$$

Equations (4.34), (4.35) and (4.36) together lead to

$$\frac{\partial MSE}{\partial \mathbf{h}} = 2\lambda \cdot (\mathbf{A}\mathbf{h} - \boldsymbol{\alpha}) + 2(1-\lambda)f(0) \cdot \mathcal{E}[e(n)\mathbf{x}(n)] \quad (4.37)$$

Using (4.30) and approximating the expectations by the unbiased estimates, we have

$$\frac{\partial MSE}{\partial \mathbf{h}} = 2\lambda \cdot [e(n)\mathbf{w}(n)] + 2(1-\lambda) \cdot f(0) \cdot [e(n)\mathbf{x}(n)] \quad (4.38)$$

If we assume  $f(0) = 1$  (if not, it can be absorbed by  $\lambda$ ), we have the steepest descent algorithm:

$$\mathbf{h}(n+1) = \mathbf{h}(n) - 2\Delta \cdot e(n) \cdot [\lambda\mathbf{w}(n) + (1-\lambda)\mathbf{x}(n)] \quad (4.39)$$

When  $\lambda=0$ , (4.39) reduces to the stochastic ZF algorithm (4.32); when  $\lambda=1$ , it reduces to the LMS algorithm (4.31).

## 2. Automatic Feedback Equalizer

### (1) Updating the Feedback Filter

It has been shown in Chapter III that the feedback equalizer has the potential of compensating for severe amplitude distortion without enhancing the noise power. But

the solution in Chapter III requires the knowledge of the channel impulse response, which is generally not available in practice. Therefore both the forward and feedback filters have to be adjusted based on the received signal and the reference signal.

Before we derive the algorithms to calculate the filter coefficients, we first establish the relationship between the feedback filter and the forward filter, given that the overall MSE at the output of equalizer is to be minimized.

First, from Figure 3-8, we establish the model for the output error, shown in Figure 4-2.

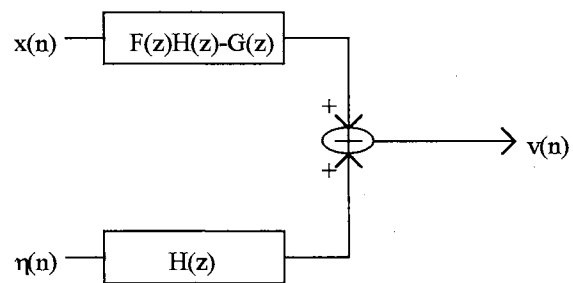


Figure 4-2. Model for the output error

Assuming that the sequence  $x(n)$  is uncorrelated, and that  $x(n)$  and  $r(n)$  are uncorrelated with each other, and for a forward filter with  $(M+1)$  taps and a feedback filter with  $N$  taps, defining

$$q'(n) = \begin{cases} \sum_{i=-M}^0 h(i)f(n-i) - g(n) & 1 \leq n \leq N \\ \sum_{i=-M}^0 h(i)f(n-i) & \text{otherwise} \end{cases} \quad (4.40)$$



then from Figure 4-2, we can evaluate the overall MSE:

$$\begin{aligned}
\mathcal{E}\{e(n)^2\} &= \mathcal{E}\{[v(n) - x(n)]^2\} \\
&= \mathcal{E}\left\{\left[\sum_j q'(j)x(n-j) + \sum_{j=-\infty}^{\infty} h(j)\eta(n-j) - x(n)\right]^2\right\} \\
&= \mathcal{E}\left\{\left[\left(\sum_j q'(j)x(n-j) - x(n)\right) + \sum_{j=-\infty}^{\infty} h(j)\eta(n-j)\right]^2\right\} \\
&= \sigma_x^2 \sum_{j \neq 0} q'^2(j) + \sigma_x^2 [q'(0) - 1]^2 + N_o \sum_j h^2(j)
\end{aligned}$$

From (4.40) and (4.41), the overall MSE can be minimized by choosing the feedback filter to eliminate the first  $N$  postcursor ISI samples, i.e., by letting  $q'(n) = 0$  for  $1 \leq n \leq N$ . This leads to a set of feedback filter coefficients

$$\begin{aligned}
\mathbf{g}(n) &= \sum_{i=-M}^0 h(i)f(n-i) \\
&= \mathbf{q}(n), \quad 1 \leq n \leq N \quad (4.42)
\end{aligned}$$

That is to say, the ISI term, hence the overall MSE, can be reduced by choosing the feedback filter to cancel the postcursor ISI within the span of it.

In practice, as in the linear equalizer cases, the channel impulse response,  $f(n)$ , is not available. Note that the right-hand side of Equation (4.42) consists of the first  $N$  samples of the combined channel impulse response following the reference position, which can be identified in a similar manner as in the linear equalizer case (Equations (4.20), (4.21)), assuming that  $x(n)$  is a signal satisfying equation (4.18).

Based on the above observations, an algorithm to update the feedback filter is proposed:

$$\mathbf{g}(n+1) = \mathbf{g}(n) + \Delta \cdot e(n) \cdot \mathbf{x}'(n) \quad (4.43)$$

where  $e(n)$  is the error at the output of the entire filter, and  $\mathbf{x}'(n)$  is the reference signal vector (which is also the transmitted signal vector):

$$\begin{aligned} e(n) &= v(n) - x(n) \\ \mathbf{x}'(n) &= \begin{bmatrix} x(n-1) \\ x(n-2) \\ \vdots \\ x(n-N) \end{bmatrix} \end{aligned} \quad (4.44)$$

Equation (4.43) is justified by the fact that

$$\begin{aligned} & \mathcal{E}\{e(n)\mathbf{x}'(n)\} \\ &= \mathcal{E}\{[v(n) - x(n)]\mathbf{x}'(n)\} \\ &= \mathcal{E}\{v(n)\mathbf{x}'(n)\} \\ &= \mathcal{E}\left\{ \left[ \sum_i q(i)x(n-i) - \sum_i g(i)x(n-i) \right] \begin{bmatrix} x(n-1) \\ x(n-2) \\ \vdots \\ x(n-N) \end{bmatrix} \right\} \\ &= \sigma_x^2 \left\{ \begin{bmatrix} q(1) \\ q(2) \\ \vdots \\ q(N) \end{bmatrix} - \begin{bmatrix} g(1) \\ g(2) \\ \vdots \\ g(N) \end{bmatrix} \right\} \end{aligned} \quad (4.45)$$

Driving  $\mathcal{E}\{e(n)\mathbf{x}'(n)\}$  to zero effectively approaches a feedback filter that satisfies Equation (4.42).

Note that Equation (4.43) does not assume any constraints on the forward filter, it just shows how the feedback filter can contribute towards the goal of minimizing the overall MSE, with the given forward filter. It is expected that the choice of the forward

filter will make both the overall MSE and the feedback filter different. As in the linear equalizer, there are several ways to update the forward filter coefficients, depending on the chosen criterion.

## (2) Automatic Zero-forcing Feedback Equalizer

The forward filter of a zero-forcing feedback equalizer strives to eliminate the precursor ISI in the span of it, as shown in (3.64). The solution, Equation (3.67), needs explicitly the channel impulse response, which is not available in practice. An alternative method has to be used.

Since our goal is to eliminate the precursor ISI, we can adjust the forward filter to approach this goal without obtaining the channel impulse response explicitly. To do this, we form the error signal from Figure 3-8,

$$\begin{aligned} e_1(n) &= y(n) - x(n) \\ &= \sum_j q(j)x(n-j) + \sum_j h(j)r(n-j) - x(n) \end{aligned} \quad (4.46)$$

Define vector

$$\mathbf{x}_v(n) = \begin{bmatrix} x(n+M) \\ \vdots \\ x(n+1) \\ x(n) \end{bmatrix} \quad (4.47)$$

Then

$$\mathcal{E}[e_1(n) \mathbf{x}_u(n)] = \sigma_x^2 \left\{ \begin{bmatrix} q(-M) \\ \vdots \\ q(-1) \\ q(0) \end{bmatrix} - \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \right\} \quad (4.48)$$

An algorithm driving  $\mathcal{E}[e_1(n) \mathbf{x}_u(n)]$  to zero effectively converges to a set of coefficients  $\mathbf{h}(n)$  ( $-M \leq n < 0$ ) that satisfy equation (3-64), i.e., resulting in the elimination of the ISI in the span of the forward filter. In practice, as before, the expectation is approximated by the instantaneous value. Based on the above observations, an algorithm is proposed as follow:

$$\mathbf{h}(n+1) = \mathbf{h}(n) - \Delta \cdot e_1(n) \mathbf{x}_u(n) \quad (4.49)$$

Equation (4.49) is similar to Equation (4.32) (the vector form stochastic ZF algorithm) for the linear ZF equalizer except that here  $\mathbf{h}$  consists of only the first  $(M+1)$  taps and  $\mathbf{x}_u(n)$  is the upper half of  $\mathbf{x}(n)$ . This is expected, because the forward filter of ZF-FE only cancels the precursor ISI. In the feedback equalizer context, we also call (4.49) the stochastic ZF algorithm. Equation (4.43) is used to update the feedback filter coefficients  $\mathbf{g}$ .

### (3) Automatic MSE Feedback Equalizer

As mentioned earlier, the feedback filter only strives to cancel the post cursor ISI at the output of the forward filter. Therefore, it does not necessarily produce global minimum MSE at its output unless the forward filter adjusts its coefficients according to

the output of the entire filter. Based on these observations, we form the augmented LMS algorithm.

From Figure 3-8, we form the error at the output of the entire filter:

$$\begin{aligned}
 e_2(n) &= v(n) - x(n) \\
 &= \mathbf{h}^T \mathbf{w}(n) - \mathbf{g}^T \mathbf{x}'(n) - x(n) \\
 &= \mathbf{h}_a^T \mathbf{w}_a(n) - x(n)
 \end{aligned} \tag{4.50}$$

where

$$\begin{aligned}
 \mathbf{h}_a &= \begin{bmatrix} \mathbf{h} \\ \mathbf{g} \end{bmatrix} \\
 \mathbf{w}_a(n) &= \begin{bmatrix} \mathbf{w}(n) \\ -\mathbf{x}'(n) \end{bmatrix}
 \end{aligned} \tag{4.51}$$

are the augmented coefficient vector and data vector, respectively. It is evident that the MSE of  $e_2(n)$  is a convex function of  $\mathbf{h}_a$ . Therefore, there is a unique value of  $\mathbf{h}_a$  that minimizes the MSE. Using the LMS algorithm as in the linear equalizer case, we have the following algorithm to update the augmented vector:

$$\mathbf{h}_a(n+1) = \mathbf{h}_a(n) - \Delta \cdot e_2(n) \cdot \mathbf{w}_a(n) \tag{4.52}$$

Equation (4.52) can be decomposed into two separate equations:

$$\mathbf{h}(n+1) = \mathbf{h}(n) - \Delta \cdot e_2(n) \cdot \mathbf{w}(n) \tag{4.53}$$

$$\mathbf{g}(n+1) = \mathbf{g}(n) + \Delta \cdot e_2(n) \cdot \mathbf{x}'(n) \tag{4.54}$$

Equation (4.54) is essentially the same as Equation (4.43), which we derived based on the observation that the feedback filter only cancels the postcursor ISI. Equation (4.53) is

similar to Equation (4.31) with two subtle differences: (1)  $\mathbf{h}$  and  $\mathbf{w}(n)$  in Equation (4.53) consist of only the first  $(M+1)$  elements of  $\mathbf{h}$  and  $\mathbf{w}(n)$  in Equation (4.31), respectively; (2)  $e_2(n)$  in Equation (4.53) is the error at the output of the entire filter (including the feedback part). The latter enables the forward filter to adjust its coefficients based on the change of the feedback filter so that together a global MSE can be approached.

#### (4) Automatic Mixed Norm Feedback Equalizer

The forward filter of the feedback equalizer that minimizes the mixed norms can be updated in a similar fashion as the linear equalizer (Equation (4.39))

$$\mathbf{h}(n+1) = \mathbf{h}(n) - \Delta \cdot [\lambda \cdot e_2(n) \cdot \mathbf{w}(n) + (1 - \lambda) \cdot e_1(n) \cdot \mathbf{x}_u(n)] \quad (4.55)$$

where

$$e_1(n) = y(n) - x(n) \quad (4.56)$$

$$e_2(n) = v(n) - x(n) \quad (4.57)$$

and  $\mathbf{x}_u(n)$  is defined in (4.47) and

$$\mathbf{w}(n) = \begin{bmatrix} \mathbf{w}(n+M) \\ \vdots \\ \mathbf{w}(n+1) \\ \mathbf{w}(n) \end{bmatrix} \quad (4.58)$$

It is evident that Equation (4.55) is the combination of the stochastic zero-forcing algorithm (4.49) and the LMS algorithm (4.53), with  $\lambda$  adjustable for different contributions from the two.

The feedback filter  $\mathbf{g}$  is also updated using (4.43).

The problem of finding the optimum  $\lambda$  remains to be solved. The “transition derivative” defined in [25] is re-written as follows in this context:

$$T_D(\lambda) = \frac{\partial \|\mathbf{q}\|_1}{\partial \lambda} \cdot \frac{\partial \|\mathbf{e}\|_1}{\partial \lambda} \quad (4.59)$$

It was suggested that  $T_D(\lambda)$  is very sensitive to the change of  $\lambda$  and can be used to choose the optimum  $\lambda$ . There are two problems in applying  $T_D(\lambda)$  in the mixed norm peak distortion problem. First, it is the equalizer impulse response  $\mathbf{h}$  instead of the combined impulse response  $\mathbf{q}$  that needs to be calculated. Therefore, it is difficult to calculate the first derivative in (4.59) numerically. Secondly, even if the first derivative in (4.59) could be calculated, the same computation has to be repeated for numerous  $\lambda$ 's so that the optimum  $\lambda$  indicated by  $T_D(\lambda)$  can be chosen. This is not practical in most applications where on-line processing of data is required. We are going to choose  $\lambda$  based on empirical evidence related to its practical implementation. We leave this topic to Chapter V.

Unlike the LMS algorithm, the convergence property of the mixed norm algorithm (4.55) is not well understood at this point. We resort to the numerical evaluation of this problem in Chapter V.

### 3. Summary

In this chapter, we studied the problem of automatic equalization by minimizing norms. This problem arises in practice, because in most applications the channel

characteristics are unknown, and only the received signal and reference signal are available.

We started with minimizing the  $L_1$  norm and reviewing Lucky's classical zero-forcing algorithm with a train of impulses as a reference signal. Then we extended these results to include the more general broadband signals because they have higher average power and hence higher signal to noise ratio for the same peak energy. This resulted in the stochastic zero-forcing algorithm. The minimization of the Mean Squared Error, an  $L_2$  norm, was then examined and the associated LMS algorithm reviewed. Algorithm to minimize the mixed norm was then proposed based on the zero-forcing and the LMS algorithms.

It was recognized in Chapter III that the feedback filter can compensate for severe amplitude distortion without enhancing noise. It was further noted that the feedback filter needs only to cancel the postcursor intersymbol interference to contribute to the global minimization of the MSE at the output of the equalizer. These observations lead to the independent update of the feedback filter. The forward filter can be updated just as a linear equalizer, depending on the underlying norm being minimized, except that only taps before the reference position need to be updated.

It was shown that Lucky's zero-forcing algorithm will converge if the initial peak distortion is less than one [39]. The stochastic ZF algorithm will behave similarly if the transmitted signal is broadband. It is also well known that the LMS algorithm will converge as long as the step size is sufficiently small. Since the mixed norm algorithm is a combination of the stochastic ZF algorithm and the LMS algorithm, it is reasonable to



expect that the mixed norm algorithm will converge if the initial peak distortion is less than one and the step size is sufficiently small. However, at this point, there is no quantitative analysis of the convergence property of the mixed norm algorithm, although the numerical evaluation provides encouraging results. Furthermore, there are no closed form expressions for the MSE of the automatic feedback equalizers at this point. Also, none of the algorithms for feedback equalizers can guarantee the equalizers obtained are stable, as commonly occurs in other adaptive IIR algorithms.

The relationship between the performance of the equalizer and the choice of algorithms, number of taps and channel characteristics to be equalized, is to be analyzed by numerical evaluation with the practical problem of television multipath cancellation in Chapter V.

We summarize the algorithms covered in this chapter in the following [42]:

A. Automatic linear equalizer:

(i) Lucky's ZF algorithm:

$$\mathbf{h}(n+1) = \mathbf{h}(n) - 2\Delta \cdot \text{sgn}[e(n)]\mathbf{x}(n)$$

(ii) Stochastic ZF algorithm:

$$\mathbf{h}(n+1) = \mathbf{h}(n) - 2\Delta \cdot e(n)\mathbf{x}(n)$$

(iii) LMS algorithm

$$\mathbf{h}(n+1) = \mathbf{h}(n) - 2\Delta \cdot e(n)\mathbf{w}(n)$$

(iv) mixed norm steepest descent algorithm (proposed)

$$\mathbf{h}(n+1) = \mathbf{h}(n) - 2\Delta \cdot e(n) \cdot [\lambda\mathbf{w}(n) + (1-\lambda)\mathbf{x}(n)]$$

B. Automatic feedback equalizer:

\*  $e_1, e_2, \mathbf{x}_u, \mathbf{x}'$  are defined in (4.56), (4.57), (4.47), (4.44),

respectively

(i) ZF feedback equalizer (proposed):

$$\mathbf{h}(n+1) = \mathbf{h}(n) - \Delta \cdot e_1(n) \mathbf{x}_u(n)$$

$$\mathbf{g}(n+1) = \mathbf{g}(n) + \Delta \cdot e_2(n) \cdot \mathbf{x}'(n)$$

(ii) MSE feedback equalizer:

$$\mathbf{h}(n+1) = \mathbf{h}(n) - \Delta \cdot e_2(n) \cdot \mathbf{w}(n)$$

$$\mathbf{g}(n+1) = \mathbf{g}(n) + \Delta \cdot e_2(n) \cdot \mathbf{x}'(n) \quad (\text{proposed})$$

(iii) mixed norm feedback equalizer (proposed):

$$\mathbf{h}(n+1) = \mathbf{h}(n) - \Delta \cdot [\lambda \cdot e_2(n) \cdot \mathbf{w}(n) + (1-\lambda) \cdot e_1(n) \cdot \mathbf{x}_u(n)]$$

$$\mathbf{g}(n+1) = \mathbf{g}(n) + \Delta \cdot e_2(n) \cdot \mathbf{x}'(n)$$

## CHAPTER V

### AUTOMATIC TELEVISION MULTIPATH CANCELLATION

#### 1. Introduction

Multipath signal propagation, or “ghosting”, has been a problem in the reception of television signal since the beginning of the television broadcasting industry. In the off-the-air broadcasting, the ghosts are caused by reflections from mountains, buildings, etc.. In cable television, the ghosts are caused by the mismatch of connector impedance.

Effort to solve the ghosting problem began a few years after television became popular, because the ghosting creates quite an annoying effect on the viewer. However, little progress had been made until the introduction of digital signal processing into this area [22]. Since then, extensive research has been done on the mechanism of the ghosting process, the structure and technique of ghost cancellers [55, 56, 57, 58, 59] and the selection of a suitable reference signal to characterize the ghosting process. [60, 61, 62, 63, 64, 65]. Until 1990, most research on the ghosting problem was done in Japan, where because of the dense population and relatively low cable television penetration, solving the ghosting problem is of great interest.

In 1989, Japan established the first ghost canceller reference (GCR) signal [60]. In 1992, the United States adopted the high energy GCR signal as standard [63]. The standard for the PAL system (the European television standard) in Europe is also near

completion [66,67]. These developments clear the way for the commercialization of ghost canceller products. At the same time, a ghost canceller for future high definition television (HDTV) is under active research [68].

Since the modulation-demodulation process in the television system is sufficiently linear, the ghosted signal can be modeled as a transmitted signal passing through a linear system. The ghost canceller is also a linear system. From this point of view, the deghosting problem is similar to the channel equalization problem.

Early research on ghost cancellation focused on the transversal filter, i.e., the linear equalizer. Later, it was realized that for sufficient coverage of the ghosts, it is necessary to use a feedback equalizer. Unfortunately, the difficulty in updating the IIR filter [69] has cast some cloud on the usefulness of the technology. Some modifications have been done to facilitate the update of the IIR filter [56]. In 1991, Winters et al. introduced the concept of time-reversal [55] where spectral factorization was used to calculate the feedback filter coefficients. Later, the author [70] proposed a structure called a virtual filter to further simplify the feedback filter update, and optimized it for the US GCR [71].

In the ghost cancellation context, the **main signal** is defined as the strongest of the received signals, The corresponding time is indicated by the peak of the cross-correlation between the received signals and the local reference signal. The received signals are then shifted so that this peak position coincides to the center tap position in the linear equalizer or the position separating the forward filter and the feedback filter in the feedback equalizer. The ghosts preceding the main signal are called the **precursor ghosts**; the ghosts following the main signal are called the **postcursor ghosts**.

It is generally recognized that the feedback portion of the ghost canceller is responsible for canceling the post-cursor ghosts. We have shown in chapter III and IV that for the norm based algorithms and reducing the overall MSE as the ultimate goal, the algorithm to update the feedback filter remains the same, independent of the way the forward filter is updated. Therefore, in this particular class of structure, it is the algorithm for updating the forward filter that makes the overall performance different.

In this chapter, we are going to evaluate the various algorithms we derived in chapters III and IV in the application of ghost cancellation. We begin by introducing the basic principle of ghost cancellation. We then analyze the requirements for the reference signal to properly characterize the ghosting process. Then we study the performance of the linear equalizer under various ghosting scenarios with zero-forcing, MSE and mixed norm criteria. This is followed by the feedback equalizers. In particular, we are going to study the effect of the choice of the parameter  $\lambda$  on the overall MSE, under various ghosting scenarios.

## **2. The Filter Structures for Ghost Cancellation**

In Chapter III and IV, for the convenience of theoretical analysis, we used the noncausal model for both linear and feedback equalizers, shown here in Figures 5-1a and 5-1b. In practice, all systems have to be causal. The time advance elements are realized by changing the relative position of the reference signal. Figures 5-2a and 5-2b show the actual implementations of Figure 5-1a and 5-1b, respectively [42], where we introduce a new set of variables for the feedback equalizer coefficients:

$$\begin{aligned}
b(0) &= h(-M) \\
b(1) &= h(-M+1) \\
&\dots \\
b(M) &= h(0) \\
&\dots \\
a(M+1) &= g(1) \\
a(M+2) &= g(2) \\
&\dots \\
a(M+N) &= g(N) \tag{5.1}
\end{aligned}$$

and for the linear equalizer coefficients:

$$\begin{aligned}
c(0) &= h(-M) \\
c(1) &= h(-M+1) \\
&\dots \\
c(M-1) &= h(-1) \\
c(M) &= h(0) \\
c(M+1) &= h(1) \\
&\dots \\
c(2M) &= h(M). \tag{5.2}
\end{aligned}$$

For the feedback equalizers (Figures 5-1a and 5-2a), the switch is set to the position “1” during the adaptation, and it is set to the position “2” after training so that the normal video signal can pass through.

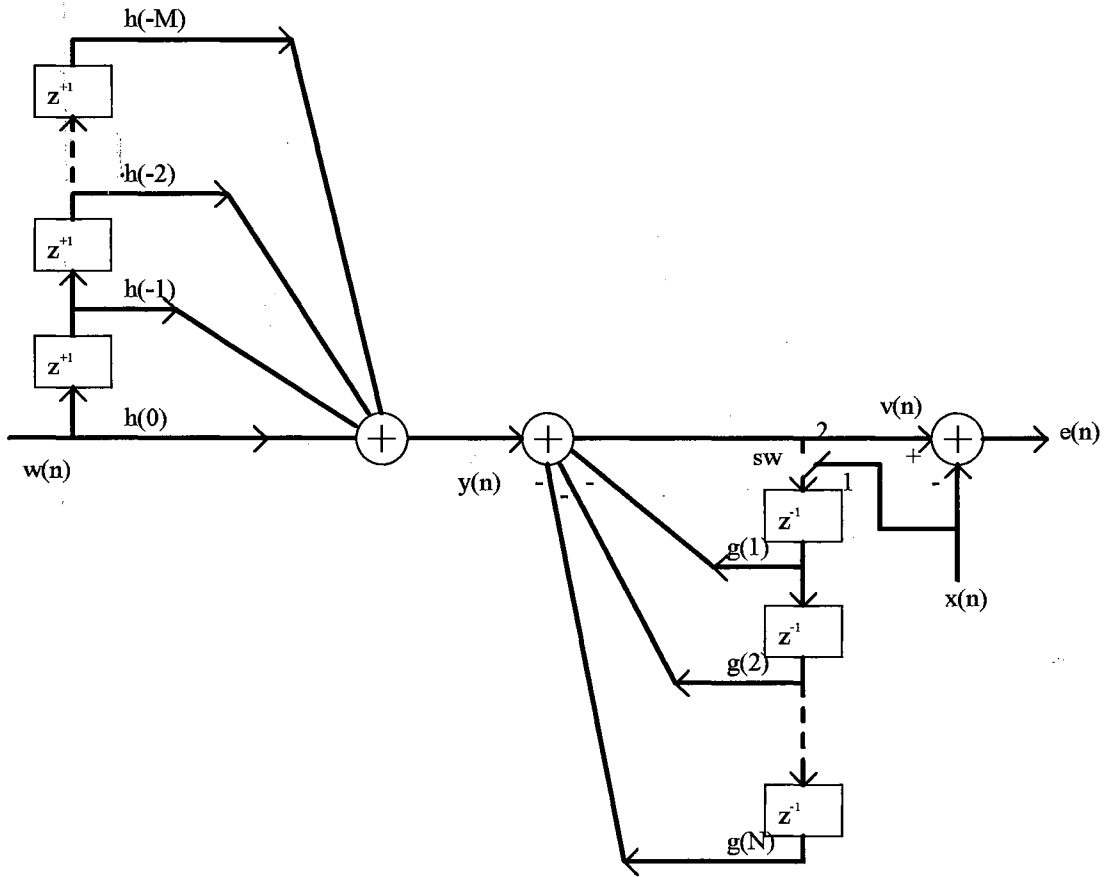


Figure 5-1a. A noncausal feedback equalizer for ghost canceler ( $M=36$ ,  $N=288$ )

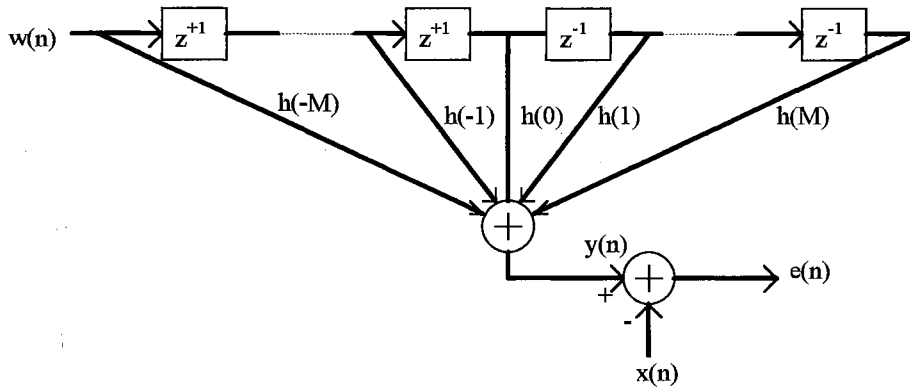


Figure 5-1b. A noncausal linear equalizer for ghost canceller ( $M=36$ )

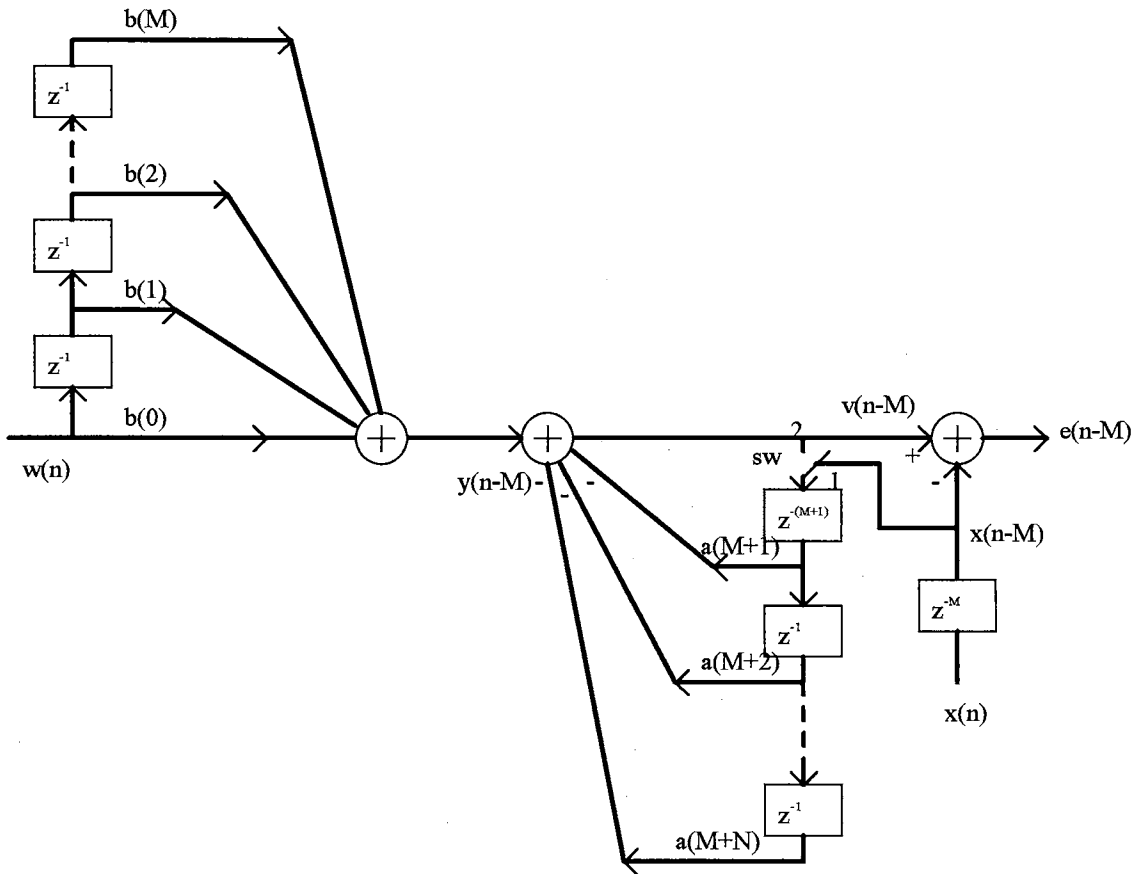


Figure 5-2a. Actual implementation of the feedback equalizer ( $M=36$ ,  $N=288$ )

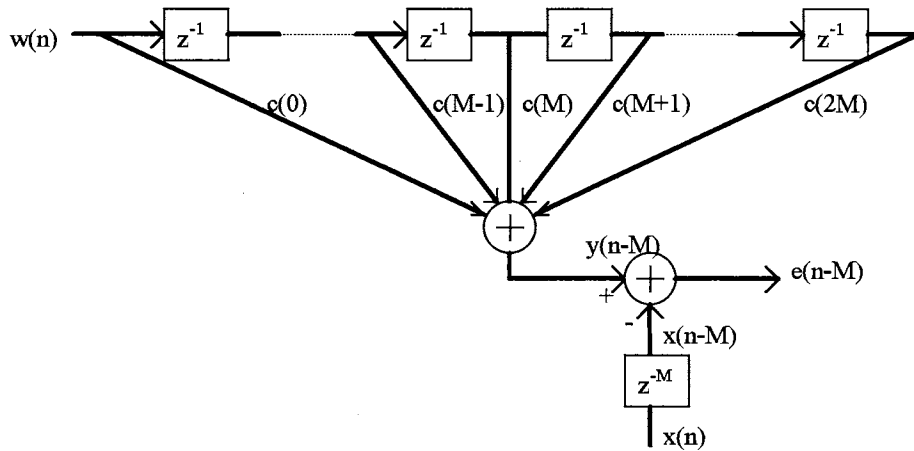


Figure 5-2b. Actual implementation of the linear equalizer ( $M=36$ )



As mentioned earlier, the advantage of the feedback structure is that wider range of ghosts can be covered with the same number of taps. In addition, as we have known from Chapters III and IV, the feedback filter can cancel severe amplitude distortion without enhancing noise. The problem with this structure is that the feedback filter brings in the stability issue. We are going to discuss this problem later in this chapter. On the other hand, the linear structure is always stable. The problem with this structure is that it will have more residuals compared with the feedback equalizer with the same number of taps, because it tries to invert the ghosting system which is typically an FIR filter.

### **3. The Ghost Cancellation Reference (GCR) Signal**

In order to characterize the ghosting process, a ghost cancellation reference (GCR) signal has to be transmitted by the television station. In the receiver, the received GCR signal is compared against the standard GCR signal stored locally. The ghosting process is then characterized, and the filter coefficients are then properly set to cancel the ghosts. In order to effectively characterize the ghosting process, the GCR signal has to have the following properties [61,62]:

(1) High energy: this is needed in order to characterize the ghosting process accurately under noisy conditions.

(2) Flat spectrum: this is important to characterize distortion in the entire frequency spectrum of the composite video bandwidth.

(3) Non-cyclicity: this is needed to detect the ghosts with arbitrary long delays.

(4)  $\sin(x)/x$  shaped autocorrelation: this is the time-domain equivalence of a flat spectrum in the frequency domain.

Besides the above properties, the GCR signal should be incorporated into the composite video signal in such a way that in the receiver the GCR signal can be extracted with minimal interference from the signals in the vicinity.

The GCR signal adopted in Japan in 1989 [72], shown in Figure 5-3, satisfies the (2) and (4) requirements. But it does not satisfy the requirements of (1) and (3). It has to be differentiated into a  $\sin(x)/x$  pulse, shown in Figure 5-4, before it can be used as a reference signal for equalizer updates. This differentiation process will enhance the noise, because it is effectively a highpass filter. In addition, the falling edge of the bar signal, after differentiation, becomes a negative pulse. This limits the maximum delay of the ghosts to  $44.7 \mu\text{s}$  in order for the reference signal to properly characterize the ghosting process.

The GCR adopted in the United States satisfies all of the above requirements. Figure 5-5, Figure 5-6 and Figure 5-6a show the GCR signal, the spectrum of the GCR signal, and the autocorrelation of it, respectively. This signal is inserted into the NTSC (National Television Standard Committee) composite signal in an 8-field sequence according to the color burst phase changes. The 8-field sequence of the GCR is:

$$S_1(+), S_2(-), S_3(+), S_4(-), S_5(-), S_6(+), S_7(-), S_8(+)$$

where “+” stands for the original GCR signal, and “-” means the GCR signal with negative polarity.  $S_i$  ( $i=1,2, \dots, 8$ ) means the GCR signal at the  $i$ -th field. At the receiver, the GCR signal can be extracted by using the following equation:

$$GCR_{average} = \frac{(s_1 - s_2 + s_3 - s_4 - s_5 + s_6 - s_7 + s_8)}{8} \quad (5.3)$$

The above averaging process not only eliminates the interference of the color bursts and horizontal synchronization signals, it also reduces the effect of the additive noise. If the signal at the line before the GCR signal stays constant in all fields, the above averaging process should also eliminate the effect of this signal. Even if this signal changes with time (such as teletext), the averaging process will reduce the effect it has on the composite GCR signal.

Under noisy conditions or when the received signal fluctuates because of the environment such as windy weather, the number of fields of GCR signals to be averaged should be increased, but it has to be a multiple of 8.

The GCR's autocorrelation is a narrow  $\sin(x)/x$  shaped pulse, which corresponds to a flat spectrum in the bandwidth of interest (the composite video bandwidth). This satisfies the requirement imposed by equation (3.3) for practical purposes. In addition, the cross-correlation between the received GCR signal and the local standard GCR signal manifests a peak, which corresponds to the location of the main (desired) signal, because the main (desired) signal is the strongest signal as defined. In addition, the polarity of this peak indicates the polarity of the received GCR signal. The combination of the position and polarity enables the averaging process of (5.3).

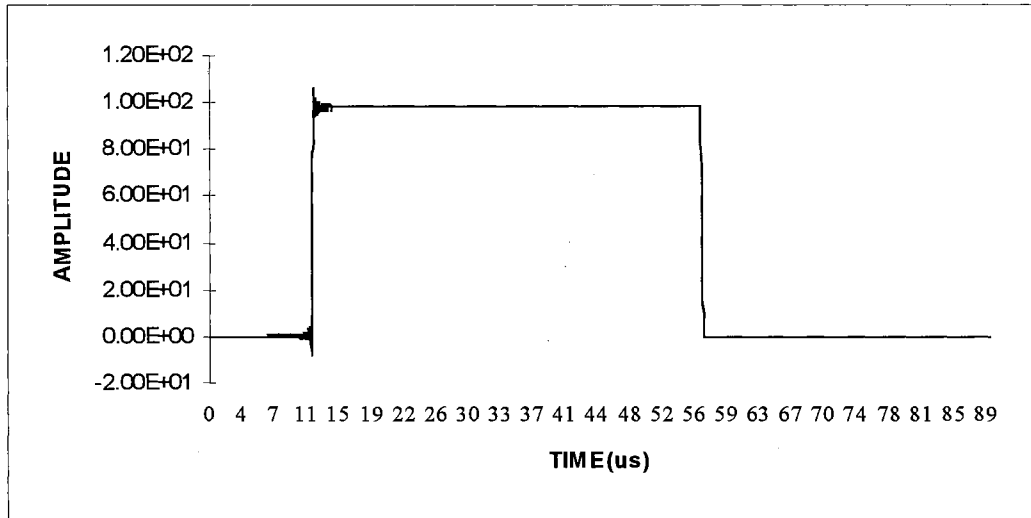


Figure 5-3. Japan standard GCR signal

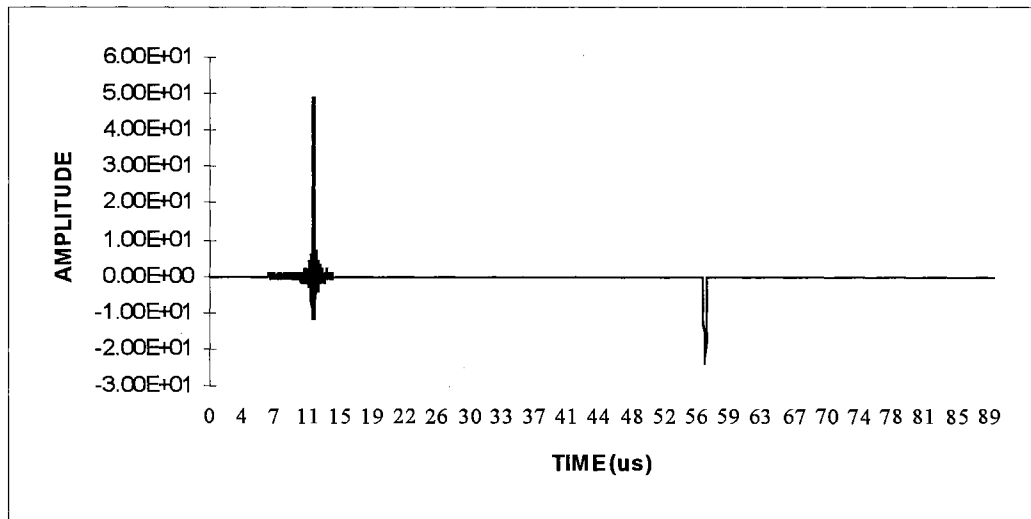


Figure 5-4. Differential Japan standard GCR signal

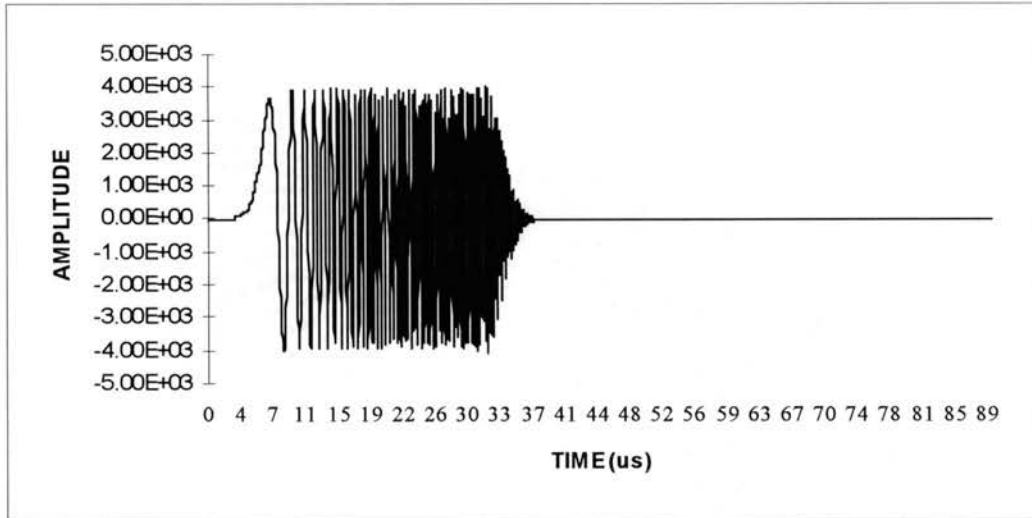


Figure 5-5. US Standard Ghost Cancellation Reference (GCR) signal

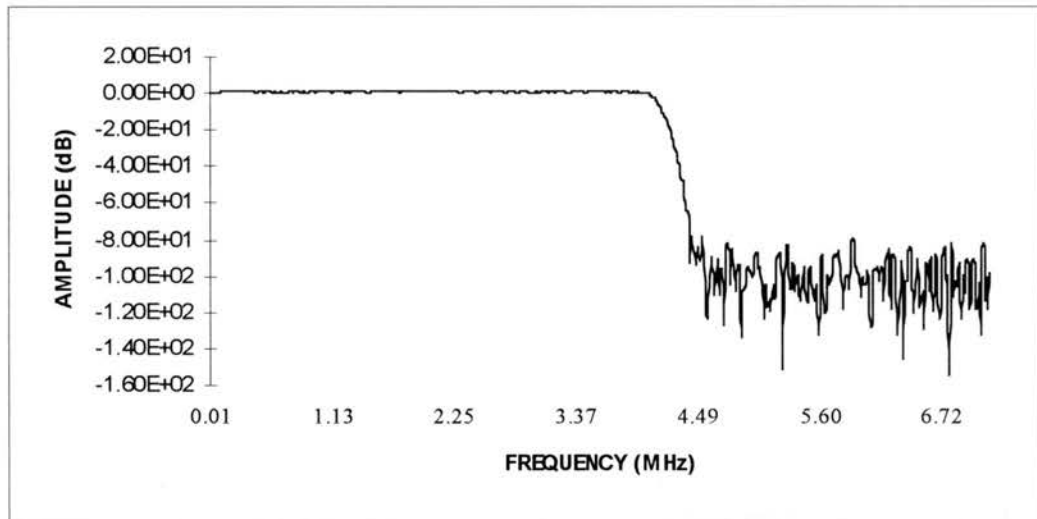


Figure 5-6. Spectrum of the US standard GCR signal

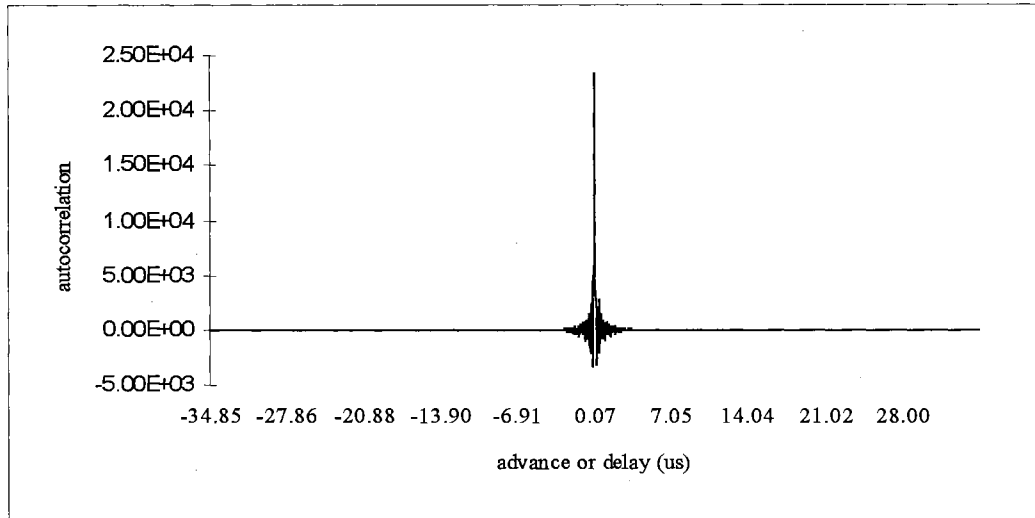


Figure 5-6a. Autocorrelation of the standard US GCR signal

#### 4. The Ghost Cancellation System

A flexible ghost cancellation system has been built for this research, shown in Figure 5-7 (part of this structure belongs to the US patent 5,321,512 [70]), where  $x(n)$  is the stored US standard GCR signal (Figure 5-5),  $w(n)$  is the received GCR signal after averaging (e.g. Figure 5-8),  $v(n)$  is the equalized GCR signal (e.g. Figure 5-10), plotted after convergence of the algorithms. There are the three switches controlling the signal path. During the GCR line and the following line, all the switches are set on the respective position “1” so that the filter can be updated. In normal video lines, all the switches are put on the respective position “2” so that the deghosted picture can be viewed in the monitor.

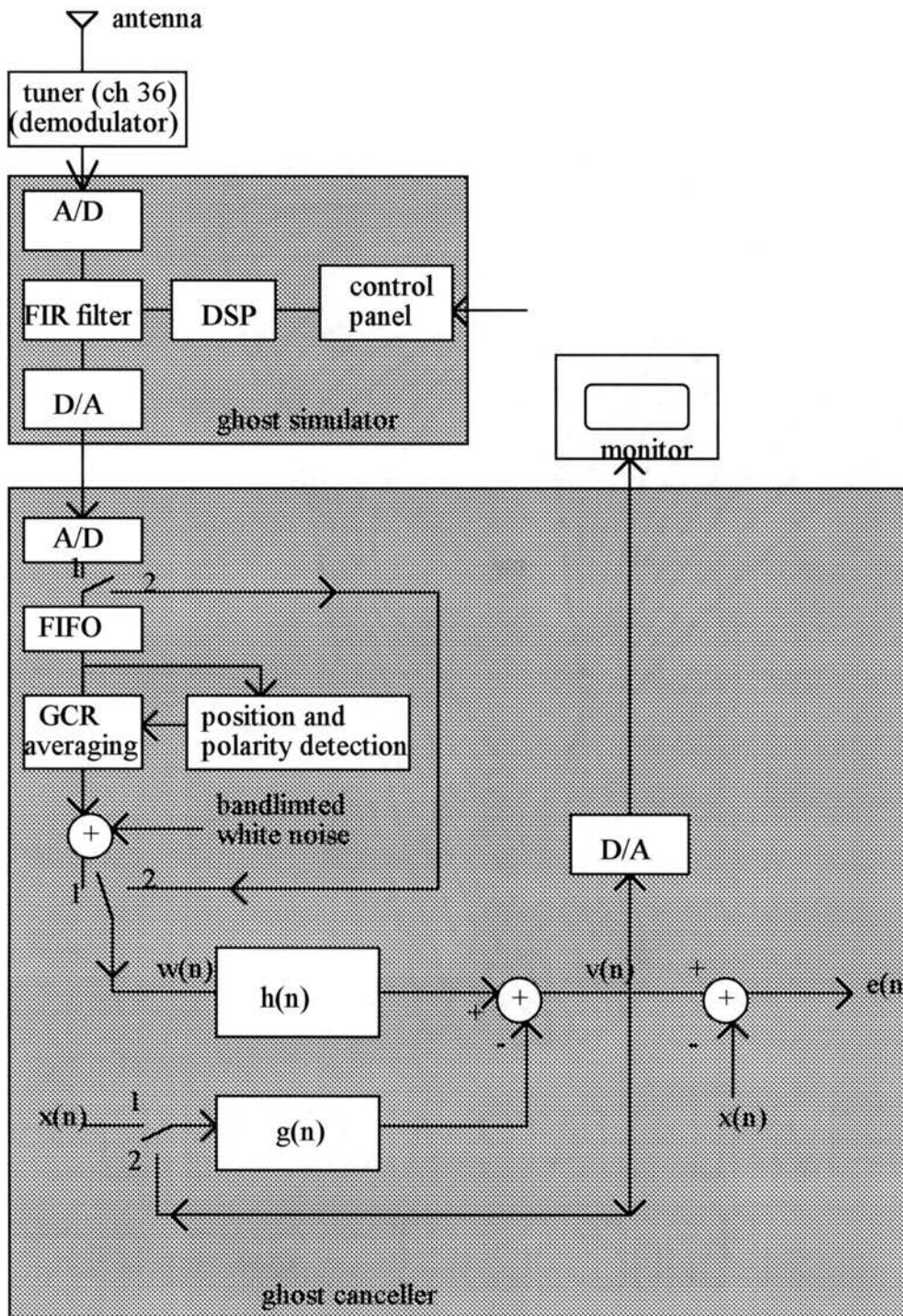


Figure 5-7. The ghost cancellation system

In the evaluation experiments conducted for this paper, the GCR signals are captured from the transmitted off-the-air signal of channel 36 in the San Francisco Bay area. The ghosts are generated through a baseband ghost generator. The noise is simulated white Gaussian noise band limited to 4.2 MHz, which is injected at the input of the ghost canceller filters. The ghost canceller is implemented by commercial digital filters with sampling rate of 14.32 MHz. The linear equalizer has a total of 72 taps with the 36th tap being the reference position. When used as a forward filter of the feedback equalizer, the taps following the 36th tap are disabled. The feedback filter (enabled only for feedback equalization) has 288 taps. All filters have 9-bit precision for data and 10-bit precision for coefficients. All internal calculations are carried out in full precision so there is no loss of precision. Data are rounded to 9 bits only when they are moved out of the filter.

The adaptation algorithms are implemented with a 24-bit fixed-point digital signal processor. The internal coefficients are stored in 24 bits. All multiply-accumulate operations needed for the convolution/correlation calculations are carried out in 56-bit precision (double precision with 8 bit sign-extension to protect from overflowing). All multiply/add operations are carried out in 56-bit precision. The final results are rounded to 24 bits when they are stored in the memory locations.

At every field of the television signal, the GCR line (line 19 in the US) and the line after are sampled at 14.32 MHz and stored in a 9-bit First-In-First-Out (FIFO) buffer. The second line signal is needed so that ghosts with delays up to one line duration (64  $\mu$ s) can be canceled, because the delayed GCR signal with long delay will be shifted to the line following it. For the best quality, the signals at the lines preceding and following the GCR



line should stay at constant waveform (which is true in most channels), so that the averaging operations (Equation (5.10)) with the help of the GCR polarities will eliminate the effect of these signals. Only the GCR signal and its ghosts will remain. If these two lines of signal are changing with time rapidly, the averaging operations will treat them as noise. Since they may not be Gaussian or may not be white at all, these changing signals will have severe adverse impact on the performance of the ghost canceller. In this case, special techniques like the one introduced in [73] can be used. Fortunately, the channel we are investigating does not have this problem. We only deal with white Gaussian noise in this report.

The ghosts are generated in baseband by a separate ghost generator. The off-the-air television signal is demodulated by a tuner. The baseband video signal is then fed to the ghost generator, which can simulate ghosts with various amplitude, delay or advance and polarities. The ghost generator is implemented by a digital FIR filter with 288 taps. The coefficients of this filter are calculated by a digital signal processor (DSP) based on the positions of the switches at the control panel. Note that this DSP is not the same as the one for the ghost canceller (they are on different boards), so the DSP for the ghost canceller is completely “blind” to the setting of the ghosting scenario. In addition, although the input signal to the ghost generator is basically ghost-free, it is not the same as the ideal GCR, because the antenna and the demodulator introduce some roll-off at the high frequency of the baseband signal. Therefore, the channel impulse response,  $f(n)$ , is the convolution of the FIR filter with an equivalent lowpass filter. In other words,  $f(n)$  is not

explicitly known in the experiments here. It is estimated by the cross-correlation between the GCR signal at the ghost canceller input and the stored standard GCR signal.

The software structure of the ghost canceller mainly consists of two parts: the foreground tasks and the background tasks. The foreground tasks are driven by interrupts. They include fetching data from FIFO buffers, all the preprocessing of the data (signal qualification, polarity detection, timing alignment, etc.), and the averaging operations. The background tasks implement the main algorithms, including convolutions, correlation, tap updates, step size controls and convergence analysis, etc..

The equalizer is updated once per 16 fields of television signal, during which a new set of data are captured. The reason to do this is that the signal is sampled at 14.32 MHz and the DSP used is at 27 MIPS (million instructions per second). But we need few thousands instruction cycles to do one update. Therefore, it is not possible to do one update per sample as required by the original algorithms. In addition, in practice, this will allow the ghost canceller to track the slowly changing ghosts (which is irrelevant to the experiments in this paper, where the ghosting settings are fixed for each experiment).

The performance indices (peak distortion, MSE) are recorded for every iteration so that the convergence process can be monitored. After 500 iterations, the algorithms are stopped and the states are stored for comparison.

Typical data collected are: the input and output of equalizer and their spectra; the estimated channel impulse response, the equalizer taps  $h(n)$  and  $g(n)$  at the end of the last iteration; the estimated peak distortion and the estimated mean square error for each iteration. For feedback equalizer, the output of the forward filter is also collected.

As mentioned earlier, the channel impulse response shown here is estimated by the cross-correlation between the received GCR signal (after averaging) and the ideal GCR signal, i.e.,

$$\hat{f}(n) = \sum_j w(j)x(j+n) \quad (5.4)$$

This is justified by the fact that

$$\begin{aligned} \mathcal{E}\{\hat{f}(n)\} &= \mathcal{E}\left\{\sum_j w(j)x(j+n)\right\} \\ &= \mathcal{E}\left\{\sum_j \left[\sum_i x(i)f(j-i)\right]x(j+n)\right\} \\ &= \sum_j \sum_i f(j-i) \cdot \mathcal{E}[x(i)x(j+n)] \\ &= \sum_j \sum_i f(j-i) \cdot \sigma_x^2 \cdot \delta(i-j-n) \\ &= \sigma_x^2 \cdot f(n) \end{aligned} \quad (5.5)$$

The MSE for the linear equalizer is calculated by

$$MSE = \sum_n [y(n) - x(n)]^2 \quad (5.6)$$

summing over the duration of the GCR, where  $y(n)$  is the output of the linear equalizer with the taps at each update. The MSE for the feedback equalizer is calculated by

$$MSE = \sum_n [v(n) - x(n)]^2 \quad (5.7)$$

The peak distortion for the linear equalizer is estimated by the cross-correlation between the error signal and the ideal GCR signal, i.e.,

$$\hat{D}_M = \sum_{n=-M}^M |q(n)| \quad (5.8)$$

where

$$\hat{q}(n) = \sum_j e(j)x(j+n) \quad (5.9)$$

This is justified by the fact that

$$\begin{aligned} \mathcal{E}\{\hat{q}(n)\} &= \mathcal{E}\left\{\sum_j e(j)x(j+n)\right\} \\ &= \mathcal{E}\left\{\sum_j [y(j) - x(j)]x(j+n)\right\} \\ &= \mathcal{E}\left\{\sum_j [\sum_i x(i)q(j-i)]x(j+n)\right\} - \sigma_x^2 \cdot \delta(n) \\ &= \sum_j \sum_i q(j-i) \cdot \sigma_x^2 \cdot \delta(i-j-n) - \sigma_x^2 \cdot \delta(n) \\ &= \sigma_x^2 \cdot q(n) - \sigma_x^2 \cdot \delta(n) \\ &= \sigma_x^2 \cdot [q(0) - 1], \quad n=0 \\ &= \sigma_x^2 \cdot q(n), \quad n \neq 0 \end{aligned} \quad (5.10)$$

Therefore,  $\hat{D}_M$  is a good estimate of  $D_M$ .

The MSE ( $\lambda$ ) is calculated by (5.6) or (5.7) at the end of the last iteration for each  $\lambda$  from 0 to 1.0 with the step size of 0.1. For this plot, the shape of the curve is more important than the absolute values, because all MSE values are scaled before they are plotted (the plotting package can only handle the number in certain range). Note that  $\lambda=0$  corresponds to the ZF equalizer, and  $\lambda=1$  corresponds to the MSE equalizer. Therefore, if this curve has a minimum value somewhere between  $\lambda=0$  and  $\lambda=1$ , we conclude that the MN equalizer outperforms the ZF and MSE equalizers in terms of reaching a smaller MSE value in the steady state.

## 5. Ghost Cancellation using Linear Equalizer

Consider the linear equalizer shown in Figure 5-2b (see also Figure 3-4). The center tap  $h_0$  is chosen as the reference position, i.e., the equalizer taps are evenly partitioned for the precursor and postcursor ghosts. The GCR signal after averaging over 16 fields is used as the received signal  $w(n)$ .

Four algorithms are used for the tap updates: the stochastic zero-forcing (4.23), the stochastic version of Lucky's ZF algorithm (4.22), the LMS (MSE) algorithm (4.31) and the mixed norm algorithm (4.39). "Gear-shifting" of the step size (larger step size at the beginning of the adaptation and smaller step size later on) is used so that fast convergence and small steady-state residual can be achieved. Three ghosting scenarios are tested:

**Case 1:** precursor and postcursor ghosts within the span of the equalizer (without noise).

Figure 5-8 shows the received GCR signal after averaging and Figure 5-8a shows the spectrum of it. Figure 5-9 shows the estimated channel impulse response in this case. Figures 5-10 and 5-11 show the impulse response and frequency response, respectively, of the linear equalizer at the last iteration. Figures 5-12 and 5-13 show the GCR signal and its spectrum, respectively, after being equalized with the stochastic ZF algorithm (equation (4.23)). Figures 5-14 to 5-17 show these for the stochastic version of Lucky's ZF algorithm (equation(4.22)). Figures 5-18 to 5-21 show the same things for the LMS algorithm (equation (4.31)). Figures 5-22 to 5-25 show these for the mixed norm algorithm (equation (4.39)) with  $\lambda = 0.5$  (arbitrary choice). Figure 5-26 shows the peak distortions for each iteration for the four algorithms. Figure 5-27 details the second part of Figure 5-

26. Figure 5-28 shows the MSEs for each iteration. Figure 5-29 details the second part of Figure 5-28. Figure 5-30 shows the MSE versus the mixed norm parameter  $\lambda$ .

It can be seen that the stochastic version of Lucky's ZF algorithm, the LMS (MSE) algorithm and the mixed norm algorithm approach similar MSE values in the steady state. But the stochastic ZF algorithm deviates from the optimum solution after about 50 iterations. The stochastic ZF algorithm converges fastest among the four algorithms in this case. The LMS algorithm is the slowest one in terms of convergence. The stochastic version of Lucky's ZF algorithm behaves abnormally on the way to convergence, and it is oscillating in the steady state. We also see that the stochastic ZF algorithm and the mixed norm algorithm indeed approach the minimum peak distortion.

**Case 2: precursor and postcursor ghosts within the span of the equalizer (SNR=30 dB).**

Figures 5-31 to 5-53 show the numerical results for this case. It can be seen that the MN algorithm with  $\lambda \approx 0.7$  yields the least MSE value in steady state, and the stochastic ZF algorithm has the highest steady state MSE value among the four algorithms. But curiously, Lucky's ZF algorithm seems to do as well as the LMS algorithm. The LMS algorithm, on the other hand, compensates for the channel distortion in a more moderate way by leaving some ISI at the equalizer output, thus having less noise enhancement. The overall MSE at the output of the MSE equalizer, including the residual ISI and the noise, is less than that of the ZF equalizer.

The linear equalizer based on the mixed norm criterion tries to compromise between the ZF equalizer and the MSE equalizer. It does not enhance the noise as much as the ZF equalizer; neither does it leave the residual ISI as much as the MSE equalizer.

It is noted from Figure 5-51 that, in spite of the higher MSE value at the end, the ZF algorithm does converge to its steady state faster than the other two algorithms. Note that all four algorithms use the same step size.

**Case 3:** postcursor ghosts out of the equalizer span (SNR=30 dB).

Figures 5-54 to 5-76 show the results for this case. The ZF equalizer is “blind” to the ghosts in this case, because the cross correlation between the reference signal and the error signal results in a series of spikes in the outside of the equalizer span, which will have no contribution to the taps updates. The end result is that the ZF equalizer will ignore these ghosts by leaving the ISI, and of course, it will not enhance the noise. (Note that the ZF equalizer enhances noise only when it tries to invert the channel with spectral nulls. This will always happen for the linear ZF equalizer with infinite length as long as the spectral nulls exist in the channel frequency response. However, this may not be the case for the linear ZF equalizer with finite length).

On the other hand, the cross correlation between the equalizer input and the error signal produces some strength within the equalizer span in addition to the spikes outside of the equalizer span. Therefore, the MSE equalizer will try its best to reduce the ISI. The end result is that the MSE equalizer will have less MSE value than the ZF equalizer in this case.

The mixed norm equalizer (with  $\lambda=0.5$ ) has the performance in between the ZF equalizer and the MSE equalizer in this case, in terms of the MSE value.

## 6. Ghost Cancellation using Feedback Equalizer

From Chapters III and IV, we know that for a feedback equalizer with an infinite number of taps for both forward and feedback filters, the MSE feedback equalizer does the best in reducing the overall MSE. However, this may not be the case for the feedback equalizer with a finite number of taps, as the one shown in Figure 5-2a.

One particular case, which turns out to be the most common case in practice, is that the postcursor ghosts fall within the span of the feedback filter. We know that the feedback filter can cancel the postcursor ghosts at the output of the forward filter exactly. If the ghosts are out of the span of the forward filter (which is true in this case since the forward filter only covers the precursor ghosts), then it would be better to leave them as they are instead of trying to cancel them with the forward filter. We also know from the previous section that the ZF linear equalizer tends to ignore the ghosts out of its span. Therefore we may anticipate that in this case the ZF feedback equalizer (equations (4.49) with equation (4.43)) may result in a lower overall MSE value than the MSE feedback equalizer (equation (4.53) with equation (4.54)).

Another case is where the precursor ghosts fall within the span of the forward filter. We know from Chapters III and IV that the feedback filter can only cancel the postcursor ghosts. Thus the forward filter is mainly responsible for canceling the precursor ghosts. On the other hand, we also know from the previous section that in the case of ghosts within the span of the linear equalizer, the MSE linear equalizer yields less MSE value than the other two linear equalizers. Therefore we may anticipate that for the case of



precursor ghosts within the span of the forward filter, the MSE feedback equalizer will yield less overall MSE value than the ZF feedback equalizer.

The case of particular interest is the one that combines the previous two cases, namely, the precursor ghosts fall within the span of the forward filter and the postcursor ghosts fall within the span of the feedback filter. Based on the discussions of the previous two paragraphs, we have reason to believe that the mixed norm feedback equalizer (equation (4.55) with equation (4.43)), with  $0 < \lambda < 1$ , may result in the least overall MSE value among the three feedback equalizers.

In what follows, we are going to present the three distinct cases mentioned above. In all cases, the signal to noise ratio (SNR) at the input of the equalizer is 30 dB.

#### **Case 1: precursor ghosts within the span of the forward filter**

Figure 5-77 shows the received GCR after averaging, and Figure 5-77a shows the spectrum of it. Figure 5-78 shows the estimated impulse response of the channel,  $f(n)$ , in this case. Figures 5-79 and 5-80 show the coefficients of the forward filter  $h(n)$  and the feedback filter  $g(n)$ , respectively, at the last iteration. Figures 5-81 and 5-82 show the signal and its spectrum, respectively, at the output of the forward filter which is updated using the stochastic ZF algorithm (Equation (4.49)). Figures 5-83 and 5-84 show the signal and its spectrum, respectively, at the output of the entire feedback equalizer, where the feedback filter  $g(n)$  is updated using (4.43). It can be seen that the feedback part does not make much an impact here, which is expected, because the feedback part cancels only the postcursor ghosts.

Figures 5-85 to 5-90 show the same things for the MSE equalizer (Equations (4.53) and (4.54)) as Figures 5-79 to 5-84 do for the ZF equalizer. Figure 5-91 shows the calculated MSEs at each iteration for both ZF and MSE equalizers. Figure 5-92 details the second part of Figure 5-91. Figure 5-93 shows the MSE at the last iteration for the mixed norm algorithm (Equation (4.55) and equation (4.43)) with various choices of the parameter  $\lambda$ .

It can be seen from Figure 5-91 and 5-93 that the MSE equalizer using the LMS algorithm outperforms the ZF equalizer in this situation, in terms of arriving at smaller amount of the MSE value in steady state. Moreover, Figure 5-93 indicates that the mixed norm equalizer can not do anything better than the MSE equalizer in this case.

**Case 2: postcursor ghost within the span of the feedback filter**

Figures 5-94 to 5-110 show the same things in this case as Figures 5-77 to 5-93 do with respect to Case 1. It can be seen from Figures 5-108 and 5-110 that the ZF equalizer produces smaller a MSE value at the end than the MSE equalizer. Figure 5-110 also shows that the mixed norm equalizer does not do better than the ZF equalizer in this case. It can also be seen that the feedback part does most of the improvement in this case, which is expected.

**Case 3: precursor and postcursor ghosts within the span of the forward filter and the feedback filter.**

Figures 5-111 to 5-127 show the same things in this case as Figures 5-77 to 5-93 do with respect to Case 1. It is interesting to see from Figure 5-127 that the mixed norm feedback equalizer yields the least MSE value at  $\lambda \approx 0.75$ . (Note that  $\lambda=0$  corresponds to

the ZF feedback equalizer;  $\lambda=1$  corresponds to the MSE feedback equalizer). In this particular channel distortion, the mixed norm feedback equalizer outperforms both the ZF and MSE counterparts in terms of arriving at smaller MSE values in the steady state. It can also be seen that both the forward filter and the feedback filter contribute to the cancellation of ghosts in this case, which is expected.

It is evident that the mixed norm equalizer will approach to different MSE values depending on the choice of the parameter  $\lambda$ . However, this relationship is highly nonlinear, as indicated by Figure 5-127, and it is very difficult to establish explicitly. Various experiments suggest that  $\lambda \approx 0.75$  generally yields better result than  $\lambda = 0$  (ZF) and  $\lambda = 1$  (MSE) in this case.

#### **Case 4. precursor and postcursor ghosts within the span of the linear equalizer.**

This case is similar to Case 3. It is conducted to evaluate the performance of the linear equalizer versus the feedback equalizer. The linear equalizer is updated using equation (4.31) (LMS algorithm); and the feedback equalizer is updated using equations (4.53) and (4.54) (also LMS algorithms).

Figures 5-128 and 5-128a show the received GCR signal and its spectrum, respectively. Figure 5-129 shows the estimated channel impulse response. Figures 5-130 and 5-131 show the coefficients and frequency response of the linear equalizer, respectively. Figures 5-132 and 5-133 show the coefficients of the forward and feedback filters, respectively, of the feedback equalizer. Figure 5-134 shows the MSE values of each iteration. It can be seen that the feedback equalizer clearly reaches a smaller MSE value than the linear equalizer in steady state in this case.

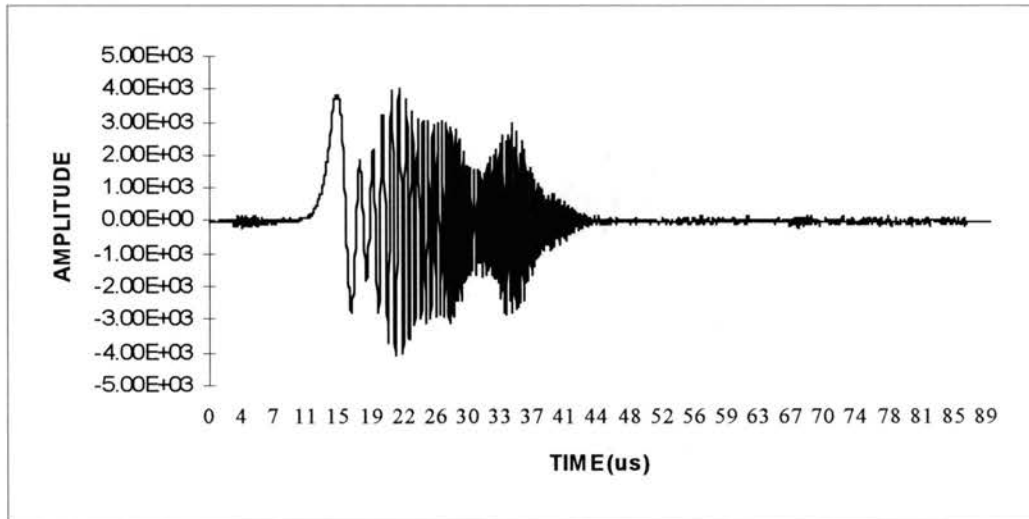


Figure 5-8. Received GCR signal for Case 1 (linear equalizer)

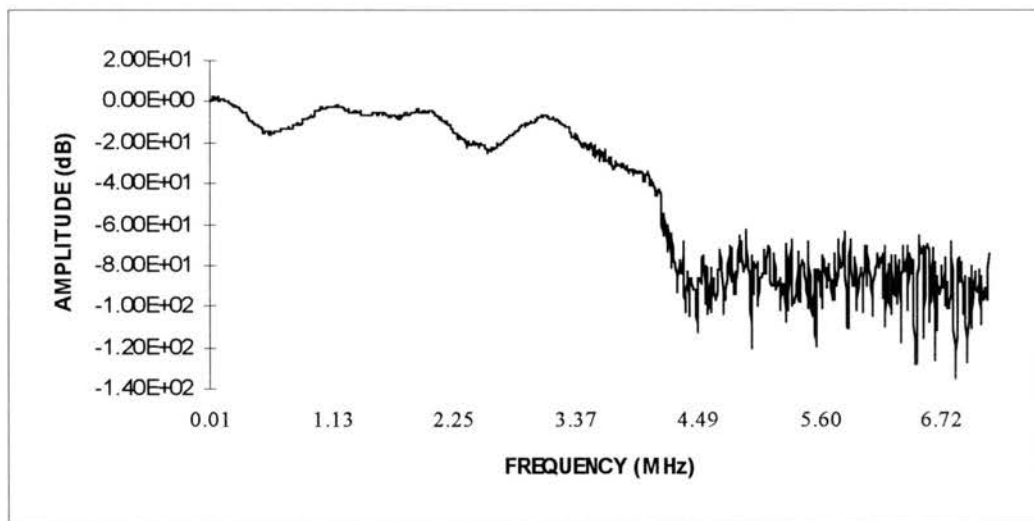


Figure 5-8a. Spectrum of the received GCR signal for Case 1 (linear equalizer)

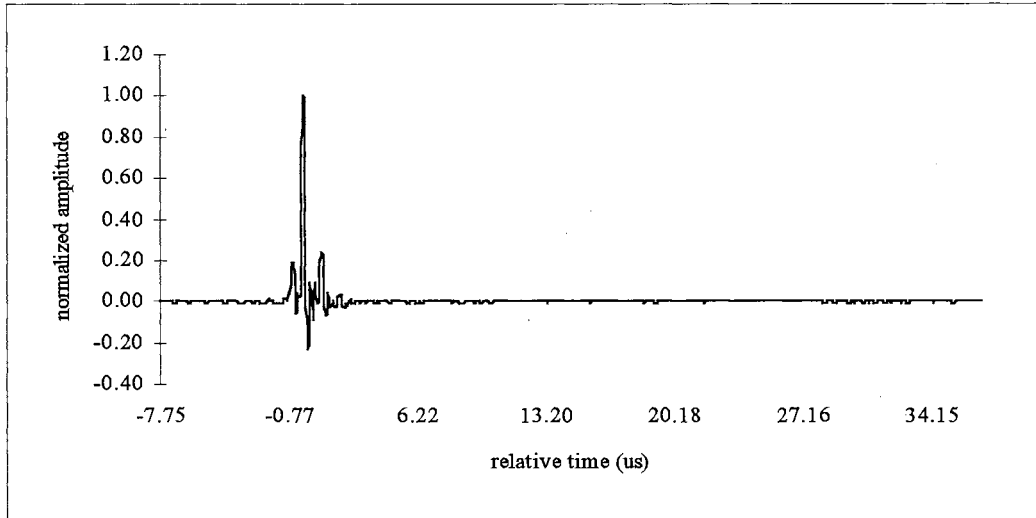


Figure 5-9. Estimated channel impulse response  $f(n)$  for Case 1 (linear)

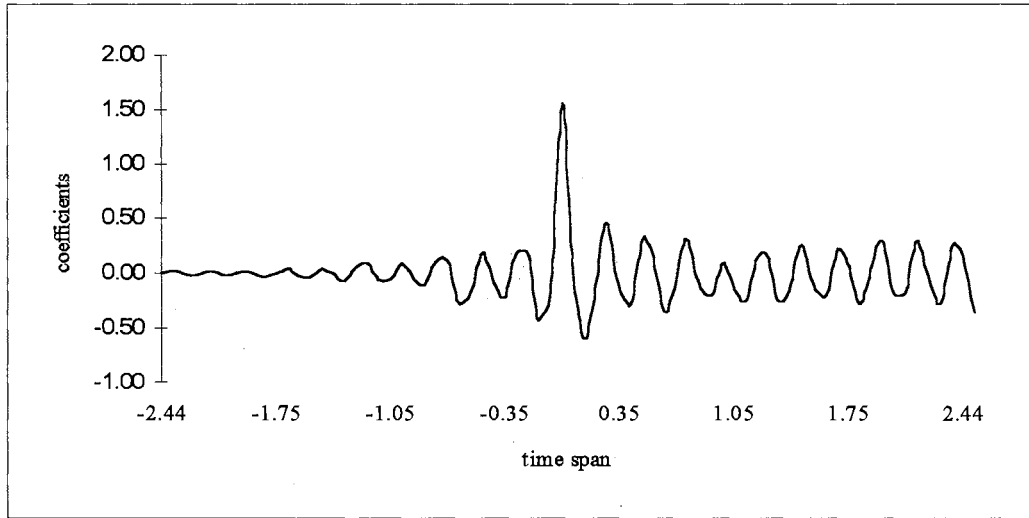


Figure 5-10. Coefficients of linear ZF equalizer for Case 1

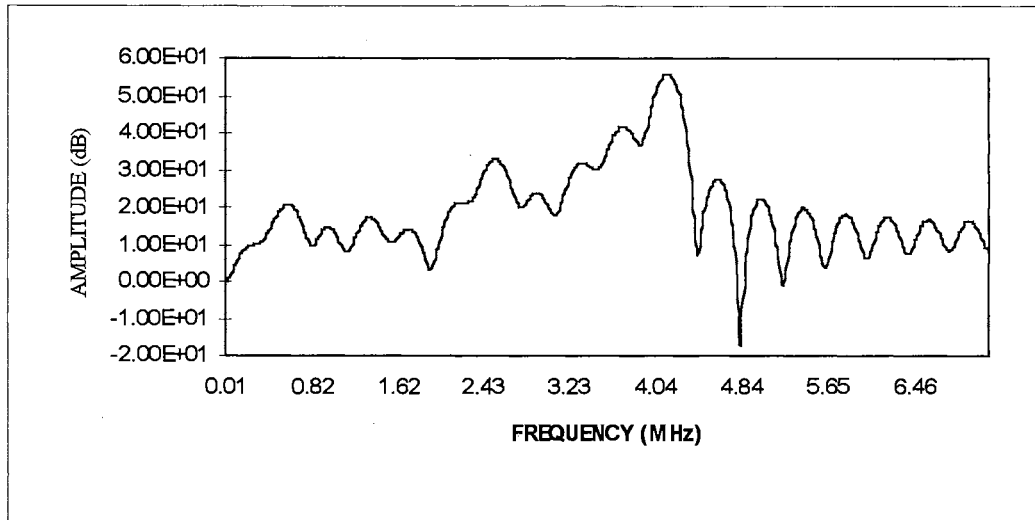


Figure 5-11. Frequency response of linear ZF equalizer for Case 1

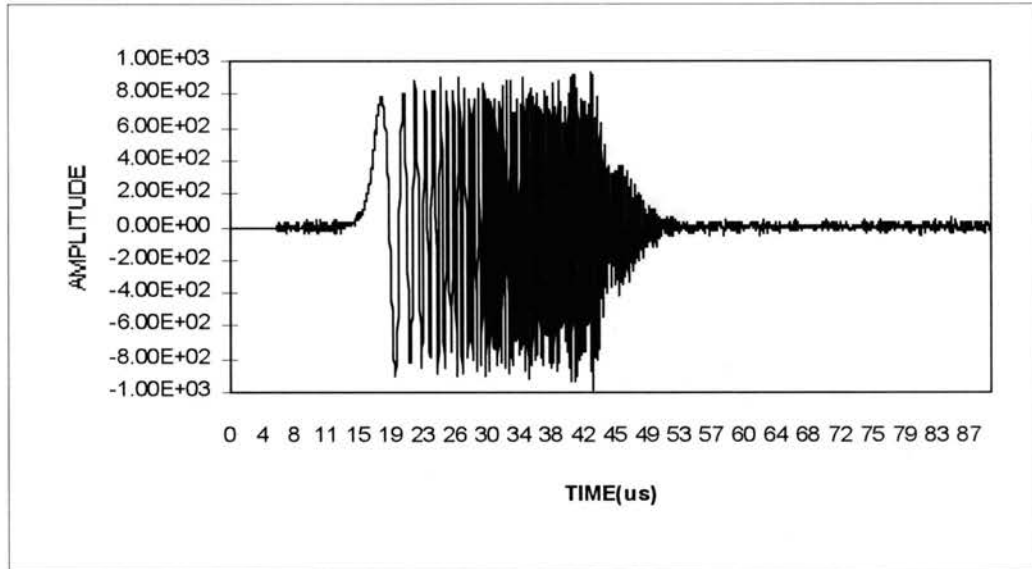


Figure 5-12. Output signal of the linear ZF equalizer for Case 1

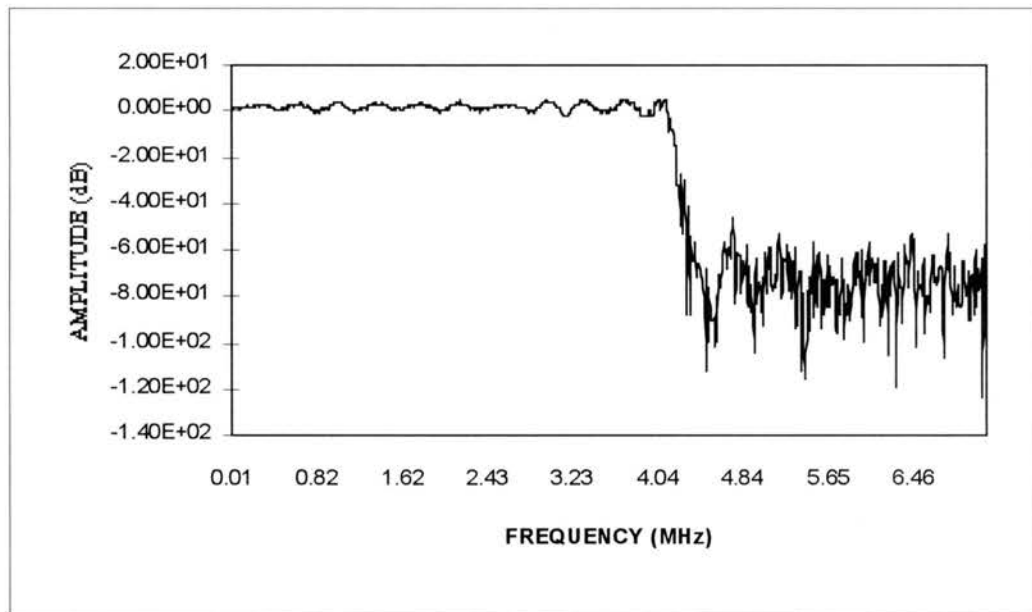


Figure 5-13. Output signal spectrum of the linear ZF equalizer for Case 1

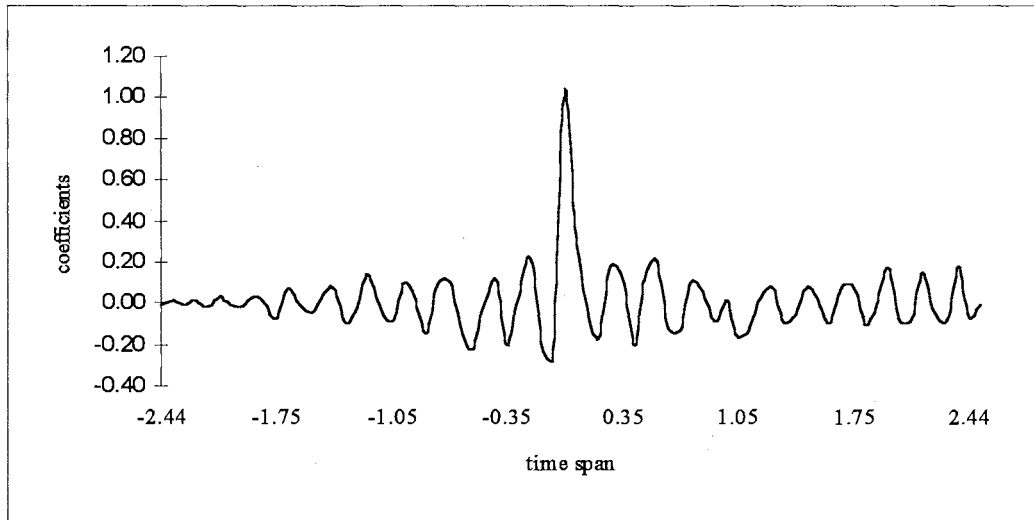


Figure 5-14. Coefficients of linear Lucky's ZF equalizer for Case 1

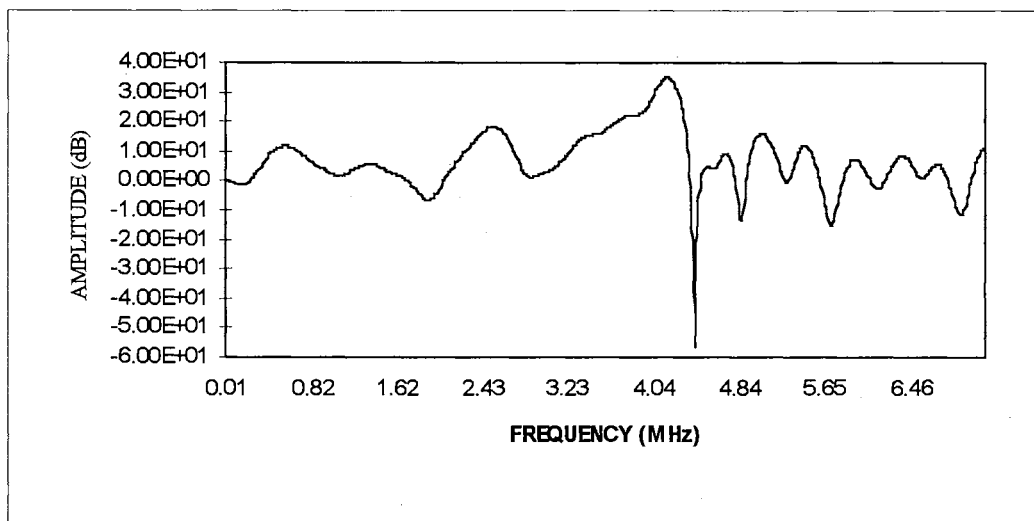


Figure 5-15. Frequency response of linear Lucky's ZF equalizer for Case 1



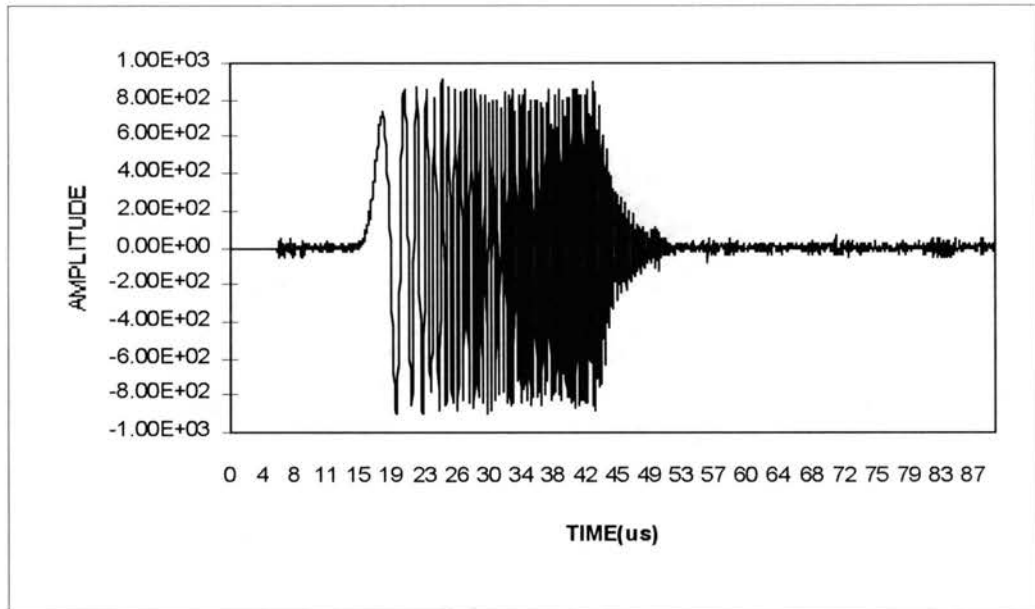


Figure 5-16. Output signal of the linear Lucky's ZF equalizer for Case 1

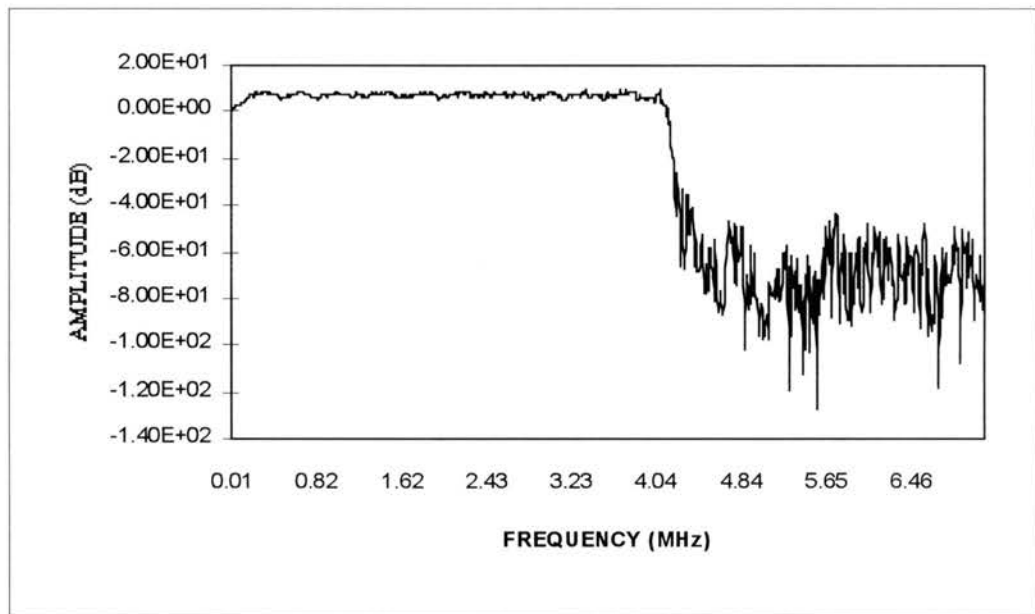


Figure 5-17. Output signal spectrum of the linear Lucky's ZF equalizer for Case 1

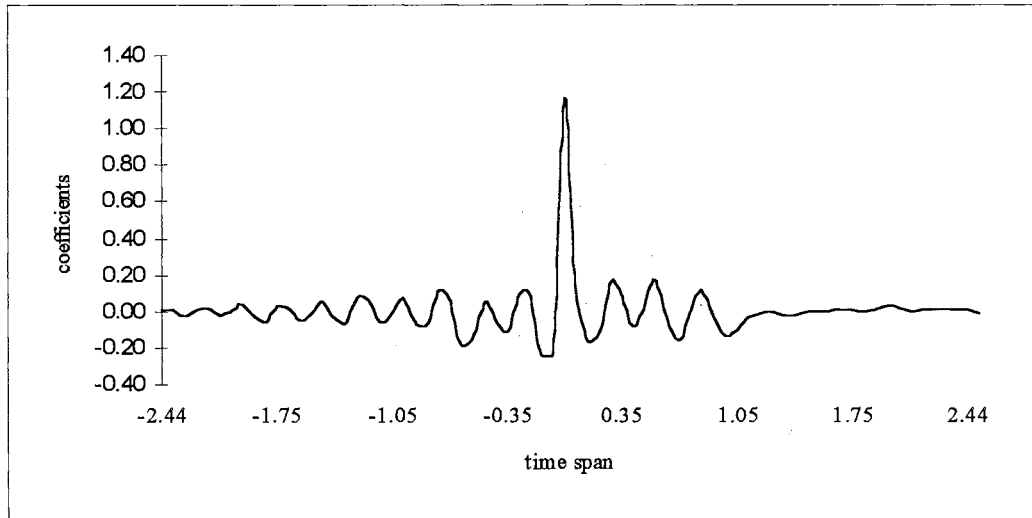


Figure 5-18. Coefficients of linear MSE (LMS) equalizer for Case 1

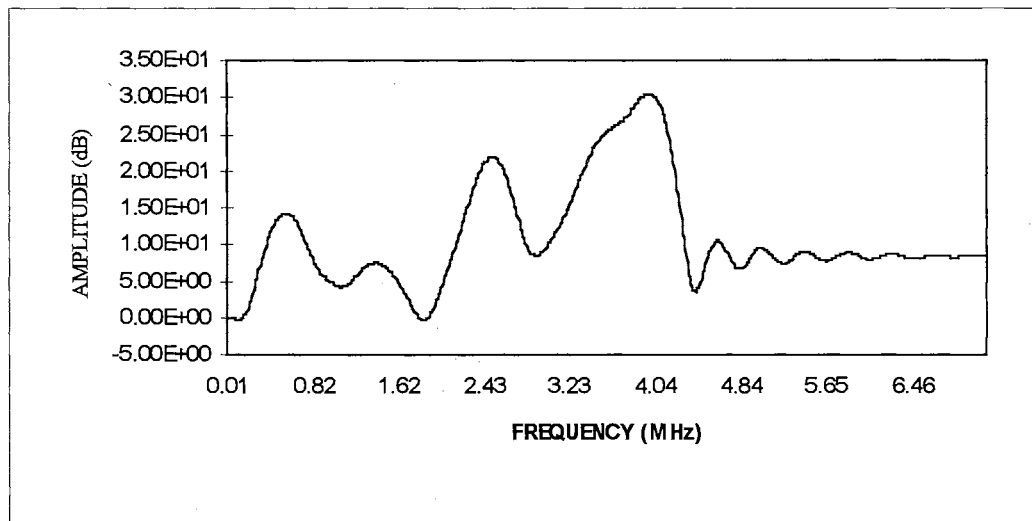


Figure 5-19. Frequency response of linear MSE (LMS) equalizer for Case 1

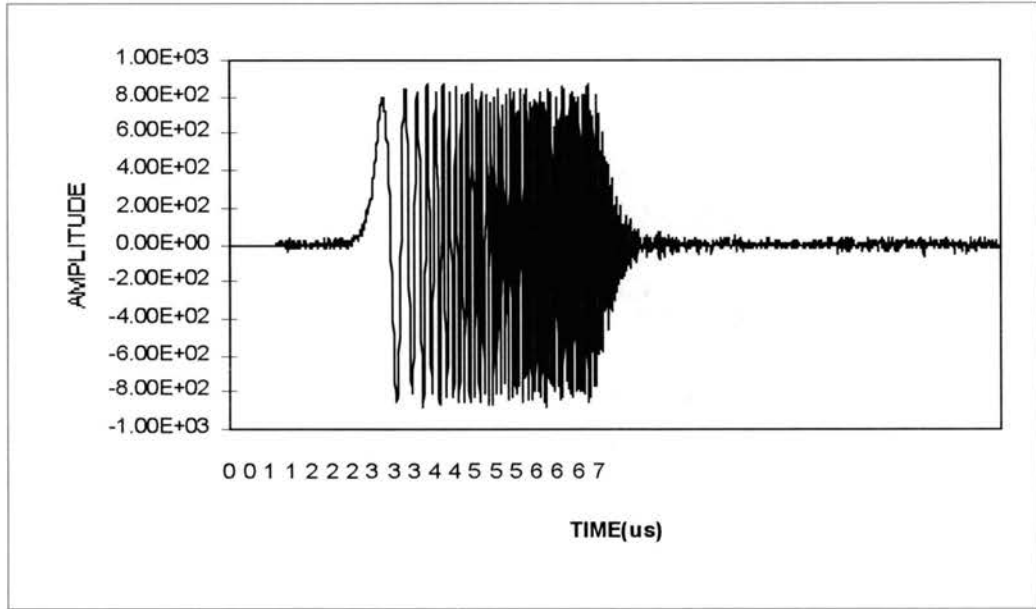


Figure 5-20. Output signal of the linear MSE (LMS) equalizer for Case 1

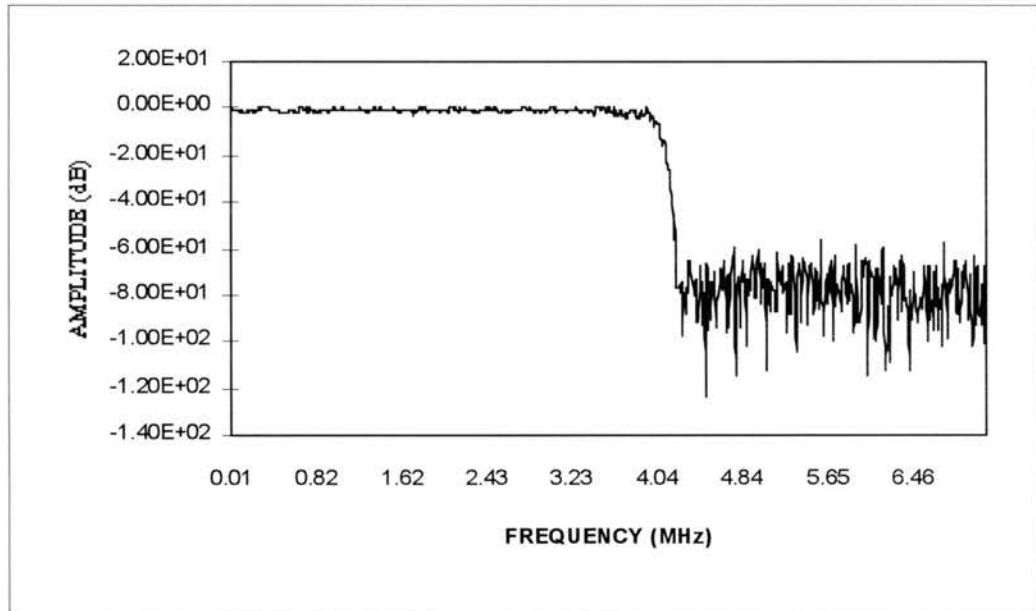


Figure 5-21. Output signal spectrum of the linear MSE (LMS) equalizer for Case 1

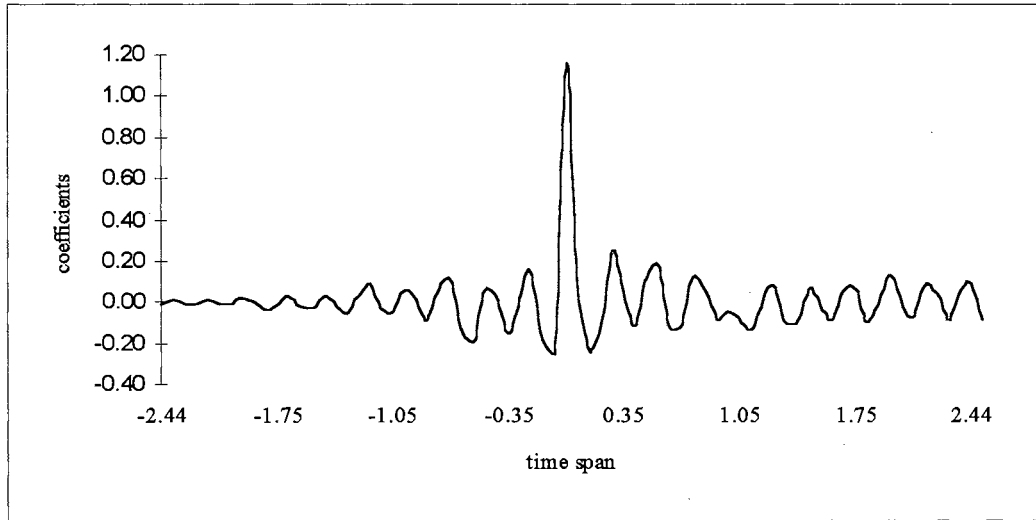


Figure 5-22. Coefficients of linear MN ( $\lambda=0.5$ ) equalizer for Case 1

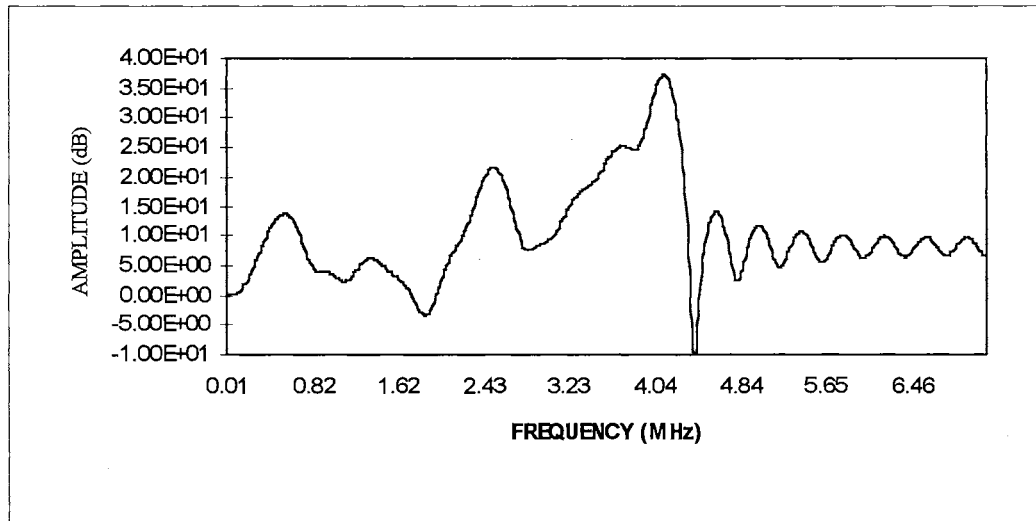


Figure 5-23. Frequency response of linear MN ( $\lambda=0.5$ ) equalizer for Case 1

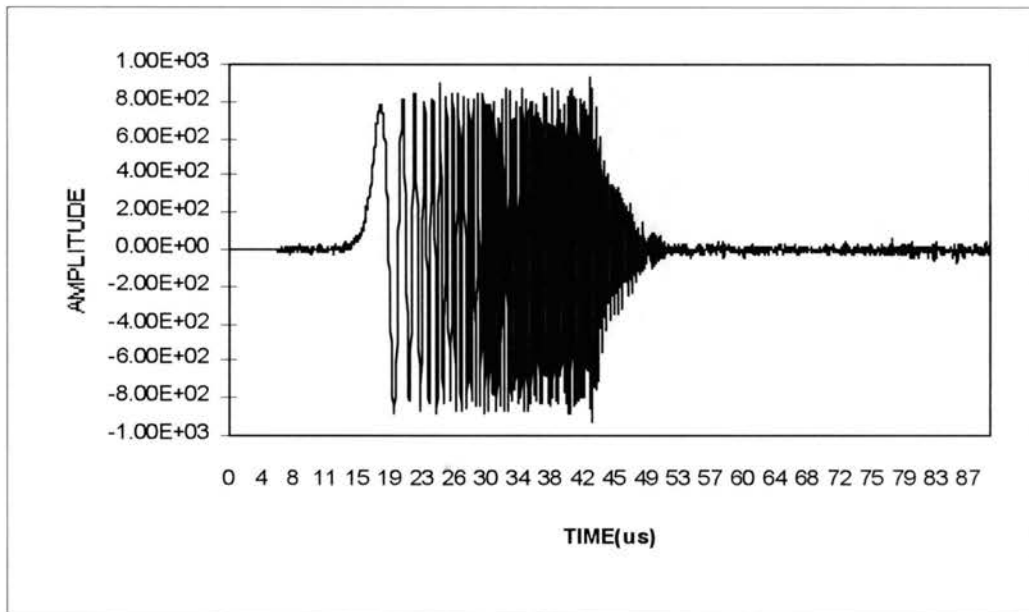


Figure 5-24. Output signal of the linear mixed norm equalizer for Case 1

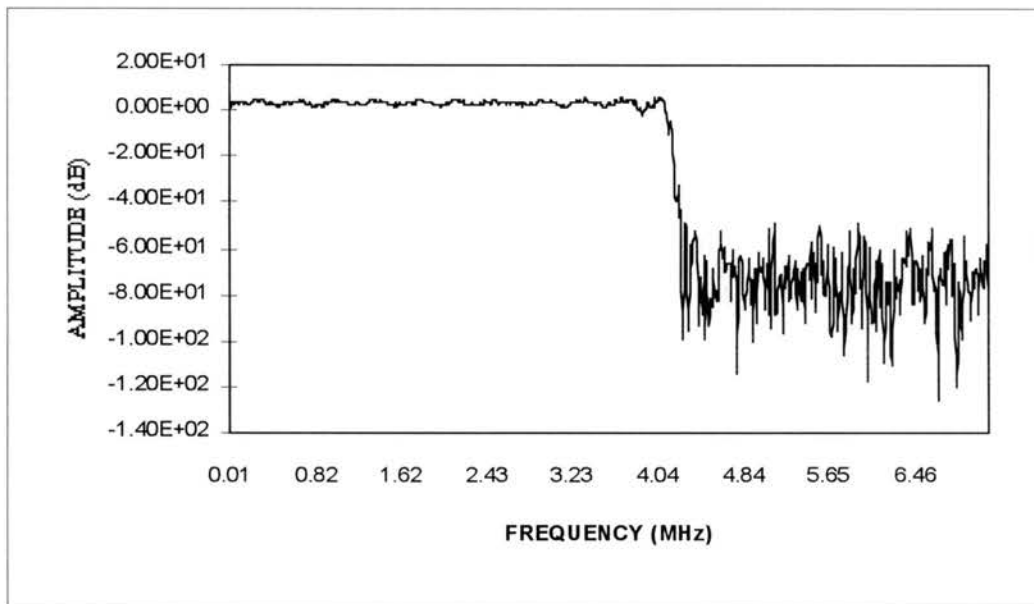


Figure 5-25. Output signal spectrum of the linear mixed norm equalizer for Case 1

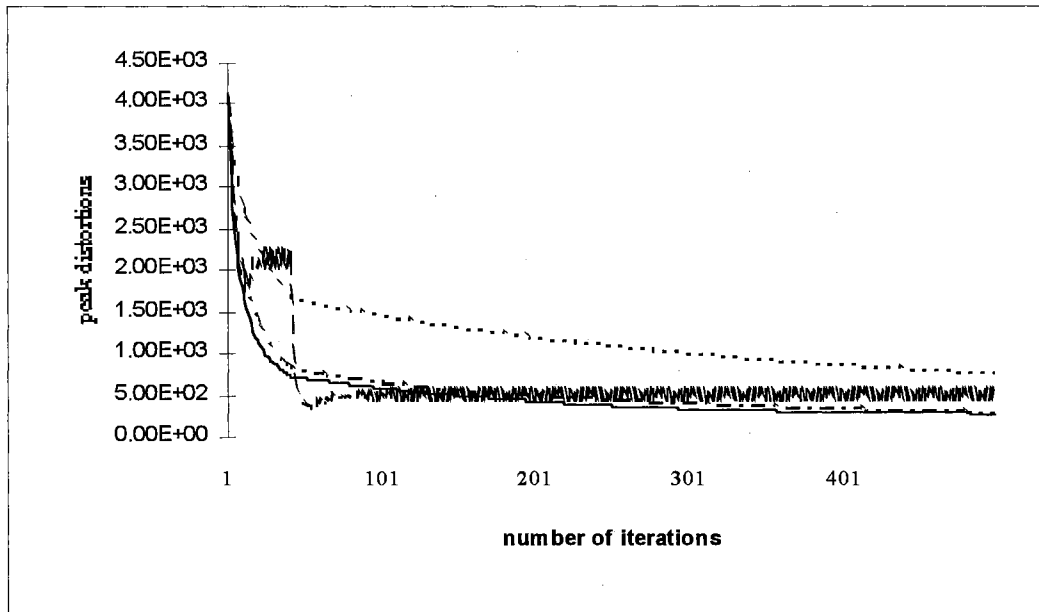


Figure 5-26. Peak distortions of linear equalizers for Case 1  
 ( ZF: ——— Lucky's ZF: - - - - - MSE: ..... MN: - . - . - )

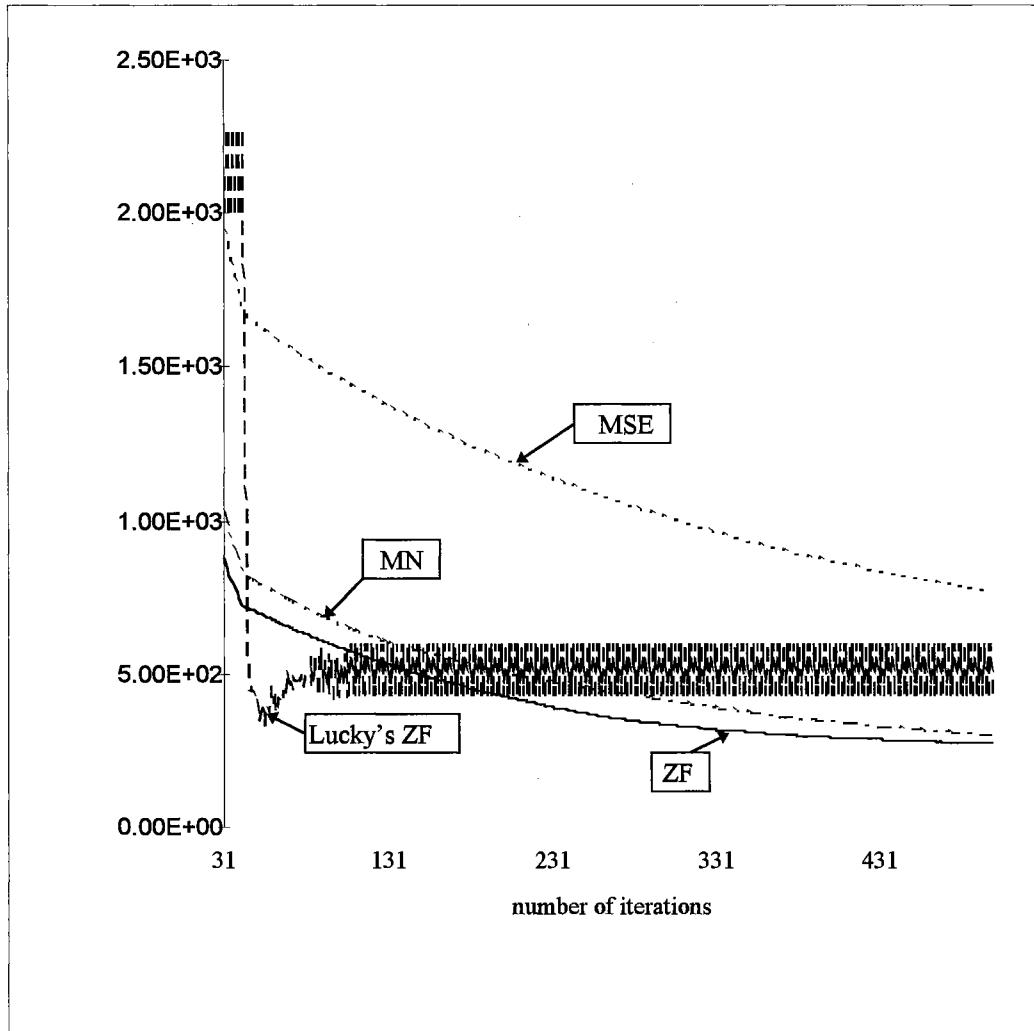


Figure 5-27. Detailed peak distortions of linear equalizers for Case 1  
 ( ZF: ——— Lucky's ZF: - - - - - MSE: ····· MN: - · - · - )

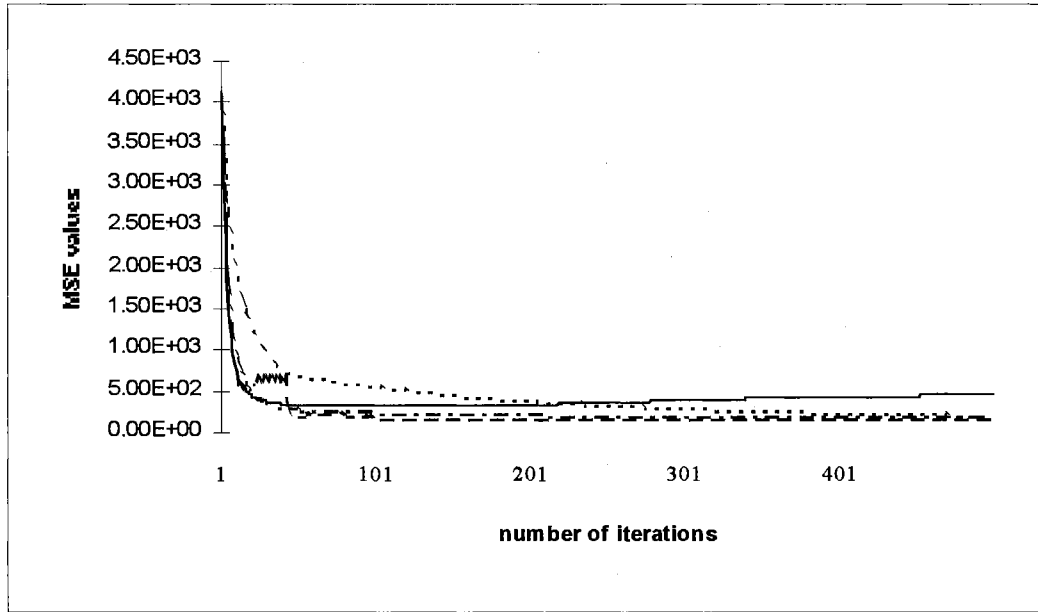


Figure 5-28. MSE of linear equalizers for Case 1

( ZF: ——— Lucky's ZF: - - - - - MSE: - · - · - MN: ····· )



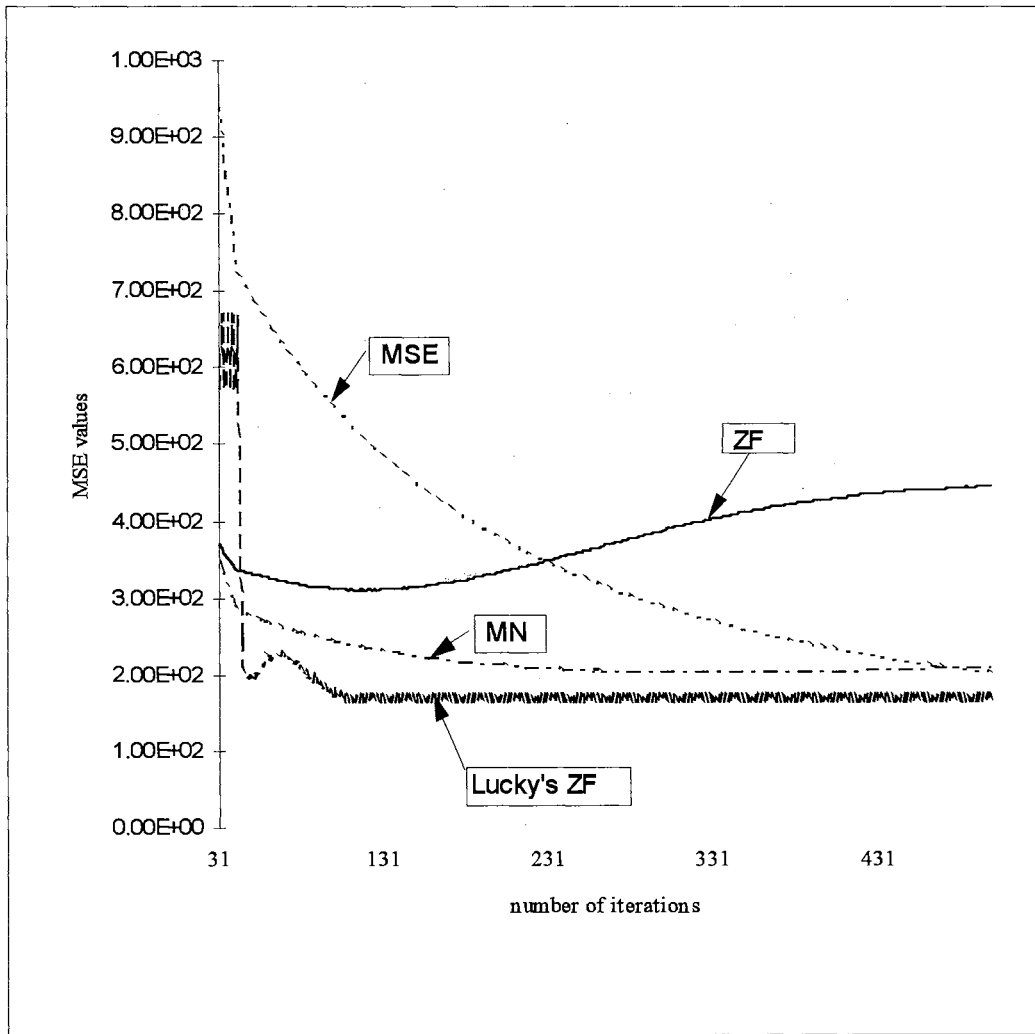


Figure 5-29. Detailed MSE of linear equalizers for Case 1  
 ( ZF: ——— Lucky's ZF: - - - - - MSE: - · - · - MN: ····· )

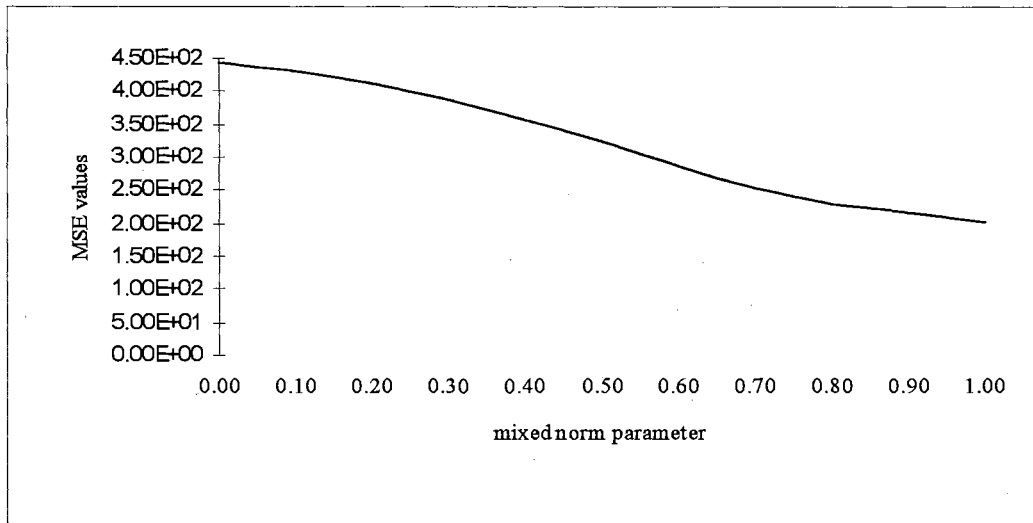


Figure 5-30. MSE values versus mixed norm parameter  $\lambda$  for Case 1 (linear)

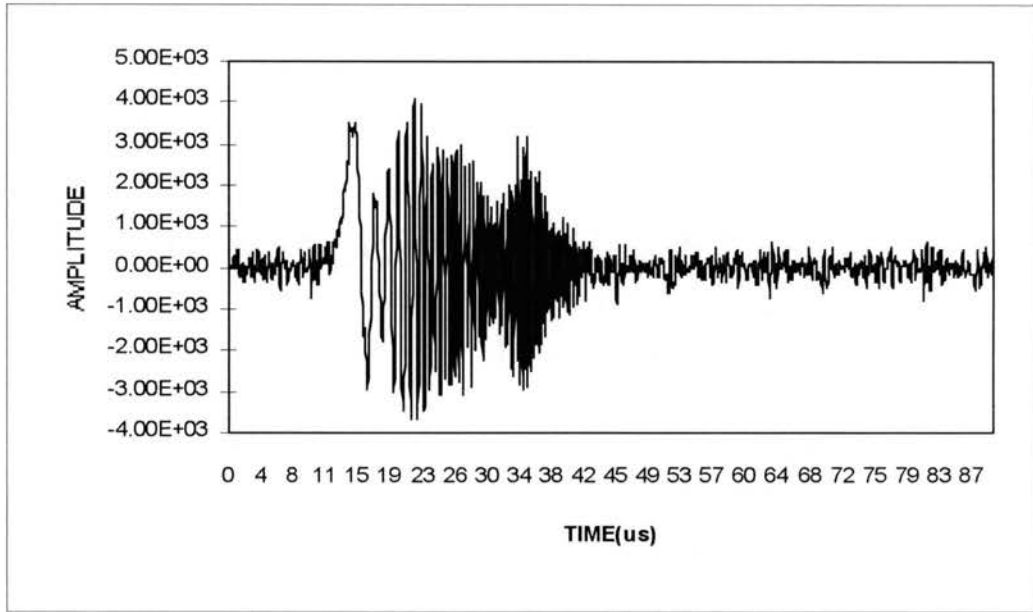


Figure 5-31. Received GCR signal for Case 2 (linear equalizer)

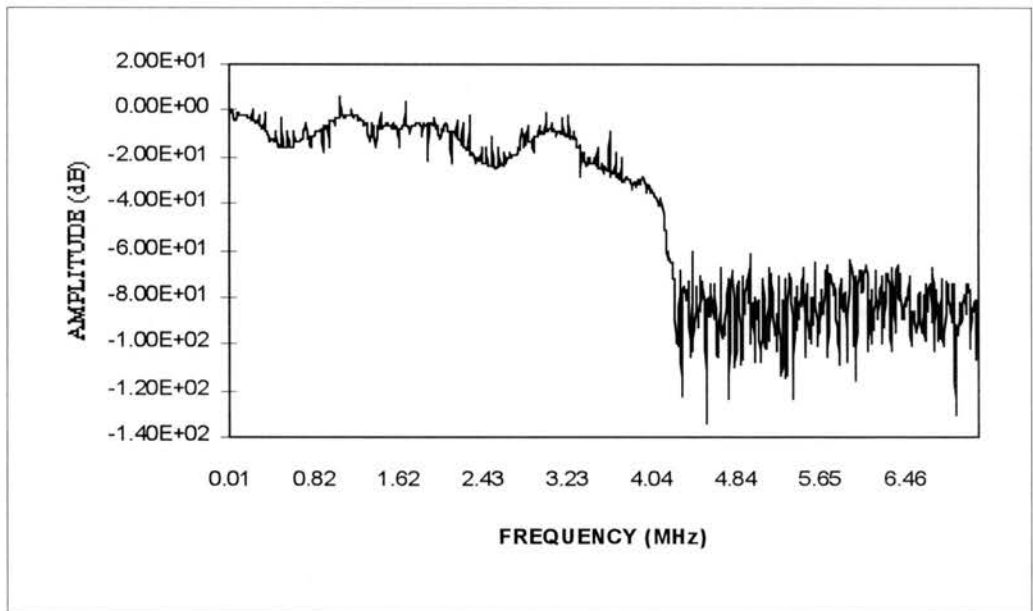


Figure 5-31a. Spectrum of the received GCR signal for Case 2 (linear equalizer)

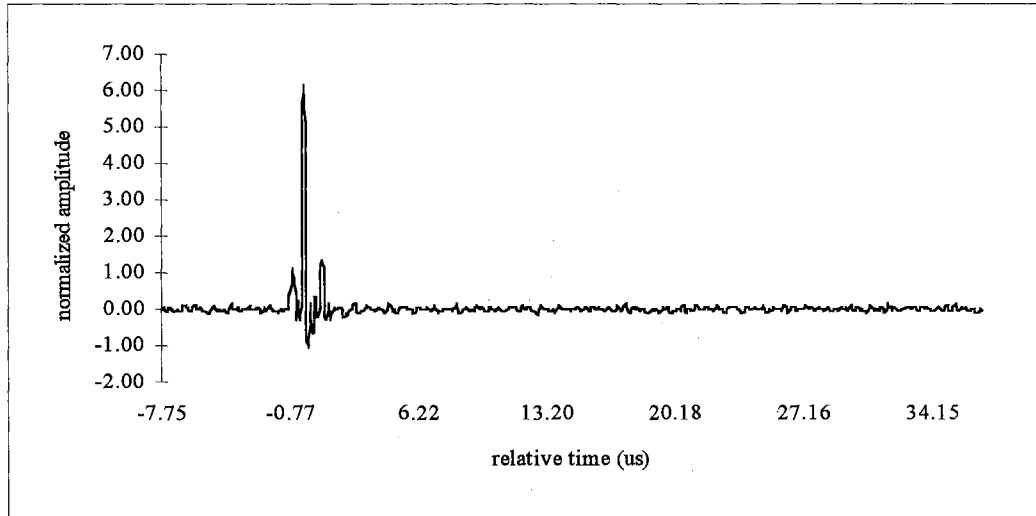


Figure 5-32. Estimated channel impulse response  $f(n)$  for Case 2 (linear)

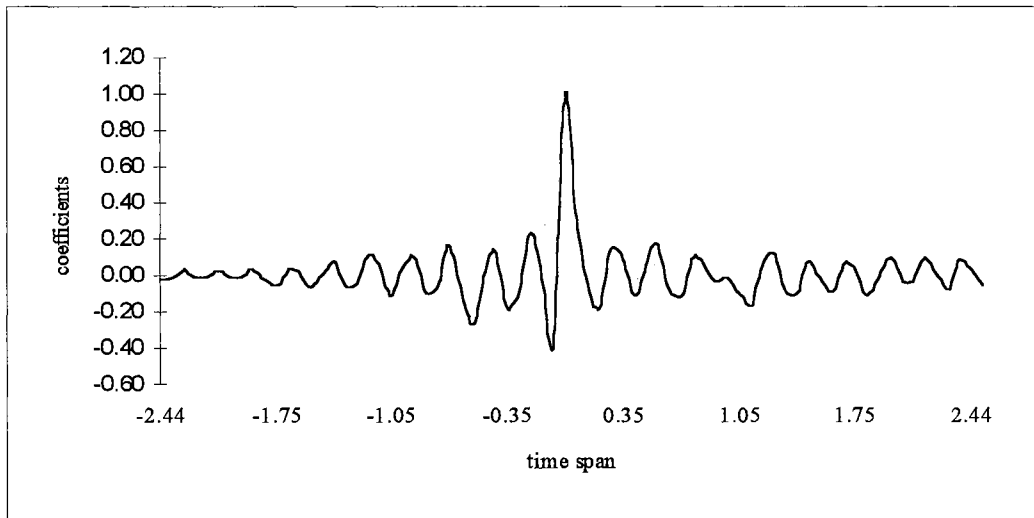


Figure 5-33. Coefficients of linear ZF equalizer for Case 2

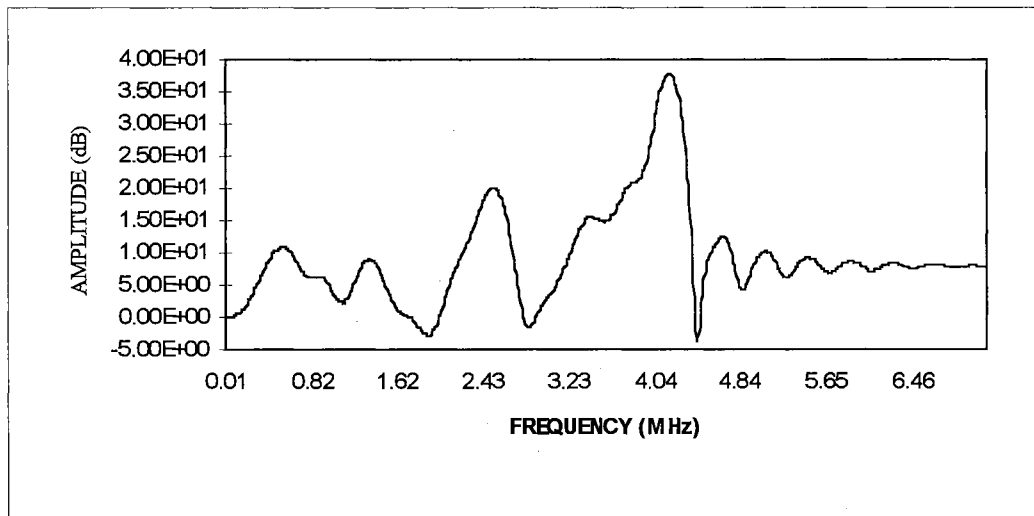


Figure 5-34. Frequency response of linear ZF equalizer for Case 2

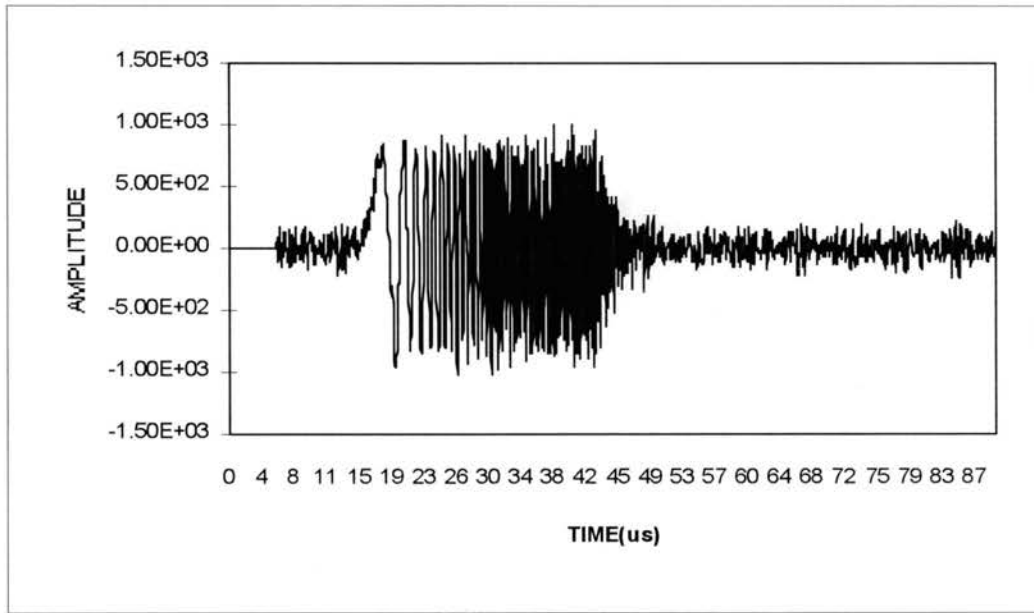


Figure 5-35. Output signal of the linear ZF equalizer for Case 2

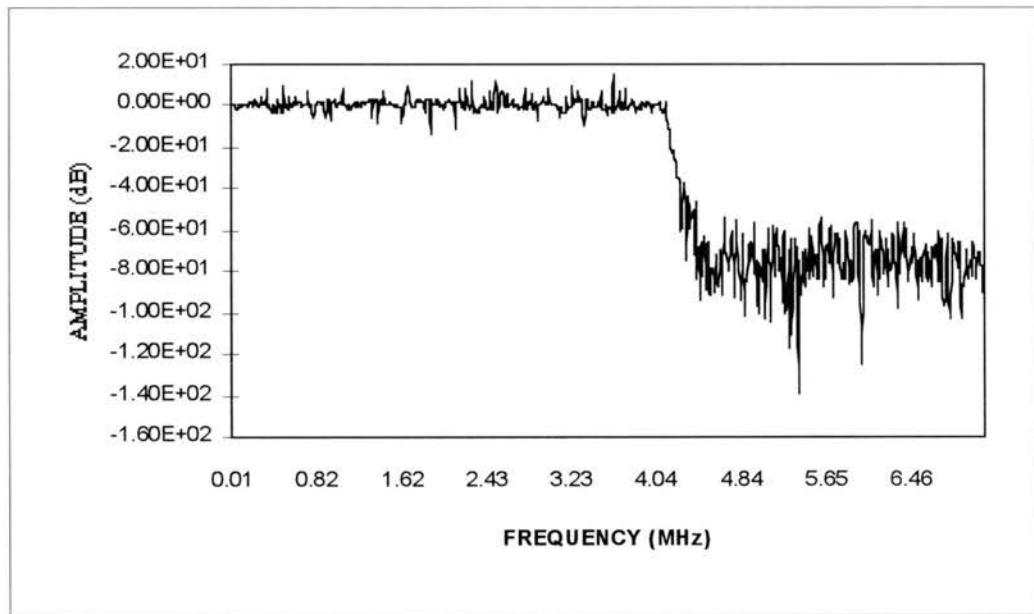


Figure 5-36. Output signal spectrum of the linear ZF equalizer for Case 2

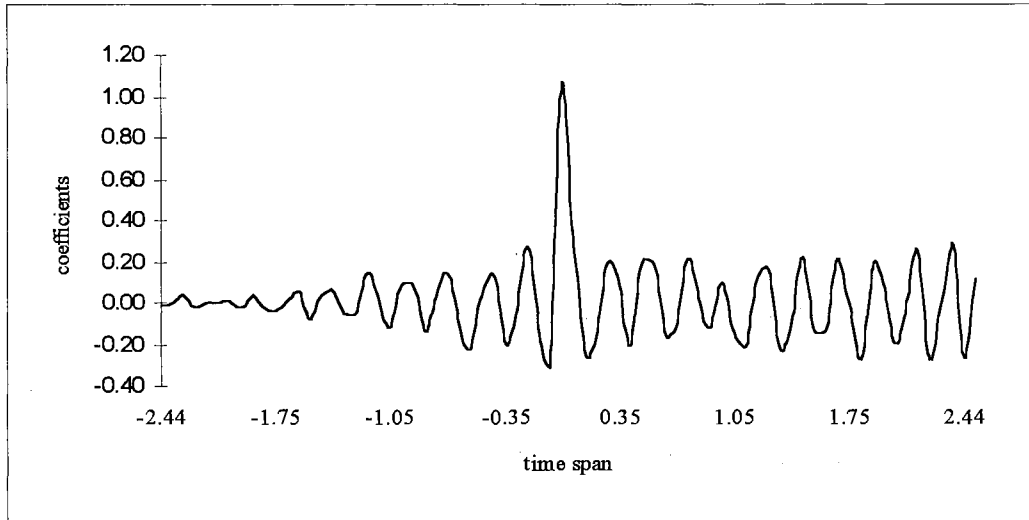


Figure 5-37. Coefficients of linear Lucky's ZF equalizer for Case 2

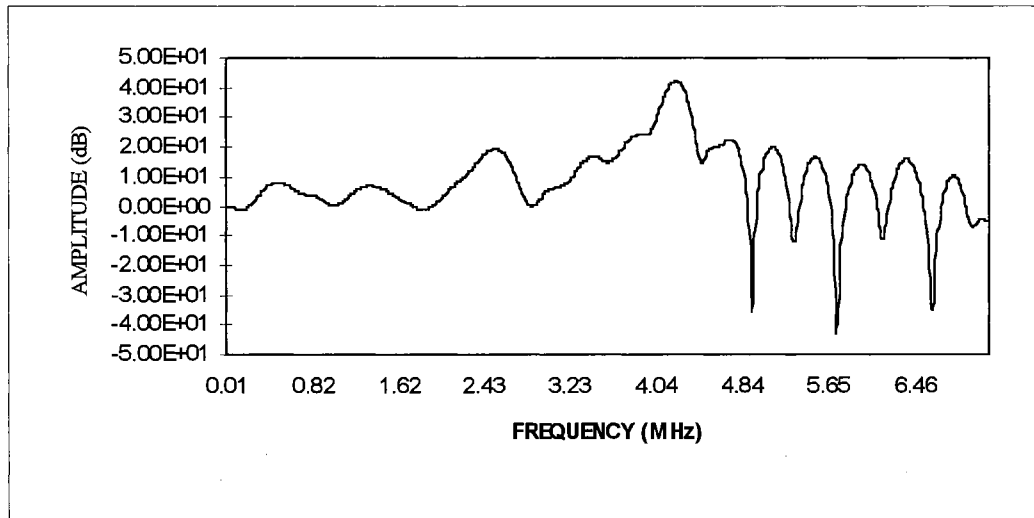


Figure 5-38. Frequency response of linear Lucky's ZF equalizer for Case 2

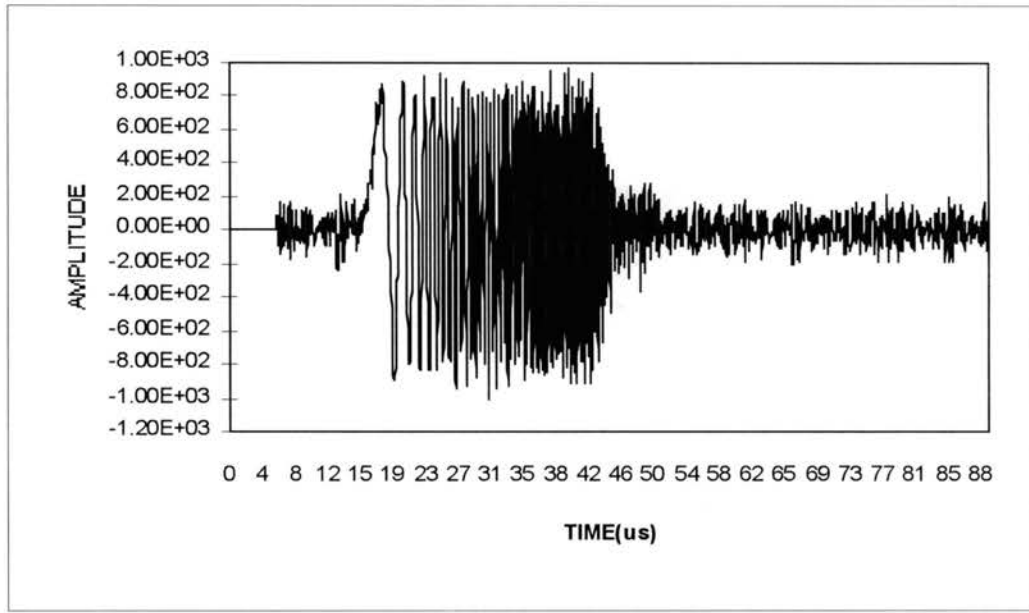


Figure 5-39. Output signal of the linear Lucky's ZF equalizer for Case 2

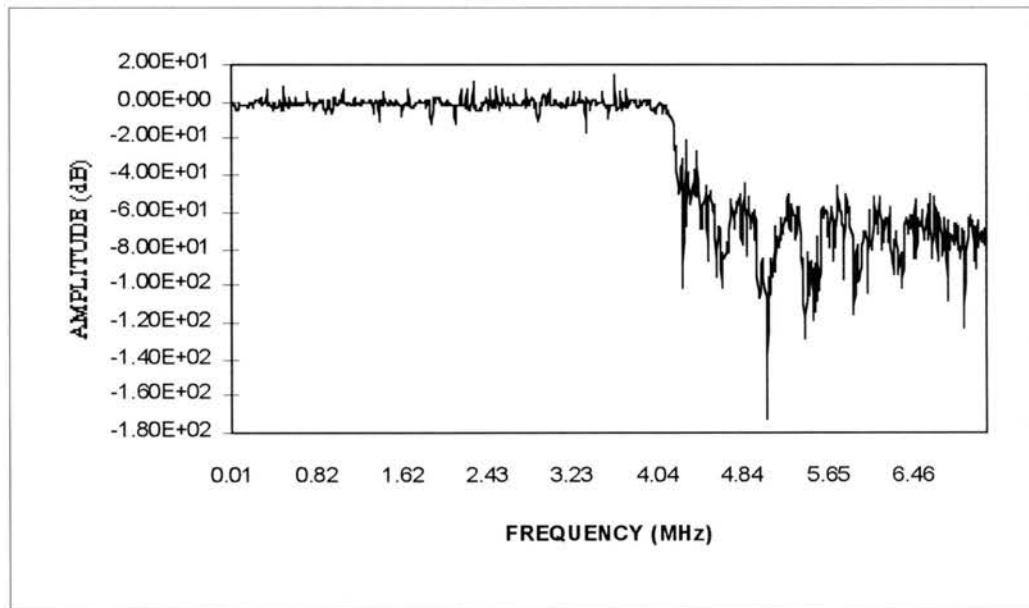


Figure 5-40. Output signal spectrum of the linear Lucky's ZF equalizer for Case 2



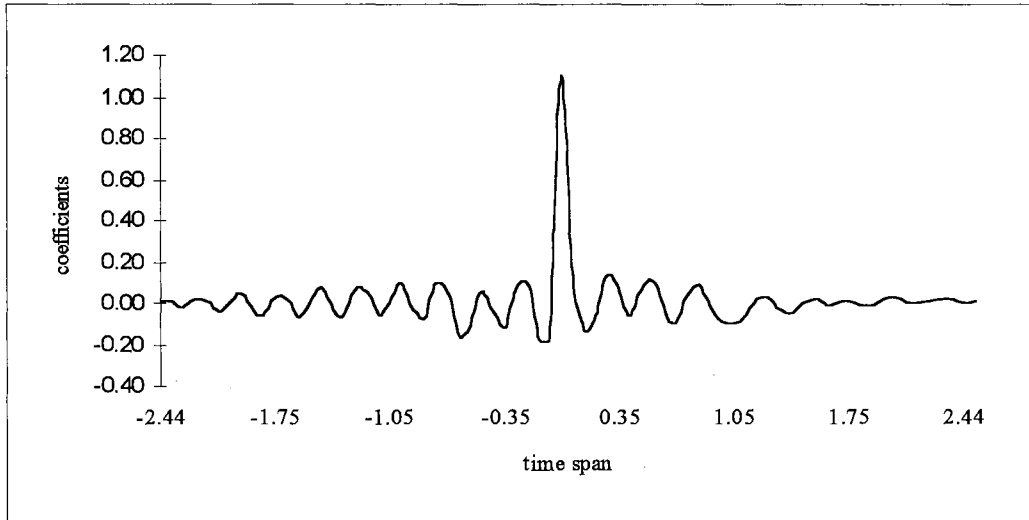


Figure 5-41. Coefficients of linear MSE (LMS) equalizer for Case 2

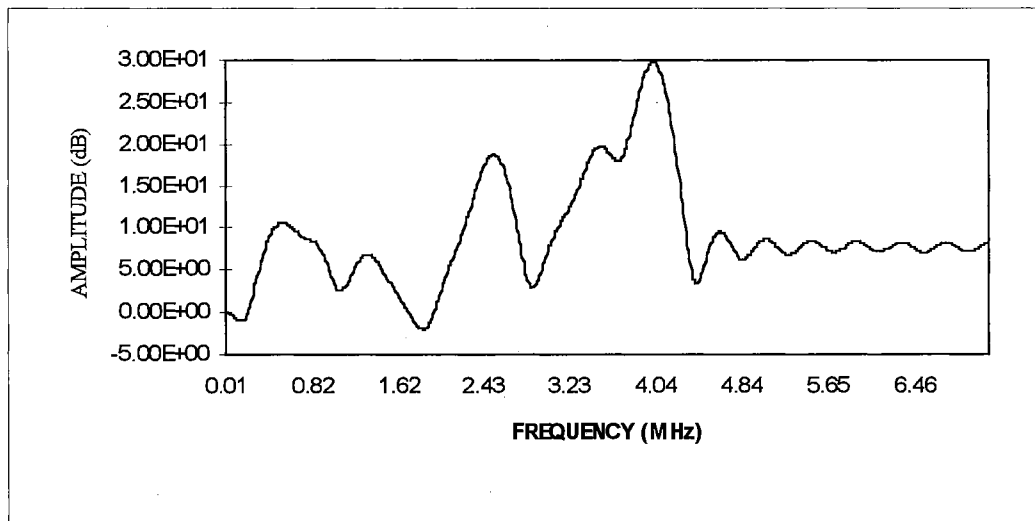


Figure 5-42. Frequency response of linear MSE (LMS) equalizer for Case 2

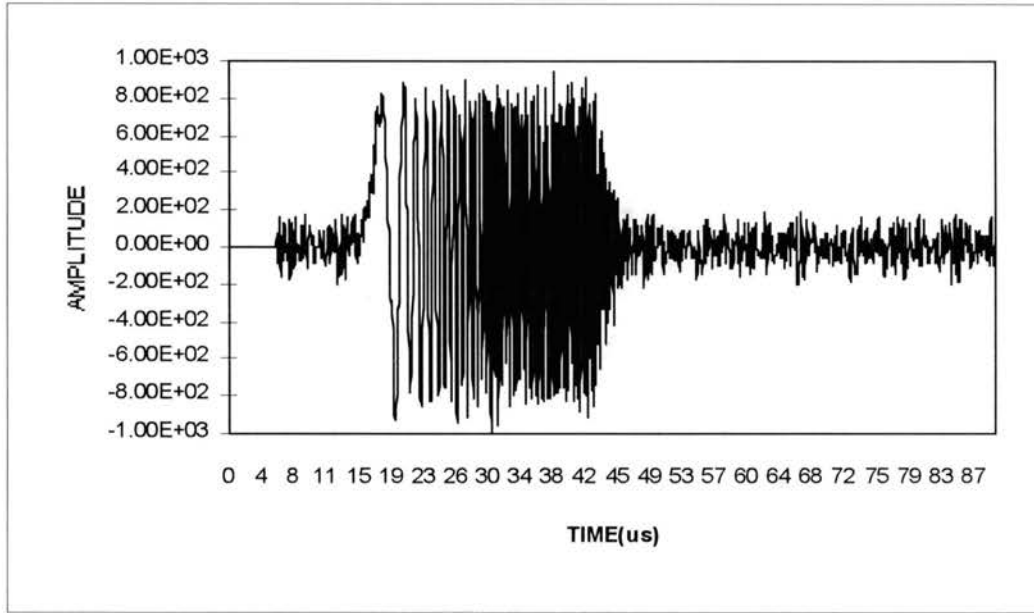


Figure 5-43. Output signal of the linear MSE (LMS) equalizer for Case 2

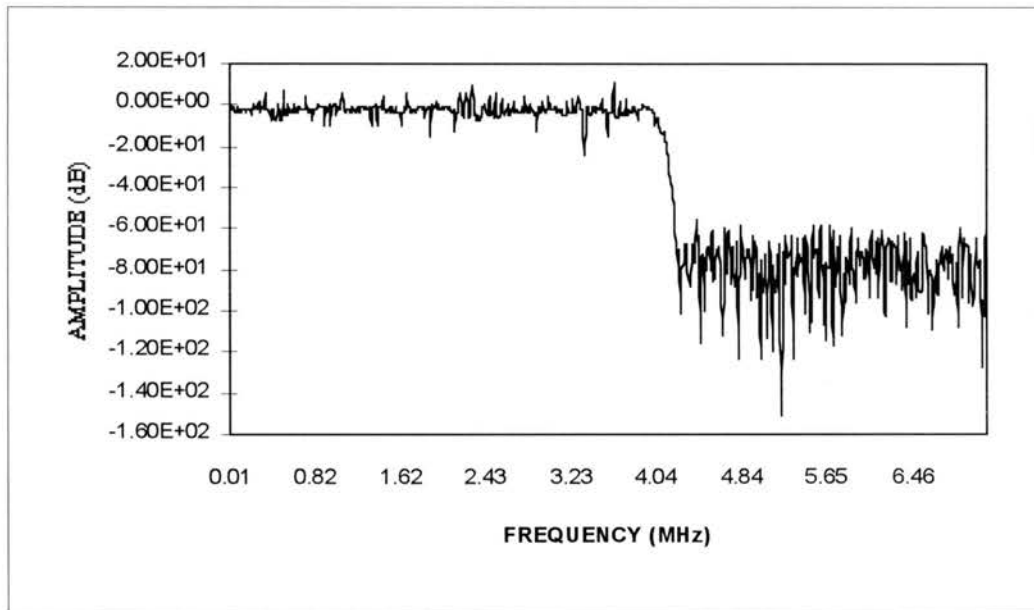


Figure 5-44. Output signal spectrum of the linear MSE (LMS) equalizer for Case 2

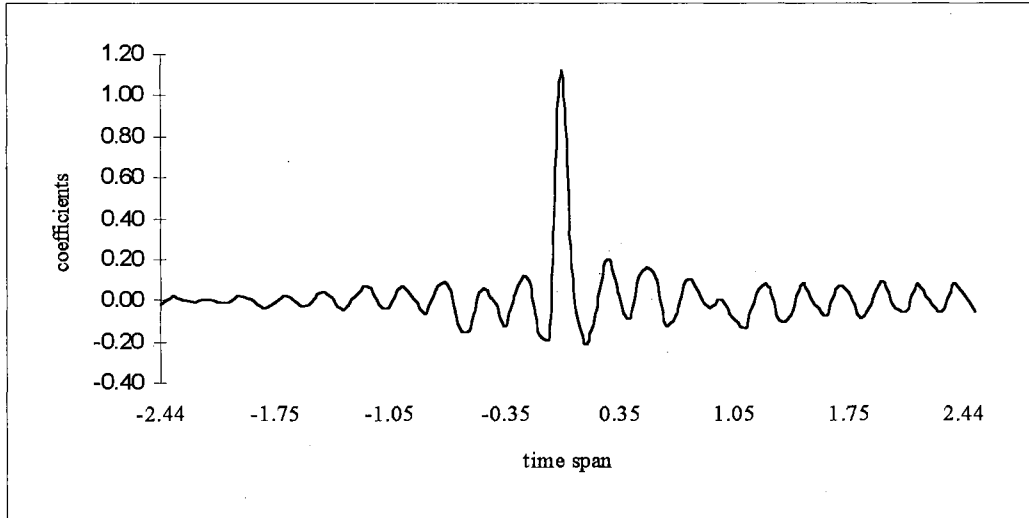


Figure 5-45. Coefficients of linear MN ( $\lambda=0.5$ ) equalizer for Case 2

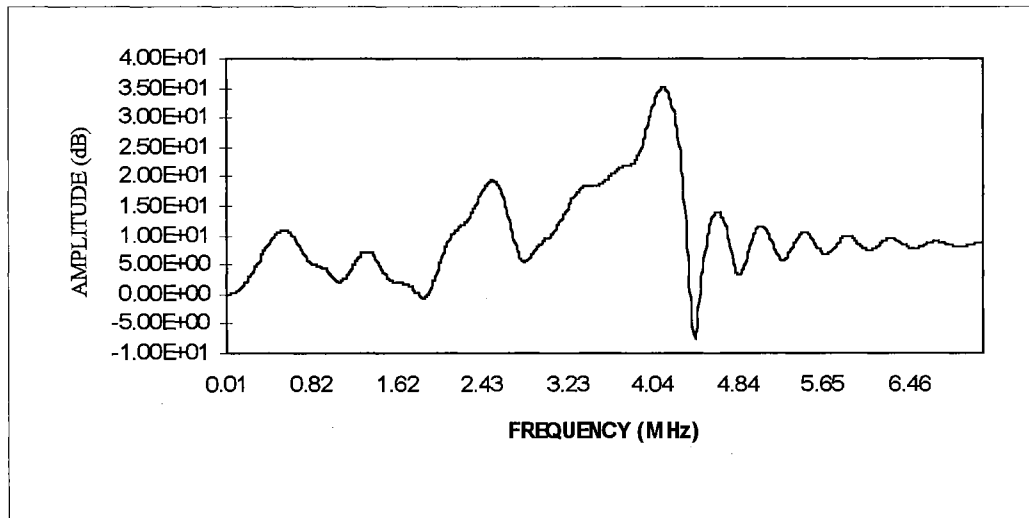


Figure 5-46. Frequency response of linear MN ( $\lambda=0.5$ ) equalizer for Case 2

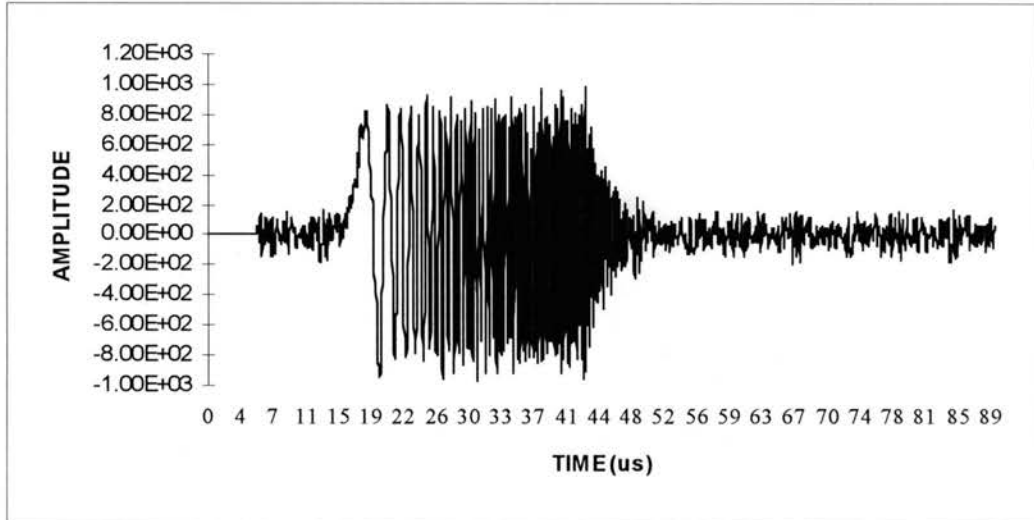


Figure 5-47. Output signal of the linear mixed norm equalizer for Case 2

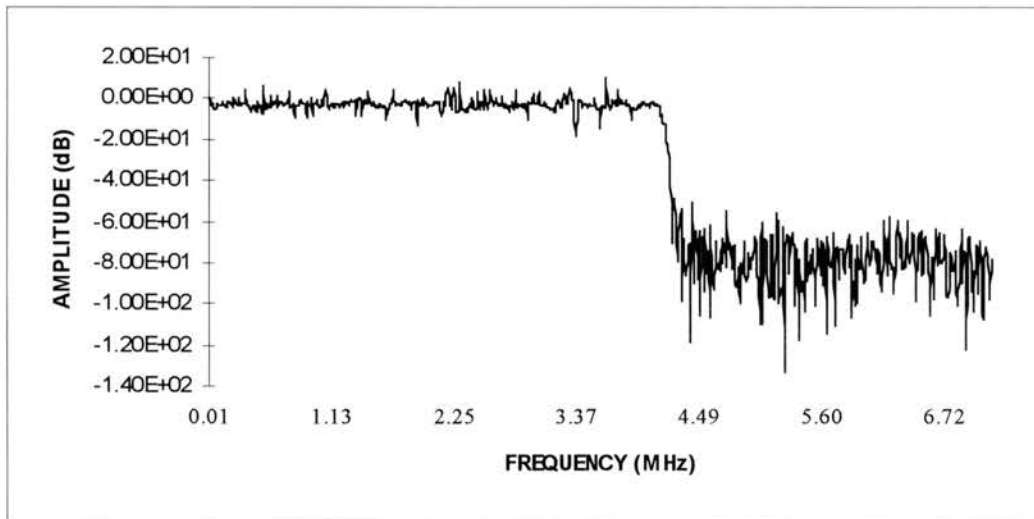


Figure 5-48. Output signal spectrum of the linear mixed norm equalizer for Case 2

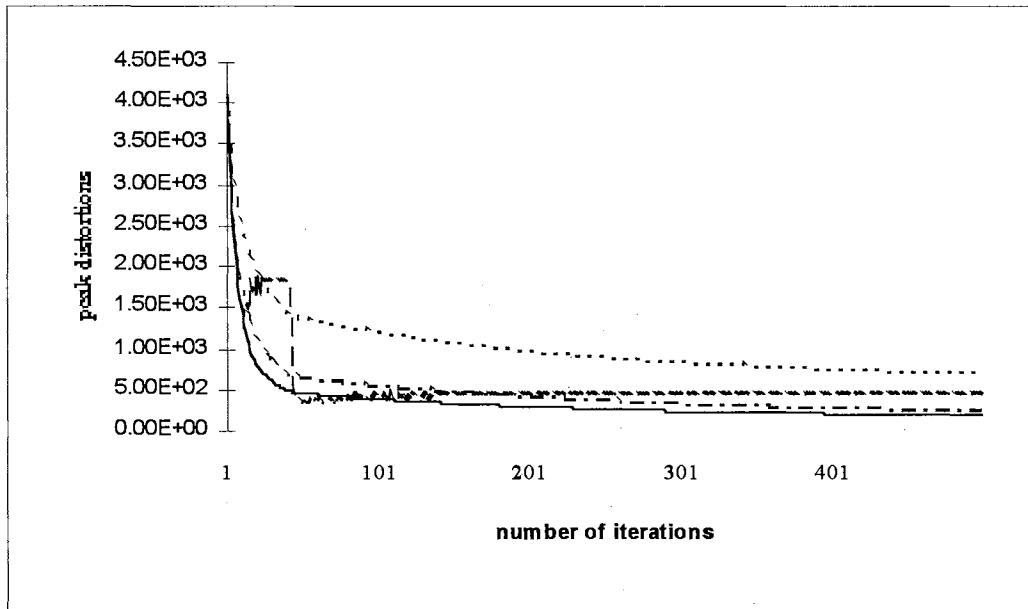


Figure 5-49. Peak distortions of linear equalizers for Case 2  
 ( ZF: ——— Lucky's ZF: - - - - - MSE: ..... MN: - . - . - )

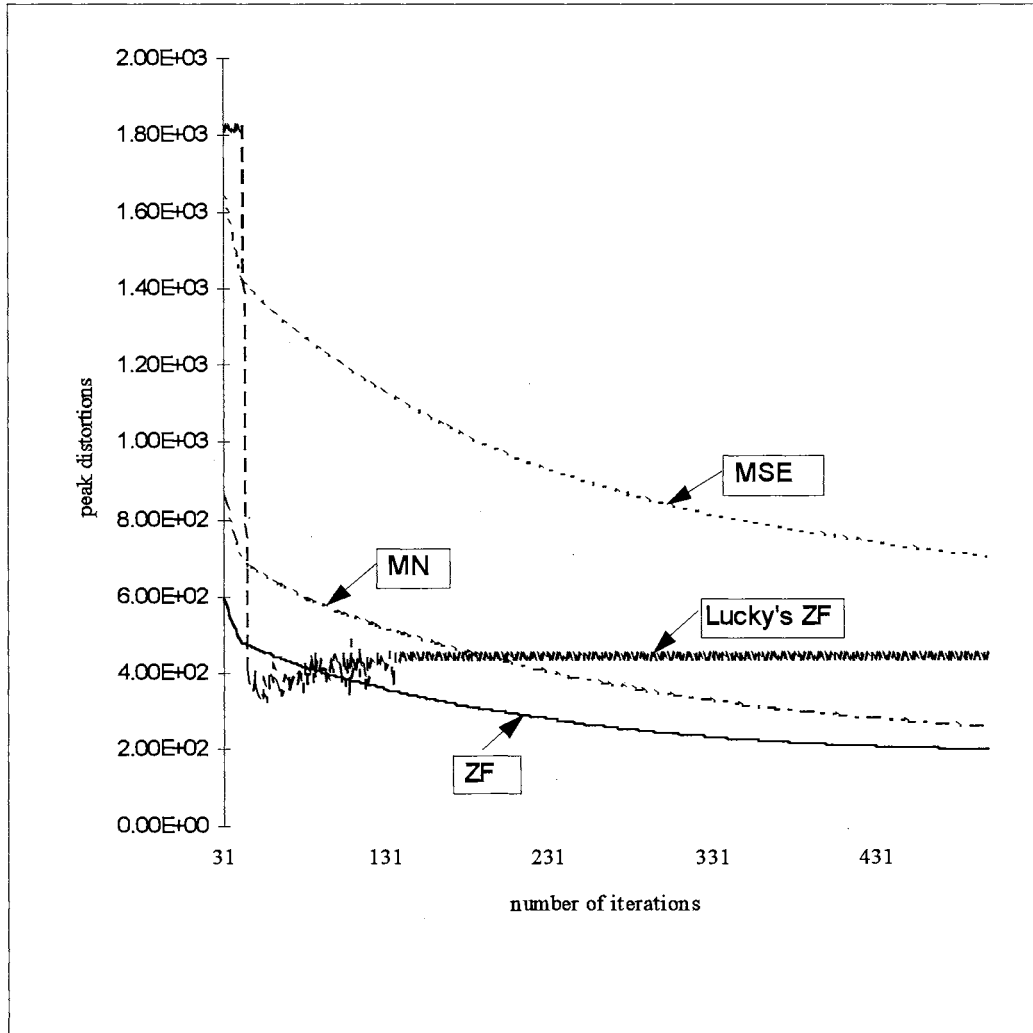


Figure 5-50. Detailed peak distortions of linear equalizers for Case 2  
 ( ZF: ——— Lucky's ZF: - - - - - MSE: - . - . - MN: . . . . . )

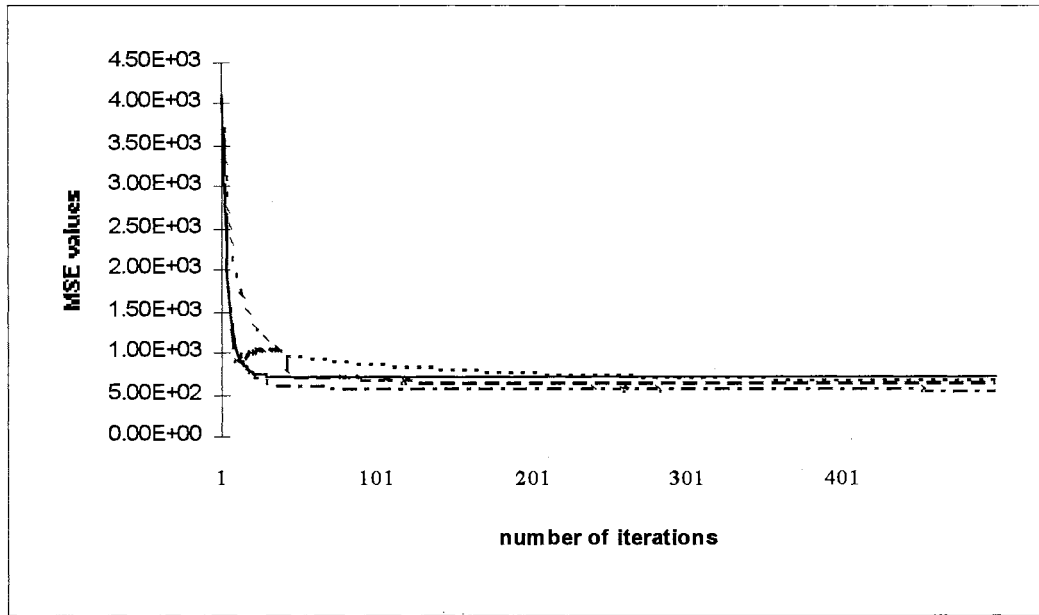


Figure 5-51. MSE of linear equalizers for Case 2  
 ( ZF: ——— Lucky's ZF: - - - - - MSE: ..... MN: - . - . - )

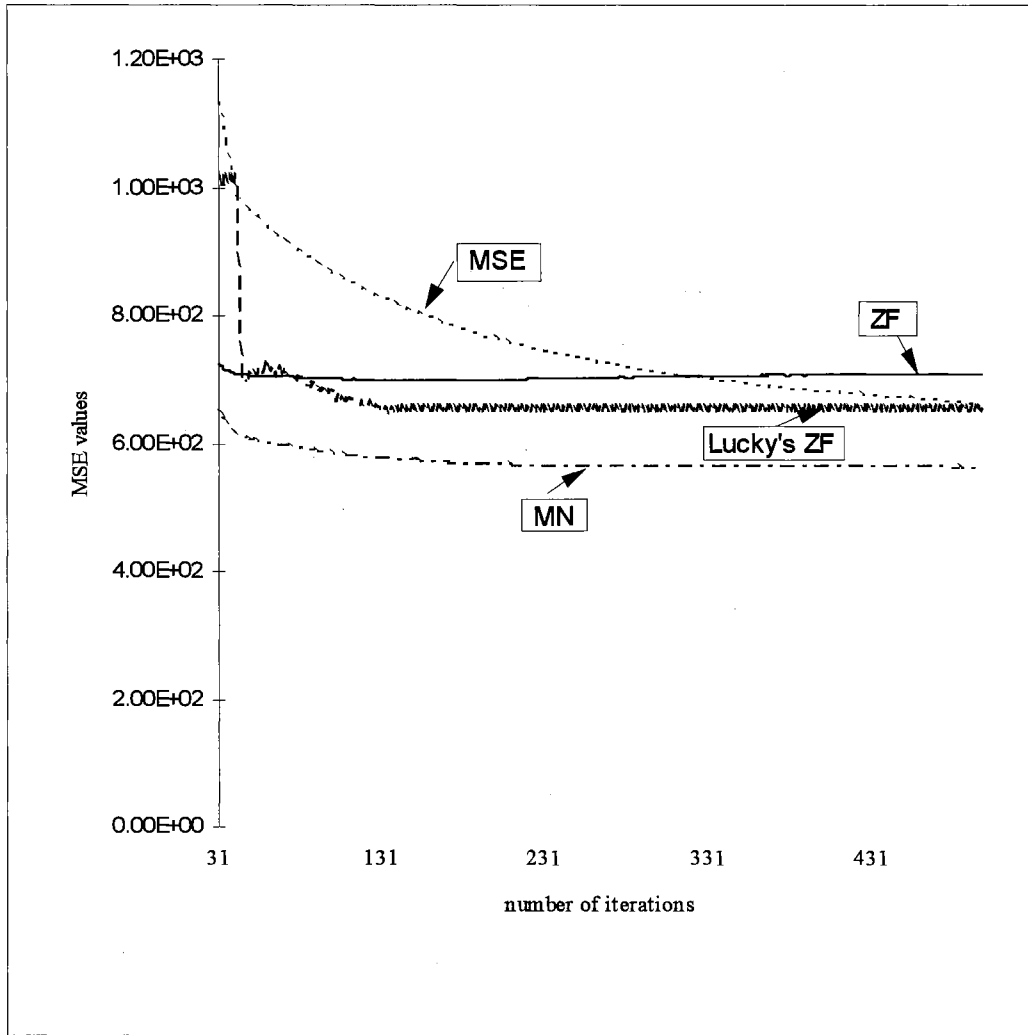


Figure 5-52. Detailed MSE of linear equalizers for Case 2  
 ( ZF: ——— Lucky's ZF: - - - - - MSE: - · - · - MN: ····· )



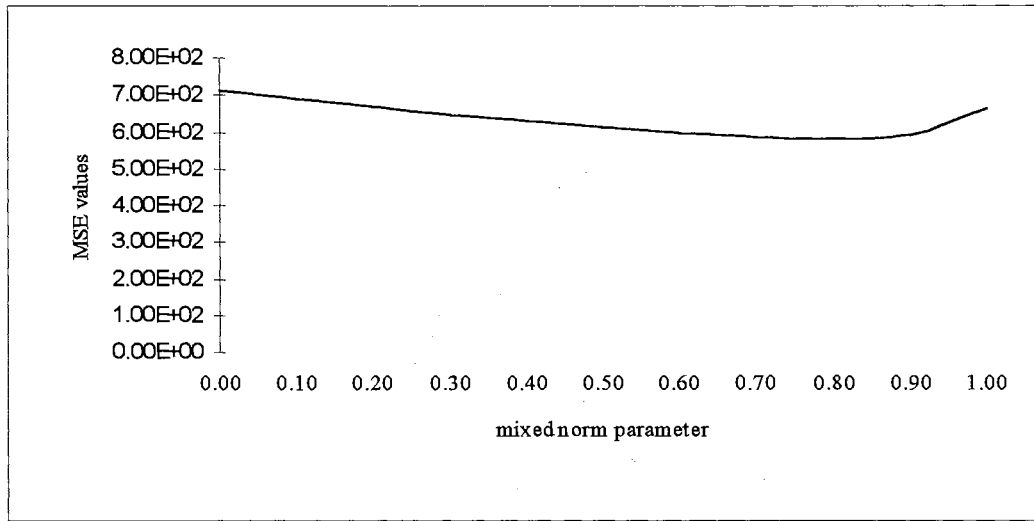


Figure 5-53. MSE values versus mixed norm parameter  $\lambda$  for Case 2 (linear)

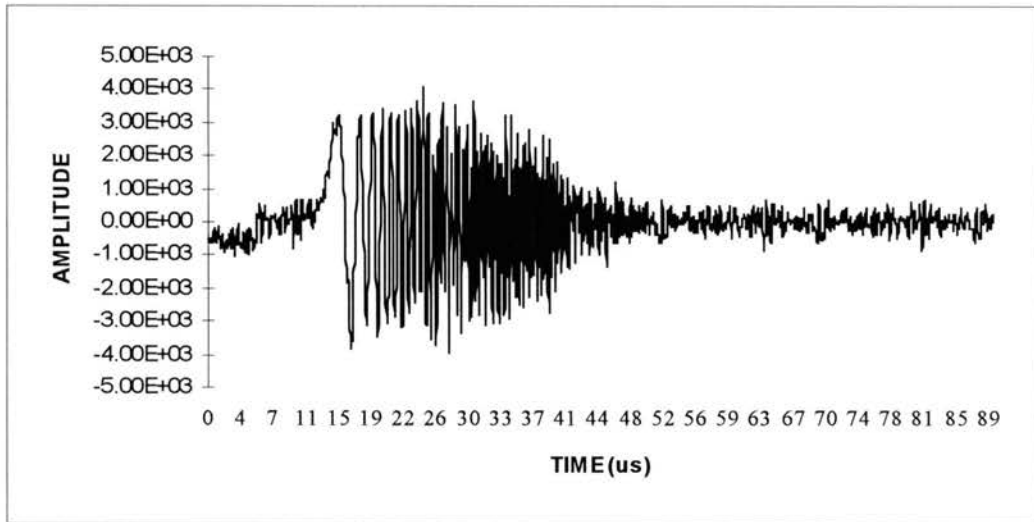


Figure 5-54. Received GCR signal for Case 3 (linear equalizer)

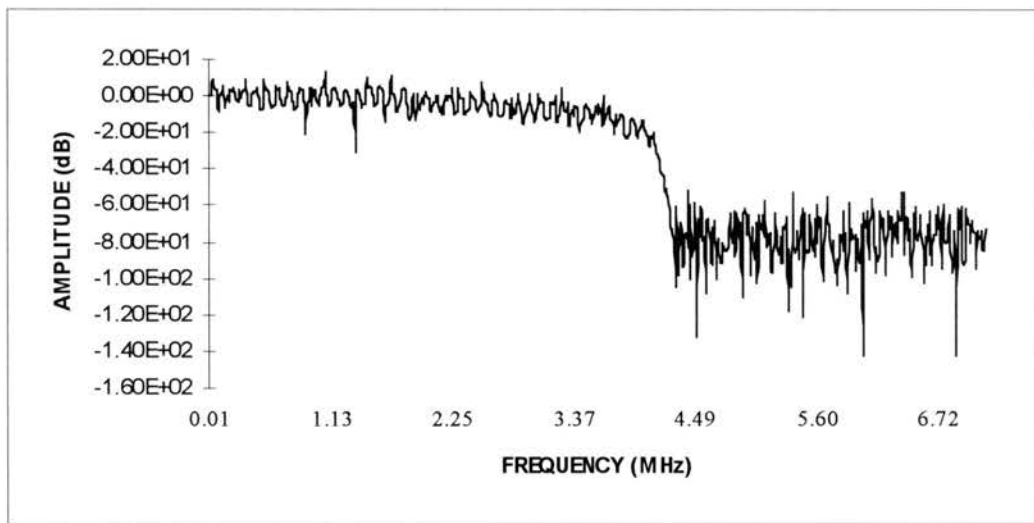


Figure 5-54a. Spectrum of the received GCR signal for Case 3 (linear equalizer)

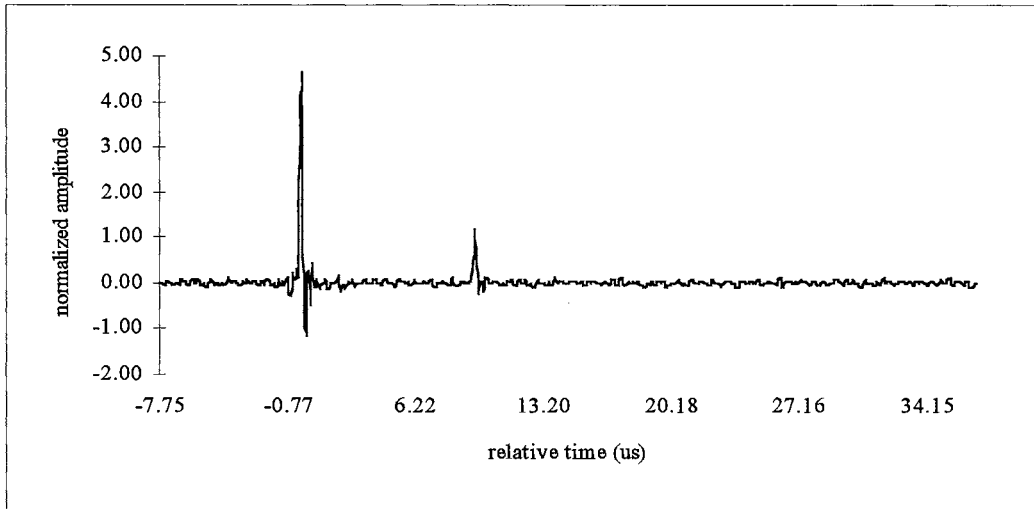


Figure 5-55. Estimated channel impulse response  $f(n)$  for Case 3 (linear)

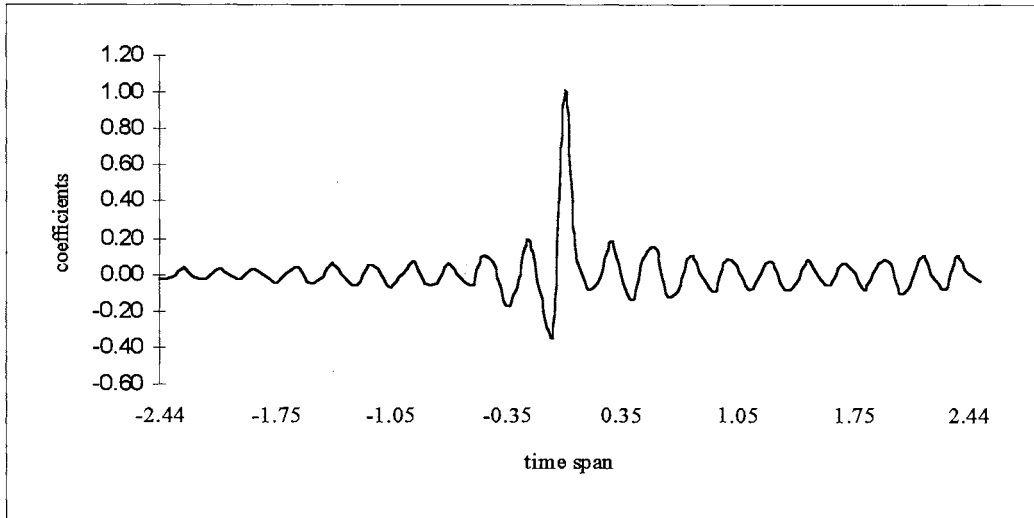


Figure 5-56. Coefficients of linear ZF equalizer for Case 3

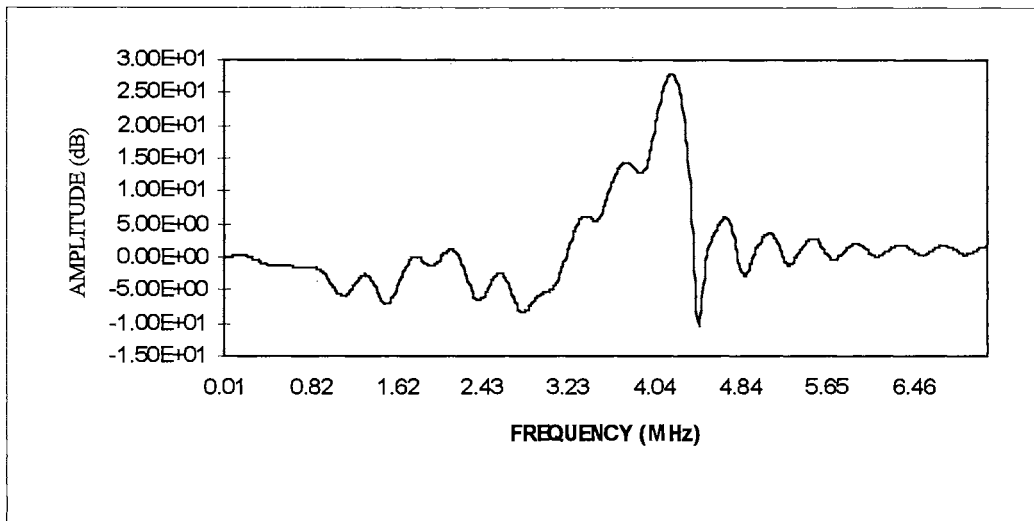


Figure 5-57. Frequency response of linear ZF equalizer for Case 3

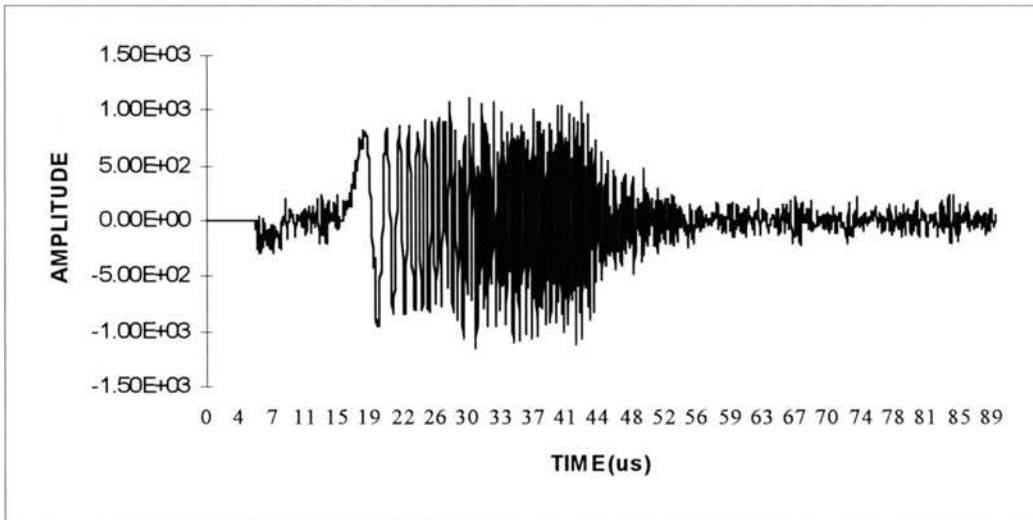


Figure 5-58. Output signal of the linear ZF equalizer for Case 3

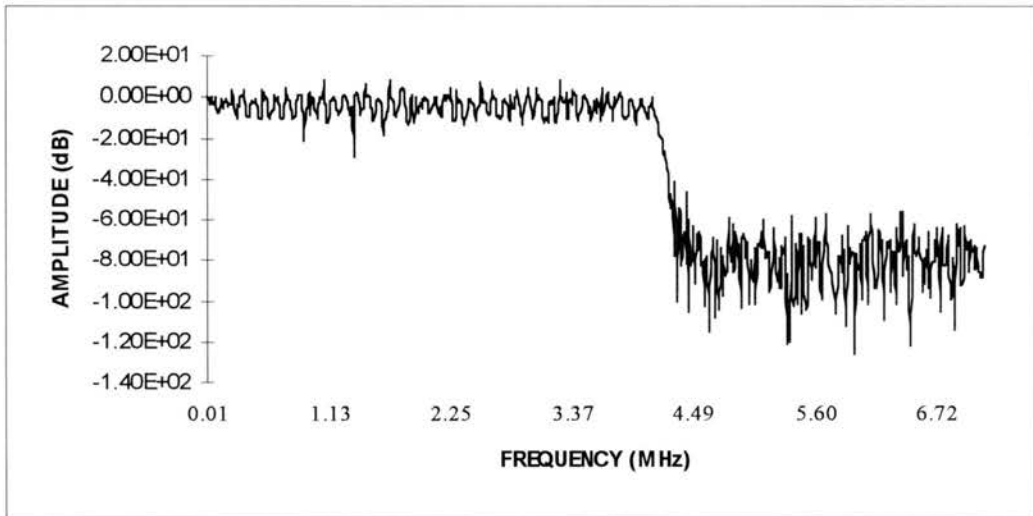


Figure 5-59. Output signal spectrum of the linear ZF equalizer for Case 3

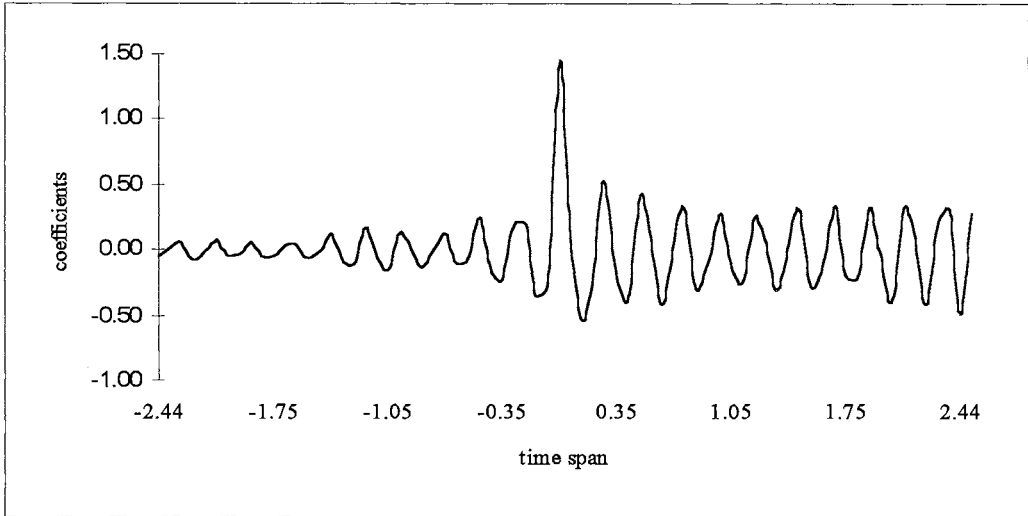


Figure 5-60. Coefficients of linear Lucky's ZF equalizer for Case 3

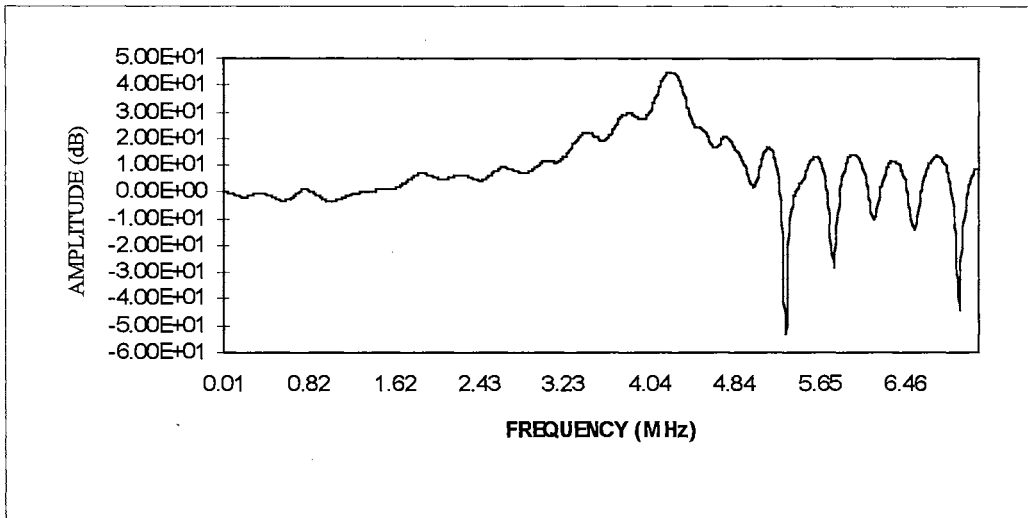


Figure 5-61. Frequency response of linear Lucky's ZF equalizer for Case 3

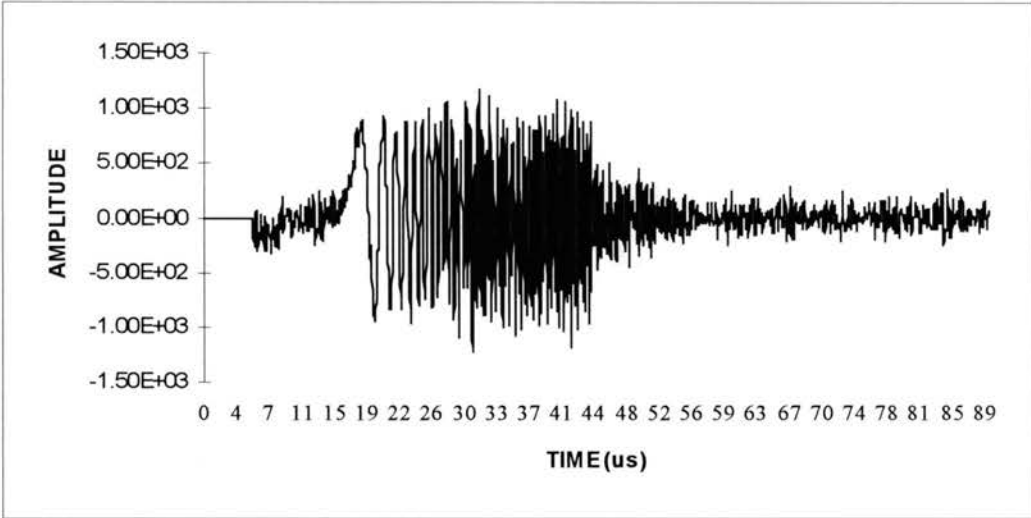


Figure 5-62. Output signal of the linear Lucky's ZF equalizer for Case 3

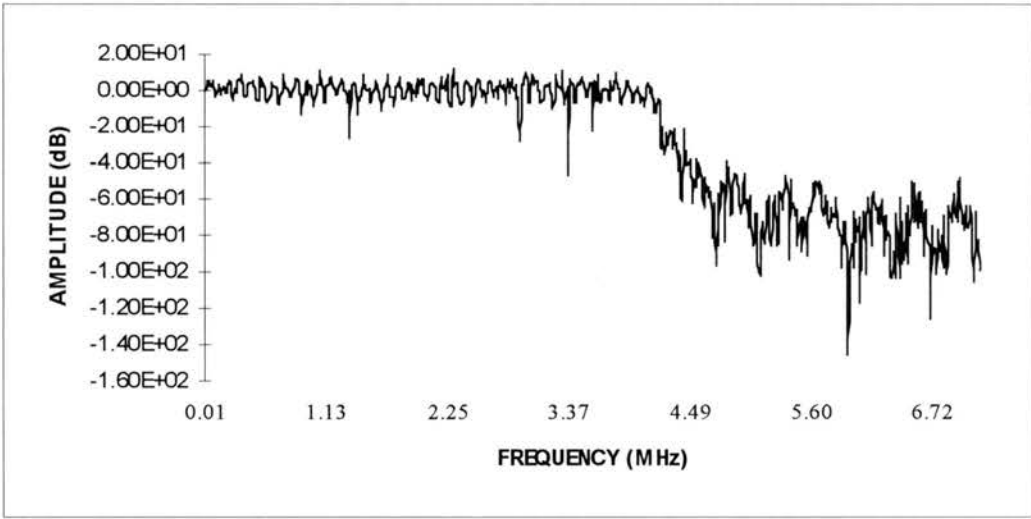


Figure 5-63. Output signal spectrum of the linear Lucky's ZF equalizer for Case 3

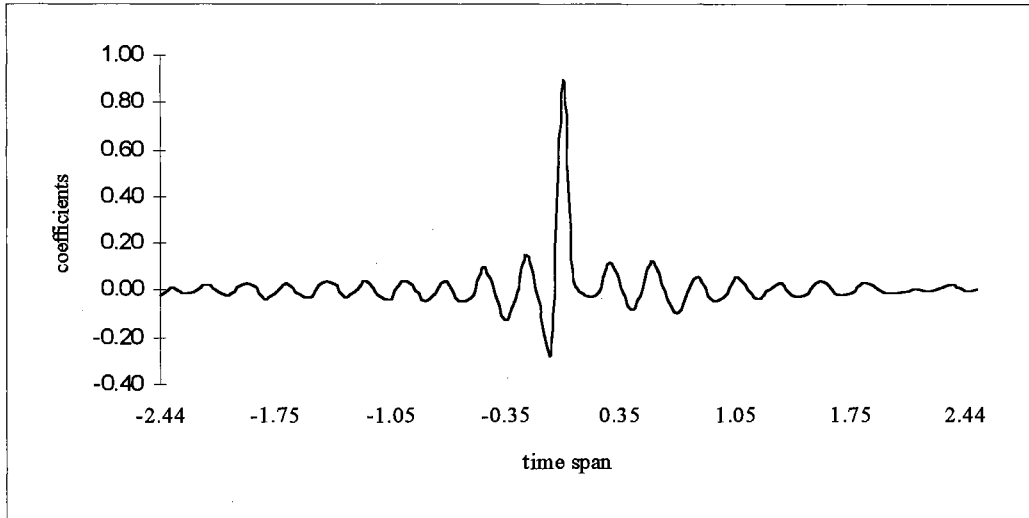


Figure 5-64. Coefficients of linear MSE (LMS) equalizer for Case 3

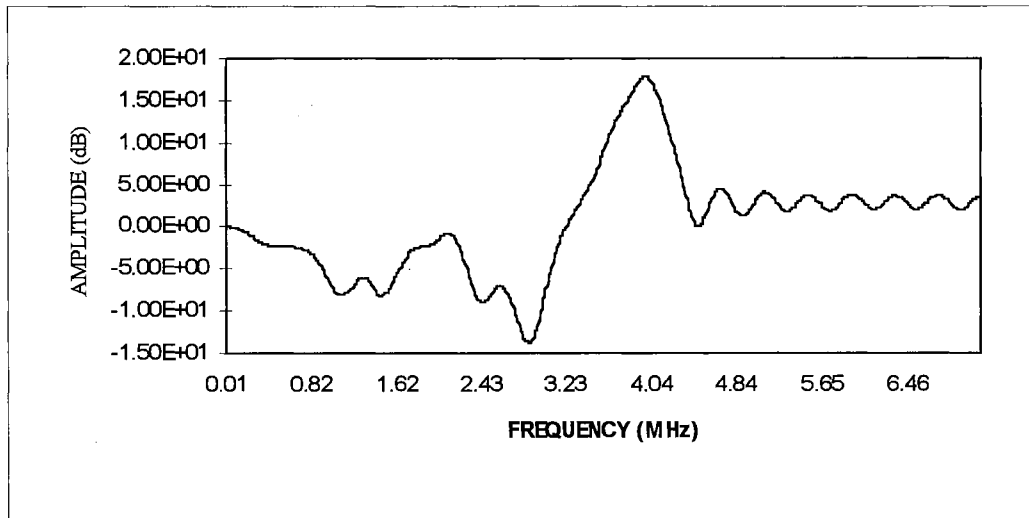


Figure 5-65. Frequency response of linear MSE (LMS) equalizer for Case 3



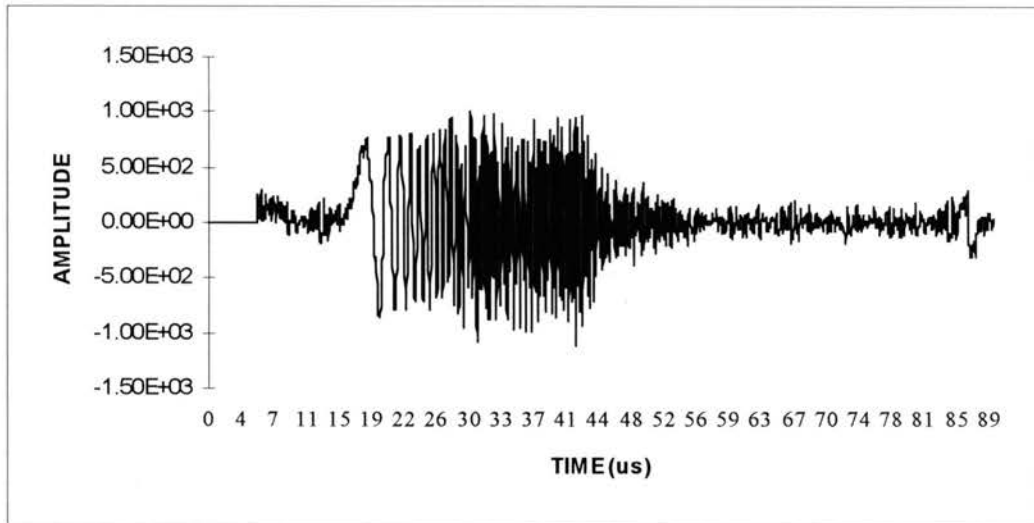


Figure 5-66. Output signal of the linear MSE (LMS) equalizer for Case 3

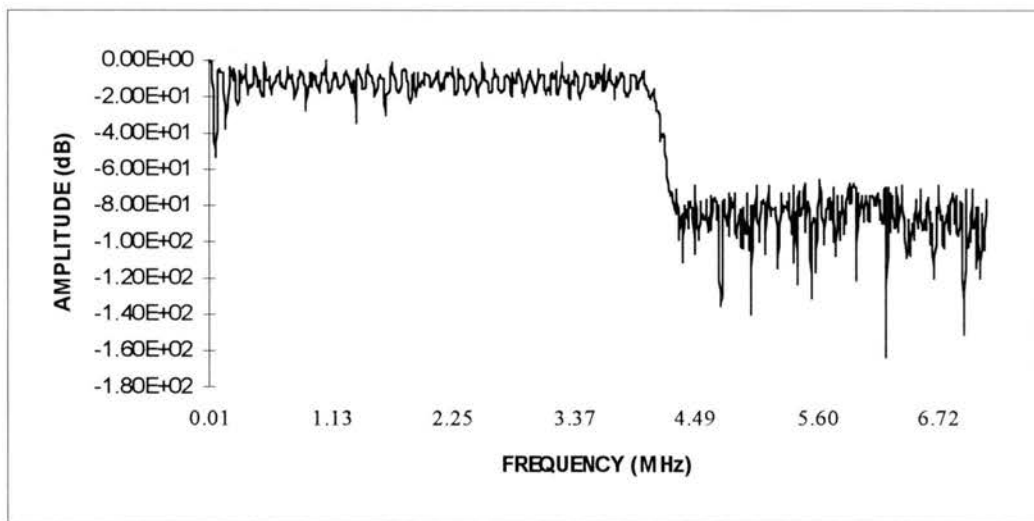


Figure 5-67. Output signal spectrum of the linear MSE (LMS) equalizer for Case 3

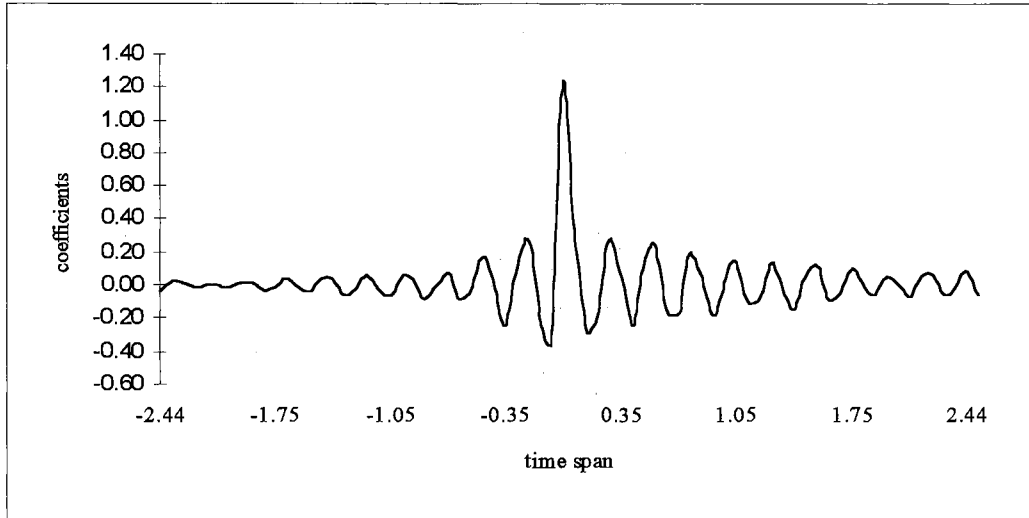


Figure 5-68. Coefficients of linear MN ( $\lambda=0.5$ ) equalizer for Case 3

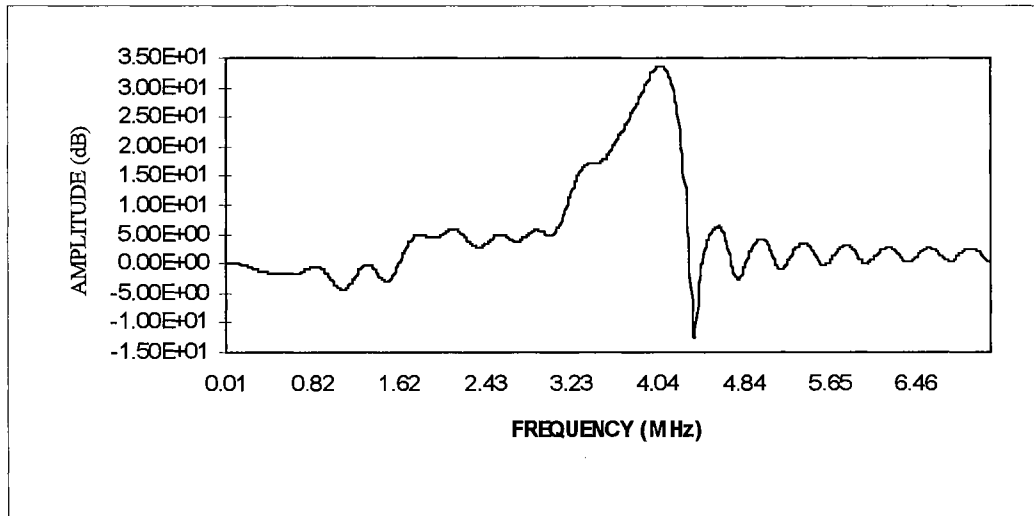


Figure 5-69. Frequency response of linear MN ( $\lambda=0.5$ ) equalizer for Case 3

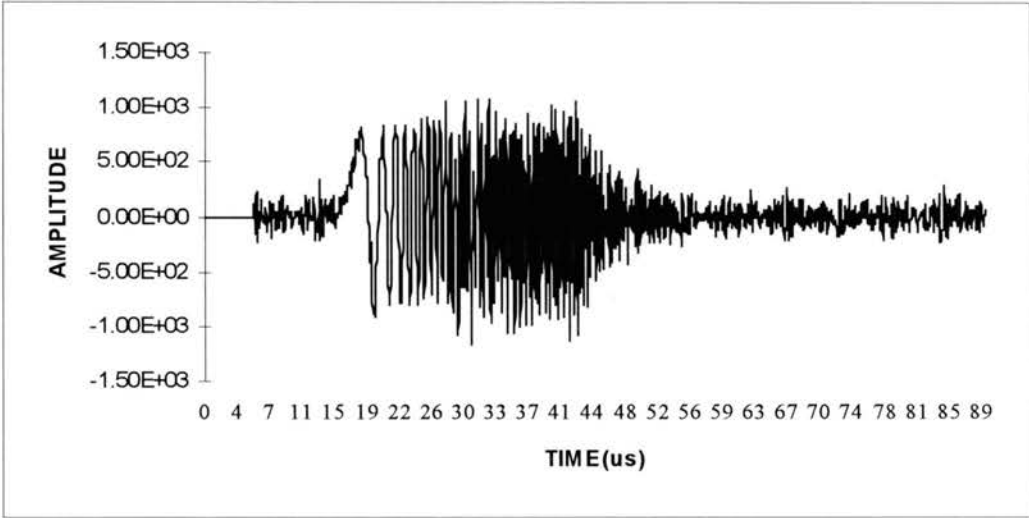


Figure 5-70. Output signal of the linear mixed norm equalizer for Case 3

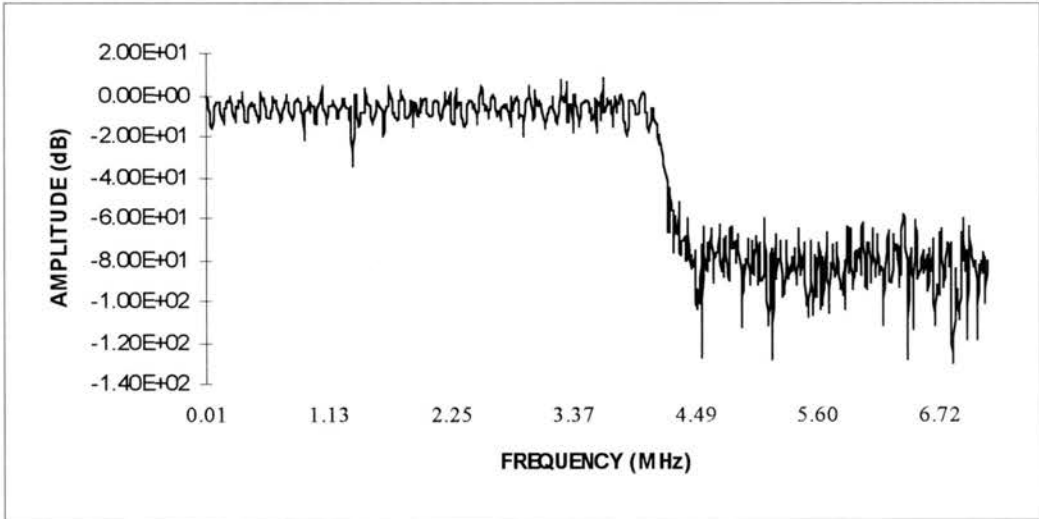


Figure 5-71. Output signal spectrum of the linear mixed norm equalizer for Case 3

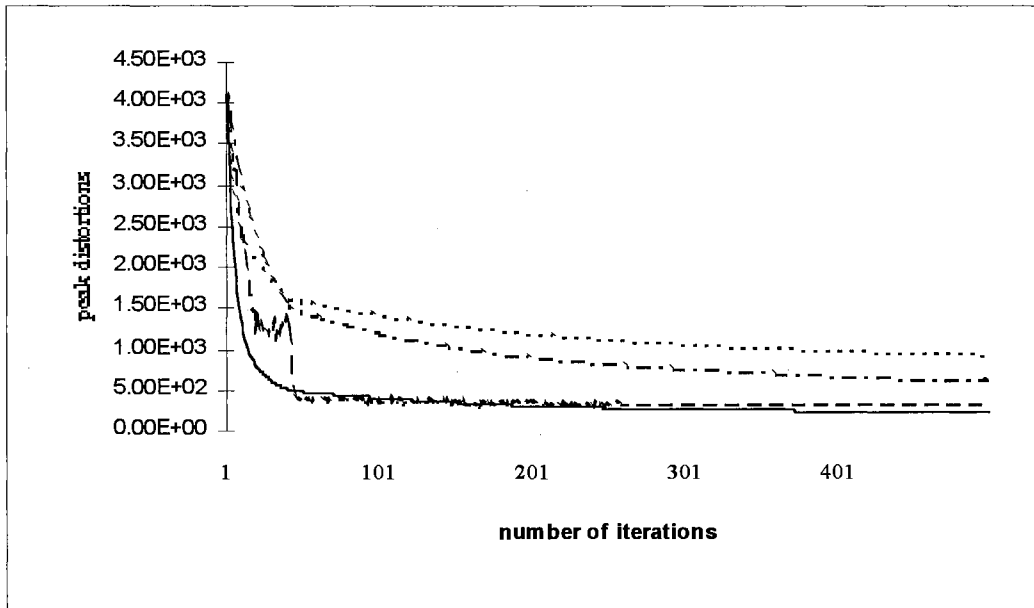


Figure 5-72. Peak distortions of linear equalizers for Case 3  
 ( ZF: ——— Lucky's ZF: - - - - - MSE: ..... MN: - . - . - )

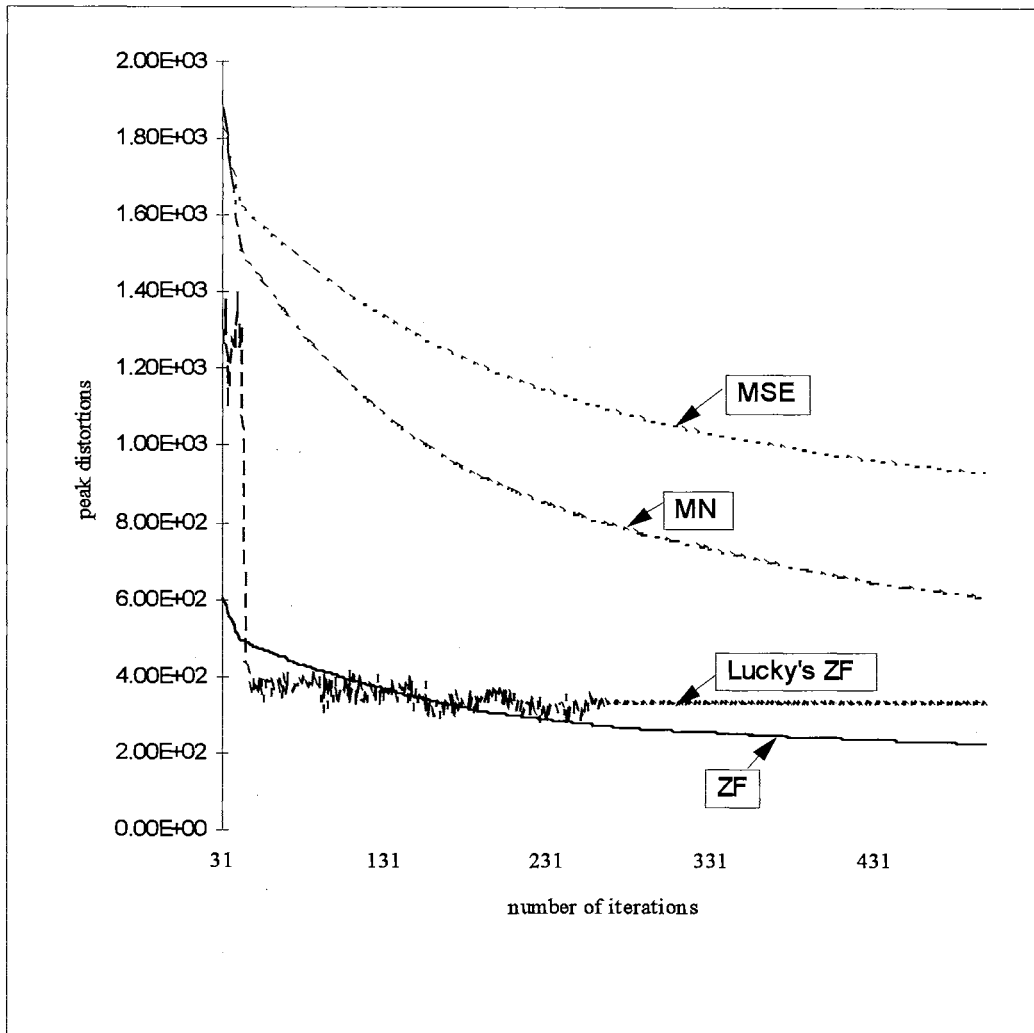


Figure 5-73. Detailed peak distortions of linear equalizers for Case 3  
 ( ZF: ——— Lucky's ZF: ..... MSE: -.-.- MN: -.-.- )

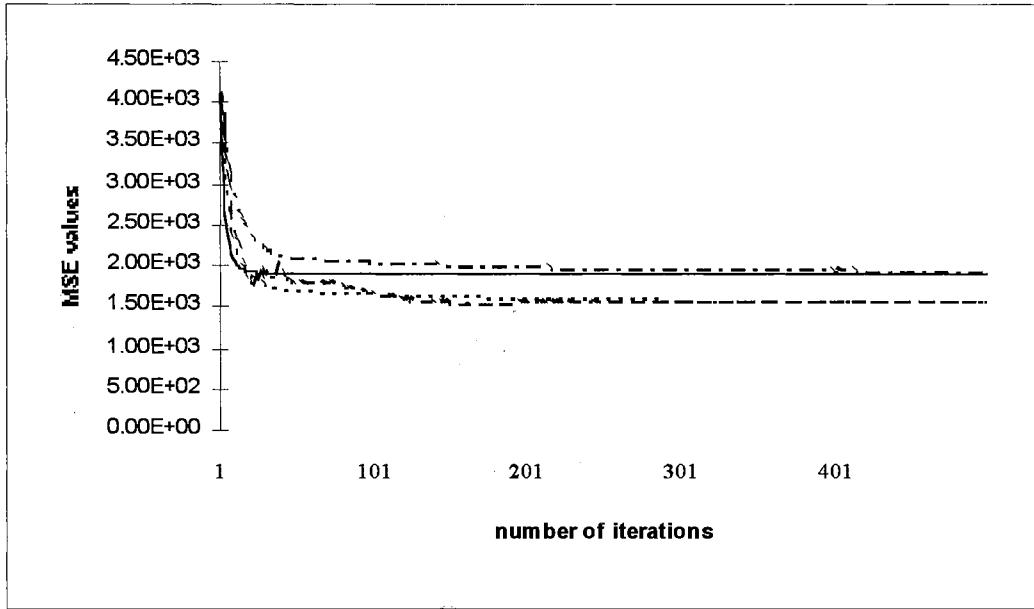


Figure 5-74. MSE of linear equalizers for Case 3

( ZF: ——— Lucky's ZF: - - - - - MSE: . . . . . MN: - . - . - )

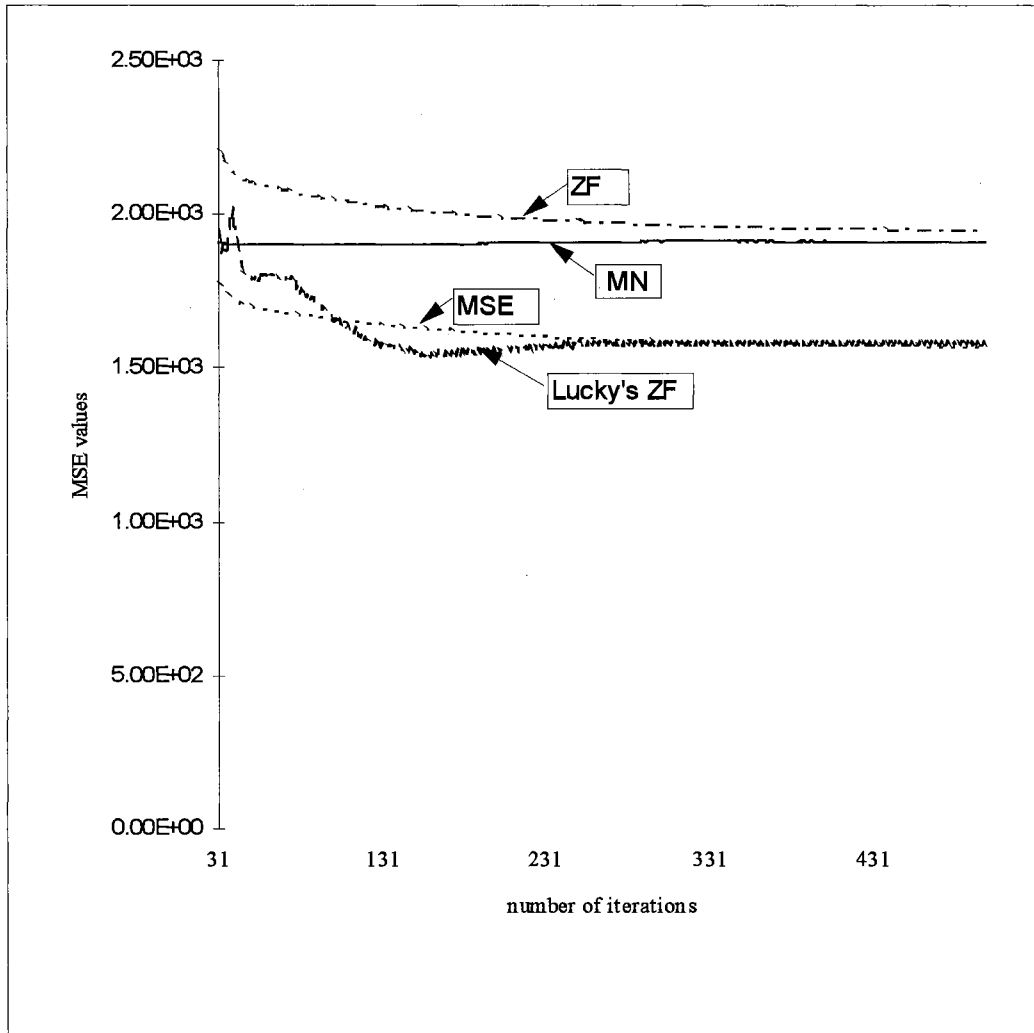


Figure 5-75. Detailed MSE of linear equalizers for Case 3  
 ( ZF: ----- Lucky's ZF: -.-.-.-.- MSE: ..... MN: \_\_\_\_\_ )

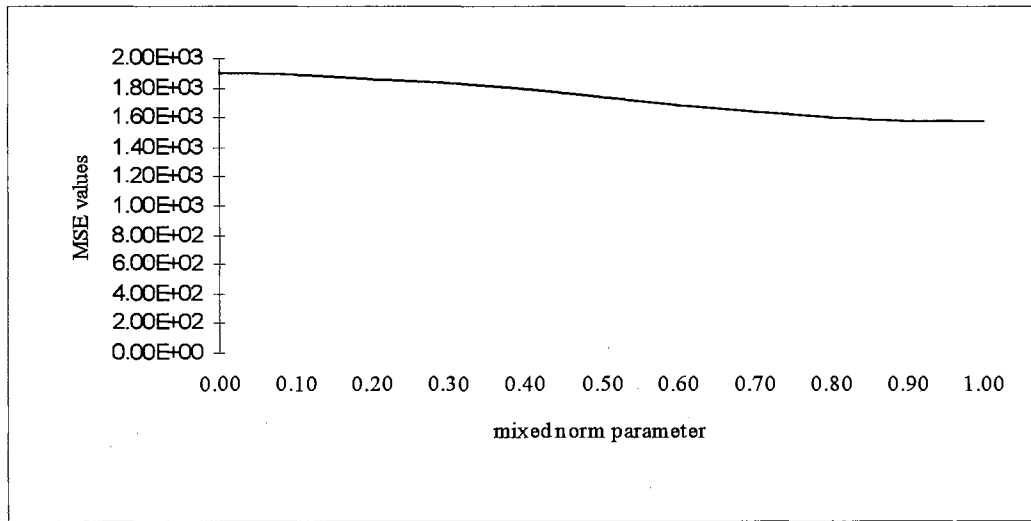


Figure 5-76. MSE values versus mixed norm parameter  $\lambda$  for Case 3 (linear)



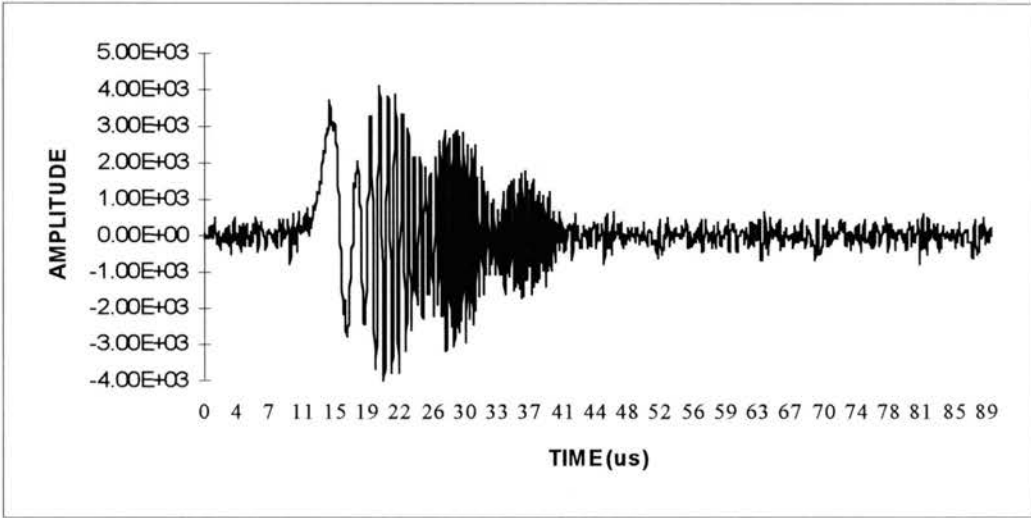


Figure 5-77. Received GCR signal for Case 1 (feedback equalizer)

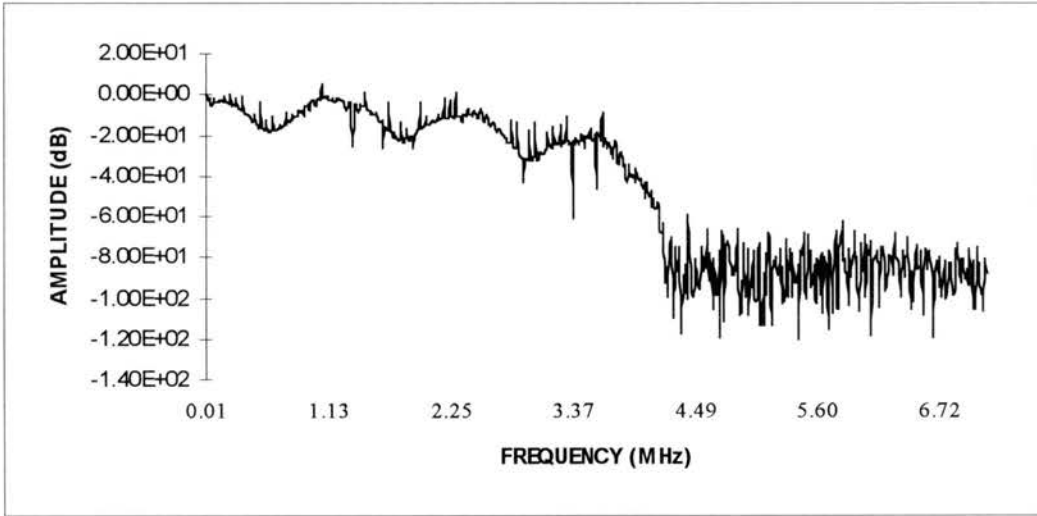


Figure 5-77a. Spectrum of the received GCR signal for Case 1 (feedback equalizer)

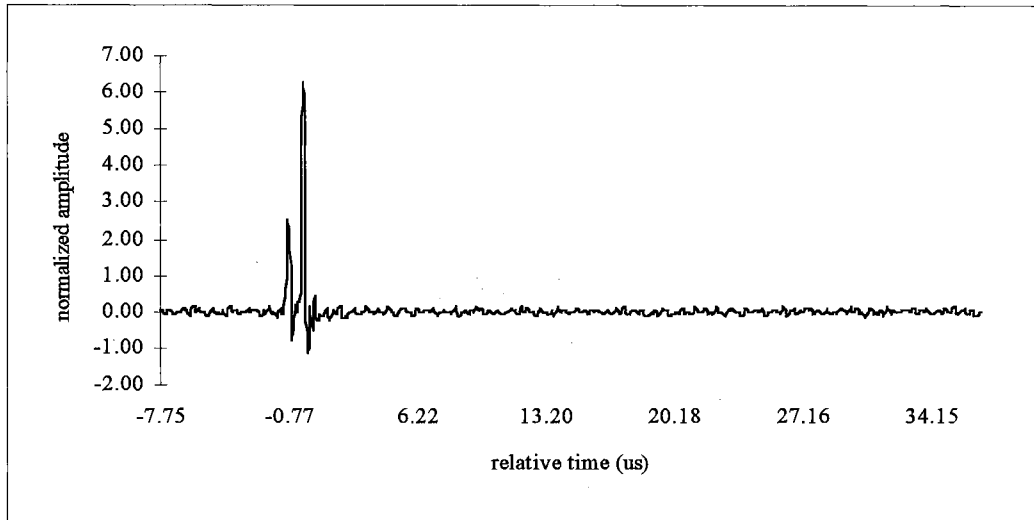


Figure 5-78. Estimated channel impulse response  $f(n)$  for Case 1 (feedback)

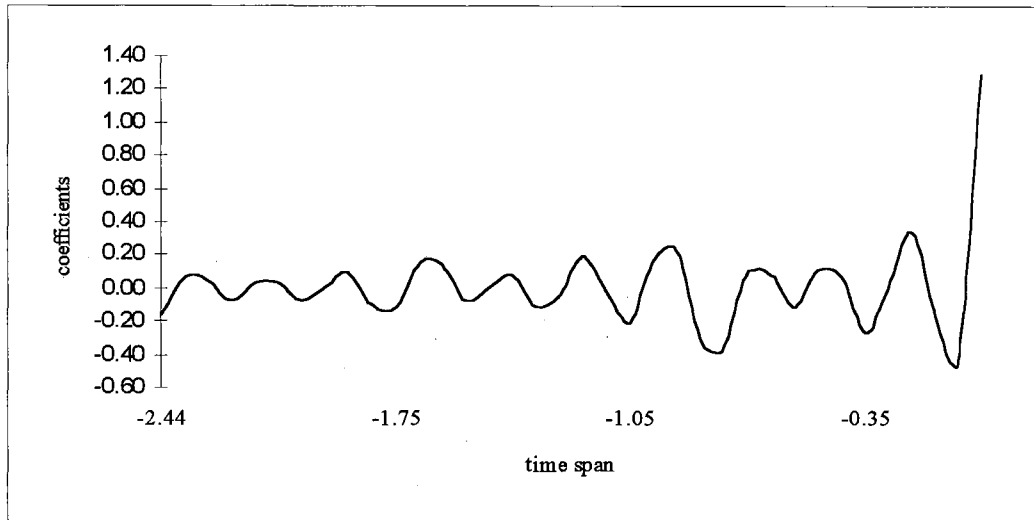


Figure 5-79. Coefficients of forward filter  $h(n)$  of ZF feedback equalizer for Case 1

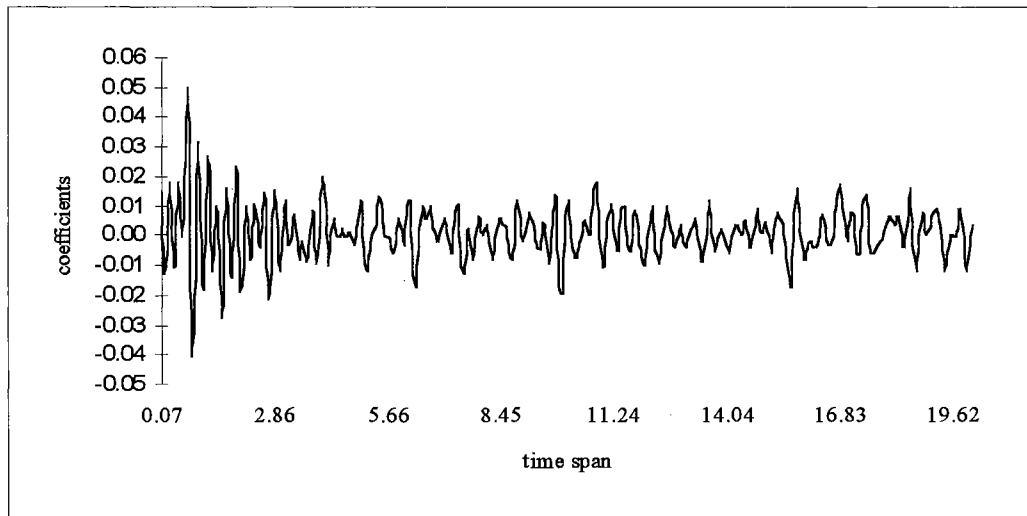


Figure 5-80. Coefficients of feedback filter  $g(n)$  of ZF feedback equalizer for Case 1

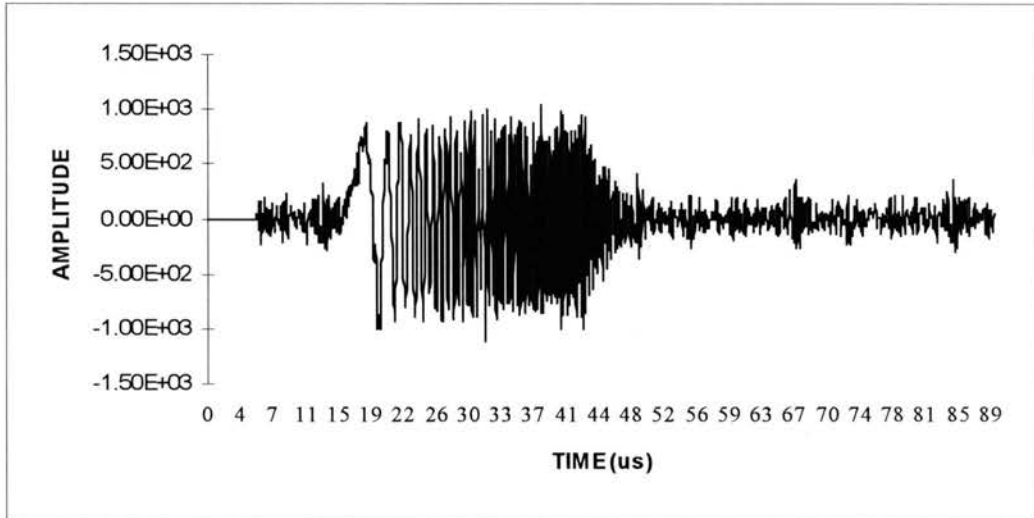


Figure 5-81. Output signal of the forward filter of the ZF feedback equalizer for Case 1

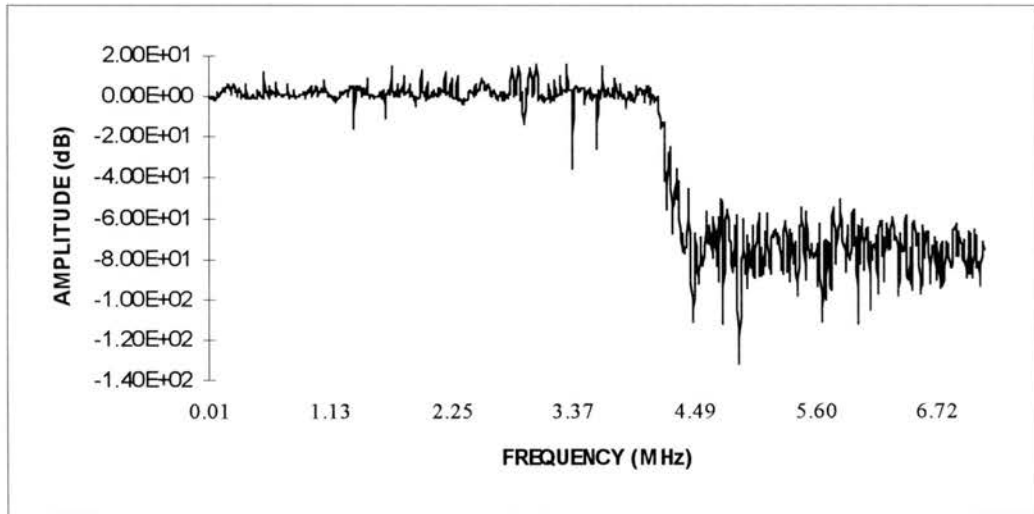


Figure 5-82. Spectrum of the forward filter output of the ZF feedback equalizer for Case 1

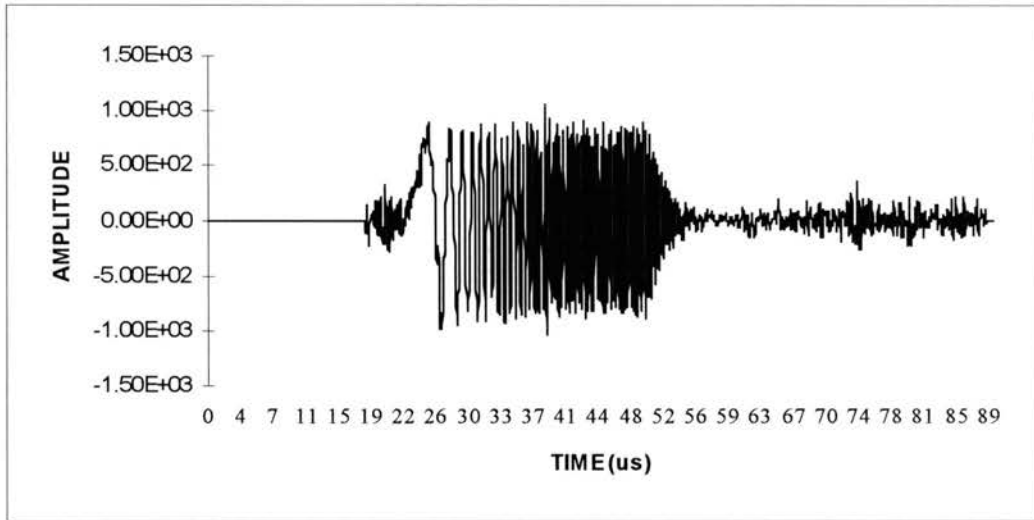


Figure 5-83. Output signal of the ZF feedback equalizer for Case 1

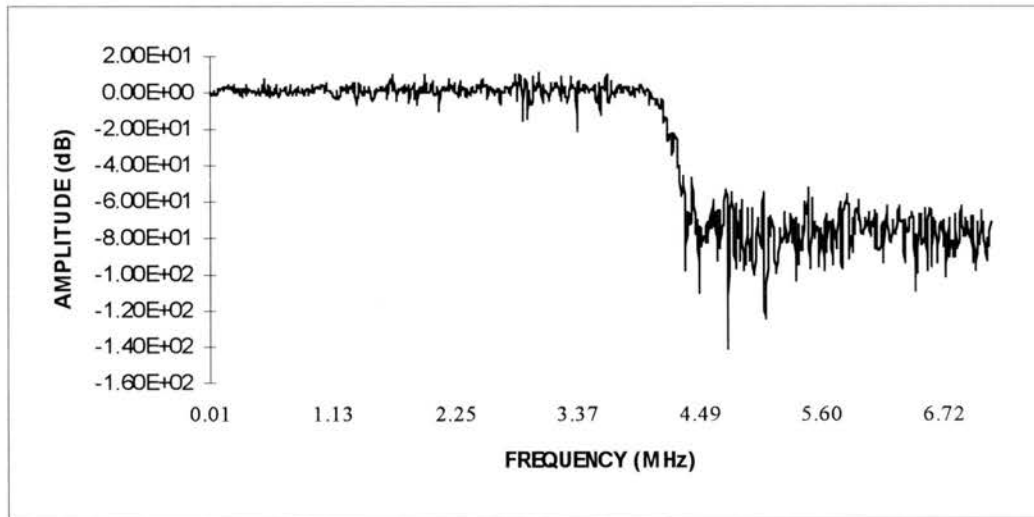


Figure 5-84. Spectrum of the output signal of the ZF feedback equalizer for Case 1

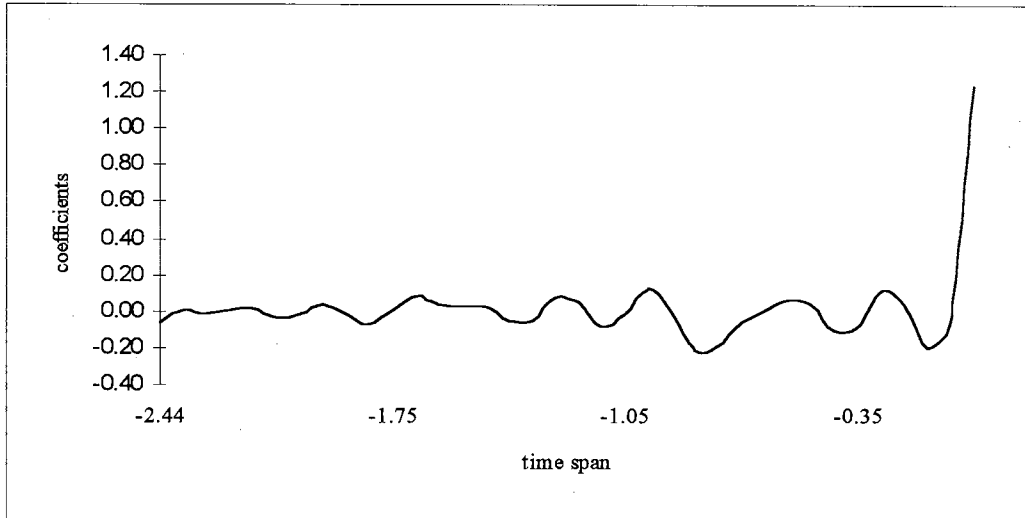


Figure 5-85. Coefficients of forward filter  $h(n)$  of MSE(LMS) feedback equalizer for Case 1

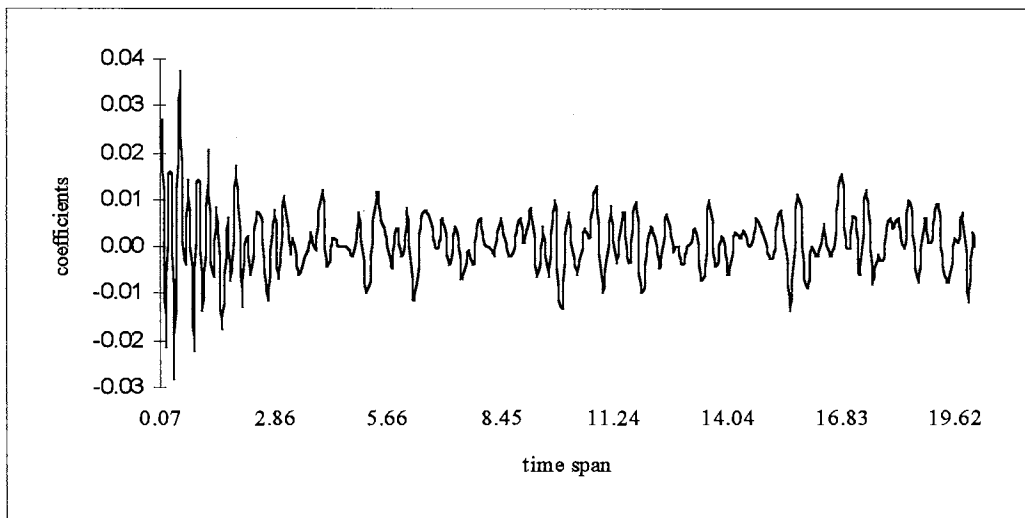


Figure 5-86. Coefficients of feedback filter  $g(n)$  of MSE(LMS) feedback equalizer for Case 1

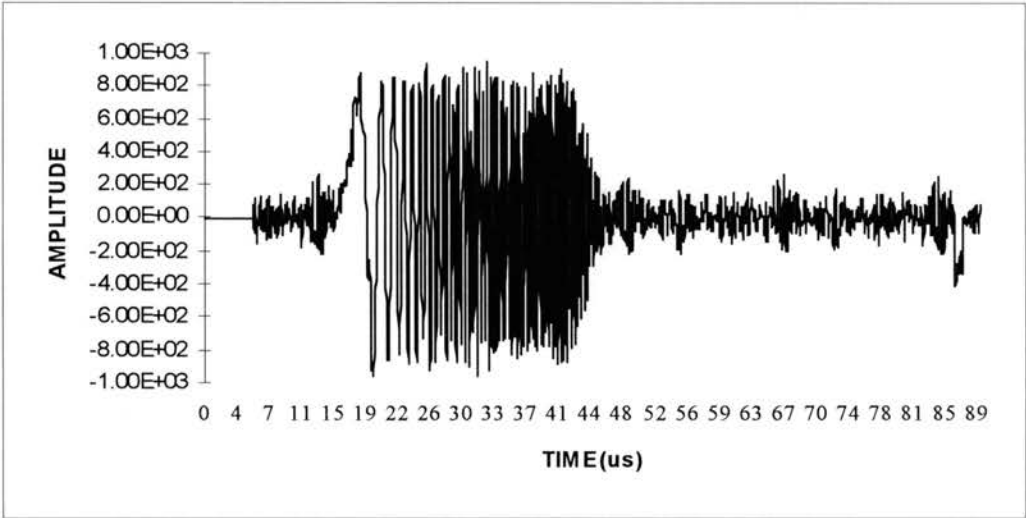


Figure 5-87. Output signal of the forward filter of the MSE(LMS) feedback equalizer for Case 1

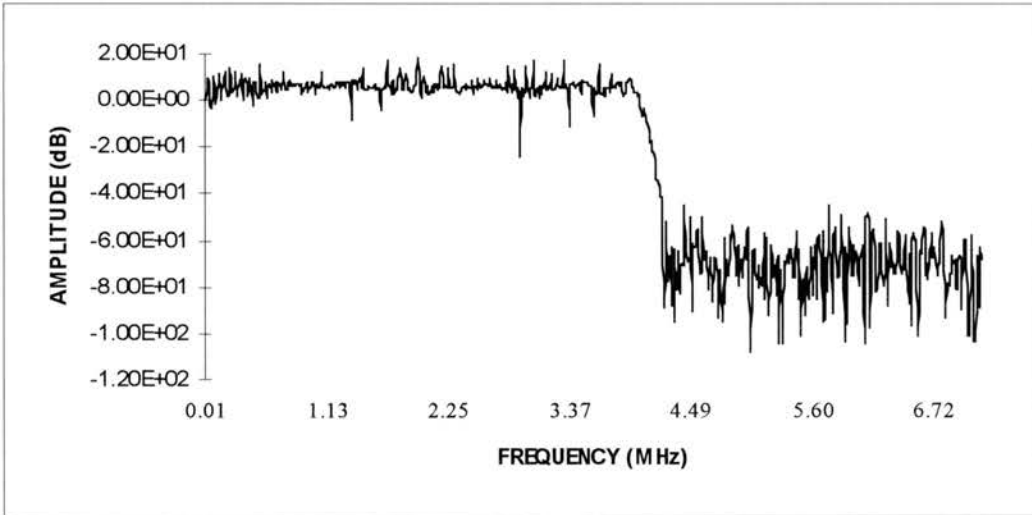


Figure 5-88. Spectrum of the forward filter output of the MSE(LMS) feedback equalizer for Case 1

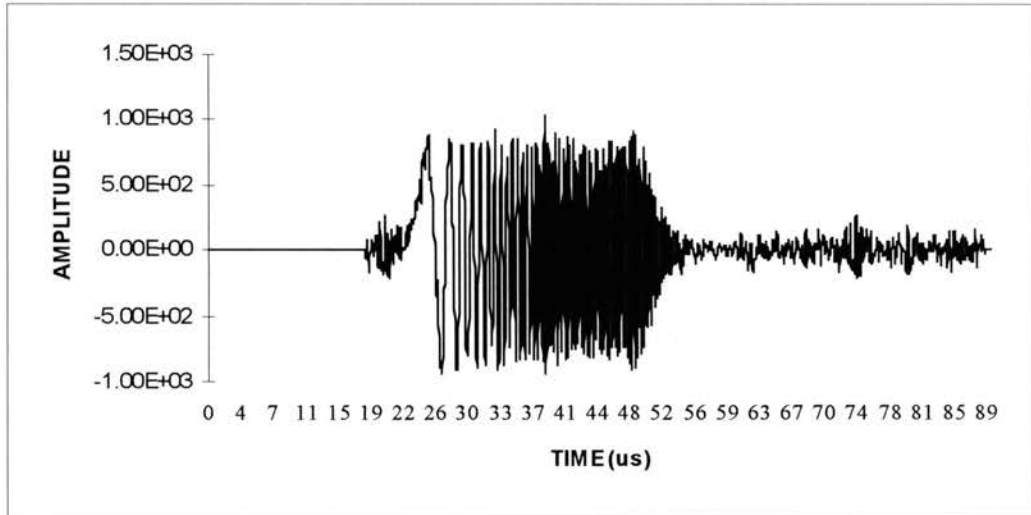


Figure 5-89. Output signal of the MSE(LMS) feedback equalizer for Case 1

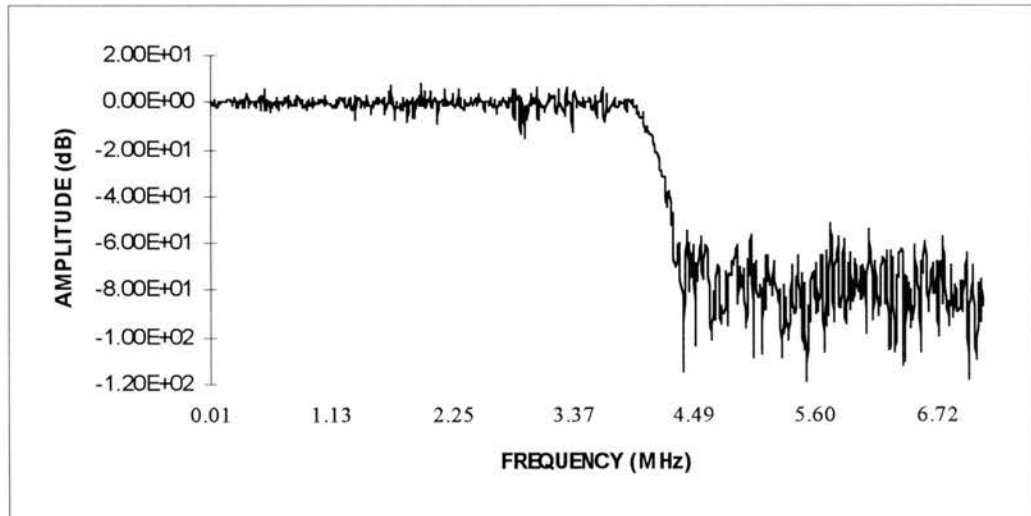


Figure 5-90. Spectrum of the output signal of the MSE(LMS) feedback equalizer for Case 1



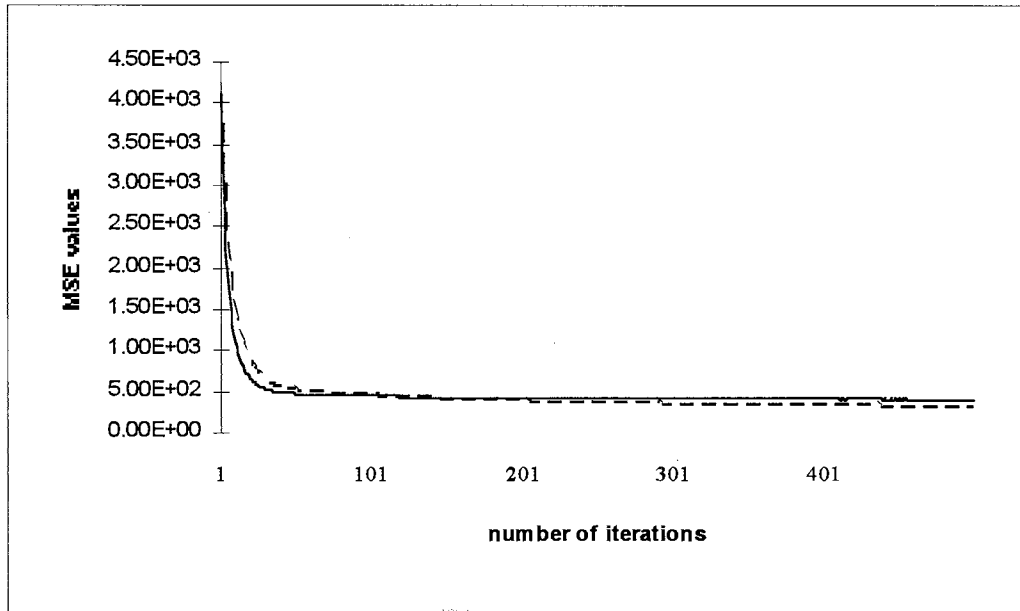


Figure 5-91. MSE of feedback equalizers for Case 1  
( ZF: — MSE: - - - )

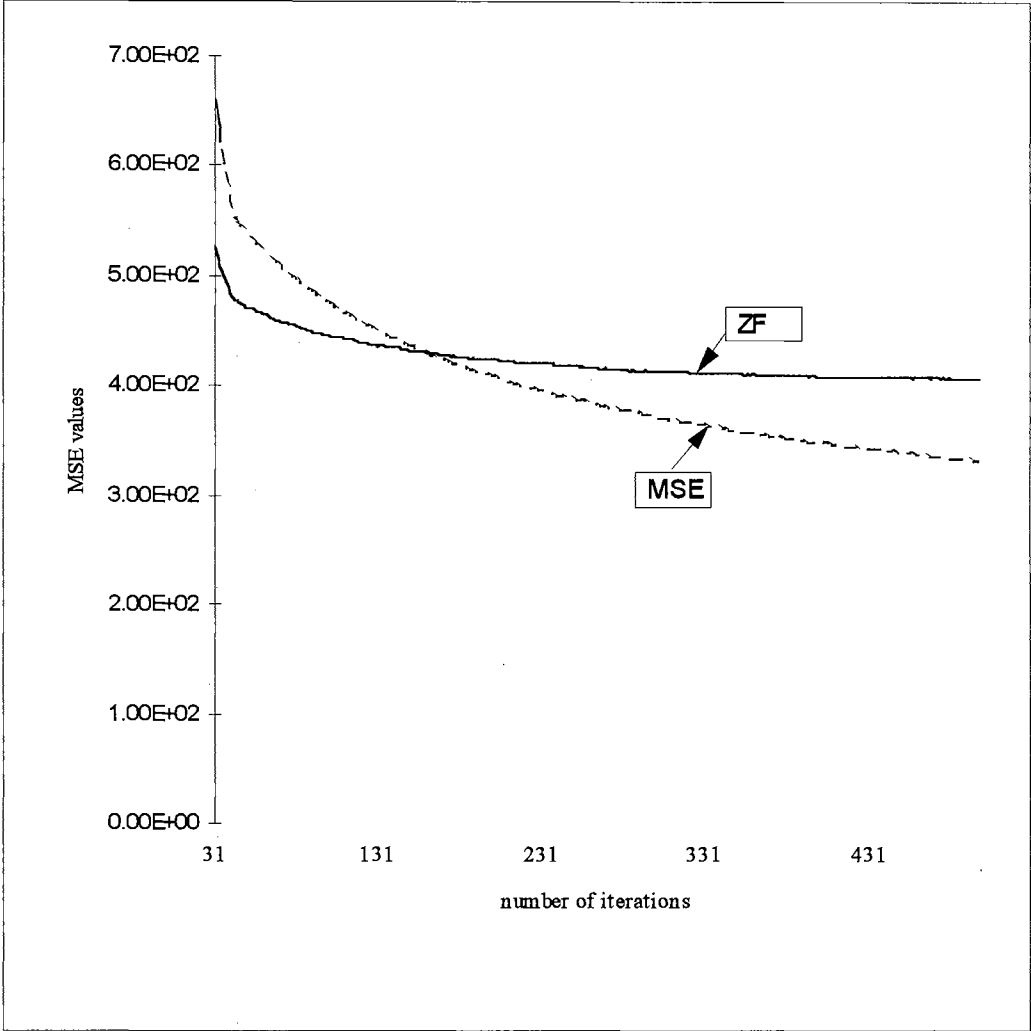


Figure 5-92. Detailed MSE of feedback equalizers for Case 1  
 ( ZF: ——— MSE: - - - - )

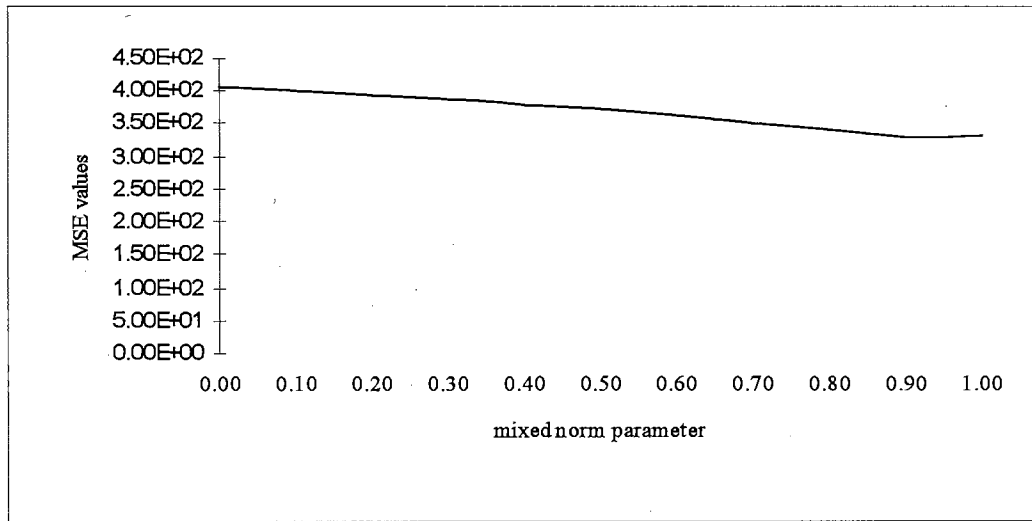


Figure 5-93. MSE values versus mixed norm parameter  $\lambda$  for Case 1 (feedback)

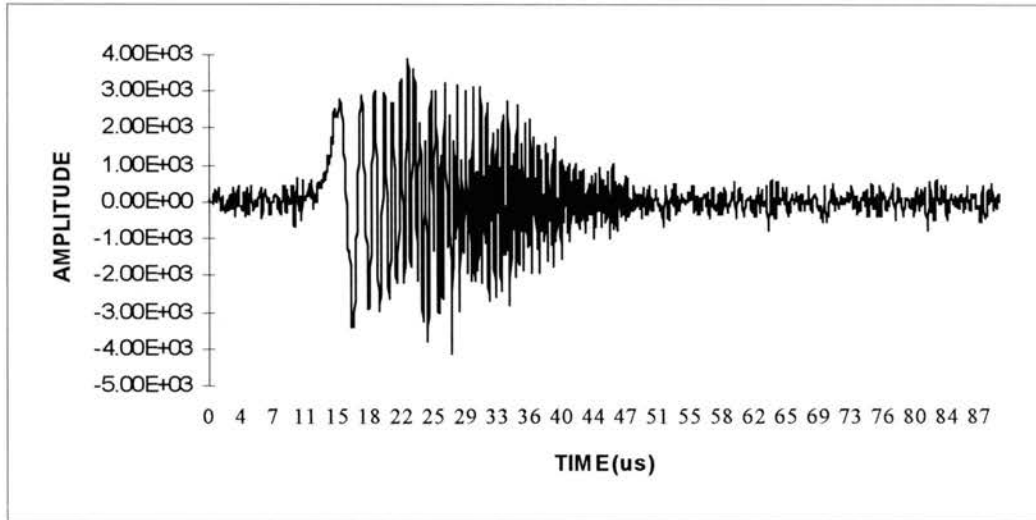


Figure 5-94. Received GCR signal for Case 2 (feedback equalizer)

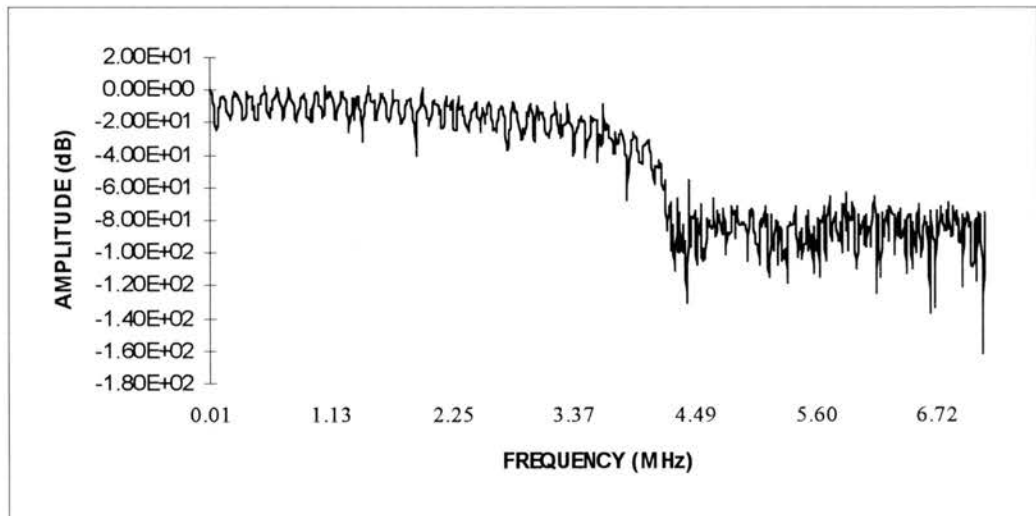


Figure 5-94a. Spectrum of the received GCR signal for Case 2 (feedback equalizer)

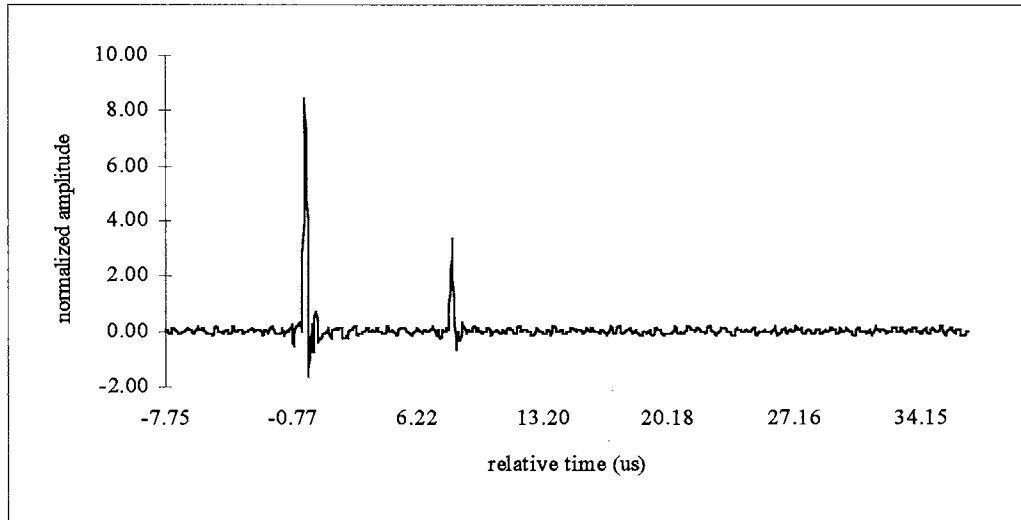


Figure 5-95. Estimated channel impulse response  $f(n)$  for Case 2 (feedback)

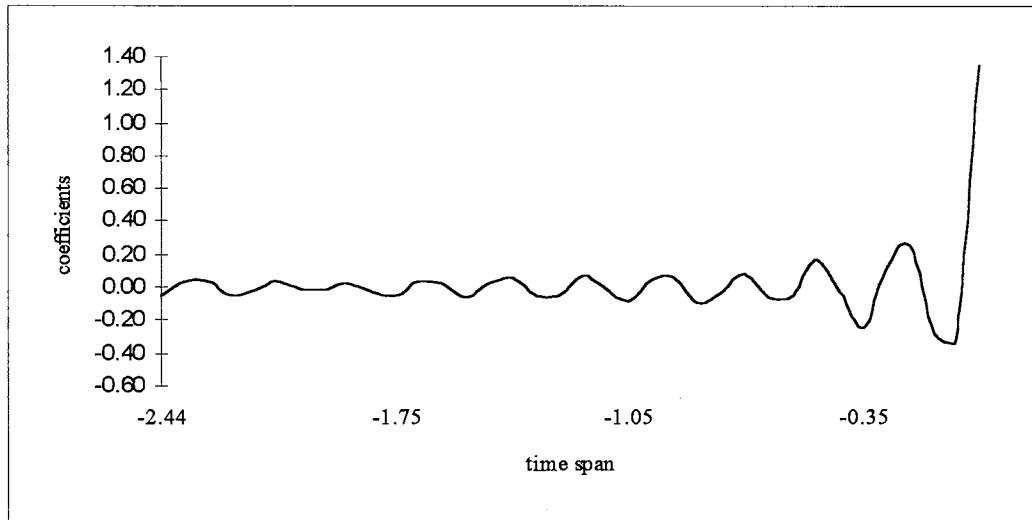


Figure 5-96. Coefficients of forward filter  $h(n)$  of ZF feedback equalizer for Case 2

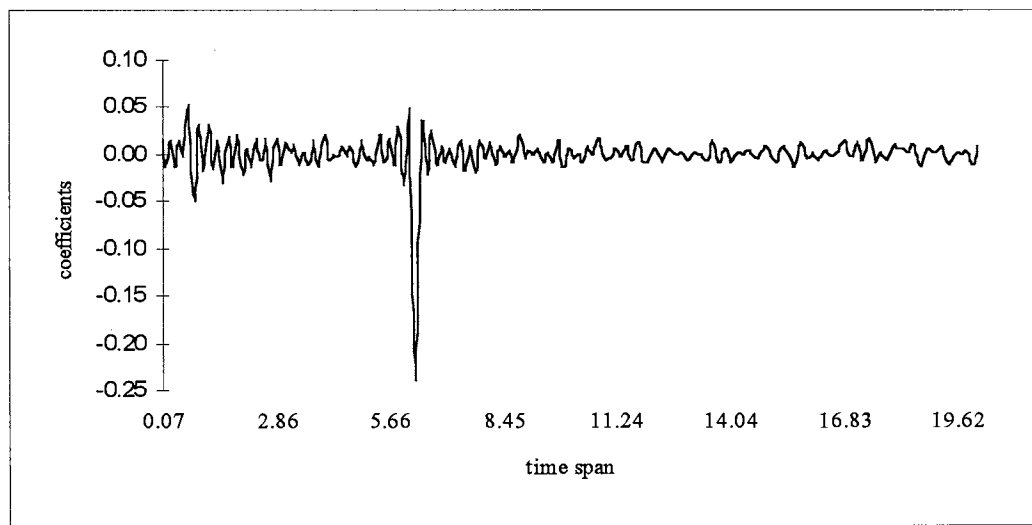


Figure 5-97. Coefficients of feedback filter  $g(n)$  of ZF feedback equalizer for Case 2

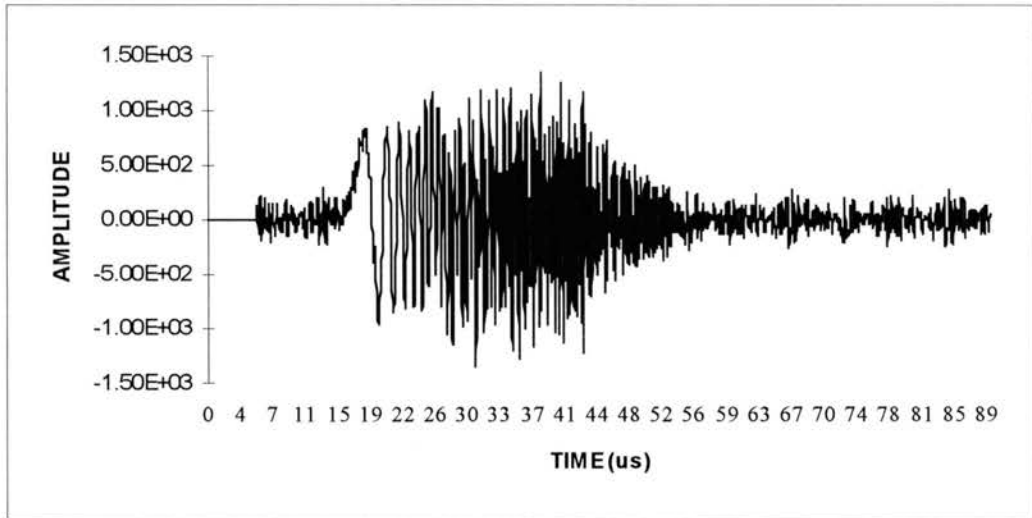


Figure 5-98. Output signal of the forward filter of the ZF feedback equalizer for Case 2

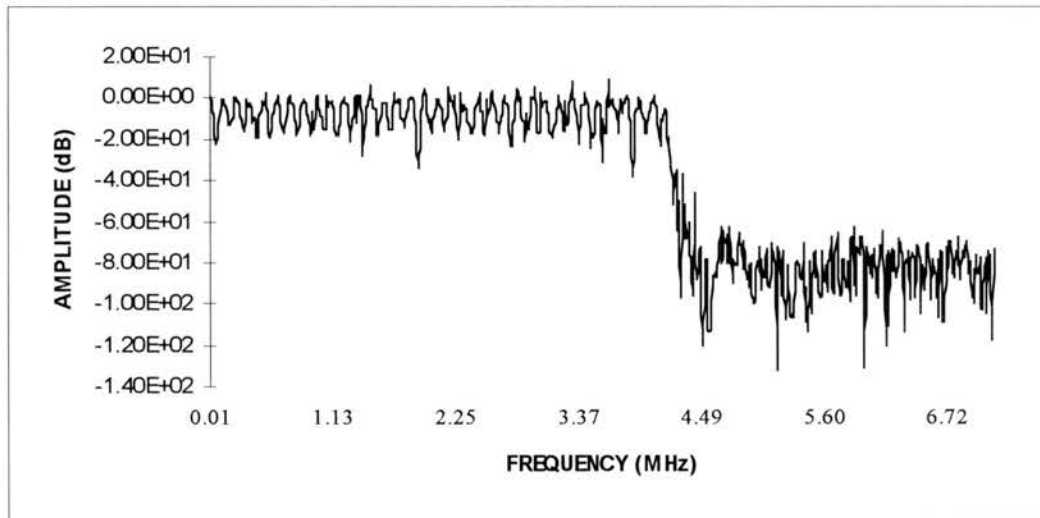


Figure 5-99. Spectrum of the forward filter output of the ZF feedback equalizer for Case 2

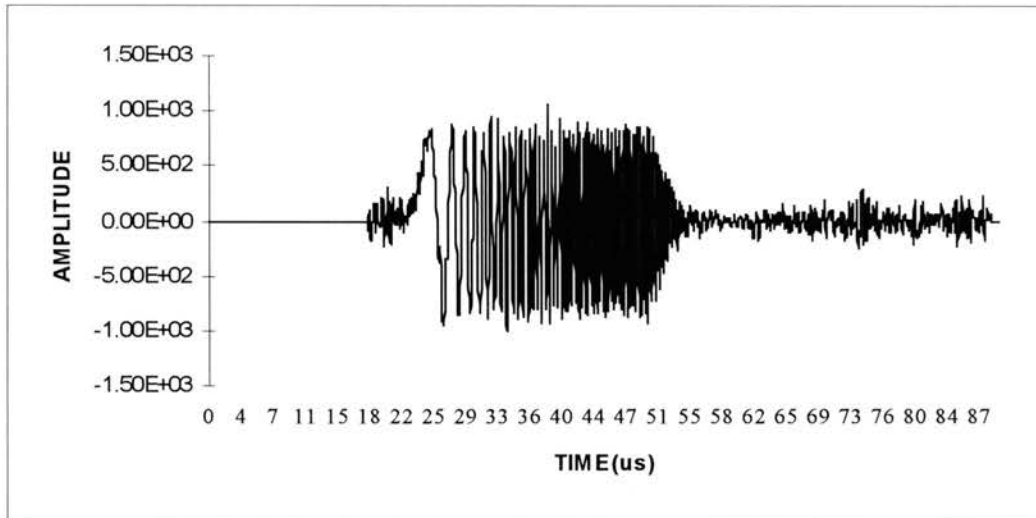


Figure 5-100. Output signal of the ZF feedback equalizer for Case 2

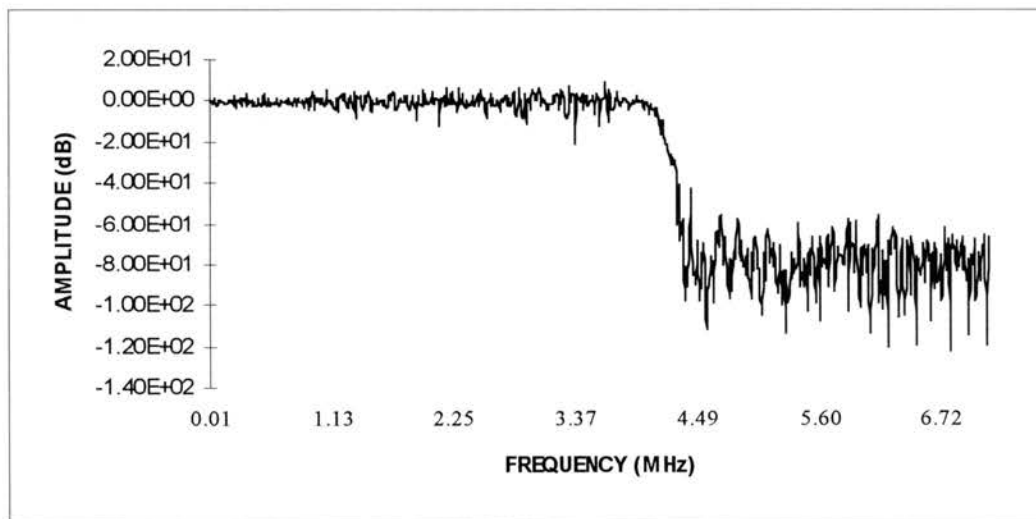


Figure 5-101. Spectrum of the output signal of the ZF feedback equalizer for Case 2



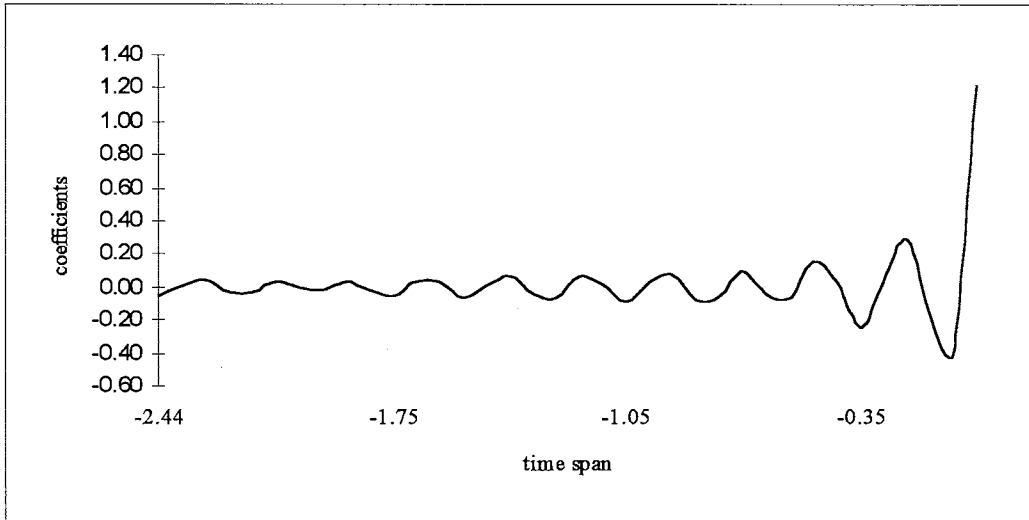


Figure 5-102. Coefficients of forward filter  $h(n)$  of MSE(LMS) feedback equalizer for Case 2

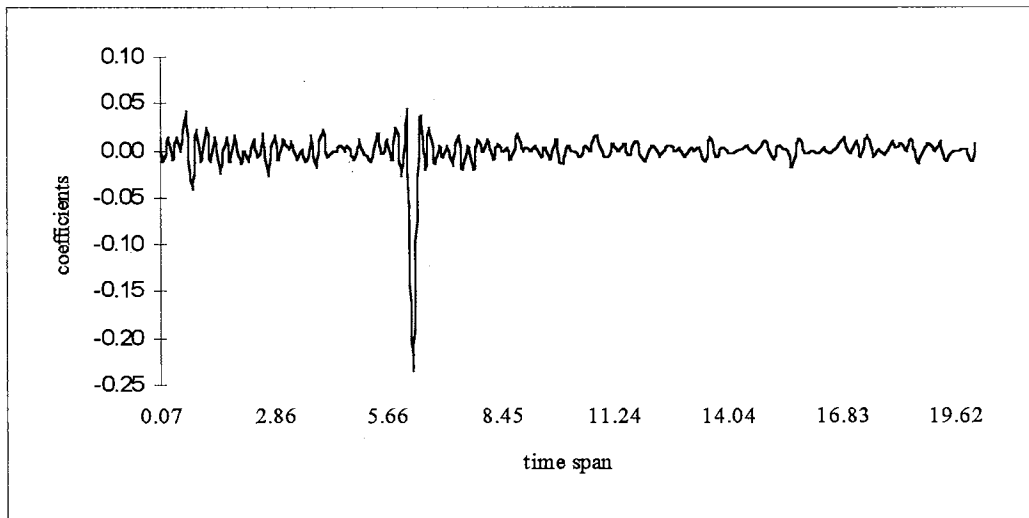


Figure 5-103. Coefficients of feedback filter  $g(n)$  of MSE(LMS) feedback equalizer for Case 2

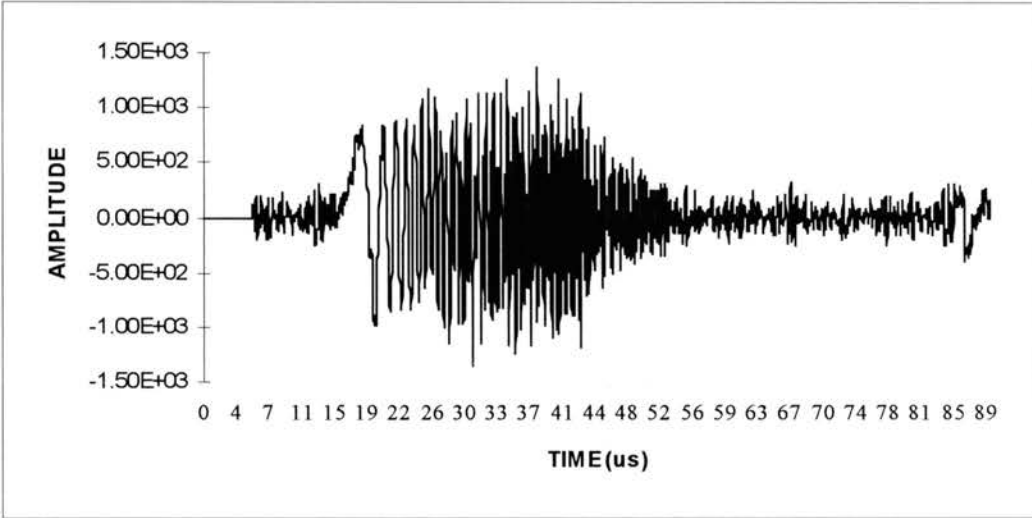


Figure 5-104. Output signal of the forward filter of the MSE(LMS) feedback equalizer for Case 2

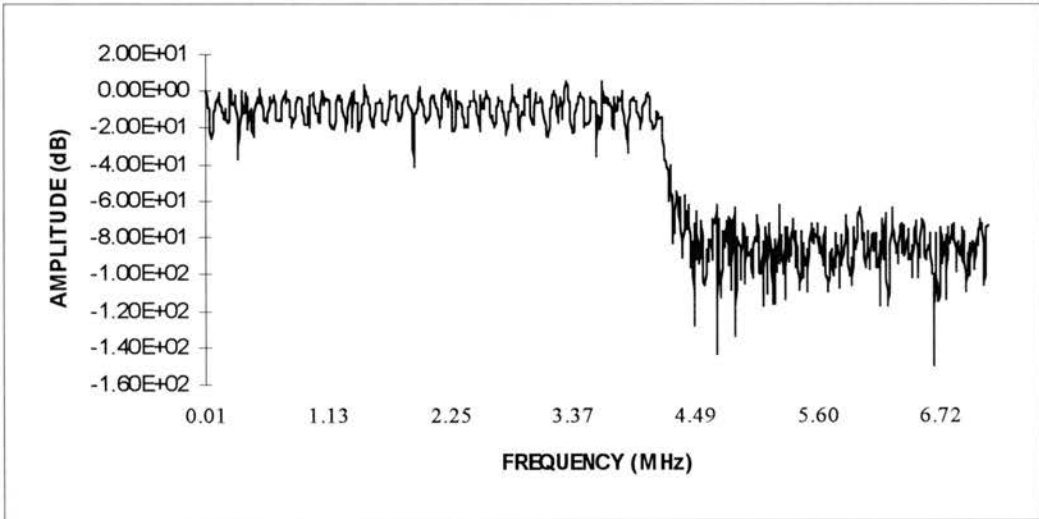


Figure 5-105. Spectrum of the forward filter output of the MSE(LMS) feedback equalizer for Case 2

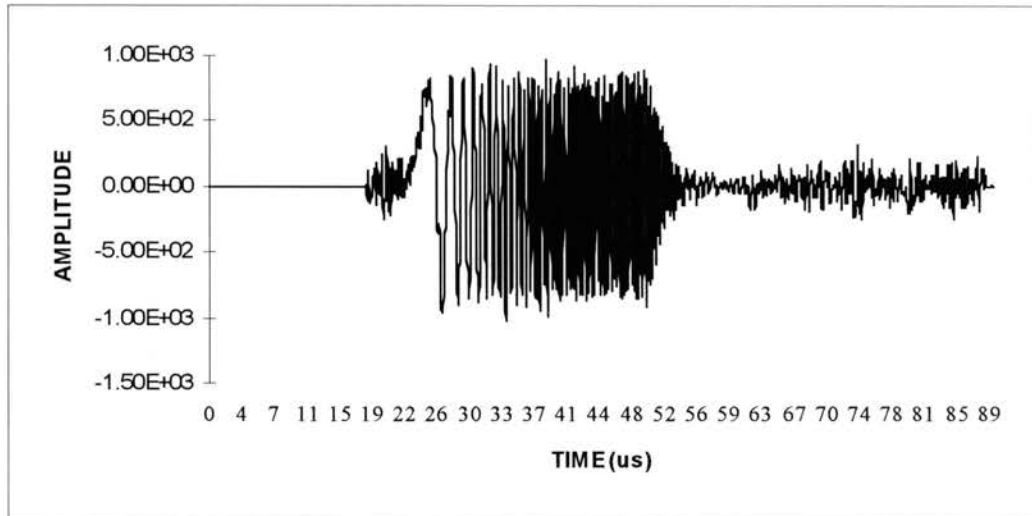


Figure 5-106. Output signal of the MSE(LMS) feedback equalizer for Case 2

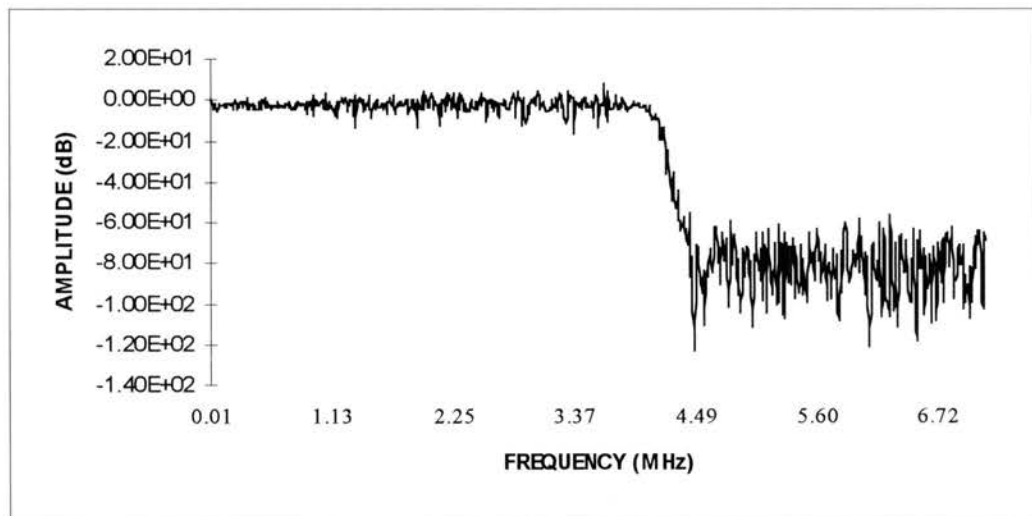


Figure 5-107. Spectrum of the output signal of the MSE(LMS) feedback equalizer for Case 2

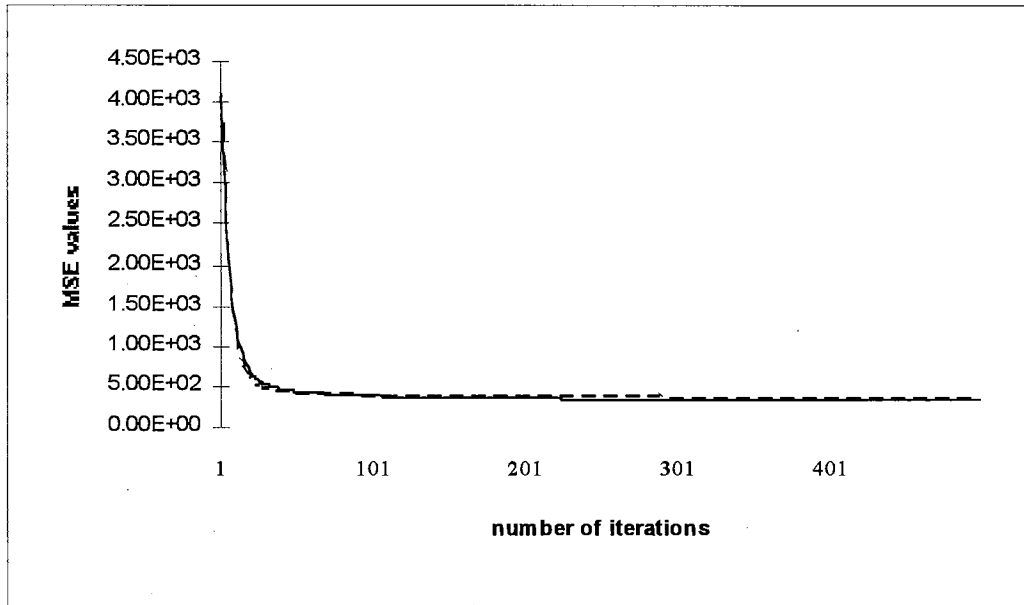


Figure 5-108. MSE of feedback equalizer for Case 2  
( ZF: ——— MSE: - - - - )

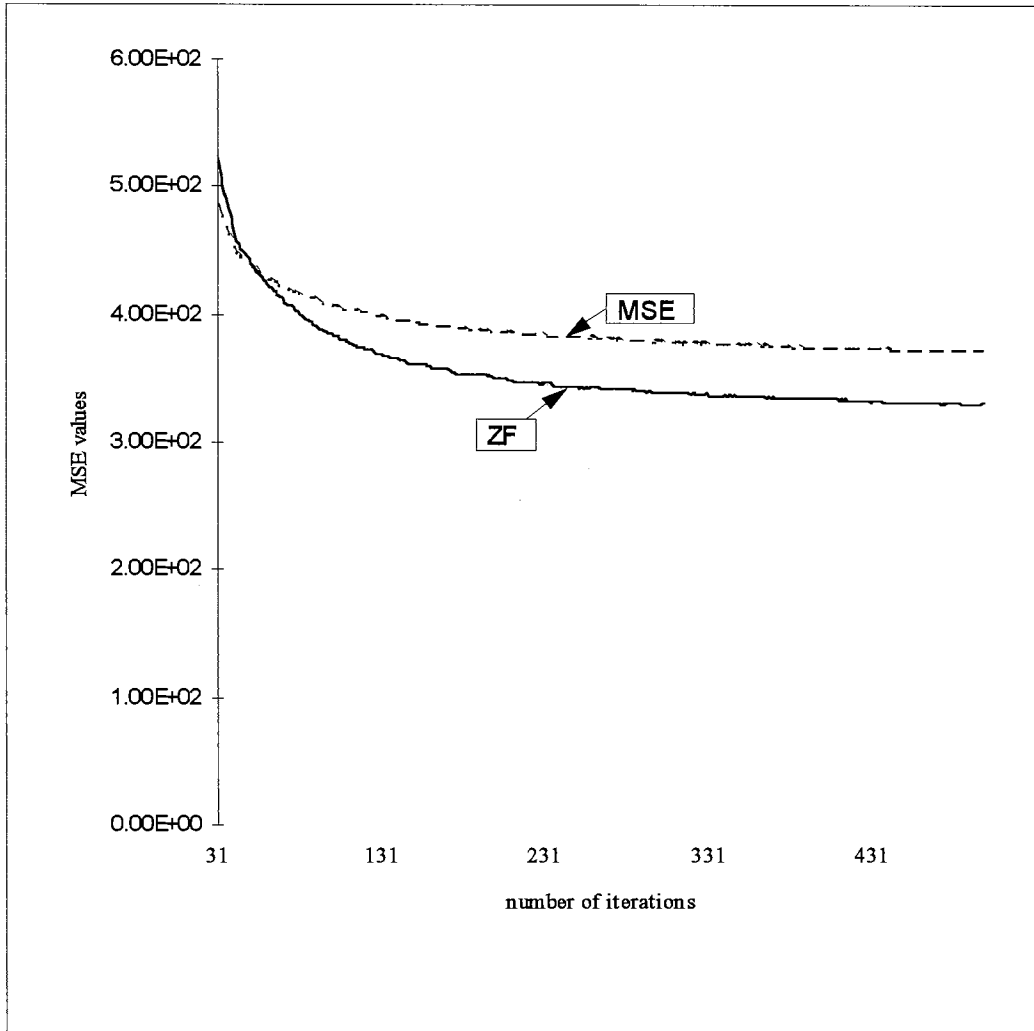


Figure 5-109. Detailed MSE values for Case 2 (feedback)  
 ( ZF: ——— MSE: - - - - )

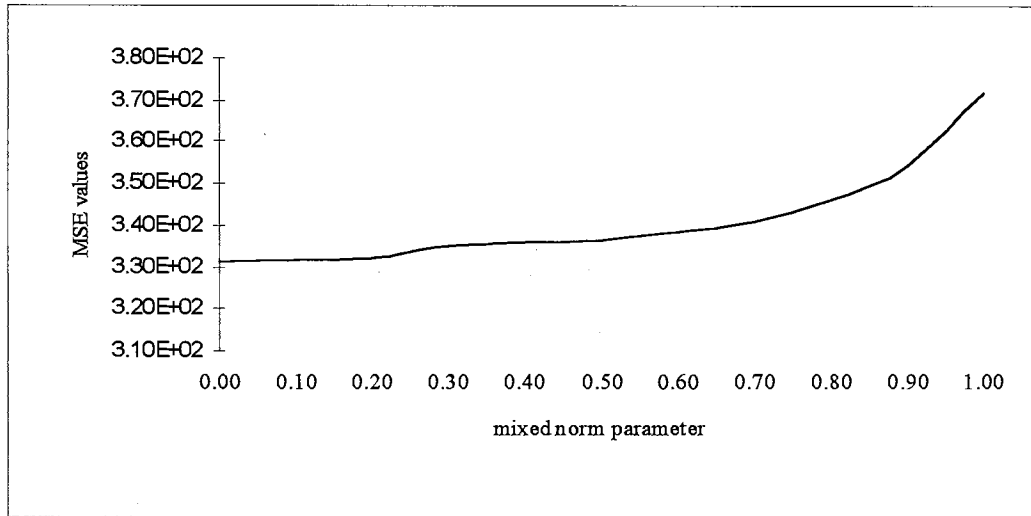


Figure 5-110. MSE values versus mixed norm parameter  $\lambda$  for Case 2 (feedback)

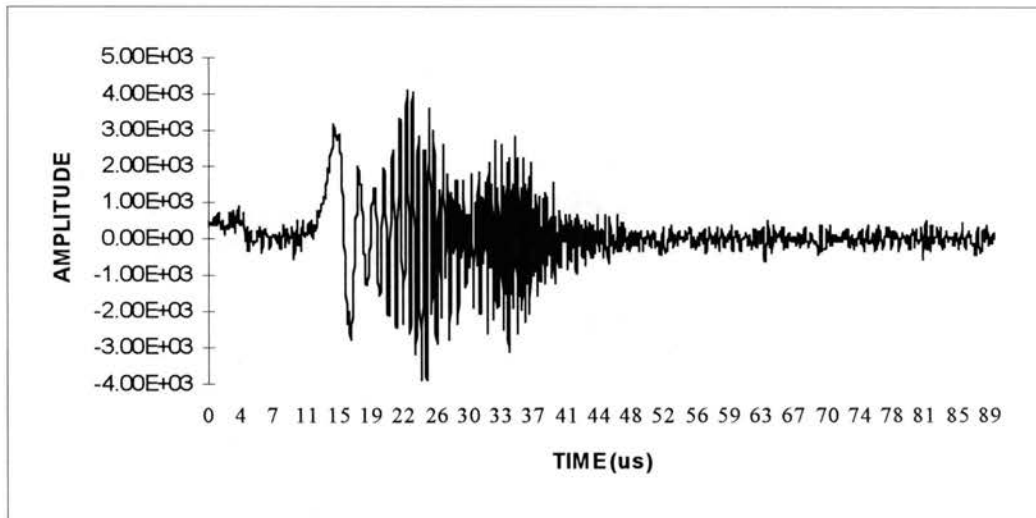


Figure 5-111. Received GCR signal for Case 3 (feedback equalizer)

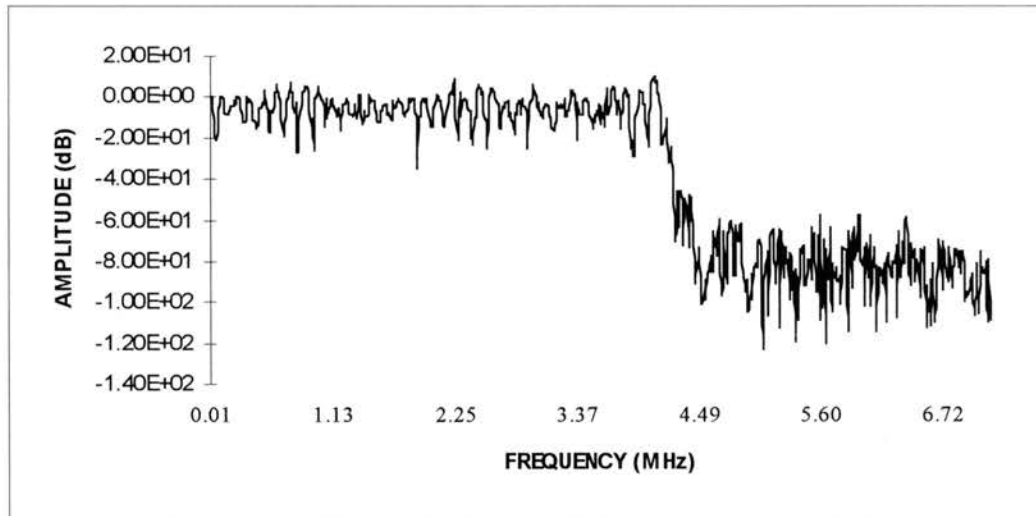


Figure 5-111a. Spectrum of the received GCR signal for Case 3 (feedback equalizer)

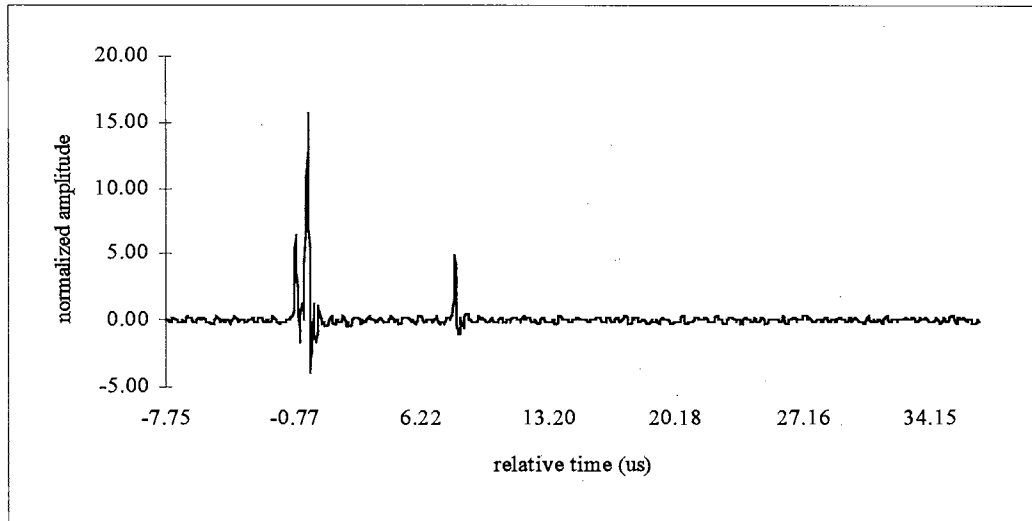


Figure 5-112. Estimated channel impulse response  $f(n)$  for Case 3 (feedback)



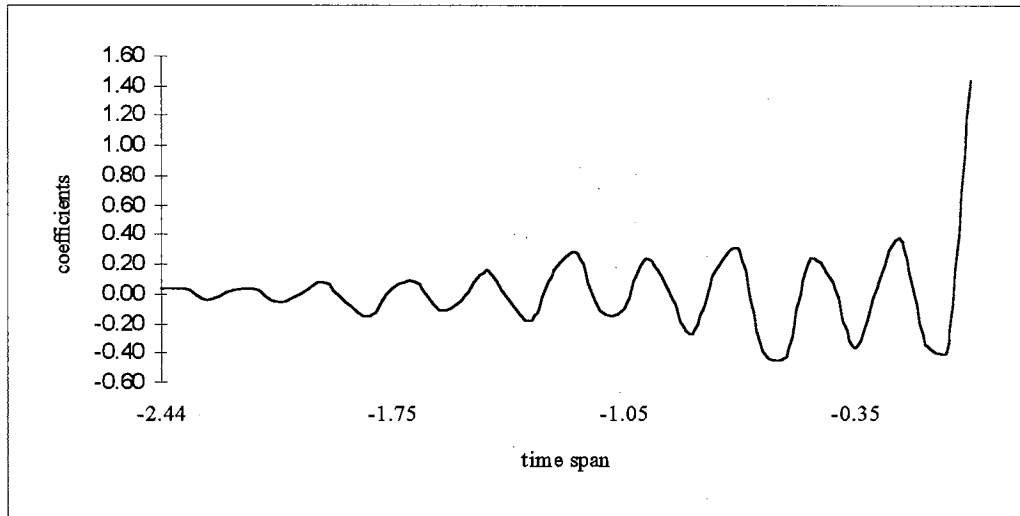


Figure 5-113. Coefficients of forward filter  $h(n)$  of ZF feedback equalizer for Case 3

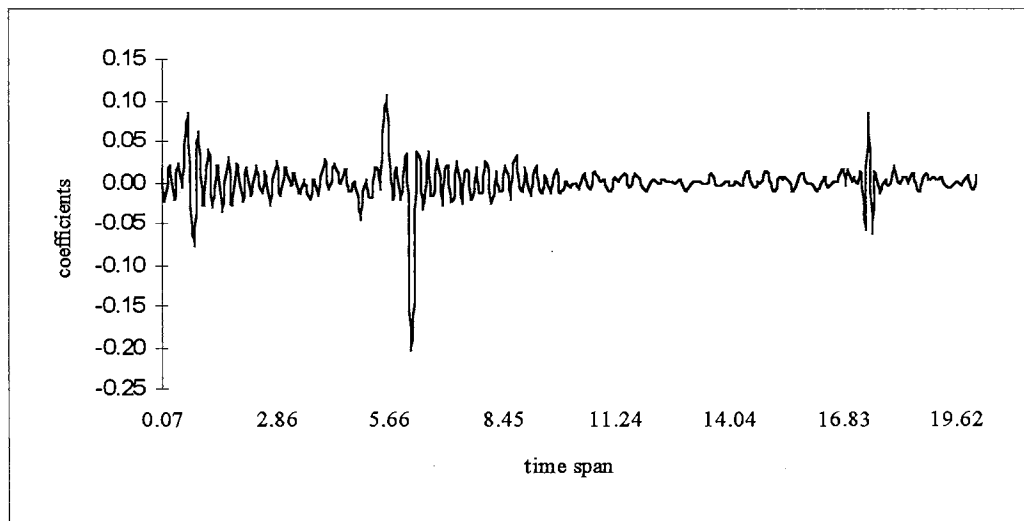


Figure 5-114. Coefficients of feedback filter  $g(n)$  of ZF feedback equalizer for Case 3

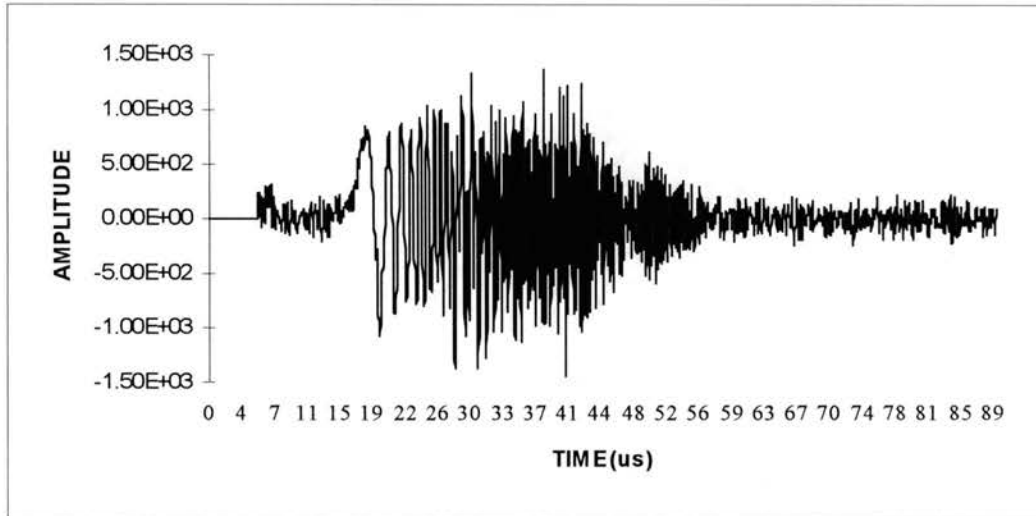


Figure 5-115. Output of the forward filter of the ZF feedback equalizer for Case 3

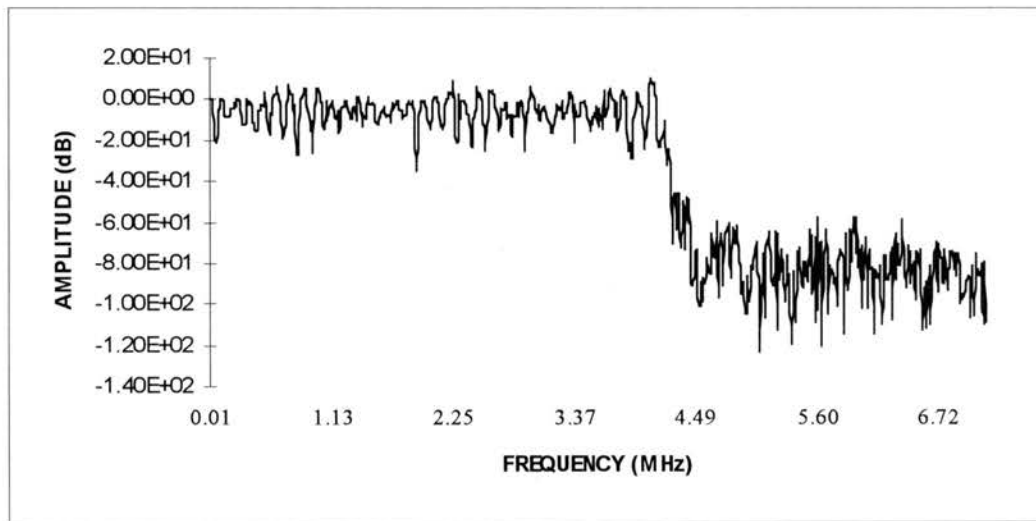


Figure 5-116. Spectrum of the forward filter output of the ZF feedback equalizer for Case 3

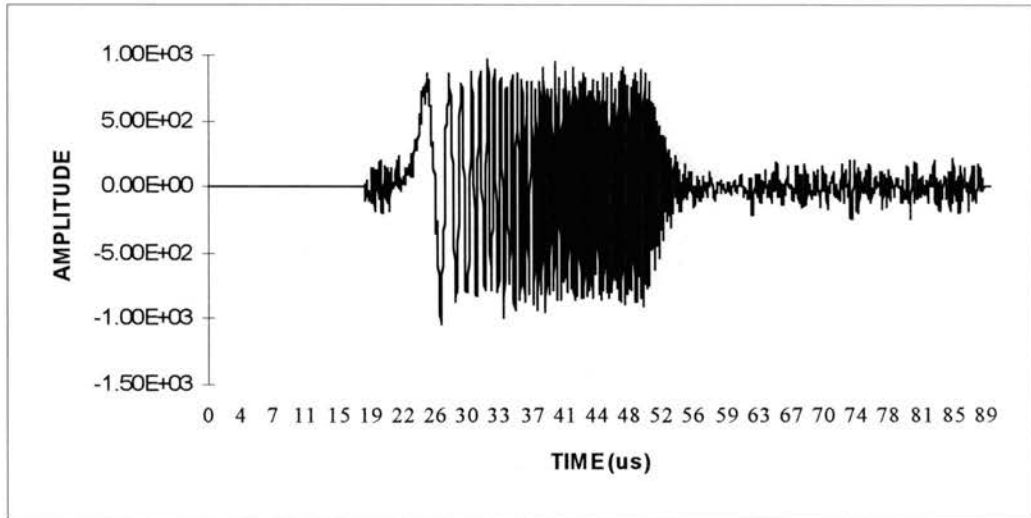


Figure 5-117. Output signal of the ZF feedback equalizer for Case 3

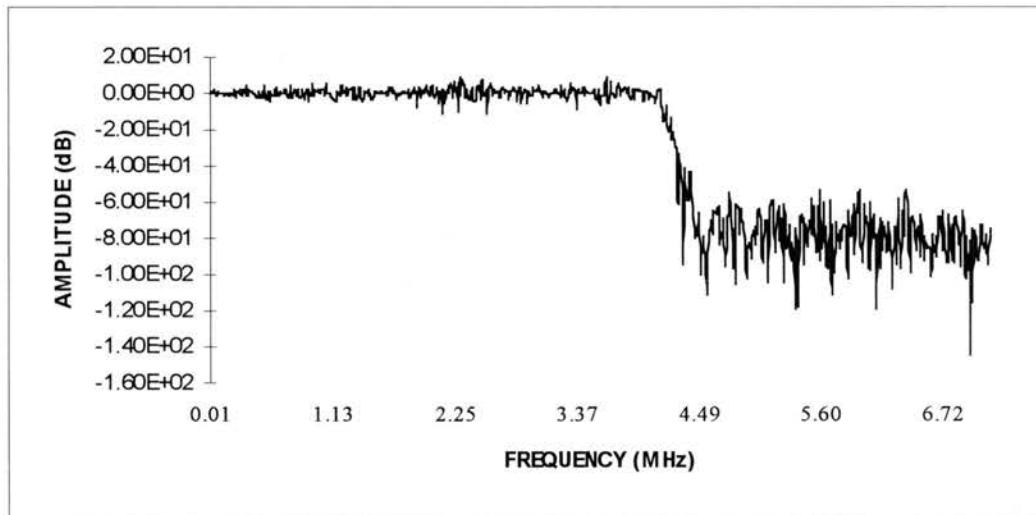


Figure 5-118. Spectrum of the output signal of the ZF feedback equalizer for Case 3

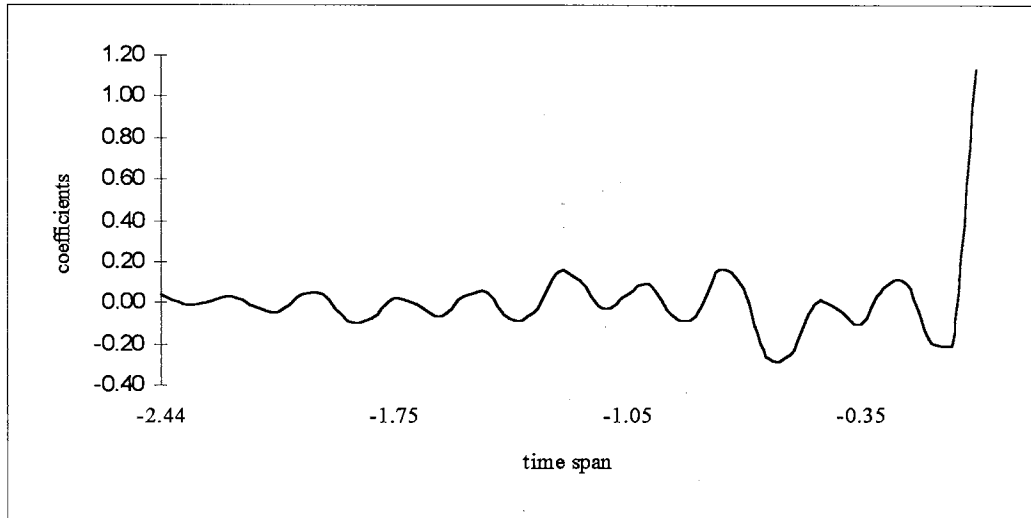


Figure 5-119. Coefficients of forward filter  $h(n)$  of MSE(LMS) feedback equalizer for Case 3

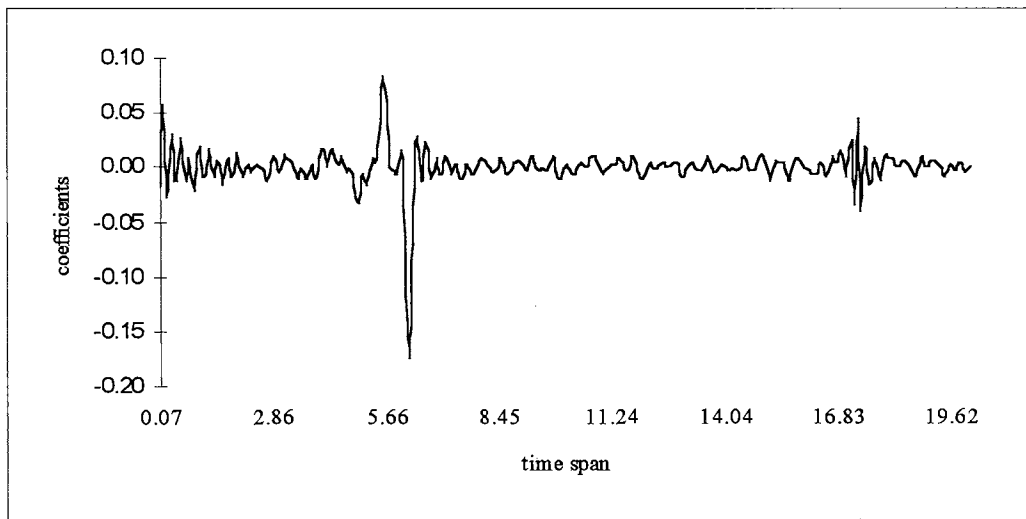


Figure 5-120. Coefficients of feedback filter  $g(n)$  of MSE(LMS) feedback equalizer for Case 3

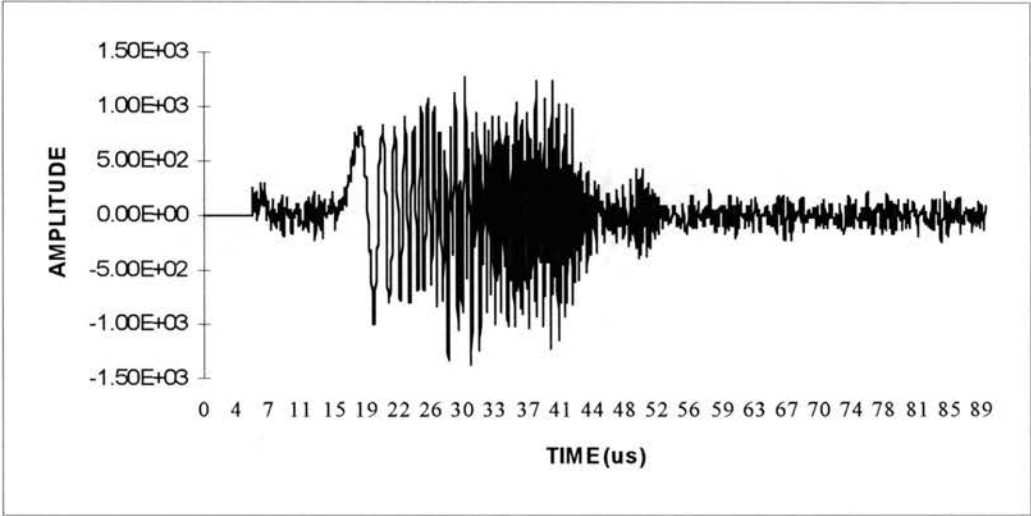


Figure 5-121. Output of the forward filter of the MSE(LMS) feedback equalizer for Case 3

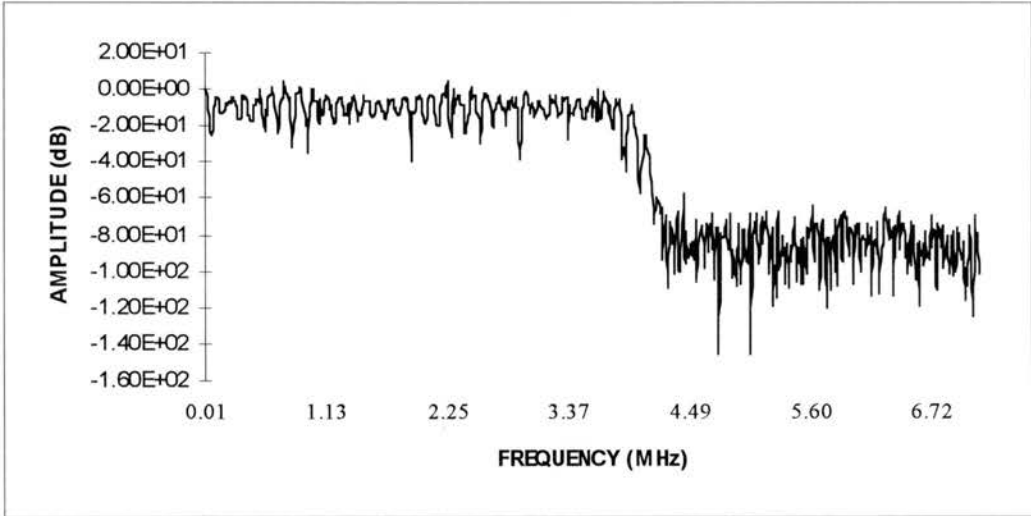


Figure 5-122. Spectrum of the forward filter output of the MSE(LMS) feedback equalizer for Case 3

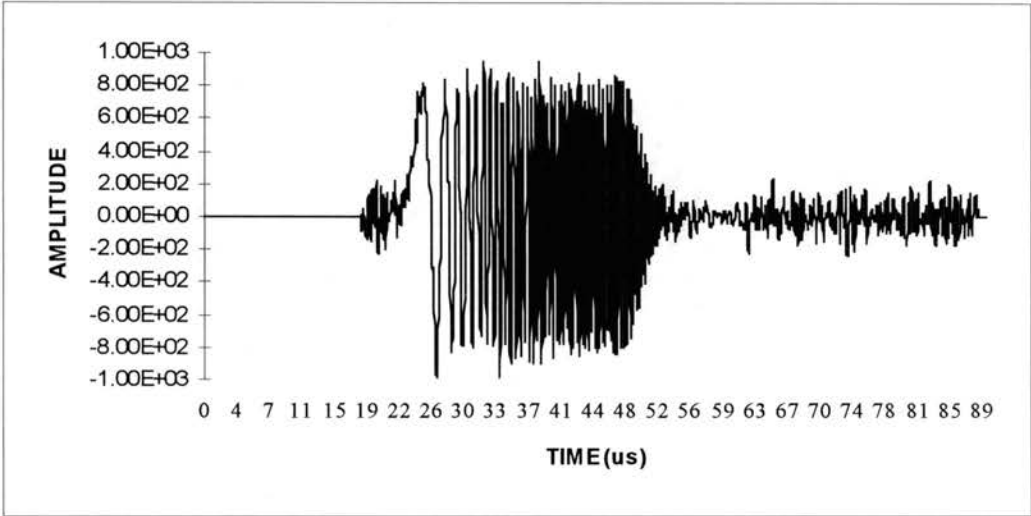


Figure 5-123. Output signal of the MSE(LMS) feedback equalizer for Case 3

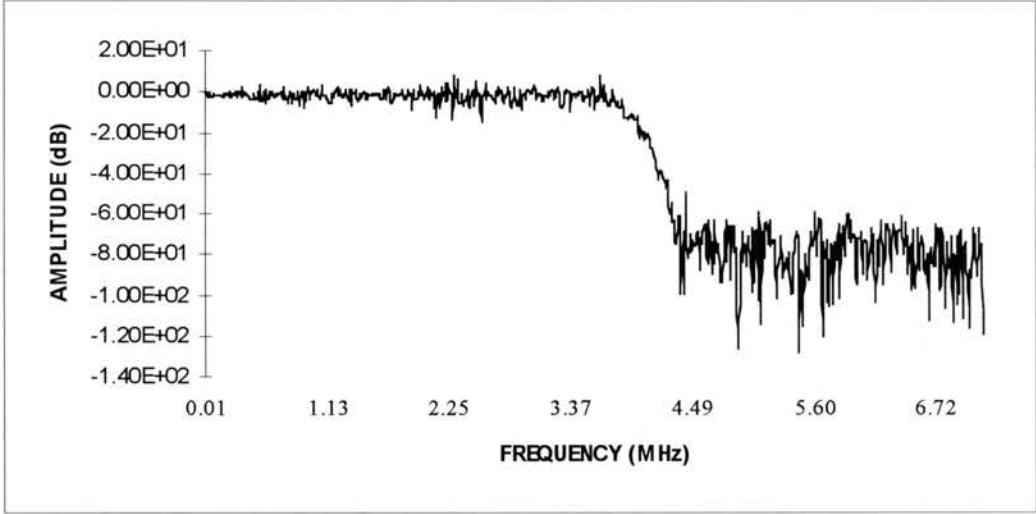


Figure 5-124. Spectrum of the output signal of the MSE(LMS) feedback equalizer for Case 3

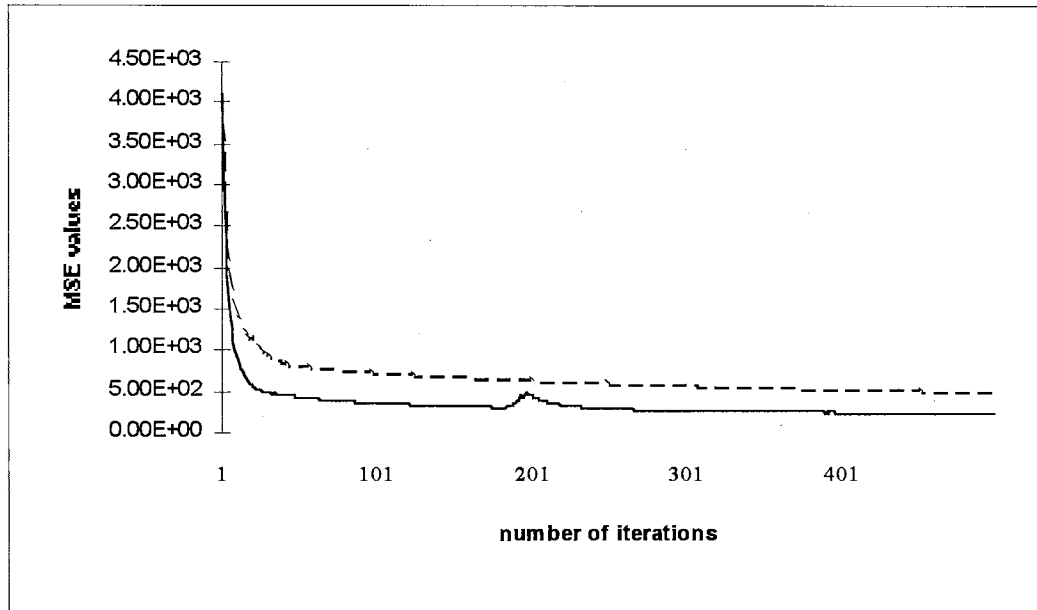


Figure 5-125. MSE of feedback equalizer for Case 3  
( ZF: — MSE: - - - )

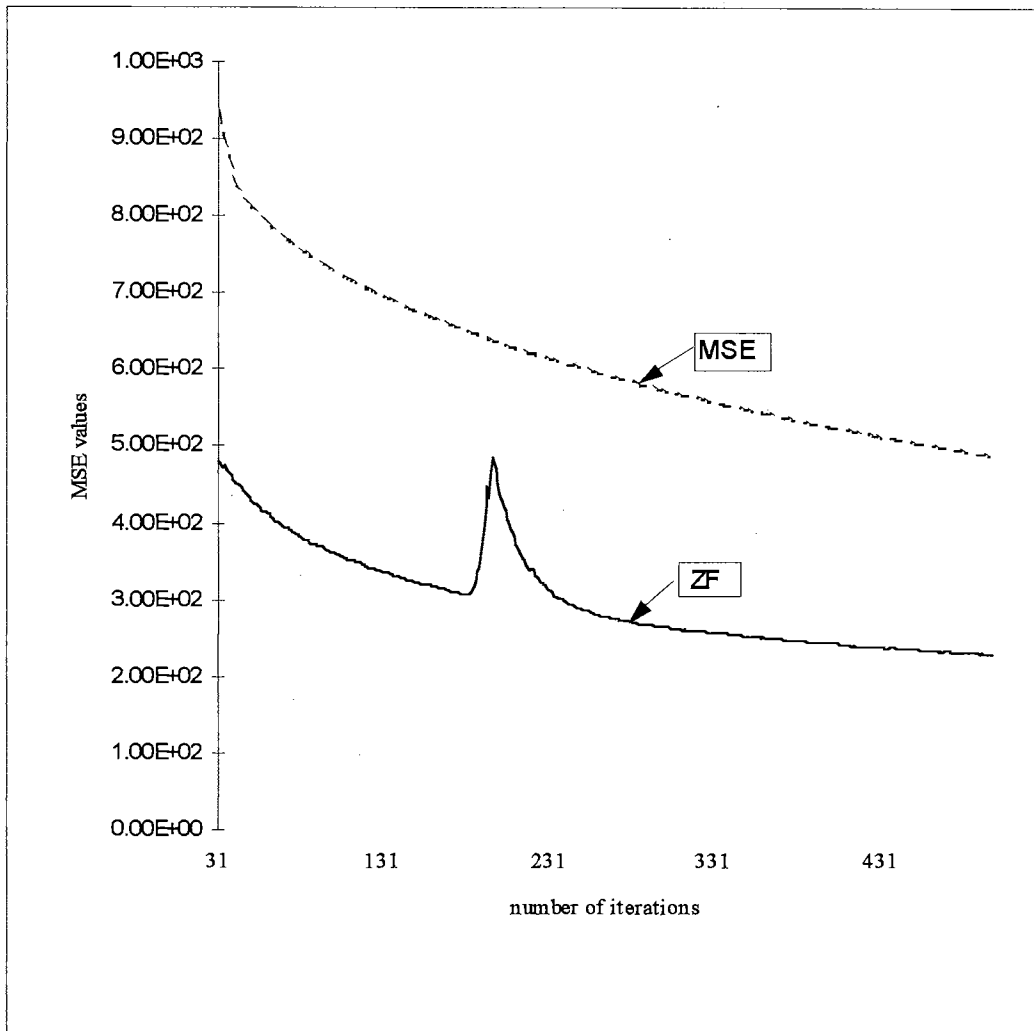


Figure 5-126. Detailed MSE values for Case 3 (feedback)  
 ( ZF: ——— MSE: - - - - )



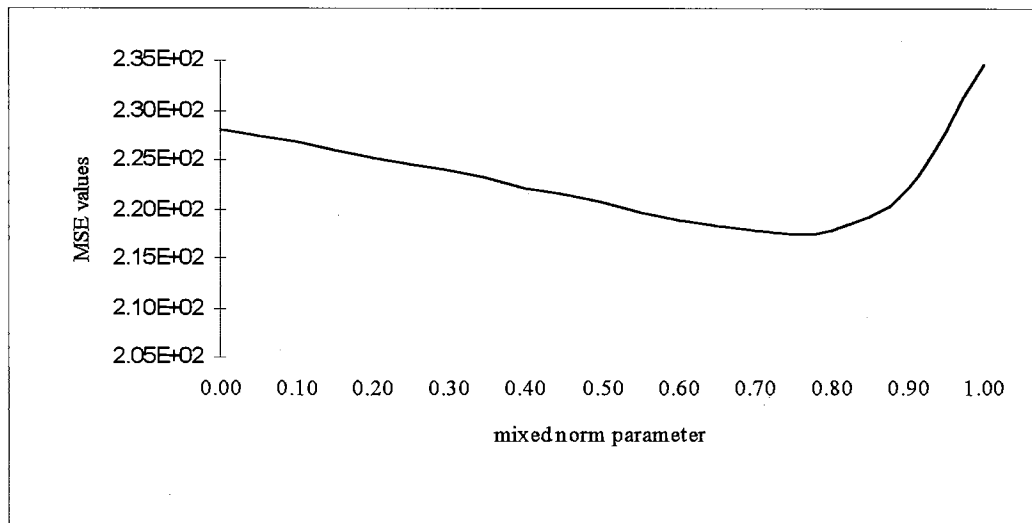


Figure 5-127. MSE values versus mixed norm parameter  $\lambda$  for Case 3 (feedback)

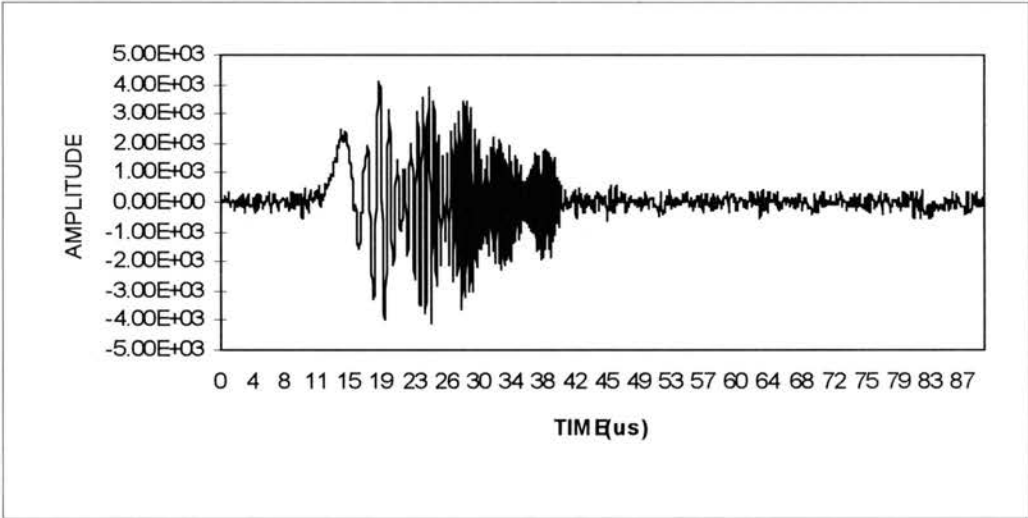


Figure 5-128. Received signal for Case 4

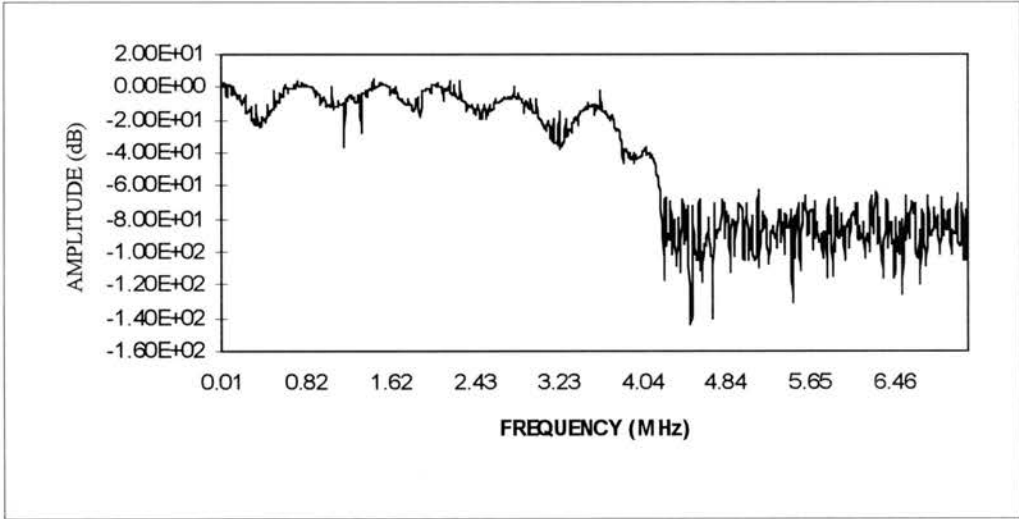


Figure 5-128a. Spectrum of the received signal for Case 4

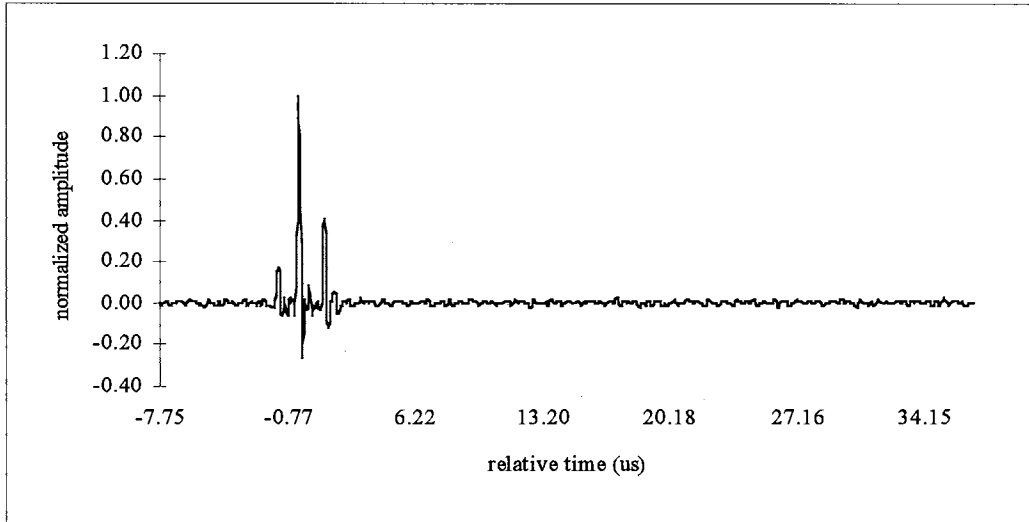


Figure 5-129. Estimated channel impulse response  $f(n)$  for Case 4 (linear and feedback)

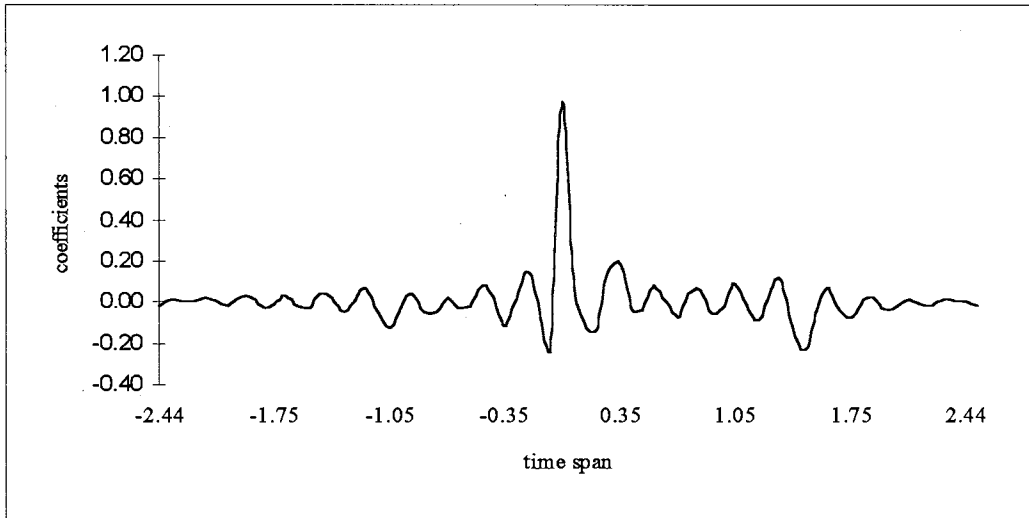


Figure 5-130. Coefficients of the linear MSE(LMS) equalizer for Case 4

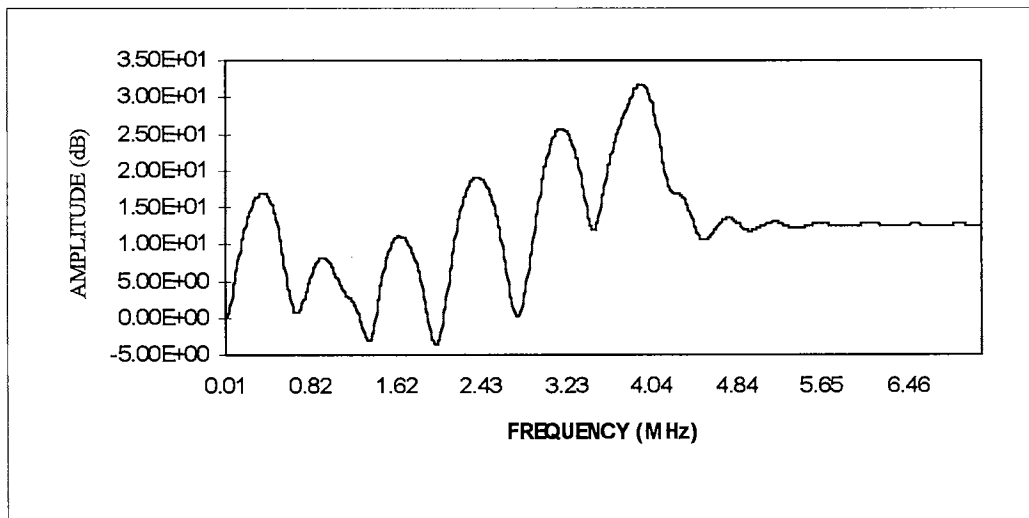


Figure 5-131. Frequency response of the linear MSE(LMS) equalizer for Case 4

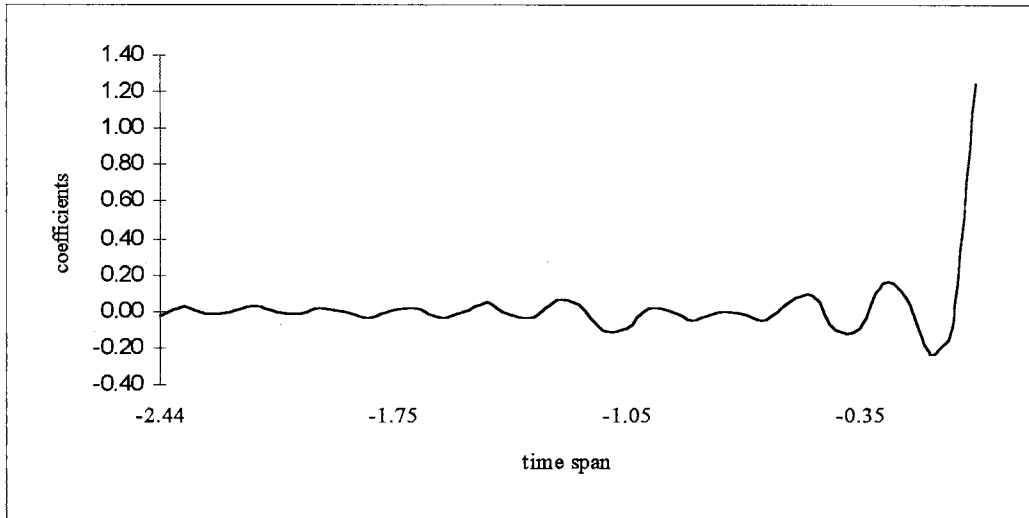


Figure 5-132. Coefficients of the forward filter  $h(n)$  of the MSE(LMS) feedback equalizer for Case 4

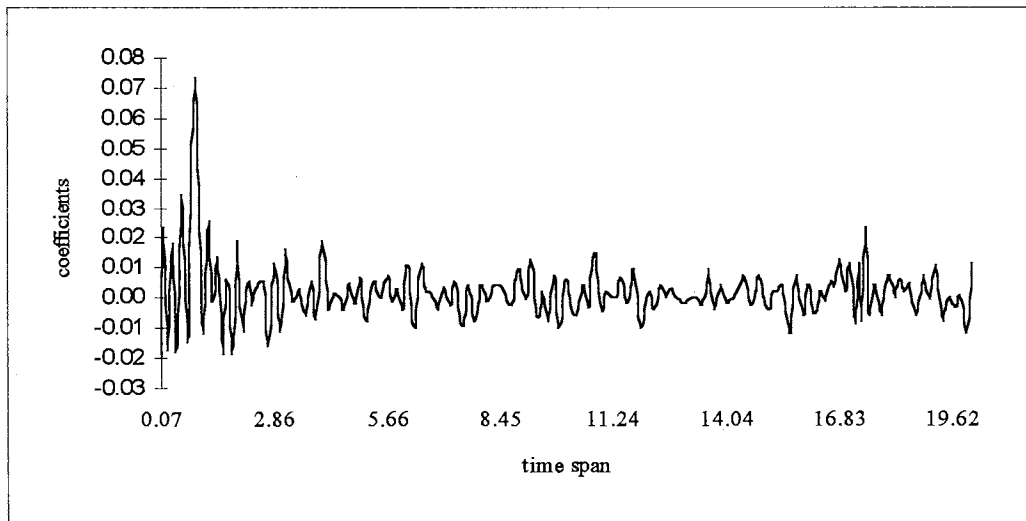


Figure 5-133. Coefficients of the feedback filter  $g(n)$  of the MSE(LMS) feedback equalizer for Case 4

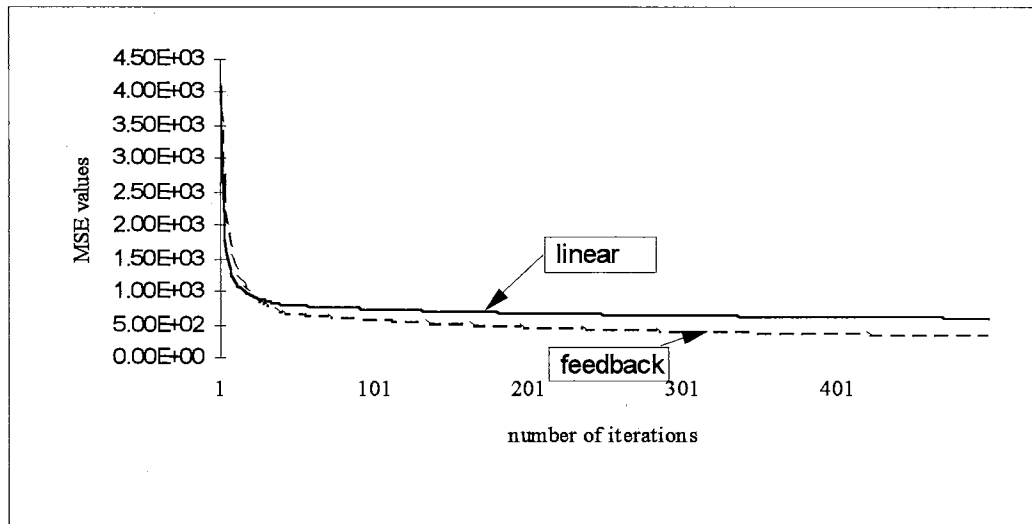


Figure 5-134. MSE values of the linear and feedback MSE(LMS) equalizers for Case 4  
 ( linear: ————— ; feedback: - - - - - )

## CHAPTER VI

### CONCLUSIONS

This paper deals with the important topic of equalization in a communications system in a broad sense. The main results obtained through this research are:

- (1) Introducing the concept of mixed norm into the area of equalization.

Channel distortion and noise are the two main factors contributing to the degradation of the communication quality. The zero-forcing (ZF) criterion attempts to eliminate the intersymbol interference (ISI) regardless of the noise environment. The mean-square-error (MSE) criterion treats the ISI and noise equally. This paper started by formulating the equalization problem in the framework of norms. Based on the Bayes theory, the maximum *a posteriori* estimation of the combined channel impulse response was obtained, resulting the mixed norm criterion. The mixed norm parameter  $\lambda$ , which is related to our belief on the prior knowledge about the channel characteristics, provides an extra degree of freedom to optimize the mixed norm for the specific application.

- (2) Developing the optimum ideal linear equalizer that minimizes the mixed norm, and obtaining an expression for the resulting mean square error (MSE).

A linear equalizer with the tapped delay line structure that minimizes the mixed norm was obtained, as a function of the channel impulse response. It was pointed out that the linear equalizer based on the zero-forcing criterion attempts to invert the channel frequency response; the linear equalizer based on the mean-square-error criterion, which is

a type of noncausal Wiener filter, strikes a balance between reducing the ISI and enhancing noise. The mixed norm linear equalizer, with the help of the mixed norm parameter, can highlight the ISI or noise.

The mean square error at the equalizer output, as a measure of merit, was obtained for each equalizer during the derivation of the equalizer itself.

(3) Developing the optimum ideal feedback equalizer that minimizes the mixed norm, and obtaining an expression for the resulting MSE.

The noise enhancement problem is inherent in the linear equalizer structure. The feedback equalizer was proposed as an alternative, with the goal of compensating severe amplitude distortion without significantly enhancing noise. The feedback portion of the feedback equalizer is designed to cancel only the postcursor ISI; while the forward filter is responsible for canceling only the precursor ISI. The relationship between the desired feedback filter and the combined channel impulse response (including the channel and the forward filter) remains the same, independent of the criterion for the forward filter, as long as the overall MSE is the goal for the entire feedback equalizer.

The ZF forward filter attempts to eliminate the precursor ISI; the MSE forward filter is a type of Wiener filter (based on the Wiener-Hopf theory); the mixed norm forward filter allows further optimization with the given channel distortion and noise environment in practice. The overall MSE is also obtained for each feedback equalizer as a measure of merit.

(4) Proposing algorithms for updating the automatic linear and feedback equalizers that minimize the mixed norms.



In most practical applications, the channel impulse response (or equivalently the channel frequency response) is not explicitly known. Therefore, an equalizer should be able to adjust itself given the received signal (and possibly a reference signal).

The original version of Lucky's ZF algorithm assumes a periodic narrow pulse as a training signal, which is not true in most practical applications today, due to the poor performance of this signal under noisy conditions. In this paper, the ZF algorithm was extended for a general broadband signal. An augmented LMS algorithm was adopted to update the LMS equalizer. A new algorithm to update the mixed norm equalizer was proposed.

Because of the availability of the local reference signal, the feedback filter is updated by a special algorithm optimized for implementation simplicity in practice.

(5) Through numerical solutions, the relationship between the parameter  $\lambda$  of the mixed norm and the performance of the mixed norm equalizer is observed.

The mixed norm parameter, related to our belief on the *a priori* knowledge about the channel characteristics under consideration, plays a key role in the performance of the mixed norm equalizer. Extensive experiments indicated that under some particular channel distortion and noise environment and the practical restriction of finite length implementation, the mixed norm equalizer can outperform both the ZF equalizer and the MSE equalizer in terms of reaching a smaller steady state mean square error.

However, the relationship between the overall MSE value and the mixed norm parameter appears to be highly nonlinear, given the same channel characteristics and noise environment.

(6) Experiments showed that in the multipath cancellation application, the introduction of the feedback structure can substantially improve the overall performance of the equalizer, in terms of reaching a smaller MSE value.

(7) New algorithms and implementations for the television multipath cancellation are obtained, resulting in two US patents and the related products.

The multipath signal propagation, or “ghosting” problem, in the television system is considered to be “the last major technological problem in the analog television system”. In fact, the technique developed here will be also useful for the future digital and high-definition television systems.

To effectively cancel the ghosts, a feedback equalizer was employed, which results in extra computational complexity. In addition, digitizing the video signal requires a very high sample rate compared with the computational capability of the digital signal processors available today. Therefore, the algorithms have to be modified and optimized to fit into today’s technology. For example, a block-based scheme instead of the conventional sample-based scheme was used to update the equalizer coefficients, with the help of a flexible hardware architecture.

Experimental results with the real life off-the-air signals indicated that the proposed algorithms, together with the hardware structure, substantially improved the video quality. A commercial product based on these techniques is its near completion.

The problems that remain to be solved are:

(1) Even though we discovered that fact that the mixed norm equalizer, with some particular mixed norm parameter, can reach a smaller overall MSE value than the ZF

equalizer and MSE equalizer under certain channel distortion and noise environment, no explicit relationship is established at this point. Further research and extensive experiments have to be done for more understanding of this relationship.

(2) Although the feedback structure substantially improves the equalizer performance, it introduces the stability issue. None of the algorithms for the feedback equalizer proposed in this paper can guarantee the stability. Further research is needed to understand the cause of the instability in the particular algorithm proposed in this paper to update the feedback portion of the equalizer.

## REFERENCES

- [1] Nyquist, H., "Certain topics in telegraph transmission theory," Trans. AIEE, Vol. 47, pp. 617-644, April 1928.
- [2] Shannon, C.E., "A mathematical theory of communication," Bell Syst. Tech. Journal, October 1948.
- [3] Lucky, R.W. et al., Principles of Data communication, McGraw-Hill, New York, 1968.
- [4] Lucky, R.W., "Techniques for adaptive equalization of digital communication systems," Bell Syst. Tech. J., Vol. 45, pp.255-286, February 1966.
- [5] Feher, K., Ed., Advanced Digital Communications, System and Signal Processing Techniques, Prentice-Hall, Englewood Cliffs, New Jersey, 1987.
- [6] Sussman, S.M., "A matched filter communication system for multipath channels," IRE Transactions on Information theory, Vol. IT-6, No. 3, pp.376-372, June 1960.
- [7] Foschini, G.J. and Salz, J., "Digital communication over fading radio channels," Bell System Technical Journal, 62, No. 2, pp. 1857-76, February 1988.
- [8] Kevehrad, M. and Salz, J., "Cross-polarization cancellation and equalization in digital transmission over dually polarized multipath fading channels," AT&T Technical Journal, Vol. 64, No. 10, December 1985.
- [9] Fenderson, G.L., et al., "Adaptive transversal equalization of multipath propagation for 16-QAM, 99-Mb/s digital radio," AT&T Technical Journal, Vol. 63, No.8, October 1984.
- [10] Cowan, C.F.N. and Grant, P.M., Ed. Adaptive Filters, Prentice-Hall, Englewood Cliffs, New Jersey, 1986.
- [11] Gitlin, D.G., et. al., Data Communications Principles, Plenum Press, New York, 1992.
- [12] Proakis, J.G., Digital Communications, 2nd Ed., McGraw-Hill, New York, 1989.

- [13] Lee, E.A. and Messerschmitt, D.G., *Digital Communications*, Kluwer Academic Publishers, Boston, 1988.
- [14] Qureshi, S.U.H., "Adaptive equalization," *Proceedings of the IEEE*, Vol. 73, No. 9, September 1985.
- [15] Proakis, J.G. and Miller, J.H., "An adaptive receiver for digital signaling through channels with intersymbol interference," *IEEE Trans. Infor. Theory*, Vol. IT-15, pp.484-497, July 1969.
- [16] Lucky, R.W. and Rudin, H.R., "An automatic equalizer for general-purpose communication channels," *Bell Syst. Tech. J.*, Vol. 46, pp.2179-2208, November 1967.
- [17] Proakis, J.G., "Advances in equalization for intersymbol interference," in *Advances in communication Systems*, Vol. 4, A.J. Viterbi, Ed., Academic Press, New York, 1975.
- [18] Cartinhour, J., *Correspondences*, April - May, 1995.
- [19] Komiya, N., "Ghost measurement by waveform synthesis," *Proceedings of the Int'l Conference on Consumer Electronics*, Chicago, IL, 1994.
- [20] Lee, M.-H., "Ghost cancellation using sparse canonical signed digit code coefficients," *Proceedings of the Int'l Conference on Consumer Electronics*, Chicago, IL, 1994.
- [21] Otsuki, K. and Imamura, M., "Development of the new ghost analyzer," *Proceedings of the Int'l Symp. on Electromagnetic Compatibility*, Miyagi, Japan, 1994.
- [22] Ciciora, W. et al., "A tutorial on ghost canceling television systems," *IEEE Trans. on Consumer Electronics*, Vol.CE-25, February 1979.
- [23] Pham, T.T. and DeFigueiredo, R.J.P., "Maximum likelihood estimation of a class of non-Gaussian densities with application to  $I_p$  deconvolution," *IEEE Trans. on Acoustics, Speech and Signal Processing*, Vol.37, No.1, January, 1989.
- [24] Alliney, S. and Ruzinsky, S.A., "An algorithm for the minimization of mixed  $L_1$  and  $L_2$  norms with Applications to Bayesian estimation," *IEEE Trans. on Signal Processing*, Vol.42, No.3, March 1994.
- [25] O'Brien, M.S. et al., "Recovery of a sparse spike time series by  $L_1$  norm deconvolution," *IEEE Trans. Signal Processing*, Vol.42, No.12, December 1994.
- [26] Ericson, T., "Structure of optimum receiving filters in data transmission systems," *IEEE Trans. Inform. Theory*, Vol. IT-17, pp.352-353, May 1971.

- [27] George, D.A. et al., "An adaptive decision feedback equalizer," IEEE Trans. Commun. Technol., Vol. COM-19, pp.281-293, June 1971.
- [28] Falconer, D.D. and Foschini, G.J., "Theory of minimum mean-square-error QAM system employing decision feedback equalization," Bell Syst. Tech. J., Vol. 53, pp.1821-1849, November 1973.
- [29] Honig, M.L. and Messerschmitt, Adaptive Filters; Structures, Algorithms, and Applications, Kluwer Academic Publications, Boston, MA, 1984.
- [30] Mosen, P., "feedback equalization for fading dispersive channels," IEEE Trans. Infor. Theory, Vol. IT-17, pp.56-64, January 1971.
- [31] Salz, J. "Optimum mean-square decision feedback equalization," Bell Syst. Tech. J., Vol. 52, pp.1341-1373, October 1973.
- [32] Salz, J., "On mean-square decision feedback equalization and timing phase," IEEE Trans. commun. Technol., Vol. COM-25, pp. 1471-1476, December 1977.
- [33] Lee, W.U. and Hill, F.S., "A maximum likelihood sequence estimator with decision feedback equalization," IEEE Trans. Commun. Technol., Vol. COM-25, pp.971-979, September 1977.
- [34] Messerschmitt, D.G., "A geometric theory of intersymbol interference: Part I and II," Bell System Tech. Journal, 50, November 1973.
- [35] Gersho, A., "Adaptive equalization of highly dispersive channel," Bell syst. Tech. J., Vol. 60, pp. 1997-2021, November 1981.
- [36] Mosen, P., "MMSE equalization of interference on fading diversity channels," IEEE Trans. Commun., Vol. COM-32, pp.5-12, January 1984.
- [37] Yarlagadda, R. et. al., "Fast algorithms for  $L_p$  deconvolution," IEEE Trans. Acoustics, Speech and Signal Processing, Vol.ASSP-33, No.1, February 1985.
- [38] Taylor, H.L., et. al., "Deconvolution with the  $L_1$  norm," Geophysics, Vol.44, No.1, January 1979.
- [39] Lucky, R.W., "Automatic equalization for digital communication," The Bell System Technical Journal, Vol.XLIV, No.4, April 1965.
- [40] Makhoul, J., "Linear prediction: a tutorial review," Proceedings of the IEEE, Vol.63, No.4, April 1975.
- [41] Yarlagadda, R., Correspondents, October, 1995.

- [42] Cartinhour, J., Correspondents, October, 1995.
- [43] Oda, H. and Sato, Y., "Blind equalization based on time-dependency principle," Proc. Int'l Symp. on Infor. Theory & Its Applications, Sydney, Australia, November 1994.
- [44] Godard, D.N., "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," IEEE Trans. Commun., Vol. COM-28, No. 11, pp.1867-1875, 1980.
- [45] Shalvi, O. and Weinstein, E., "New criteria for blind deconvolution of nonminimum phase systems (channels)," IEEE Trans. Infor. Theory, Vol. IT-36, No. 2, pp. 312-321, 1990.
- [46] Nikias, C.L. and Rayhuveer, M.R., "Bispectrum estimation: a digital signal processing framework," Proc. IEEE, Vol. 75, No. 7, pp.869-891, 1987
- [47] Barrodale, I. and Roberts, F.D.K., "An improved algorithm for discrete  $L_1$  linear approximation," SIAM J. Numer. Anal., Vol.10., No.5, October 1973.
- [48] Bartels, R.H. et. al. , "Minimization techniques for piecewise differentiable functions: the  $L_1$  solution to an overdetermined linear system," SIAM J. Numer. Anal., Vol.15, No.2, April 1978.
- [49] Bloomfield, P and Steiger, W.L, "Least absolute deviation curve-fitting," SIAM J. Sc. Sta. Comput., 1980, pp.290-301.
- [50] Bloomfield, P, Least Absolute Deviations Regression, New York: John Wiley, 1968.
- [51] Ruzinsky, S.A. and Olsen, E.T., "Strong consistency of the LAD ( $L_1$ ) estimator of parameters of stationary autoregressive processes with zero mean," IEEE Trans. Acoustics, Speech and Signal Processing, Vol.ASSP-37, No.4, April 1989.
- [52] Schlossmacher, E.J., "An iterative technique for absolute deviations curve fitting," Journal of the American Statistical Association, Applications Section, Vol.68, No. 344,, December 1973.
- [53] Fair, R.C. and Peck, J.K., "A note on an iterative technique for absolute deviations curve-fitting," Tech. Rep., Dept. Economics, Yale Univ., 1974.
- [54] Huber, P.J. and Dutter, R., "Numerical solution of robust regression problems," in COMPSAT 1974, Proc. Symp. Computational Statistics, G. Bruchmann, Ed. Wien, Germany: Phisika Verlag, 1974.

- [55] Winters, J. et al., "Ghost cancellation of analog TV signals: with applications to IDTV, EDTV and HDTV," IEEE Trans. on Circuits and Systems for Video Technology, Vol. 1, No. 1, March 1991.
- [56] Makino, S. et al., "Novel automatic ghost canceller," IEEE Trans. Consumer Electronics, Vol. CE-26, August 1980.
- [57] Huang, J., "A ghost cancellation system for the NTSC television," IEEE Transactions on Consumer Electronics, Vol.39, No.4, November 1993.
- [58] Johnson, L. et al, "Low cost stand-alone ghost cancellation system," Proceedings of the Int'l Conference on Consumer Electronics, Chicago, IL, 1994.
- [59] Kim, K.B., "New ghost cancellation system," Proceedings of the Int'l Conference on Consumer Electronics, Chicago, IL, 1994.
- [60] Matsuura, S. et al., "Development of a ghost canceling technology for TV broadcasting," NAB Engineering Conference Proceedings, 1990.
- [61] Koo, D., "Developing a new class of high energy ghost cancellation reference signals," Proceedings of the Int'l Conference on Consumer Electronics, Rosemont, IL, June, 1992.
- [62] Koo, D., "Properties and Applications of the new ghost cancellation reference signal," Proceedings of the Int'l Conference on Consumer Electronics, Rosemont, IL, June, 1993.
- [63] United States Advanced Television Systems Committee, Ghost Canceling Reference Signal for NTSC, 1992.
- [64] Fiallos, E. et al., "Enhanced ghost cancellation reference signal," IEEE Trans. on Consumer Electronics, Vol. 40, No. 3, pp.640-644, August 1994.
- [65] Kyeong, S.Y., "New GCR signal and its Applications," IEEE Trans. Consumer Electronics, Vol. 40, No. 4, pp. 852-860, November 1994.
- [66] Gardiner, P.N. et al., "Ghost cancellation for 625-line systems," Proceedings of the Int'l Broadcasting Convention, Amsterdam, Netherland, 1994.
- [67] Ellis, R.J.G., "PALplus project: conception to introduction," Proceedings of the Int'l Broadcasting Convention, Amsterdam, Netherlands, 1994.
- [68] Wu, W.-R., "Blind ghost cancellation for HDTV," Proceedings of the National Science Council, ROC, Part A: Physical Science and Engineering, Vol. 18, No. 3, pp. 281-288, May 1994.



- [69] Shynk, J.J., "Adaptive IIR filtering," IEEE ASSP Magazine, April 1989.
- [70] Huang, J., "Ghost signal cancellation system using feedforward and feedback filters for television signals," US Patent 5,321,512, June 1994.
- [71] Huang, J. and Casey, R., "Ghost signal cancellation system for television signals," US Patent 5,386,243, January 1995.
- [72] Miyazawa, H. et al., "Development of a ghost cancel reference signal for TV broadcasting," IEEE Trans. on Broadcasting, Vol. 33, No.4, December 1989.
- [73] Greenberg, C.B., "Ghost cancellation system for the US standard GCR," IEEE Trans. Consumer Electronics, Vol. 39, No.4, November 1993.
- [74] Bloomfield, P. and Steiger, W.L., Least Absolute Deviations, Boston-Basel: Birkhauser, 1983.
- [75] Gentle, J.E., et. al., "On the least absolute values estimation," Commun. Statist. - Theor. Meth., A6(9), pp.839-845, 1977.
- [76] Papoulis, A, Probability, Random Variables, and Stochastic Process, 2nd Edition, McGraw-Hill Book Company, New York, 1984.
- [77] Nehari, Z., Conformal Mapping, Dover, New York, 1952.
- [78] Zadeh, L.A. and Desoer, C.A., Linear System Theory: A State Space Approach, McGraw-Hill, New York, 1963.
- [79] Widrow, B. and Hoff, M., Jr., "Adaptive switching circuits," IRE WESCON Convention Record, Pt. 4, pp.96-104, 1960.
- [80] Bednar, J.B., Yarlagadda, R. and Watt, T., " $L_1$  deconvolution and its application to seismic signal processing," Proc. of Int'l Conf. Acoust. Speech and Signal Processing, 1985.
- [81] Denoel, E. and Solvay, J.-P., "Linear prediction of speech with least absolute error criterion," IEEE Trans. Acoustics, Speech and Signal Processing, Vol. ASSP, No.6, December 1985.
- [82] Lansford, J. and R. Yarlagadda, "Adaptive  $L_p$  approach to speech coding," ICASSP-88, pp.335-338, New York, 1988.
- [83] Schroeder, J. and Yarlagadda, R., "Linear predictive spectral estimation via  $L_1$  norm," Signal Processing 17 (1989) 19-29.

- [84] Schroeder, J. and Yarlagadda, R., "Two dimensional linear predictive estimation via  $L_1$  norm," EUSIPCO-86, pp.247-250, 1986.
- [85] Cichocki, A. and Lobos, T., "Artificial neural networks for real-time estimation of basic waveforms of voltage and currents," IEEE Trans. on Power Systems, Vol. 9, No. 2, pp. 612-618, May 1994.
- [86] Narula, S.C. and Korhonen, P.J., "Multivariate multiple linear regression based on the minimum sum of absolute errors criterion," European Journal of Operational Research, Vol. 73, No. 1, pp.70-75, February 1994
- [87] Fiodorov, E.D., "Least absolute values estimation: computational aspects," IEEE Transactions on Automatic Control, Vol. 39, No.3, pp. 629-630, March 1994
- [88] Klein, A. and Baier, P.W., "Linear unbiased estimation in mobile radio systems applying CDMA," IEEE Journal on Selected Areas in Communications, Vol. 11, No.7, pp. 1058-1066, September 1993.
- [89] Soliman, S.A. et al., "New dynamic filter based on least absolute value algorithm for on-line tracking of power system harmonics," IEE Proceedings Generation, Transmission and Distribution, Vol. 142, No. 1, pp. 37-44, January 1995.
- [90] Hawley, R.W. and Gallagher, N.C. Jr., "On Edgeworth's method for minimum absolute error linear regression," IEEE Trans. on Signal Processing, Vol. 42, No. 8, pp. 2045-2054, January 1995.
- [91] Rinne, J. and Renfors, M., "Equalization of orthogonal frequency division multiplexing signals," Proceedings of the 1994 IEEE Global Telecom. Conference - GLOBECOM'94, Part 1, San Francisco, 1994.
- [92] Chiani, M., "Erasures and error propagation in decision-feedback equalization," Proceedings of the Int'l Symp. on Information Theory & Its Applications, Part 2, Sydney, Australia, 1994.
- [93] Lee, G.H. et al., "Modification of the reference signal for fast convergence in LMS-based adaptive equalizers," IEEE Trans. on Consumer Electronics, Vol. 40, No. 3, pp. 645-654, August 1994.
- [94] Wang, J.D. et al., "Training signal and receiver design for multipath channel characterization for TV broadcasting," IEEE Trans. Consumer Electronics, Vol. 36, No. 4, November 1990.
- [95] Strauss, H., "Best  $L_1$ - approximation," Journal of Approximation Theory 41, pp.297-308, 1984.

- [96] Sunwoo, H.S. and Kim, B.C., " $L_1$  estimation for the simple linear regression model," *Commun. Statist. - Theory Meth.*, 16(6), pp.1903-1715, 1987.
- [97] Dielman, T.E., "Least absolute value estimation in regression models: an annotated bibliography," *Commun. Statist. - Theor. Meth.*, 13(4), pp.513-541, 1984.
- [98] Harter, H.L., "Nonuniqueness of least values regression," *Commun. Statist. - Theor. Mech.*, A6(9), pp.829-838, 1977.
- [99] Heathcote, C.R., Welsh, A.H., "The robust estimation of autoregressive process by functional least squares," *Journal of Applied Probability*, 20, pp.737-753, 1983.
- [100]. An, H.-Z. and Chen, Z.-G., "On the convergence of LAD estimates in autoregression with infinite variance," *Journal of Multivariate Analysis*, 12, pp.335-345, 1982.
- [101] Nyquist, H., "The optimal  $L_p$  norm estimator in linear regression models," *Commun. Statist. - Theor. Meth.*, 12(21), pp.2511-2524, 1983.
- [102] Farebrother, R.W., "Unbiased  $L_1$  and  $L_\infty$  estimation," *Commun. Statist. - Theor. Meth.*, 14(8), pp.1941-1962, 1985.
- [103] Sposito, V.A., et. al., "Useful generalized properties of  $L_1$  estimators," *Commun. Statist. - Theor. Meth.*, A9(12), pp.1309-1315, 1980.
- [104] El-Attar, R.A., et. al. , "An algorithm for  $l_1$ -norm minimization with application to nonlinear  $l_1$ -approximation," *SIAM J. Numer. Anal.*, Vol.16, No.1, February 1979.
- [105] Kroo, A., "Error estimations for deviation of best uniform, discrete and  $L_q$ -approximations," *SIAM J. Numer. Anal.*, Vol. 18, No.5, October 1981.
- [106] Hald, J. and Madsen, K., "Combined LP and quasi-Newton methods for nonlinear  $l_1$  optimization ," *SIAM J. Numer. Anal.*, Vol.22, No.1, February 1985.
- [107] Shanno, D.F., and Rocke, D.M., "Numerical methods for robust regression: linear models," *SIAM J. Sci. Stat. Comput.*, Vol. 7, No.1, January 1986.
- [108] Osborne, M.R. and Watson, G.A., "An analysis of the total approximation problem in separable norms, and an algorithm for the total  $l_1$  problem," *SIAM J. Sci. Stat. Comput.*, Vol.6, No.2, April 1985.
- [109] Money, A.H. et. al., "The linear regression model:  $L_p$  norm estimation and the choice of p," *Commun. Statist. - Simula.. Computa.*, 11(1), pp.89-109, 1982.

- [110] Sposito, V.A. et. al. "On the efficiency of using the sample kurtosis in selecting optimal  $L_p$  estimators," *Commun. Statist. - Simula. Computa.*, 12(3), pp.265-272, 1983.
- [111] Sposito, V.A., "On unbiased  $L_p$  regression estimators," *Journal of the American Statistical Association, Theory and Methods Section*, Vol.77, No. 371, September 1982.
- [112] Sielken, R.L., and Hartley, H.O., "Two linear programming algorithms for unbiased estimation of linear models," *Journal of the American Statistical Association, Theory and Methods Section*, Vol.68, No. 343, September 1973.
- [113] Forsythe, A.B., "Robust estimation of straight line regression coefficients by minimizing  $p^{\text{th}}$  power deviations," *Technometrics*, Vol.14, No.1, February 1972.
- [114] Taylor, H.L., et. al., "Deconvolution with the  $L_1$  norm," *Geophysics*, Vol.44, No.1, January 1979.
- [115] Clark, D.I. and Osborne, M.R., "A discrete algorithm for minimizing polyhedral convex functions," *SIAM J. Sci. Sta. Comput.*, Vol.4, No.4, December 1983.
- [116] Clark, D.I., "The mathematical structure of Huber's M-estimator," *SIAM J. Sci. Stat. Comput.*, Vol.6, No.1, January 1985.
- [117] Clark, D.I. and Osborne, M.R., "Finite algorithms for Huber's m-estimator " *SIAM J. Sci. Sta. Comput.*, Vol.7, No.1, January 1986.
- [118] Li, S. and Dickinson, B.W., "Application of lattice filter to robust estimation of AR and ARMA models," *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol.36, No.4, April 1988.
- [119] Lee, C.-H., "On robust linear prediction of speech," *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol.36, No.5, May 1988.
- [120] Martin, R.D. and Thomson, D.J., "Robust-resistant spectrum estimation," *Proceedings of the IEEE*, Vol.70, No.9, September 1982.
- [121] Mathews, V.J. and Cho S.H., "Improved convergence analysis of stochastic gradient adaptive filters using the sign algorithm," *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol.ASSP-35, No.4, April 1987.
- [122] Holland, P.W. and Welsch, R.E., "Robust regression using iteratively reweighted least-squares," *Commun. Statist. - Theor. Meth.*, A6(9), 813-827, 1977.

- [123] Armstrong, R.D. and Frome, E.L., "A comparison of two algorithms for absolute deviation curve fitting," *Journal of the American Statistical Association, Applications Section*, Vol.71, June 1976.
- [124] Karst, O.J., "Linear curve fitting using least deviations," *Journal of the American Statistical Association*, March 1958.
- [125] Miller, J.H. and Thomas, J.B., "Detectors for discrete-time signals in non-Gaussian noise," *IEEE Trans. Information Theory*, Vol.IT-18, No.2, March 1972.
- [126] Martin, R.D., and Schwartz, S.C., "Robust detection of known signal in nearly Gaussian noise," *IEEE Trans. Information Theory*, Vol.IT-17, No.1, January 1971.
- [127] Kassam, S.A. and Thomas, J.B., "Asymptotically robust detection of a known signal in contaminated non-Gaussian noise," *IEEE Trans. Information Theory*, Vol.IT-22, No.1, January 1976.
- [128] Kassam, S.A. and Poor, H.V., "Robust techniques for signal processing: a survey," *Proceedings of the IEEE*, Vol.73, No.3, March 1985.
- [129] Kirlin, R.L. and Moghaddamjoo, A., "Robust adaptive Kalman filtering for systems with unknown step inputs and non-Gaussian measurement errors," *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol.ASSP-34, No.2, April 1986.
- [130] Rao, B.D. and Hari, K.V.S., "Statistical performance analysis of the minimum-norm method," *Int'l Conference on Acoustics, Speech and Signal Processing*, San Diego, 1989.
- [131] Bassett, G., Jr. and Koenker, R., "Asymptotic theory of least absolute error regression," *Journal of the American Statistical Association, Theory and Method Section*, Vol.73, No.363, September 1978.
- [132] Hartley, G., *Linear programming*, Addison-Wesley, 1963.
- [133] Ward, R., "An on-line adaptation for discrete  $L_1$  linear estimation," *IEEE Trans. Automatic Control*, Vol.AC-29, No.1, January 1984.
- [134] Sherman, S., "Non-mean-square error criteria," *IRE Trans. Information Theory*, September 1958.
- [135] Li, W.K., "ARMA modeling with non-Gaussian innovations," *Journal of Time Series Analysis*, Vol.9, No.2, 1988.
- [136] Scales, J.A., et al., "Fast  $L_p$  solution of large, sparse, linear systems: application to seismic travel time tomography," *Journal of Computational Physics* 75, pp.314-333, 1988.

- [137] Sengupta, D. and Kay, S., "Efficient estimation of parameters for non-Gaussian autoregressive processes," IEEE Trans. Acoustics, Speech and Signal Processing, Vol.37, No.6, June 1989.
- [138] Barrodale, I., "On computing best  $L_1$  approximations," Proceedings of the Symposium on Approximation Theory (Edited by A Talbot), Lancaster, July 1969.
- [139] Walden, A.T., "Non-Gaussian reflectivity, entropy, and deconvolution," Geophysics, Vol.50, No.12, pp.2862-2888, December 1985.
- [140] Wesolowsky, G.O., "A new descent algorithm for least absolute vales regression problem," Commun. Statist. - Simula. Computa., B10(5), pp.479-491, 1981.
- [141] Gunin, R., "Numerical algorithms for solving nonlinear  $L_p$ -norm estimation problems: Part I - a first-order gradient algorithm for well-conditioned small residual problems," Commun. Statist. - Simula., 15(3), pp.801-813, 1986.
- [142] Bershad, N.J., "On the probability density function of LMS adaptive filter weights," IEEE Trans. Acoustics, Speech and Signal Processing, Vol.ASSP-37, No.1, January 1989.
- [143] Widrow, B., et al., "Stationary and nonstationary learning characteristics of LMS adaptive filter," Proceedings of the IEEE, Vol.64, No.8, August 1976.
- [144] Masreliez, C.J., "Approximate non-Gaussian filtering with linear state and observation relations," IEEE Trans. Automatic Control, February 1975.
- [145] Masreliez, C.J. and Martin, R.D., "Robust Bayesian estimation for linear model and robustifying the Kalman filter," IEEE Trans. Automatic Control, Vol.AC-22, No.3, June 1977.
- [146] Scales, J.A. and Gersztenkorn, A., "Robust methods in inverse theory," Inverse Problems 4, pp.1071-1091, 1988.
- [147] Shao, M. and Nokias, C.L., "Signal processing with fractional lower order moments: stable process and their applications," Proceedings of the IEEE, Vol. 81, No.7, pp.986-1010, July 1993.
- [148] Abend, K. and Fritchman, "Statistical detection for communication channels with intersymbol interference," Proceedings of the IEEE, Vol. 58, pp.779-785, May 1970.
- [149] Belfiore, C.A. and Park, J.H. Jr., "Decision feedback equalization," Proceedings of the IEEE, Vol.67, pp.1143-1156, August 1979.

- [150] Biglieri, E. et al., "Adaptive cancellation of nonlinear intersymbol interference for voiceband data transmission," IEEE Journal on Selected Areas of Commun., Vol. SAC, pp.765-777, September 1984.
- [151] Duttweiler, D.L. et al., "An upper bound on the error probability in decision-feedback equalization," IEEE Trans. Inform. Theory, Vol. IT-20, pp.490-497, July 1974.
- [152] Falconer, D.D. and Salz, J., "Optimal reception of digital data over the Gaussian channel with unknown delay and phase jitters," IEEE Trans. Inform. Theory, Vol. IT-23, pp. 117-126, January 1977.
- [153] Falconer, D.D., "Adaptive reference echo cancellation," IEEE Trans. Commun., Vol. COM-30, pp. 2083-2094, September 1982.
- [154] Gitlin, R.D. et al., "On the design of gradient algorithms for digitally implemented adaptive filters," IEEE Trans. Circuits Theory, Vol. CT-20, pp. 125-136, March 1973.
- [155] Gitlin, R.D. and Weinstein, S.B., "On the required tap-weight precision for digitally-implemented adaptive mean-squared equalizers," Bell syst. Tech. J., Vol. 58, pp. 301-321, February 1979.
- [156] Grenander, N. and Szego, G., Toeplitz Forms and Their Applications, Univ of Calif. Press, Berkeley, CA, 1958.
- [157] Johnson, C.R. Jr., "Adaptive IIR filtering: current results and open issues," IEEE Trans. Inform. Theory, Vol. IT-30, pp. 237-250, March 1984.
- [158] Macchi, O. and Eweda, E., "Convergence analysis of self-adaptive equalizers," IEEE Trans. Inform. Theory, Vol. IT-30, pp. 161-176, March 1984.
- [159] Mazo, J.E., "Analysis of decision-directed equalizer convergence," Bell Syst. Tech. J., Vol. 59, pp. 1857-1876, December 1980.
- [160] Mazo, J.E., "Echo cancellation in speech and data transmission," IEEE J. Selected Areas Commun., Vol. SAC-2, pp.283-296, March 1984.
- [161] Milewski, A., "Periodic sequences with optimal properties for channel estimation and fast start-up equalization," IBM J. Res. Develop., Vol. 27, pp.426-431, September 1983.
- [162] Qureshi, S.U.H., and Newhall, E.E., "An adaptive receiver for data transmission over time-dispersive channels," IEEE Trans. Inform. Theory, Vol. IT-19, pp. 448-457, July 1973.

[163] Siller, C.A. Jr., "Multipath propagation," IEEE Commun. Mag., Vol. 22, pp. 6-15., February 1984.

[164] Ungerboeck, G., "Theory on the speed of convergence in adaptive equalizers for digital communication," IBM J. Res. Devel., Vol. 16, pp. 546-555, November 1972.

[165] Weinstein, S.B., "A baseband data-driven echo canceller for full-duplex transmission on two-wire circuits," IEEE Trans. Commun., Vol. COM-25, pp. 654-666, July 1977.

[166] Monsen, P., "Theoretical and measured performance of a DFE modem on a fading multipath channel," IEEE Trans. Commun., Vol. COM-25, pp. 1144-1153, October 1977.



## APPENDIX A

### THE $L_1$ AND $L_2$ NORMS AND THEIR PROPERTIES

#### 1. Linear Norm Space

Before we start our discussion of the  $L_1$  norm, we review some of the basic concepts of “linear norm space”.

**Definition:** Let  $\mathbf{E}$  be a real linear space. If for every element  $\mathbf{x} \in \mathbf{E}$ , there exists a real number  $\|\mathbf{x}\|$  such that

$$\begin{aligned} \|\mathbf{x}\| &\geq 0; \|\mathbf{x}\| = 0 \text{ if and only if } \mathbf{x} = \mathbf{0}; \\ \|\alpha\mathbf{x}\| &= |\alpha| \cdot \|\mathbf{x}\|, \text{ for any real number } \alpha; \\ \|\mathbf{x} + \mathbf{y}\| &\leq \|\mathbf{x}\| + \|\mathbf{y}\| \text{ for any } \mathbf{x}, \mathbf{y} \in \mathbf{E} \end{aligned} \tag{A.1}$$

then  $\mathbf{E}$  is called a **real linear norm space**, or simply **real norm space**.  $\|\mathbf{x}\|$  is called the **norm of  $\mathbf{x}$** .

From the definition, we know that the  $M$ -dimensional Euclidean space  $R^M$  is a norm space.

**Definition:** Let  $\mathbf{E}$  be a linear space, and let  $\mathbf{V}$  be a subset of  $\mathbf{E}$ . If for any two points  $\mathbf{x}, \mathbf{y}$  in  $\mathbf{V}$ , the segment connecting the two points

$$\{ \tau \cdot \mathbf{x} + (1 - \tau) \cdot \mathbf{y}, \text{ where } \tau \text{ is real, } 0 \leq \tau \leq 1 \} \tag{A.2}$$

is also in  $\mathbf{V}$ , then  $\mathbf{V}$  is called a **convex set**. (A.2) is called the **convex combination of  $\mathbf{x}$  and  $\mathbf{y}$** .

From the definition, the set of all non-negative real numbers is a convex set.

**Definition:** a function  $d(\cdot)$  in the norm space  $\mathbb{E}$  is a mapping from a convex set  $V$  to the set of the non-negative real numbers. The function is said to be **convex** if the function of the convex combination of  $\mathbf{x}, \mathbf{y} \in V$  is not larger than the convex combination of the functions of  $\mathbf{x}$  and  $\mathbf{y}$ , i.e.

$$d\{\tau \cdot \mathbf{x} + (1 - \tau) \cdot \mathbf{y}\} \leq \tau \cdot d(\mathbf{x}) + (1 - \tau) \cdot d(\mathbf{y}) \quad (\text{A.3})$$

**Definition:** Let  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M$  be  $M$  vectors in the linear sub-space  $V$ . If any vector  $\mathbf{y} \in V$  can be expressed as a linear combination of  $\mathbf{x}_i$  ( $i=1, 2, \dots, M$ )

$$\mathbf{y} = a_1 \cdot \mathbf{x}_1 + a_2 \cdot \mathbf{x}_2 + \dots + a_M \cdot \mathbf{x}_M \quad (\text{A.4})$$

then we say  $\mathbf{x}_i$  ( $i=1, 2, \dots, M$ ) spans  $V$ .

**Definition:** A **convex hull** of a set of  $M$  points in a convex space  $V$  is the smallest convex polygon  $S$  for which each point in  $V$  is either on the boundary of  $S$  or in its interior.

Figure A-1 shows a convex hull in the two dimensional Euclidean space.

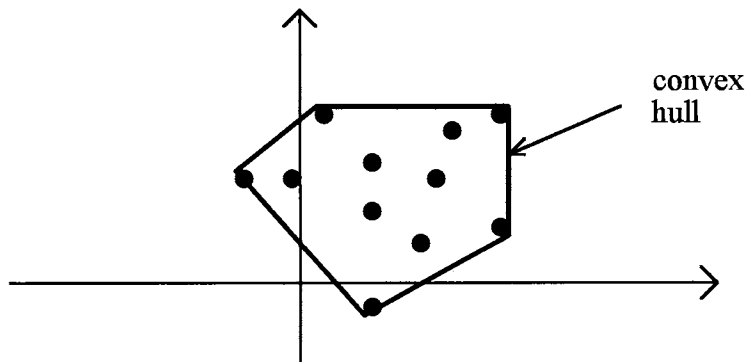


Figure A-1. Two dimensional convex hull

## 2. The $L_1$ Norm and Its Properties

Let  $p=1$  in the  $d_p(\mathbf{c})$  defined in (2.4), we have the  $L_1$  norm

$$d_1(\mathbf{c}) = \sum_{i=1}^N |r_i(\mathbf{c})| \quad (\text{A.5})$$

Before we formalize the properties of the  $L_1$  norm, we summarize the basic problems of the  $L_1$  solution to the equations of (2.1), in the following:

- a. find all points  $\mathbf{c}$  which satisfy  $M$  out of the  $N$  equations in (2.1) by checking all the combinations;
- b. evaluate  $d_1(\mathbf{c})$  for each of these points;
- c. if there is a points  $\mathbf{c}_i$  such that  $d_p(\mathbf{c}_i) < d_p(\mathbf{c}_j)$  for all  $j \neq i$ , then  $\mathbf{c}_i$  is the unique solution;
- d. the solution may not be unique.

The computation involved in step (a) may be enormous. The properties of the  $L_1$  norm in the following will help to better understand the  $L_1$  norm and its solution set, leading to the alternative algorithms to find the  $L_1$  solution.

### (1) Continuity and Convexity of $L_1$ Norm

**Property 1:**  $d_1(\mathbf{c})$  is continuous and convex

**Proof:** Assume

$$\mathbf{a} = [a_1 \ a_2 \ \dots \ a_M]^T$$

$$\begin{aligned}
\mathbf{b} &= [\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_M]^T \\
\mathbf{X} &= [x_{ij}]_{N \times M} \\
\mathbf{x}_i &= [x_{i1} \ x_{i2} \ \dots \ x_{iM}]^T \\
\mathbf{y} &= [y_1 \ y_2 \ \dots \ y_N]^T
\end{aligned} \tag{A.6}$$

then

$$\begin{aligned}
|d_1(\mathbf{a}) - d_1(\mathbf{b})| &= \left| \sum_{i=1}^N |r_i(\mathbf{a})| - \sum_{i=1}^N |r_i(\mathbf{b})| \right| \\
&= \left| \sum_{i=1}^N \left| y_i - \sum_{j=1}^M a_j x_{ij} \right| - \sum_{i=1}^N \left| y_i - \sum_{j=1}^M b_j x_{ij} \right| \right| \\
&\leq \sum_{i=1}^N \left| \{y_i - \sum_{j=1}^M a_j x_{ij}\} - \{y_i - \sum_{j=1}^M b_j x_{ij}\} \right| \\
&= \sum_{i=1}^N \left| \sum_{j=1}^M (a_j - b_j) x_{ij} \right| \\
&= \sum_{i=1}^N |(\mathbf{a} - \mathbf{b})^T \mathbf{x}_i|
\end{aligned} \tag{A.7}$$

If  $\mathbf{a} \rightarrow \mathbf{b}$ , then  $|d_1(\mathbf{a}) - d_1(\mathbf{b})| \rightarrow 0$ . Therefore  $d_1(\mathbf{c})$  is continuous.

The convexity can be shown as follows:

Assume  $\tau \in [0, 1]$ , then

$$\begin{aligned}
d_1(\tau \cdot \mathbf{a} + (1 - \tau) \cdot \mathbf{b}) &= \sum_{i=1}^N |r_i(\tau \cdot \mathbf{a} + (1 - \tau) \cdot \mathbf{b})| \\
&= \sum_{i=1}^N \left| y_i - \sum_{j=1}^M [\tau \cdot a_j + (1 - \tau) \cdot b_j] x_{ij} \right| \\
&= \sum_{i=1}^N \left| \tau \{y_i - \sum_{j=1}^M a_j x_{ij}\} + (1 - \tau) \{y_i - \sum_{j=1}^M b_j x_{ij}\} \right| \\
&= \sum_{i=1}^N |\tau \cdot r_i(\mathbf{a}) + (1 - \tau) \cdot r_i(\mathbf{b})| \\
&\leq \tau \sum_{i=1}^N |r_i(\mathbf{a})| + (1 - \tau) \sum_{i=1}^N |r_i(\mathbf{b})| \\
&= \tau \cdot d_1(\mathbf{a}) + (1 - \tau) \cdot d_1(\mathbf{b})
\end{aligned} \tag{A.8}$$

Therefore,  $d_1(\mathbf{c})$  is convex. ■

## (2) Convexity of the $L_1$ Solution Set

The  $L_1$  solution set  $S$  is defined as the set of all vectors that minimize the  $L_1$  norm.

Formally,

$$S = \{ \mathbf{c} \in R^M : d_1(\mathbf{c}) \leq d_1(\mathbf{a}) \text{ for all } \mathbf{a} \in R^M \} \quad (\text{A.9})$$

**Property 2:**  $S$  is non-empty and bounded.

The proof of this property can be found in [74].

**Property 3:**  $S$  is convex.

**Proof:** Assume  $\mathbf{a}, \mathbf{b} \in S$ , then  $d_1(\mathbf{a}) = d_1(\mathbf{b}) = \min$ . Assume  $\tau \in [0, 1]$ , then

$$\begin{aligned} d_1(\tau \cdot \mathbf{a} + (1-\tau) \cdot \mathbf{b}) &= \sum_{i=1}^N |y_i - [\tau \cdot \mathbf{a} + (1-\tau) \cdot \mathbf{b}]^T \mathbf{x}_i| \\ &= \sum_{i=1}^N |\tau \{y_i - \mathbf{a}^T \mathbf{x}_i\} + (1-\tau) \{y_i - \mathbf{b}^T \mathbf{x}_i\}| \\ &\leq \tau \sum_{i=1}^N |y_i - \mathbf{a}^T \mathbf{x}_i| + (1-\tau) \sum_{i=1}^N |y_i - \mathbf{b}^T \mathbf{x}_i| \\ &= \tau \cdot d_1(\mathbf{a}) + (1-\tau) \cdot d_1(\mathbf{b}) \\ &= d_1(\mathbf{a}) \end{aligned} \quad (\text{A.10})$$

Since  $d_1(\mathbf{a}) = \min$ ,  $d_1(\tau \cdot \mathbf{a} + (1-\tau) \cdot \mathbf{b}) = \min$ . Therefore,

$$\tau \cdot \mathbf{a} + (1-\tau) \cdot \mathbf{b} \in S \quad (\text{A.11})$$

and  $S$  is convex. ■

From property 2,  $S$  has at least one element. If  $S$  has more than one element, by convexity, it will have infinite number of elements. Therefore, the  $L_1$  solution may not be

unique. Although  $S$  may have infinite number of elements, it has certain structure. This structure is characterized by some special elements of  $S$  called the **extreme points**.

Define an index set

$$Z = \{i: r_i(\mathbf{c}) = 0\} \quad (\text{A.12})$$

A point  $\mathbf{c} \in R^M$  is said to be **extreme** if  $\{\mathbf{x}_i, i \in Z\}$  spans  $R^M$ . In other words, any other vectors in  $R^M$  can be expressed as a linear combination of  $\{\mathbf{x}_i, i \in Z\}$ . Therefore, the number of elements in  $Z$ , denoted as  $|Z|$ , has to be greater than or equal to  $M$ . That is to say, an extreme point  $\mathbf{c}$  must satisfy at least  $M$  out of the  $N$  ( $N \geq M$ ) equations in (2.5).

An extreme point may or may not be an  $L_1$  solution. If it is an  $L_1$  solution, then it can not be expressed as a convex combination of other  $L_1$  solutions (see [74] for the proof). On the other hand, if  $\mathbf{c} \in S$  is not extreme, then it is a convex combination of the other points in  $S$ . Therefore,  $S$  has the following property:

**Property 4:**  $S$  is the convex hull of the finite set of its extreme points.

The proof of this property can be found in [74].

**Example A-1.** Assume we have a set of linear equations with  $M=1$ ,  $N=2$ :

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \mathbf{c} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

The  $L_1$  norm of the residual vector is

$$\begin{aligned} d_1(\mathbf{c}) &= |r_1(\mathbf{c})| + |r_2(\mathbf{c})| \\ &= |1 - c| + |2 - c| \end{aligned}$$

Figure A-2 shows  $d_1(\mathbf{c})$  versus  $c$ .

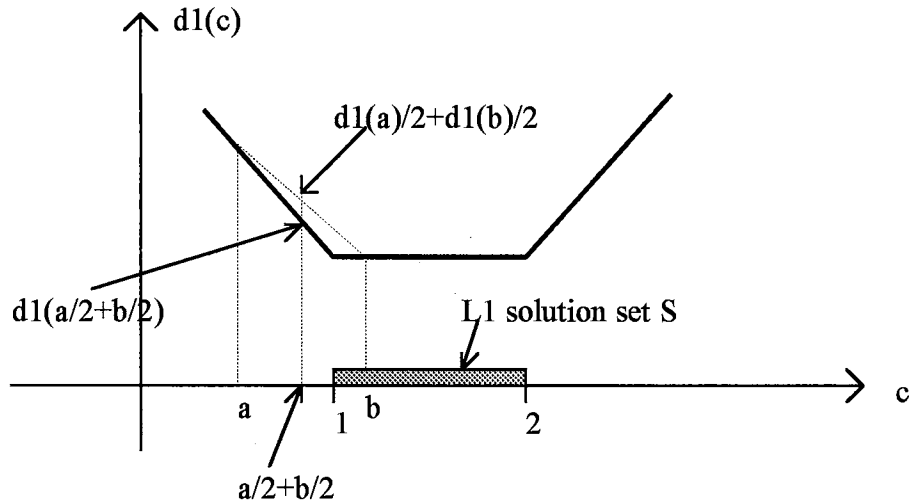


Figure A-2.  $d_1(c)$  versus  $c$  for example A-1

From the figure, we can see that

(a)  $d_1(c)$  is piece-wise continuous;

(b) for a point between  $a$  and  $b$ , say,  $a/2 + b/2$ , we have

$d_1(a/2 + b/2) < d_1(a)/2 + d_1(b)/2$ , therefore  $d_1(c)$  is convex;

(c) the  $L_1$  solution set is  $S = \{c: 1 \leq c \leq 2\}$ . It is non-empty and bounded;

(d)  $S$  is convex, because for any  $c_1, c_2 \in [1, 2]$  and  $\tau \in [0, 1]$ , we have

$(\tau \cdot c_1 + (1 - \tau) \cdot c_2) \in [1, 2]$ ;

(e)  $S$  is a line which is a convex hull in one-dimensional space.

From this example, we can see that the  $L_1$  solution may not be unique unless an additional constraint is imposed. One of the solution is the  $L_1/L_2$  estimator [75], or restricted least square, which finds the solution that minimizes the  $L_2$  norm within the  $L_1$  solution set.

**Example A-2.** Assume we have the same set of linear equations as in example A-

1. The  $L_2$  objective function

$$\begin{aligned} d_2(c) &= (1-c)^2 + (2-c)^2 \\ &= 2\{(c-3/2)^2 + 1/4\} \end{aligned}$$

Figure A-3 shows  $d_2(c)$  together with  $d_1(c)$ .

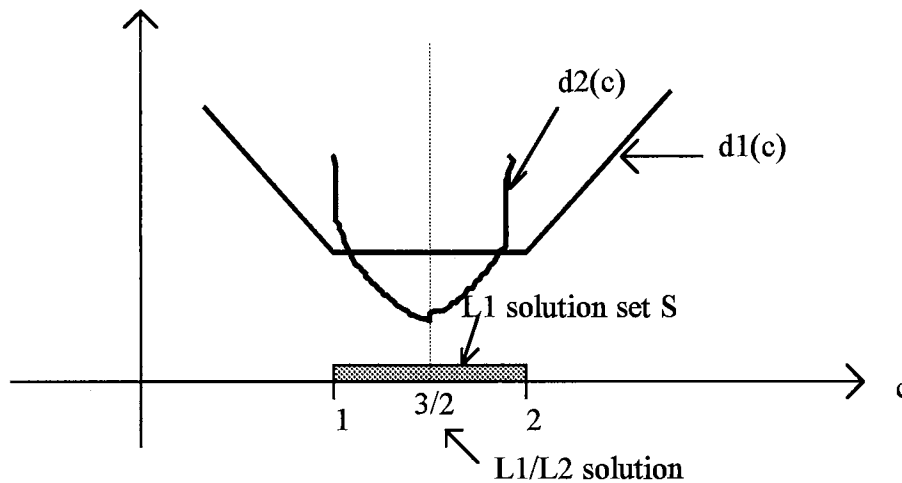


Figure A-3.  $d_1(c)$  and  $d_2(c)$  versus  $c$  for example A-2

We can see that the value of  $c$  within the  $L_1$  solution set that minimizes the  $d_2(c)$  is  $3/2$ .

Therefore,  $3/2$  is the  $L_1/L_2$  solution.

In this example, the  $L_1/L_2$  solution is also the global  $L_2$  solution. This is not always true. Sometimes, the global  $L_2$  solution may not be in the  $L_1$  solution set. In this case, the  $L_1/L_2$  solution is just the one that has least squared error within the  $L_1$  solution set.

### (3) Directional Derivative

The directional derivative of  $d_1(c)$  in the direction  $\mathbf{J}$  is defined as



$$d'_1(\mathbf{c}, \mathbf{J}) = \lim_{\tau \rightarrow 0} \{d_1(\mathbf{c} + \tau \cdot \mathbf{J}) - d_1(\mathbf{c})\} / \tau \quad (\text{A.13})$$

**Property 5:** If  $d'_1(\mathbf{c}, \mathbf{J})$  is non-negative for all directions  $\mathbf{J} \in R^M$  ( $\mathbf{J} \neq \mathbf{0}$ ), then  $\mathbf{c}$  is in  $S$ .

**Proof:** By contradiction. Otherwise, there will be a vector  $\mathbf{a} \in R^M$  ( $\mathbf{a} \neq \mathbf{c}$ ) with  $d_1(\mathbf{c}) > d_1(\mathbf{a})$ , and  $d'_1(\mathbf{c}, \mathbf{J})$  in the direction  $\mathbf{J} = \mathbf{a} - \mathbf{c}$  will be negative. ■

**Corollary:**  $\mathbf{c}$  is the unique minimizer of  $d_1(\mathbf{c})$  if and only if  $d'_1(\mathbf{c}, \mathbf{J})$  is positive for all  $\mathbf{J} \in R^M$ .

### 3. The $L_2$ Norm and Its Properties

The  $L_2$  norm is defined as in (2.2) when  $p=2$ . As mentioned earlier, for a non-negative argument, the function  $(\cdot)^{1/2}$  is a monotonically increasing function. Therefore, we concentrate on the function

$$d_2(\mathbf{c}) = \sum_{i=1}^N |r_i(\mathbf{c})|^2 \quad (\text{A.14})$$

where  $r_i(\mathbf{c})$  is the element of the residual vector

$$\mathbf{r}(\mathbf{c}) = \mathbf{y} - \mathbf{X}\mathbf{c} \quad (\text{A.15})$$

From (A.14) and (A.15), we have the vector expression of  $d_2(\mathbf{c})$ :

$$\begin{aligned} d_2(\mathbf{c}) &= \mathbf{r}^T(\mathbf{c})\mathbf{r}(\mathbf{c}) \\ &= (\mathbf{y} - \mathbf{X}\mathbf{c})^T (\mathbf{y} - \mathbf{X}\mathbf{c}) \\ &= (\mathbf{y}^T - \mathbf{c}^T \mathbf{X}^T)(\mathbf{y} - \mathbf{X}\mathbf{c}) \\ &= \mathbf{y}^T \mathbf{y} - 2\mathbf{y}^T \mathbf{X}\mathbf{c} + \mathbf{c}^T \mathbf{X}^T \mathbf{X}\mathbf{c} \end{aligned} \quad (\text{A.16})$$

From (A.16), we have the following important property of the  $L_2$  norm:

**Property:**  $d_2(\mathbf{c})$  and hence  $\|\mathbf{r}(\mathbf{c})\|_2$  is a strictly convex function of  $\mathbf{c}$ .

**Proof:** To prove that  $d_2(\mathbf{c})$  is a strictly convex function of  $\mathbf{c}$ , we need to show that for any real N-dimensional vectors  $\mathbf{c}_a$  and  $\mathbf{c}_b$ , and parameter  $0 \leq \lambda \leq 1$ , we have

$$d_2[\lambda \cdot \mathbf{c}_a + (1-\lambda)\mathbf{c}_b] \leq \lambda d_2(\mathbf{c}_a) + (1-\lambda)d_2(\mathbf{c}_b) \quad (\text{A.17})$$

and the equality holds only when  $\mathbf{c}_a = \mathbf{c}_b$ . In order to do that, we expand the left-hand side of (A.17) according to (A.16):

$$\begin{aligned} & d_2[\lambda \cdot \mathbf{c}_a + (1-\lambda)\mathbf{c}_b] \\ &= \mathbf{y}^T \mathbf{y} - 2\mathbf{y}^T \mathbf{X}[\lambda \cdot \mathbf{c}_a + (1-\lambda)\mathbf{c}_b] \\ &+ [\lambda \cdot \mathbf{c}_a + (1-\lambda)\mathbf{c}_b]^T \mathbf{X}^T \mathbf{X}[\lambda \cdot \mathbf{c}_a + (1-\lambda)\mathbf{c}_b] \\ &= \mathbf{y}^T \mathbf{y} - 2\lambda \mathbf{y}^T \mathbf{X} \mathbf{c}_a - 2(1-\lambda) \mathbf{y}^T \mathbf{X} \mathbf{c}_b \\ &+ \lambda^2 \mathbf{c}_a^T \mathbf{X}^T \mathbf{X} \mathbf{c}_a + 2\lambda(1-\lambda) \mathbf{c}_a^T \mathbf{X}^T \mathbf{X} \mathbf{c}_b + (1-\lambda)^2 \mathbf{c}_b^T \mathbf{X}^T \mathbf{X} \mathbf{c}_b \end{aligned} \quad (\text{A.18})$$

The right-hand side of (A.17) is expanded as

$$\begin{aligned} & \lambda d_2(\mathbf{c}_a) + (1-\lambda)d_2(\mathbf{c}_b) \\ &= \mathbf{y}^T \mathbf{y} - 2\lambda \mathbf{y}^T \mathbf{X} \mathbf{c}_a - 2(1-\lambda) \mathbf{y}^T \mathbf{X} \mathbf{c}_b \\ &+ \lambda \mathbf{c}_a^T \mathbf{X}^T \mathbf{X} \mathbf{c}_a + (1-\lambda) \mathbf{c}_b^T \mathbf{X}^T \mathbf{X} \mathbf{c}_b \end{aligned} \quad (\text{A.19})$$

Canceling common terms in (A.18) and (A.19), we need to show that

$$\begin{aligned} & \lambda^2 \mathbf{c}_a^T \mathbf{X}^T \mathbf{X} \mathbf{c}_a + 2\lambda(1-\lambda) \mathbf{c}_a^T \mathbf{X}^T \mathbf{X} \mathbf{c}_b + (1-\lambda)^2 \mathbf{c}_b^T \mathbf{X}^T \mathbf{X} \mathbf{c}_b \\ & \leq \lambda \mathbf{c}_a^T \mathbf{X}^T \mathbf{X} \mathbf{c}_a + (1-\lambda) \mathbf{c}_b^T \mathbf{X}^T \mathbf{X} \mathbf{c}_b \end{aligned} \quad (\text{A.20})$$

or

$$0 \leq \lambda(1-\lambda)[\mathbf{c}_a^T \mathbf{X}^T \mathbf{X} \mathbf{c}_a - 2\mathbf{c}_a^T \mathbf{X}^T \mathbf{X} \mathbf{c}_b + \mathbf{c}_b^T \mathbf{X}^T \mathbf{X} \mathbf{c}_b] \quad (\text{A.21})$$

Since  $0 \leq \lambda \leq 1$ , except for the trivial cases where  $\lambda=0$  or  $\lambda=1$  (then (A.17) is trivially true),

we always have  $\lambda(1-\lambda) > 0$ . Dividing (A.21) by  $\lambda(1-\lambda)$  and combining the quadratic forms

yields

$$0 \leq (\mathbf{c}_a - \mathbf{c}_b)^T \mathbf{X}^T \mathbf{X} (\mathbf{c}_a - \mathbf{c}_b) \quad (\text{A.22})$$

or

$$0 \leq [\mathbf{X}(\mathbf{c}_a - \mathbf{c}_b)]^T [\mathbf{X}(\mathbf{c}_a - \mathbf{c}_b)] \quad (\text{A.23})$$

The right hand side of (A.23) is the  $L_2$  norm of the vector  $\mathbf{X}(\mathbf{c}_a - \mathbf{c}_b)$ , and by definition, is always non-negative. The equality of (A.23) will hold only when  $\mathbf{c}_a - \mathbf{c}_b = \mathbf{0}$ , i.e.,  $\mathbf{c}_a = \mathbf{c}_b$ .

Therefore, we have proved that (A.17) is true, and  $d_2(\mathbf{c})$  is strictly convex. Since the function  $(\cdot)^{1/2}$  is also convex, we have that the  $L_2$  norm of  $\mathbf{r}(\mathbf{c})$  is also strictly convex.

The strict convexity of  $d_2(\mathbf{c})$  indicates that unlike the case of  $d_1(\mathbf{c})$  there is a **unique**  $\mathbf{c}$  that minimizes  $d_2(\mathbf{c})$ . This will allow us to use the steepest descent algorithm to find that unique  $\mathbf{c}$ .

## APPENDIX B

### WIENER FILTER AND THE MSE EQUALIZER

Since both the Wiener filter and the MSE equalizer are based on the minimum mean squares (MSE) criterion, they bear a lot of similarity. Indeed, as will be shown in the following, the MSE equalizer is a kind of Wiener filter in special case.

#### 1. Noncausal Wiener Filter and the Linear MSE Equalizer

Consider the smoothing problem in the following [76]. We wish to estimate the present value of a sequence  $u(n)$  given the values of  $w(i)$  for every  $i$  from  $-\infty$  to  $\infty$ :

$$w(n) = u(n) + \eta(n) \quad (\text{B.1})$$

where  $\eta(n)$  is a stationary random sequence representing the noise in the measurement.

The desirable estimate is formed as a linear function of the value of  $w(i)$ :

$$\hat{u}(n) = \sum_{j=-\infty}^{\infty} c(j)w(n-j) \quad (\text{B.2})$$

where  $c(j)$  is the impulse response of a time-invariant noncausal system. We want to find the  $c(j)$  such that the mean square value of the estimation error

$$\mathcal{E}\{[u(n) - \hat{u}(n)]^2\}$$

is minimum.

From the orthogonality principle [76], we know that the estimation error so obtained is orthogonal to the  $w(i)$  for all  $i$ , i.e.,

$$\mathcal{E}\{[u(n) - \hat{u}(n)]w(i)\} = 0 \quad \text{for all } i, \quad (\text{B.3})$$

From (B.2) and (B.3) and by setting  $i=n-m$ , we have

$$\mathcal{E}\{[u(n) - \sum_{j=-\infty}^{\infty} c(j)w(n-j)]w(n-m)\} = 0 \quad (\text{B.4})$$

or

$$R_{uw}(m) = \sum_{j=-\infty}^{\infty} c(j)R_w(m-j) \quad (\text{B.5})$$

where  $R_{uw}$  is the cross correlation between  $u(n)$  and  $w(n)$  and  $R_w$  is the autocorrelation of  $w(n)$ . Taking the discrete Fourier transform of both sides of (B.5), we obtain

$$S_{uw}(\omega) = C(\omega)S_w(\omega) \quad (\text{B.6})$$

where  $S_{uw}(\omega)$  is the cross-spectral density of  $u(n)$  and  $w(n)$ , and  $S_w(\omega)$  is the power spectral density of  $w(n)$ . From (B.6), we have

$$C(\omega) = \frac{S_{uw}(\omega)}{S_w(\omega)} \quad (\text{B.7})$$

The linear system  $C(\omega)$  in (B.7) is called the **noncausal Wiener filter**.

If the signal  $u(n)$  and noise  $\eta(n)$  are uncorrelated, i.e.,

$$\mathcal{E}\{u(n)\eta(n)\} = 0 \quad (\text{B.8})$$

and  $\eta(n)$  is white noise with zero mean and spectral density of  $N_0$ , i.e.,

$$\mathcal{E}\{\eta^2(n)\} = N_0 \cdot \delta(n) \quad (\text{B.9})$$

then

$$\begin{aligned} S_{uw}(\omega) &= S_u(\omega) \\ S_w(\omega) &= S_u(\omega) + N_0 \end{aligned} \quad (\text{B.10})$$

From (B.7) and (B.10), we have

$$C(\omega) = \frac{S_u(\omega)}{S_u(\omega) + N_0} \quad (\text{B.11})$$

Now we come to the equalization problem. If we model the channel as a linear system with an impulse response of  $f(n)$  and the transmitted signal  $x(n)$ , then the received signal without noise is

$$u(n) = f(n) * x(n) \quad (\text{B.12})$$

where “ $*$ ” stands for the convolution. Furthermore, as we have done in Chapter III (equation (3.3)), we assume that the transmitted signal  $x(n)$  is uncorrelated at different sampling instances, i.e.,

$$\mathcal{E}\{x(n)x(n-m)\} = \sigma_x^2 \cdot \delta(n) \quad (\text{B.13})$$

Then from (B.12) and (B.13) and the fact that the Fourier transform of the autocorrelation function is the power spectral density, we obtain

$$S_u(\omega) = |F(\omega)|^2 \cdot \sigma_x^2 \quad (\text{B.14})$$

From (B.11) and (B.14), we have

$$C(\omega) = \frac{|F(\omega)|^2}{|F(\omega)|^2 + \frac{N_0}{\sigma_x^2}} \quad (\text{B.15})$$

$C(\omega)$  of (B.15) is a filter to estimate  $u(n)$  from  $w(n)$ . From the estimate of  $u(n)$ , we can estimate  $x(n)$  if we know the channel impulse response  $f(n)$ . To do so, we take the Discrete Fourier transform of both sides of (B.12), resulting in

$$U(\omega) = F(\omega) X(\omega)$$

or

$$X(\omega) = \frac{U(\omega)}{F(\omega)} \quad (\text{B.16})$$

Therefore, from  $w(n)$ , we can estimate the transmitted signal  $x(n)$  by using a “compound filter”

$$\begin{aligned} H(\omega) &= \frac{C(\omega)}{F(\omega)} \\ &= \frac{|F(\omega)|^2 / F(\omega)}{|F(\omega)|^2 + \frac{N_0}{\sigma_x^2}} \\ &= \frac{F^*(\omega)}{|F(\omega)|^2 + \frac{N_0}{\sigma_x^2}} \end{aligned} \quad (\text{B.17})$$

Comparing (B.17) with (3.20), we can see that this “compound filter” is indeed a linear MSE equalizer. That is to say, the linear MSE equalizer with infinite length is a special noncausal Wiener filter.

## 2. Prediction and the MSE Feedback Equalizer

In Chapter III, we derived the MSE feedback equalizer based on the linear prediction model. An alternative method is to use the Wiener-Hopf theory, as shown in the following.

In Chapter III, we have shown that the best choice of the feedback portion to minimize the overall MSE is to set the feedback filter coefficients to the corresponding postcursor sample values of the combined channel impulse response, i.e.,

$$g(j) = q(j) \quad \text{for } j = 1, 2, 3, \dots \quad (\text{B.18})$$

With this choice, from (3.48), we know that the overall MSE is

$$MSE = \sigma_x^2 \cdot \sum_{j=-\infty}^{-1} q^2(j) + \sigma_x^2 [q(0) - 1]^2 + N_0 \cdot \sum_{j=-\infty}^0 h^2(j) \quad (\text{B.19})$$

where  $h(j)$  is the coefficient of the forward filter (note that  $h(j)=0$  for  $j > 0$ ),  $\sigma_x^2$  is the variance of the transmitted signal  $x(n)$  and  $N_0$  is the power spectral density of the noise.

(B.19) can be rearranged as

$$\frac{MSE}{\sigma_x^2} = 1 + \sum_{j=-\infty}^0 q^2(j) - 2q(0) + \frac{N_0}{\sigma_x^2} \sum_{j=-\infty}^0 h^2(j) \quad (\text{B.20})$$

Since

$$q(j) = \sum_{m=-\infty}^0 h(m) f(j-m) \quad (\text{B.21})$$

where  $f(j)$  is the channel impulse response, taking the first variation of (B.20) with respect to  $h(n)$  yields

$$f(-n) = \frac{N_0}{\sigma_x^2} h(n) + \sum_{m=-\infty}^0 q(m) f(m-n) \quad (\text{B.22})$$

(B.21) can be rearranged as

$$\begin{aligned} h(n) &= \frac{\sigma_x^2}{N_0} \left\{ f(-n) - \sum_{m=-\infty}^0 q(m) f(m-n) \right\} \\ &= \sum_{m=-\infty}^0 \rho(m) f(m-n) \\ &= \rho(n) * f(-n) \end{aligned} \quad (\text{B.23})$$

where



$$\begin{aligned}
\rho(m) &= \frac{\sigma_x^2}{N_0} [1 - q(0)], & m = 0 \\
&-\frac{\sigma_x^2}{N_0} q(m), & m \leq -1 \\
&0 & m > 0
\end{aligned} \tag{B.24}$$

Equation (B.23) indicates that the forward filter  $h(n)$  of the MSE feedback equalizer consists of a matched filter,  $f(-n)$ , followed by a **one-sided** (anticausal) tapped delay line with weights of  $\rho(m)$ . This structure will allow the forward filter to minimize the mean square precursor ISI and the filtered noise, leaving only the postcursor to be compensated for by the feedback portion.

(B.24) does not provide the coefficients  $\rho(m)$ , since  $q(m)$  is a function of  $h(m)$  which is the unknown. To find  $\rho(m)$ , we define the autocorrelation function of  $f(-n)$

$$R_f(k) = R_f(-k) = \sum_n f(-n)f(k-n) \tag{B.25}$$

Multiplying both sides of (B.22) by  $f(k-n)$  and summing over  $n$  yields

$$\sum_n f(-n)f(k-n) = \frac{N_0}{\sigma_x^2} \sum_n h(n)f(k-n) + \sum_{m=-\infty}^0 q(m) \sum_n f(m-n)f(k-n) \tag{B.26}$$

From (B.25) and (B.26), we obtain

$$R_f(k) = \frac{N_0}{\sigma_x^2} q(k) + \sum_{m=-\infty}^0 q(m)R_f(m-k) \tag{B.27}$$

(B.27) is the well-known Wiener-Hopf equations [76]. Solving these equations is beyond the scope of this paper. Using the method introduced in [11], one can show that the discrete Fourier transform of the sequence  $q(n)$  ( $-\infty \leq m \leq 0$ ) is

$$\begin{aligned}
Q(\omega) &= \sum_{n=-\infty}^0 q(n)e^{jn\omega} \\
&= 1 - \frac{N_0 / \sigma_x^2}{\Theta^-(\omega)\gamma_0}
\end{aligned}
\tag{B.28}$$

where the power spectral density is factored according to

$$\begin{aligned}
\Theta(\omega) &= \Theta^+(\omega)\Theta^-(\omega) \\
&= S_f(\omega) + \frac{N_0}{\sigma_x^2} \\
&= \sum_{n=-\infty}^{\infty} \varphi_n e^{jn\omega}
\end{aligned}
\tag{B.29}$$

with the one-sided transforms

$$\begin{aligned}
\Theta^+(\omega) &= \sum_{n=0}^{\infty} \gamma_n e^{jn\omega} \\
\Theta^-(\omega) &= \sum_{n=-\infty}^0 \gamma_n e^{jn\omega}
\end{aligned}
\tag{B.30}$$

and  $S_f(\omega)$  is the power spectral density of the sequence  $f(-n)$ .

## APPENDIX C

### CLOSED FORM EXPRESSIONS OF MSE FOR FEEDBACK EQUALIZERS WITH INFINITE LENGTH

#### 1. Residue Theorem and Jensen's Formula [77]

If an analytic function  $f(z)$  is single-valued in a domain  $D$  and is regular there except at a point  $p_1$  of  $D$ , then  $f(z)$  may be expanded in the vicinity of  $p_1$  with Laurent series

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - p_1)^n \quad (\text{C.1})$$

The coefficient  $a_{-1}$  is called the **residue** of  $f(z)$  at the singular point  $z = p_1$ ,

$$R_{p_1} = a_{-1} = \frac{1}{2\pi j} \int_{c_1} f(u) du \quad (\text{C.2})$$

Where  $c_1$  is a closed contour surrounding  $z = p_1$  and, except at  $z = p_1$ ,  $f(z)$  is regular within and on  $c_1$ . Then we have the following **residue theorem**:

If  $f(z)$  is single-valued and regular within and on the closed contour  $c$  except at the  $n$  singular points  $p_1, p_2, \dots, p_N$ , then,

$$\int_c f(u) du = 2\pi j \sum_{i=1}^N R_{p_i} \quad (\text{C.3})$$

Where  $R_i$  ( $i=1, \dots, N$ ) is the residue of  $f(z)$  at the singular point  $p_i$ .

While the residue theorem establishes the relationship between the contour integral of a function and its poles, the following **Jensen's formula** [78] establishes the relationship between the contour integral of a function and its zeros:

If  $f(z)$  is **analytic** within and on the closed contour  $c$ , and  $f(z)$  has  $M$  zeros  $z_1, z_2, \dots, z_m$  inside  $c$ , and  $f(0) \neq 0$ , then

$$\int_c |f(z)| dz = 2\pi j \left[ \ln |f(0)| - \sum_{i=1}^M \ln |z_i| \right] \quad (C.4)$$

## 2. Evaluation of MSE for Feedback Equalizers

We chose the unit circle as the contour  $c$  along which we are going to integrate. In Equation (3-56), we have assumed that the linear predictor  $1+G(z)$  is causal, and has all its poles and zeros inside the unit circle. Therefore,  $1+G^*(1/z^*)$  has all its poles and zeros outside the unit circle, and hence  $1+G^*(1/z^*)$  is analytic inside the unit circle. In addition,  $1+G^*(1/z^*)$  is anti-causal. It can be expanded with only positive power of  $z$ . Therefore  $[1+G^*(1/z^*)]|_{z=0}=1$ .

Now we let  $f(z) = 1+G^*(1/z^*)$  in (A4), then the first term of (C.4) is 0. The second term is also 0 because  $1+G^*(1/z^*)$  has no zeros inside the unit circle. Therefore

$$T \int_{-1/2T}^{1/2T} \ln |1+G^*(e^{j2\pi fT})| df = 0 \quad (C.5)$$

Furthermore,

$$|1+G(e^{j\omega T})| = |1+G^*(e^{j\omega T})| \quad (C.6)$$

Taking logarithm of both sides of (3-56) and integrating along the unit circle lead to

$$\begin{aligned}
& \ln \left| \frac{\mathcal{E}^2(ZF - FE)}{N_o} \right| + T \int_{-1/2T}^{1/2T} \ln |F(e^{j2\pi f T})|^2 df \\
&= T \int_{-1/2T}^{1/2T} \ln |1 + G^*(e^{j2\pi f T})|^2 df = 0
\end{aligned} \tag{C.7}$$

Therefore

$$\begin{aligned}
\mathcal{E}^2(ZF - FE) &= N_o \exp \left\{ -T \int_{-1/2T}^{1/2T} \ln |F(e^{j2\pi f T})|^2 df \right\} \\
&= N_o \exp \left\{ T \int_{-1/2T}^{1/2T} \ln [1 / |F(e^{j2\pi f T})|^2] df \right\}
\end{aligned} \tag{C.8}$$

Similarly, taking the logarithm of both sides of (3-74) and integrating along the unit circle lead to

$$\mathcal{E}^2(MSE - FE) = N_o \exp \left\{ T \int_{-1/2T}^{1/2T} \ln [1 / (|F(e^{j2\pi f})|^2 + N_o / \sigma_x^2)] df \right\} \tag{C.9}$$

and doing the same thing on (3-88) leads to

$$\mathcal{E}^2(MN - FE) = N_o \exp \left\{ T \int_{-1/2T}^{1/2T} \ln [1 / (|F(e^{j2\pi f})|^2 + \lambda N_o / \sigma_x^2)] df \right\} \tag{C.10}$$

Since the functions  $\exp(\cdot)$  and  $\ln(\cdot)$  are both monotonically increasing functions, it is obvious from (C.8), (C.9) and (C.10) that

$$\mathcal{E}^2(MSE - FE) \leq \mathcal{E}^2(MN - FE) \leq \mathcal{E}^2(ZF - FE) \tag{C.11}$$

## APPENDIX D

### COMPARISON OF MSE BETWEEN LINEAR AND FEEDBACK EQUALIZERS

It has been shown in [77] that for a function  $f(z)$  which is analytic within and on the closed contour  $c$ ,

$$\exp \left[ \int_c f(z) dz \right] \leq \int_c \exp [f(z)] dz \quad (D.1)$$

and the equality is established when  $f(z)$  is constant along the contour  $c$ .

If we let  $c$  be the unit circle and

$$f(z) = \ln[1/|F(z)|^2] \quad (D.2)$$

Then from (D.1), we have

$$\exp \left[ T \int_{-1/2T}^{1/2T} \ln[1/|F(e^{j2\pi f})|^2] df \right] \leq T \int_{-1/2T}^{1/2T} 1/|F(e^{j2\pi f})|^2 df \quad (D.3)$$

From (3-10) and (3-62), we have

$$\mathcal{E}^2(ZF - FE) \leq \mathcal{E}^2(ZF - LE) \quad (D.4)$$

Similarly, we have

$$\mathcal{E}^2(MSE - FE) \leq \mathcal{E}^2(MSE - LE) \quad (D.5)$$

and

$$\mathcal{E}^2(MN - FE) \leq \mathcal{E}^2(MN - LE) \quad (D.6)$$

## APPENDIX E

### AN ALTERNATIVE MODEL OF THE GHOSTING PROBLEM

In the ghost cancellation literature, the ghosting process is often modeled according to the causality [57]. The overall ghosting channel is modeled as an FIR filter with transfer function

$$F(z) = K(z) G(z) \quad (\text{E.1})$$

where  $K(z)$  is an anti-causal system and  $G(z)$  is a causal system.

The ghost canceling system is basically an inverse system of the ghosting system. In the postcursor case, the ghost is the time-delayed (and possibly phase-shifted) version of the main signal:

$$u(n) = x(n) + g_1 \cdot x(n-1) + g_2 \cdot x(n-2) + \dots + g_N \cdot x(n-N) \quad (\text{E.2})$$

The transfer function of the ghosting system is

$$G(z) = 1 + g_1 \cdot z^{-1} + g_2 \cdot z^{-2} + \dots + g_N \cdot z^{-N} \quad (\text{E.3})$$

The inverse filter is an IIR filter with transfer function of

$$\begin{aligned} T(z) &= \frac{1}{G(z)} \\ &= \frac{1}{1 + g_1 \cdot z^{-1} + g_2 \cdot z^{-2} + \dots + g_N \cdot z^{-N}} \end{aligned} \quad (\text{E.4})$$

In the precursor case, the ghost is the time-advanced (and possibly phase-shifted) version of the main signal:

$$u(n) = k_0 \cdot x(n) + k_1 \cdot x(n+1) + k_2 \cdot x(n+2) + \dots + k_P \cdot x(n+P) \quad (\text{E.5})$$

The transfer function of the ghosting system is

$$K(z) = k_0 + k_1 \cdot z^1 + k_2 \cdot z^2 + \dots + k_P \cdot z^P \quad (\text{E.6})$$

The inverse filter is a physically unrealizable IIR filter:

$$\begin{aligned} H(z) &= \frac{1}{K(z)} \\ &= \frac{1}{k_0 + k_1 \cdot z^1 + k_2 \cdot z^2 + \dots + k_P \cdot z^P} \end{aligned} \quad (\text{E.7})$$

This filter can be approximated by a long FIR filter:

$$\begin{aligned} H(z) &= \frac{1}{k_0} \left\{ 1 - \left( \frac{k_1}{k_0} z^1 + \dots + \frac{k_P}{k_0} z^P \right) + \left( \frac{k_1}{k_0} z^1 + \dots + \frac{k_P}{k_0} z^P \right)^2 - \dots \right\} \\ &\approx h_0 + h_{-1} \cdot z^1 + h_{-2} \cdot z^2 + \dots + h_{-M} \cdot z^M \end{aligned} \quad (\text{E.8})$$

where  $h_i$  ( $i = -M, \dots, -1, 0$ ) can be calculated by re-arranging the terms in the polynomial.

The combined ghost canceling system is an IIR filter with both forward section  $H(z)$  and feedback section  $1/G(z)$ .



2  
VITA

Jinshi Huang

Candidate for the Degree of

Doctor of Philosophy

**Thesis:** MIXED NORM EQUALIZATION WITH APPLICATIONS  
IN TELEVISION MULTIPATH CANCELLATION

**Major Field:** Electrical Engineering

**Biographical:**

**Personal Data:** Born in Guangdong, China, On October 9, 1965, the son of Huixiong Huang and Meixiang Ye.

**Education:** Received Bachelor of Science degree in Electrical Engineering from Wuhan Technical University of Surveying and Mapping, Wuhan, China in July 1985; received Master of Science in Electrical Engineering from Oklahoma State University, Stillwater, Oklahoma in December 1988. Completed the requirements for the Doctor of Philosophy degree with major in Electrical Engineering at Oklahoma State University in December 1995.

**Experience:** Employed by Oklahoma State University, School of Electrical and Computer Engineering as a graduate research assistant, from January 1988 to May 1990; employed by Phylon Communications, Inc., Fremont, California as a member of technical staff, from May 1990 to October 1992; employed by Zoran Corporation, Santa Clara, California as a senior applications engineer from October 1992 to December 1993; employed by Exar Corporation, San Jose, California as a senior design engineer, from December 1993 to June 1994; employed by Sierra Semiconductor Corporation, San Jose, California as a senior software engineer, from June 1994 to present.