Non-Wiener Effects in Narrowband Interference Mitigation Using Adaptive Transversal Equalizers

By

Takeshi Ikuma

Dissertation submitted to the Faculty of the
Bradley Department of Electrical and Computer Engineering,
Virginia Polytechnic Institute and State University
In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

Advising Committee
Dr. A. A. (Louis) Beex, Chair
Dr. Joseph A. Ball Dr. William T. Baumann
Dr. Pushkin Kachroo Dr. William H. Tranter

April 10, 2007
Blacksburg, Virginia

Keywords: LMS algorithm, NLMS algorithm, Mean weight analysis, MSE analysis, Butterweck expansion, Transfer function approximation

© 2007 Takeshi Ikuma. All rights reserved.
Non-Wiener Effects in Narrowband Interference Mitigation Using Adaptive Transversal Equalizers

Takeshi Ikuma

(ABSTRACT)

The least mean square (LMS) algorithm is widely expected to operate near the corresponding Wiener filter solution. An exception to this popular perception occurs when the algorithm is used to adapt a transversal equalizer in the presence of additive narrowband interference. The steady-state LMS equalizer behavior does not correspond to that of the fixed Wiener equalizer: the mean of its weights is different from the Wiener weights, and its mean squared error (MSE) performance may be significantly better than the Wiener performance. The contributions of this study serve to better understand this so-called non-Wiener phenomenon of the LMS and normalized LMS adaptive transversal equalizers.

The first contribution is the analysis of the mean of the LMS weights in steady state, assuming a large interference-to-signal ratio (ISR). The analysis is based on the Butterweck expansion of the weight update equation. The equalization problem is transformed to an equivalent interference estimation problem to make the analysis of the Butterweck expansion tractable. The analytical results are valid for all step-sizes. Simulation results are included to support the analytical results and show that the analytical results predict the simulation results very well, over a wide range of ISR.

The second contribution is the new MSE estimator based on the expression for the mean of the LMS equalizer weight vector. The new estimator shows vast improvement over the Reuter-Zeidler MSE estimator. For the development of the new MSE estimator, the transfer function approximation of the LMS algorithm is generalized for the steady-
state analysis of the LMS algorithm. This generalization also revealed the cause of the breakdown of the MSE estimators when the interference is not strong, as the assumption that the variation of the weight vector around its mean is small relative to the mean of the weight vector itself.

Both the expression for the mean of the weight vector and for the MSE estimator are analyzed for the LMS algorithm at first. The results are then extended to the normalized LMS algorithm by the simple means of adaptation step-size redefinition.
To my mother, Misako, and

In loving memory of my late father, Kenji
TABLE OF CONTENTS

Table of Figures ..........................................................................................................................vii
Acknowledgements ..................................................................................................................... x
Chapter 1  Introduction .......................................................................................................... 1
  1.1.  Literature Review ........................................................................................................... 3
        1.1.1.  Brief history of equalizer-based narrowband interference suppression .... 3
        1.1.2.  Recent literature analyzing non-Wiener effects of adaptive equalizers .... 4
        1.1.3.  Non-Wiener effects in other adaptive filter applications ....................... 5
  1.2.  Contributions ................................................................................................................ 6
  1.3.  Organization ................................................................................................................... 7
Chapter 2  Adaptive Equalizers and Their Non-Wiener Behaviors ................................. 8
  2.1.  Adaptive Equalizer with Narrowband Interference ............................................... 9
  2.2.  Input Signals and Their Stochastic Properties ........................................................... 11
  2.3.  Adaptive Algorithms and Corresponding Wiener Filter ...................................... 16
  2.4.  Observation of the Phenomenon ............................................................................... 18
        2.4.1.  Sinusoidal interference ............................................................................. 19
        2.4.2.  Other narrowband interference scenarios .............................................. 25
  2.5.  Summary ....................................................................................................................... 29
Chapter 3  Analysis of Mean Weight Behavior ................................................................. 30
  3.1.  Transformation of Equalization Problem ................................................................. 31
  3.2.  Butterweck LMS Weight Expansion ........................................................................ 33
  3.3.  Mean of Butterweck Partial Weights ....................................................................... 37
  3.4.  Mean of LMS Weights .............................................................................................. 45
        3.4.1.  Mean of LMS weight vector and its magnitude-phase decomposition .. 45
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.2. Expression for individual weights</td>
<td>47</td>
</tr>
<tr>
<td>3.5. LMS Weight Behavior with Large ISR</td>
<td>50</td>
</tr>
<tr>
<td>3.6. Application to NLMS Algorithm</td>
<td>52</td>
</tr>
<tr>
<td>3.7. Numerical Illustrations</td>
<td>53</td>
</tr>
<tr>
<td>3.8. Summary</td>
<td>57</td>
</tr>
<tr>
<td>Chapter 4 Analysis of Mean Squared Error Behavior</td>
<td>58</td>
</tr>
<tr>
<td>4.1. Estimation of Mean Square Error Based on Transfer Function Approximation of the LMS Algorithm</td>
<td>58</td>
</tr>
<tr>
<td>4.2. Estimated LMS MSE for the Equalization Problem</td>
<td>65</td>
</tr>
<tr>
<td>4.3. Observations on the Analytical MSE Expression</td>
<td>71</td>
</tr>
<tr>
<td>4.4. Application to the NLMS Adaptive Equalizer</td>
<td>73</td>
</tr>
<tr>
<td>4.5. Numerical Illustrations</td>
<td>74</td>
</tr>
<tr>
<td>4.6. Summary</td>
<td>85</td>
</tr>
<tr>
<td>Chapter 5 Conclusions and Future Directions</td>
<td>87</td>
</tr>
<tr>
<td>5.1. Conclusions</td>
<td>87</td>
</tr>
<tr>
<td>5.2. Future Directions of the Research</td>
<td>88</td>
</tr>
<tr>
<td>Appendix A Derivation of (3-12)</td>
<td>90</td>
</tr>
<tr>
<td>Appendix B Expression for Mean of Weight Magnitude</td>
<td>92</td>
</tr>
<tr>
<td>Appendix C Matrix Inverse of (3-72)</td>
<td>95</td>
</tr>
<tr>
<td>References</td>
<td>97</td>
</tr>
<tr>
<td>Vita</td>
<td>100</td>
</tr>
</tbody>
</table>
Table of Figures

Figure 2-1. Transversal equalization problem with narrowband interference.................. 9
Figure 2-2. Simplified adaptive equalization model used in the analyses....................... 11
Figure 2-3. Wiener filter interpretation as two filters in parallel with an output
               gain.......................................................................................................................... 18
Figure 2-4. MSEs of Wiener equalizer and NLMS adaptive equalizer (in steady
               state) with varying step-size and fixed equalizer using the final NLMS
               weights.......................................................................................................................... 19
Figure 2-5. Scatter plots of equalizer outputs for each case in Figure 2-4. NLMS
               and fixed NLMS weight equalizers use unit step-size............................................ 20
Figure 2-6. Instantaneous squared errors associated with Figure 2-5 (250 samples
               shown). ......................................................................................................................... 21
Figure 2-7. The unit-step-size NLMS weights (‘•’) and Wiener weights (‘x’) associated with Figure 2-5 (numbers indicate the weight indices). Each
               cluster contains 20,000 observations of each weight and radial grid
               lines are separated by 0.2π rad............................................................................. 22
Figure 2-8. Time-domain behavior of NLMS weight $\hat{w}_{n}$ in Figure 2-7............... 22
Figure 2-9. Estimated PSD of NLMS weight $\hat{w}_{n}$ in Figure 2-7. Welch’s method is
               applied to 20,000-sample data using 1,024-point Blackman window
               with 50% overlap. The single peak—at DC—is not shown here. ..................... 23
Figure 2-10. MSE, as a function of ISR, of Wiener equalizer and NLMS adaptive
               equalizers with several different step-sizes. ........................................................... 23
Figure 2-11. MSE of NLMS adaptive equalizer and Wiener equalizer as functions of
               NBI frequency............................................................................................................. 24
Table of Figures

Figure 2-12. MSE of NLMS adaptive equalizers and Wiener equalizer as functions of the pole magnitude of the AR(1) interference .......................................................... 26

Figure 2-13. MSE of Wiener and NLMS adaptive equalizers with two equi-power sinusoidal interferences as a function of the frequency of the second sinusoid. The first sinusoid frequency is fixed at \(0.2\pi\). ........................................... 27

Figure 2-14. MSE of Wiener equalizer and NLMS adaptive equalizers as a function of transmitted signal power in the first ray of the two-ray multi-path channel ........................................................................................................ 29

Figure 3-1. Systems view of the Butterweck expansion of the LMS weight update equation ...................................................................................................................... 34

Figure 3-2. LMS steady-state weight behavior (10,000 samples), analytical mean, and corresponding Wiener weights on the complex plane ........................................... 54

Figure 3-3. Magnitudes of the mean LMS weights as a function of step-size (a) on a linear scale, with the Wiener solution, and (b) on a log scale .................. 55

Figure 3-4. Magnitude of mean LMS weights as a function of ISR. Step-size is varied as a function of ISR ................................................................................................. 56

Figure 3-5. Magnitude of mean NLMS weights (\(\mu = 1\)) as a function of ISR .......... 57

Figure 4-1. Two-channel generation of the fixed filter error signal ....................... 66

Figure 4-2. Relative MSE as a function of step-size \((M = 51, \text{ISR} = 20 \text{ dB, SNR} = 25 \text{ dB})\). ................................................................................................................................. 76

Figure 4-3. Relative MSE as a function of step-size \((M = 3, \text{ISR} = 20 \text{ dB, SNR} = 25 \text{ dB})\). ................................................................................................................................. 77

Figure 4-4. MSE as a function of number of equalizer taps (a) using its optimal step-size (b) \((\text{ISR} = 20 \text{ dB, SNR} = 25 \text{ dB})\) ............................................................................................................... 78

Figure 4-5. MSE as a function of SNR (a) using its optimal step-size (b) \((M = 51, \text{ISR} = 20 \text{ dB})\). ................................................................................................................................. 79

Figure 4-6. MSE as a function of ISR (a) using its optimal step-size (b) \((M = 51, \text{SNR} = 25 \text{ dB})\). ................................................................................................................................. 81

Figure 4-7. Relative LMS MSE as a function of ISR with optimal step-sizes in Figure 4-6b \((M = 51, \text{SNR} = 25 \text{ dB})\). ............................................................................................................... 82
Figure 4-8. Relative LMS MSE surfaces as functions of ISR and normalized step-size: (a) simulation and (b) new model ($M = 51$, SNR = 25 dB, relative MSE above 10 dB clipped). ................................................................. 83

Figure 4-9. Relative NLMS MSE surfaces as functions of ISR and step-size: (a) simulation and (b) new model ($M = 51$, SNR = 25 dB, relative MSE above 10 dB clipped). ................................................................. 85
Acknowledgements

First and foremost, I would like to express my most sincere gratitude to my advisor, Dr. A. A. (Louis) Beex for his guidance and... patience. His deep knowledge and comprehension of our discipline, signal processing, helped me immensely to build my own over eight years as a member of his DSP Research Laboratory. I also thank him for providing opportunities to work for multiple funded research projects over time, exposing me to many different facets of the vast field of signal processing and financially enabling me to continue pursuing the degree. Lastly, he is as much of a friend to me as a great mentor; I always enjoyed being in his company with our non-research related activities from everyday conversation to attending musical performances.

I would like to extend my appreciation to the other members of my Ph.D. advisory committee—Dr. Joe Ball, Dr. Bill Baumann, Dr. Pushkin Kachroo, and Dr. Bill Tranter—for their time and effort to review my work.

My fiancée, Laura. Her presence made my time through the graduate study so much more fun, comforting, and loving. Not only is she the best thing that ever happened to me, but also I’m grateful to have her by my side when we graduate together with our Ph.D.s.

Last but not least, I am deeply indebted to my parents Misako and Kenji for their support and understanding of having me far away from home, especially through my father’s medical hardship and eventual passing. Dad, when I look at this dissertation in the future, I will always associate it with your smile.
CHAPTER 1  INTRODUCTION

Adaptive transversal equalizers play an important role in digital communication, especially in mobile applications where the characteristics of the communication channel are unknown and time-varying. The primary functions of equalizers are to estimate the channel characteristics and to undo its effect—intersymbol interference—on the received communication signal [1 Ch.1]. While there are several types of equalizers, the one of our interest is the transversal (or linear, i.e. linear in structure) equalizer, which linearly combines the received signal samples to produce the output. Moreover, in most applications, the characteristics of the communication channel are not known a priori, and some channel impairments are time-varying, such as fading multipath. In such channels, equalizers must be adapted to counteract such unknown and time-varying channel characteristics [1 Ch.11, 2].

In addition to its primary role to mitigate intersymbol interference, equalizers can also act as general interference mitigation devices. The focus of this research is this secondary feature of the adaptive transversal equalizer, specifically the mitigation of narrowband interferences. The narrowband interference has been receiving a good amount of attention in the past three decades to battle both intentional and unintentional interferences. A prime example of an intentional narrowband interference is narrowband jamming in tactical communications. Another example, of the intentional variation, of narrowband interference is in an overlay communication system where a new spread-spectrum service is installed in the same frequency band as an existing narrowband service. Many techniques to mitigate narrowband interference were proposed over the years, as excellently summarized by Poor [3], Laster and Reed [4], and Milstein [5].
Interestingly, over the three decades of research efforts by many researchers in the area of narrowband interference suppression, very little was said about the operation of these equalizers in the presence of narrowband interference. One of the reasons given by Monsen [6], although the topic is on parallel channel equalization and correlated interference, seems to apply to the single channel equalization scenario here. That is, large interferers pose a problem in acquisition and tracking of the communication signal. Nonetheless, when adaptive transversal equalizers are used to mitigate narrowband interference (even with limited interference power to assure proper receiver operation), these adaptive equalizers are found to behave out of the norm expected from conventional adaptive filtering theory, as first observed by North, Axford, and Zeidler [7].

In conventional adaptive filtering theory, the performance of adaptive filters—that are designed to minimize squared error—is assessed by the corresponding Wiener (or minimum mean square error) filter of the same structure [8]. Moreover, in case of the algorithm of our interest—the least mean square (LMS) algorithm—its performance is subject to additional misadjustment error due to the weight adaptation. Hence, the Wiener filter is asserted as the optimal filter, and its performance is expected to be the lower bound to the performance of the LMS filter. Consequently, the LMS algorithm is usually applied with small adaptation step-size so as to minimize the misadjustment error, i.e., to approach the performance of the Wiener filter. North et al. observed the probability-of-error performance of the LMS equalizer to be better than the corresponding Wiener equalizer [7].

This dissertation contributes to the efforts to explain this unexpected behavior of adaptive transversal equalizers. The behavior is referred to in the literature as non-Wiener or nonlinear effects. The term “non-Wiener” is used to reflect that adaptive filters with such effect do not follow the behavior of the corresponding Wiener filter, as its performance can exceed that of the Wiener filter; the term “nonlinear” is used to indicate that the better performance of the adaptive linear equalizer is attained due to the nonlinear nature of the adaptive algorithms. The previous work in this area of research is reviewed in Section 1.1. The brief descriptions of the contributions from this
study are then given in Section 1.2. Lastly, Section 1.3 contains the mapping of the rest of the dissertation.

1.1. Literature Review

We first briefly review the research efforts in narrowband interference suppression using equalizers, leading up to the discovery of the non-Wiener effect [7]. Then, the recent literature geared to analyze the non-Wiener effects in this adaptive equalizer application is reviewed, followed by the existing analyses of non-Wiener effects observed in other adaptive filtering applications.

1.1.1. Brief history of equalizer-based narrowband interference suppression

In one of the pioneering works in narrowband interference suppression research, Li and Milstein [9] proposed linear prediction (one-sided) and smoothing (two-sided) error filters for use in narrowband interference mitigation. This technique is non-data-aided (i.e., the only required knowledge of the desired signal is that the signal is effectively white relative to the filter) and simply tries to minimize the output power. These structures resemble the transversal equalization structure but actually differ in two respects. First, the reference input tap is not weighted. Second, its output power is minimized as opposed to the power of the error between its output and the expected signal. In their subsequent paper, Li and Milstein [10] improved the performance of [9] by introducing decision feedback. The additional important change is the use of the error signal between the filter output and the desired signal (i.e., the most recent decision). While their system is converging to the decision-feedback equalizer [1], it still lacks the weighting of the reference feed-forward tap.

Niger and Vandamme [11] are the first researchers, to our best knowledge, who explicitly studied equalizer performance under the influence of narrowband interference (along with the multipath effect). They presented the analytically approximated bit-error-rate performance for transversal and decision-feedback (Wiener) equalizers for both $T$-spaced (symbol-duration tap spacing) and fractional $T/2$-spaced configurations. Davis and Milstein [12] also showed the effectiveness of a
transversal equalizer for narrowband interference mitigation in a direct-sequence spread-spectrum system. Hasan, Lee, and Bhargava [13] modified the two-sided transversal filter of Li and Milstein [9] to have an adjustable reference tap weight, effectively forming a transversal equalizer even though the authors do not state that in the correspondence.

North, Axford, and Zeidler [7] are the first to report an adaptive transversal equalizer benefiting from the non-Wiener effect of the least mean squares (LMS) algorithm [8] when the received signal is dominantly contaminated by narrowband interference. In their study, the steady-state performance of adaptive equalizers is experimentally measured, via Monte Carlo simulations. They observed the performance in terms of probability of symbol error and notch bandwidth. The fixed Wiener filter is shown to place a notch at the interference frequency. Although the notch eliminates the interference, at the same time, it introduces intersymbol interference to the desired signal. Hence, they employed notch bandwidth as one of the performance measures. In their sinusoidal interference scenarios (both with and without multipath) the LMS algorithm with larger step-size resulted in better performance (i.e., a narrower notch) than with the smaller step-size LMS or Wiener filters. They also illustrated that the time-varying misadjustment filter (that is, the remainder of the LMS weights after removing the Wiener weights) produced an output that effectively fills the notch created by the Wiener filter.

1.1.2. Recent literature analyzing non-Wiener effects of adaptive equalizers

There are several studies that followed the paper by North, Axford, and Zeidler [7]. First, Reuter and Zeidler [14] analyzed the LMS and normalized LMS (NLMS) equalizer behaviors for both sinusoidal and narrowband first-order autoregressive interference processes, using the so-called transfer function approximation of the LMS algorithm [15]. Their analysis has shown good (but not exact) agreement with experimental results.

In the search for an explanation for the non-Wiener effects, Beex and Zeidler [16, 17] modeled the ideal interference cancellation with a transversal equalizer as a two-channel Wiener filter with the interference sample given as the secondary channel
input, leading to a time-varying mechanism as the likely source. Although this structure has no practical use because it requires a priori knowledge of the interference, it leads to the time-varying weight target that the NLMS algorithm tracks. The manifold of time-varying Wiener solutions is identified.

Finally, Batra et al. [18] compared, through simulation, the performances of the NLMS fractionally-spaced transversal equalizer with time-varying benefit, the NLMS fractionally-spaced decision-feedback equalizer, and the Wiener fractionally-spaced equalizer. This paper points out that the adaptive transversal equalizer suffers from an extremely slow convergence rate when the equalizer needs to perform both multipath equalization and mitigation of a strong narrowband interference component.

1.1.3. Non-Wiener effects in other adaptive filter applications

Similar non-Wiener phenomena are observed in other applications of the LMS algorithm such as adaptive noise (interference) cancellation [19-24], chirped sinusoid tracking [25], and biomedical signal processing [26, 27]. The inadequacy of traditional analysis methods such as the independence theory and the small-step-size theory to predict the non-Wiener performance prompted researchers to pursue other analysis methods.

An analytical technique based on the transfer function approach was first proposed by Glover [28] and later generalized by Clarkson and White [15]. This technique has been used in numerous applications to analyze non-Wiener behavior [23, 24, 26, 27], including—as mentioned earlier—by Reuter and Zeidler [14]. This method models the LMS algorithm as a linear time-invariant filter (assuming that the input is a periodic signal) which explicitly relates the desired signal input to the LMS error signal. This technique, however, does not extend to general random process inputs as noted by Butterweck [29].

Douglas and Pan [30] derived an exact expectation method to predict statistical behavior of the tap-delay-line based LMS filter. Their method has shown to accurately evaluate the mean squared error of the LMS filter for a wide range of step-sizes by constructing and evaluating a set of linear equations. The number of equations, however, increases drastically as the number of taps increases, resulting in
computational intractability for a filter with a large number of taps.

Beex and Zeidler [17] identified the optimal two-channel Wiener filter, of which the secondary channel bears the information that the time-varying LMS weights attempt to carry. This two-channel filter structure coincides with the LMS performance bound recognized by Quirk, Milstein, and Zeidler [31]. Through the notion of a linking sequence, which connects the implicit second channel with the available single-channel input, Beex and Zeidler established a manifold of single-channel time-varying Wiener filters, one of which the LMS and normalized LMS (NLMS) algorithms tries to track. For the adaptive noise cancellation application, they identified the unique time-varying solution that the NLMS algorithm tracks and subsequently derived the exact steady-state behaviors of the NLMS weights for arbitrary step-size.

1.2. Contributions

This dissertation contains two major contributions towards comprehending the non-Wiener effects in LMS and NLMS transversal equalization in the presence of sinusoidal interference, a limiting case for general narrowband interference. In addition, it includes new results from numerical simulation, in which an additional non-Wiener effect is observed when there are two sinusoidal interferers of close frequencies.

The first contribution of this work is the discovery of the difference in the mean of the adapted weights in steady state and the corresponding Wiener filter weights, and the derivation of the analytical expression for the mean of the weights under the assumption of large interference-to-signal ratio. The analysis utilizes the expansion of the weight update equation, proposed by Butterweck [32].

The second contribution is the enhancement of the mean-squared-error model of Reuter and Zeidler [14]. This contribution includes two parts. First, the transfer-function technique [15] is generalized to be free of the initial condition on the weight vector, and the steady-state mean-squared-error estimate is re-derived based on the generalized transfer function method with the reference taken in steady-state. This analysis revealed the implicit small weight fluctuation assumption that was previously ignored. This assumption is held responsible for the observed breakdown of the MSE estimators as
interference weakens. Secondly, the Reuter-Zeidler MSE model is improved using the new expression for the mean of the steady-state LMS and NLMS weights. The new MSE model has shown to be very accurate for large interference.

1.3. Organization

The main section of the dissertation spans three chapters. Chapter 2 introduces the adaptive transversal equalization problem with narrowband interference and defines the input signals and their properties. This chapter also includes the observation of the simulation results to gain insights into the problem. Chapter 3 contains the analysis of the mean of the LMS and NLMS weights. This expression for the mean of the weights is then used in Chapter 4 for the analysis of the mean squared error. Both Chapter 3 and Chapter 4 contain numerical illustrations to validate the expressions that were developed. Finally, Chapter 5 concludes the work and suggests possibilities for future extensions.
CHAPTER 2 ADAPTIVE EQUALIZERS AND THEIR NON-WIENER BEHAVIORS

To establish a solid basis for the research, this chapter defines the equalization system under study and presents simulation results that reveal, through numerical simulation, the benefit of the non-Wiener effect in narrowband interference mitigation using a (large step-size) adaptive transversal equalizer.

Section 2.1 first presents the complete adaptive equalization problem with additive interference, introducing the signals involved and the structure of the transversal equalizer. Then, the actual problem of interest—a reduced version of the full equalization problem—is presented. The simplified problem limits the interference to complex sinusoidal interference and neglects the intersymbol interference caused by the channel and decision errors in the output of the receiver decision device. The above simplifications are used to focus on understanding the interaction between the narrowband interference and adaptation.

All the input signals are formally introduced in Section 2.2, with special attention paid to the second moments of these signals. The two adaptive algorithms of interest to us—the LMS and NLMS algorithms—are defined in Section 2.3 along with the corresponding Wiener filter solution.

Finally, Section 2.4 presents a set of simulation results, demonstrating the benefit of using adaptive equalizers. The MSE performance of the adaptive equalizer is compared to that of the corresponding Wiener equalizer. Non-sinusoidal interference scenarios are included to show that the non-Wiener effect is not strictly limited to cases with sinusoidal interference.
2.1. Adaptive Equalizer with Narrowband Interference

Figure 2-1 depicts the general formulation of the adaptive equalization problem with a band-limited linear filter channel subject to narrowband interference. The system takes three input signals: the transmitted discrete-time symbol sequence $x_n$, the additive narrowband interference $i_n$, and the additive noise $n_n$. The transmitted signal $x_n$ is subject to channel imperfection, which is modeled as a linear time-invariant filter with unit pulse response $h_n$. The contaminated transmitted signal is denoted by $\tilde{x}_n$.

![Figure 2-1. Transversal equalization problem with narrowband interference.](image)

All the analyses and simulations in this study deal with discrete-time complex-valued signals, that is, the complex baseband signals that are ideally sampled at the symbol rate of the communication signal. Furthermore, a sample of the transmitted signal exclusively determines a communication symbol transmitted at that time. In other words, the sample is assumed to be drawn from a finite set of complex values that represents all possible communication symbols. Any effects due to intermediate systems (e.g., the transmitter digital-to-analog converter, the receiver sampler, and discrete-time noise-whitening filter) are modeled lumped together as “the channel”.

The transversal equalizer structure is determined by two parameters: the number of input taps $M$ and the delay $\Delta$ in the desired signal with respect to the most recent input sample $u_n$. Given the received signal $u_n \triangleq \tilde{x}_n + n_n + i_n$, the equalizer forms the
input vector \( \mathbf{u}_n = [u_n \ u_{n-1} \ \cdots \ u_{n-M+1}]^T \in \mathbb{C}^M \), where \((\cdot)^T\) is the transpose operator. Subsequently, the components of \( \mathbf{u}_n \) are denoted by \( x_n, \hat{n}_n, \) and \( i_n \). The delay \( \Delta \) in the desired signal must be chosen to be less than \( M \); typically, \( \Delta \approx M/2 \) so that the desired signal aligns with the middle of the input vector. The adaptive equalizer has a tunable weight vector \( \mathbf{w}_n = [\hat{w}_{0,n} \ \hat{w}_{1,n} \ \cdots \ \hat{w}_{M-1,n}]^T \in \mathbb{C}^M \), and the equalizer output \( y_n \) is computed by

\[
y_n = \hat{\mathbf{w}}_n^H \mathbf{u}_n
\]

where \((\cdot)^H\) is a Hermitian (conjugate) transpose operator. The adaptation methods for the weights are covered later in Section 2.3.

A practical adaptive equalizer typically has two operational modes (indicated by the switch in Figure 2-1): training and decision-directed. During the training period, the transmitter sends a known finite-length training symbol sequence to allow the equalizer to initialize its weights. After the training period, the equalizer adaptation mechanism is switched to the decision-directed mode, in which the estimated symbol \( \hat{x}_{n-\Delta} \) — chosen by the receiver decision device — is used for adaptation.

Since the focus of this study is to understand the system behavior due to narrowband interference, we will use a simplified version of the conventional equalization problem that is described above. Specifically, the channel is assumed to be ideal, i.e., \( \hat{x}_n = x_n \), causing no intersymbol interference. As a result, the received signal is now found to be \( u_n = x_n + \hat{n}_n + i_n \). Also, the equalizer is fixed to operate in the training mode so that the equalizer has exact knowledge of what is transmitted at all times. This is essentially equivalent to the equalizer being in the decision-directed mode with very few decision errors. The resulting equalization problem is depicted in Figure 2-2.

Another restriction that we applied in our analyses is on the type of narrowband interference. There are three basic types of narrowband interference [3]: tonal signals, narrowband stochastic processes, and narrowband digital communication signals. Tonal signals are composed of multiple pure (complex) sinusoids, which in practice are introduced by tone jammers and harmonic interference phenomena. Entropic
Chapter 2: Adaptive Equalizers and Their Non-Wiener Behaviors

narrowband stochastic processes, such as autoregressive processes, are often used to model non-deterministic interferences. Lastly, the narrowband communication signal appears as narrowband interference to spread-spectrum service that resides in the same frequency band as the narrowband service. For the analyses in this study, we only consider the limiting case, of having a complex sinusoidal interference as the narrowband interference.

2.2. Input Signals and Their Stochastic Properties

This section defines all the signals that appear in Figure 2-2. These definitions are used throughout the subsequent analyses. Since the subsequent analyses make extensive use of the second moments of the input processes, this section covers various forms of second moments, e.g., auto- and cross-correlation functions of the signals and correlation matrices of processes with arbitrary delays.

All three input signals — $x_n$, $i_n$ and $\tilde{n}_n$ — are modeled as complex random processes that are zero-mean, wide-sense-stationary, mean-ergodic, and proper [33]. In addition, the information-bearing transmitted signal $x_n$ is assumed to be a white process with power $\sigma_x^2$. Moreover, $x_n$ is drawn from a finite set of information symbols which are mapped symmetrically on the complex plane, and the symbols are drawn randomly.

Figure 2-2. Simplified adaptive equalization model used in the analyses.
with equal probability. The additive noise process \( \tilde{n}_n \) is also assumed white, but with power \( \sigma_{\tilde{n}}^2 \). The additive sinusoidal interference \( i_n \) possesses average power \( \sigma_i^2 \) and frequency \( \omega_i \) in radian/sample (rad/S) and is modeled as:

\[
i_n = \sigma_i e^{j(\omega_n + \phi)}. \tag{2-2}\]

where \( j = \sqrt{-1} \). The phase \( \phi \) is randomly drawn from \([0, 2\pi)\) radians but is fixed in each realization. All three input signals are mutually uncorrelated.

The above statistical properties are sufficient conditions for the non-Wiener phenomenon to occur. Some of the conditions, such as whiteness and stationarity, are introduced purely to make the analyses tractable. On the other hand, the phenomenon is insensitive to the exact nature—e.g., type of modulation or the underlying probability density function—of both \( x_n \) and \( \tilde{n}_n \).

We denote the autocorrelation function of a signal \( s_n \) as \( r_{ss,l} \triangleq E\{s_n s_{n-l}^*\} \). Accordingly, the autocorrelation functions of the input signals are defined as follows:

\[
r_{xx,l} = \sigma_x^2 \delta_l \tag{2-3}\]

\[
r_{\tilde{n}\tilde{n},l} = \sigma_{\tilde{n}}^2 \delta_l \tag{2-4}\]

\[
r_{ii,l} = \sigma_i^2 e^{j\omega_l} \tag{2-5}\]

Here, \( (\cdot)^* \) is the complex conjugation operator, and \( \delta_l \) is the Kronecker delta function, which is 1 if \( l = 0 \) and 0 otherwise. From (2-3), (2-4), (2-5), and the uncorrelatedness assumption, we derive the autocorrelation function of the input to the equalizer to be

\[
r_{uu,l} = r_{xx,l} + r_{\tilde{n}\tilde{n},l} + r_{ii,l}
= (\sigma_x^2 + \sigma_{\tilde{n}}^2) \delta_l + \sigma_i^2 e^{j\omega_l} \tag{2-6}\]

The properness of the input processes plays a crucial role in Chapter 3. For a wide-sense-stationary, zero-mean complex signal, call it \( s_n \), to be proper, its pseudo-auto correlation function, which is defined as \( \tilde{r}_{ss,l} \triangleq E\{s_n s_{n-l}^*\} \), must vanish identically [33].
For the white communication signal to be proper, its power $\sigma_n^2$ must be equally distributed between its real and imaginary components. To show this, we first observe that the (conventional) correlation function expands to

$$r_{xx,l} = E[(x_{r,n} + jx_{i,n})(x_{r,n-l} - jx_{i,n-l})]$$

$$= (E[x_{r,n}x_{r,n-l}] + E[x_{i,n}x_{i,n-l}]) + j(E[x_{r,n}x_{i,n-l}] - E[x_{i,n}x_{r,n-l}])$$

(2-7)

where $x_n$ is expanded to its real component $x_{r,n}$ and imaginary component $x_{i,n}$. For (2-3) (whiteness) to hold, the cross-correlation function of the real and imaginary components, $E[x_{r,n}x_{i,n-l}]$, must be an even function so that the imaginary component of (2-7) vanishes. We then observe that the pseudo-correlation function similarly expands to

$$\tilde{r}_{xx,l} = E[(x_{r,n} + jx_{i,n-l})(x_{r,n} + jx_{i,n-l})]$$

$$= (E[x_{r,n}x_{r,n-l}] - E[x_{i,n}x_{i,n-l}]) + j(E[x_{r,n}x_{i,n-l}] + E[x_{i,n}x_{r,n-l}])$$

(2-8)

For $x_n$ to be proper, the imaginary term of (2-8) must vanish, but from (2-7) we know that $E[x_{r,n}x_{i,n-l}]$ is an even function. Consequently, $E[x_{r,n}x_{i,n-l}] = 0$ to satisfy both conditions. Finally, to satisfy (2-3) and for the real part of the pseudo-autocorrelation in (2-8) to vanish, the autocorrelation functions of the real and imaginary components must satisfy

$$E[x_{r,n}x_{r,n-l}] = E[x_{i,n}x_{i,n-l}]$$

$$= \frac{\sigma_n^2}{2} \delta_l$$

(2-9)

The same conditions hold for the proper additive noise $\tilde{n}_n$, i.e., its power $\sigma_n^2$ must be equally distributed between its real and imaginary components. Lastly, the properness of the interference $i_n$ is observed straight from the definition. Using (2-2), its pseudo-correlation function is found to be
\[ \tilde{r}_{ii,l} = E \left[ \{ \sigma_i e^{j[(\omega_i n + \phi)]} \} \{ \sigma_i e^{j[(\omega_i (n-l) + \phi)]} \} \right] = \sigma_i^2 e^{-j\omega_i l} E \left[ e^{j2(\omega_i n + \phi)} \right] \quad (2-10) \]

Because \( \phi \) is uniformly distributed between 0 and \( 2\pi \), the expectation in (2-10) vanishes for all \( n \); consequently, the pseudo-autocorrelation in (2-10) vanishes for all \( l \).

Next, the sinusoidal process \( i_n \) possesses a useful property. Since it is deterministic in each realization, we have

\[ i_n i_{n-l}^* = \left[ \sigma_i e^{j(\omega_i n + \phi)} \right] \left[ \sigma_i e^{-j(\omega_i (n-l) + \phi)} \right] = \sigma_i^2 e^{j\omega_i l} = r_{ii,l} \quad (2-11) \]

In other words, two samples (or just one sample for \( r_{ii,0} \)) are sufficient to determine the exact autocorrelation function value of \( i_n \) (sinusoidal process with fixed amplitude and frequency) for the lag specified by the delay between these two samples. No ensemble-averaging (or time-averaging) is necessary.

Next, we establish the correlation matrices involving the input signals. Following (2-3), the cross-correlation matrix between \( x_n \) and its delayed version \( x_{n-l} \) is

\[ E \left[ x_n x_{n-l}^H \right] = \begin{cases} \sigma_x^2 Z^l, & \text{if } 0 \leq l < M \\ \sigma_x^2 N^l, & \text{if } -M < l < 0 \\ 0, & \text{otherwise} \end{cases} \quad (2-12) \]

The matrix \( Z \in \mathbb{R}^{M \times M} \) is the lower shift matrix, which has ones on the subdiagonal and zeros elsewhere, and \( N \triangleq Z^T \) is the upper shift matrix. The all-zero matrix is represented by \( O \in \mathbb{R}^{M \times M} \). The correlation matrix of \( \tilde{i}_n \) has the same structure:

\[ E \left[ \tilde{i}_n \tilde{i}_{n-l}^H \right] = \begin{cases} \sigma_{\tilde{i}}^2 Z^l, & \text{if } 0 \leq l < M \\ \sigma_{\tilde{i}}^2 N^l, & \text{if } -M < l < 0 \\ 0, & \text{otherwise} \end{cases} \quad (2-13) \]

From (2-5), the cross-correlation matrix between \( i_n \) and its delayed version \( i_{n-l} \) is
\[ E[i_n^H i_{n-l}^H] = \sigma_i^2 e^{j\omega_l} ee^H \]  \hfill (2-14)

where

\[ e \triangleq \begin{bmatrix} 1 & e^{-j\omega_1} & \ldots & e^{-j\omega_{M-1}} \end{bmatrix}^T \]  \hfill (2-15)

Using (2-12), (2-13), and (2-14) for \( l = 0 \), and employing the uncorrelatedness assumption, the input correlation matrix evaluates to

\[ R \triangleq E[u_n u_n^H] = (\sigma_x^2 + \sigma_n^2) \mathbf{I} + \sigma_i^2 ee^H \]  \hfill (2-16)

Using the Woodbury identity [34 Appendix A], the inverse of \( R \) is found to be

\[ R^{-1} = \frac{1}{\sigma_x^2 + \sigma_n^2} \left( \mathbf{I} - \frac{\sigma_i^2}{\lambda_{\text{max}}} ee^H \right) \]  \hfill (2-17)

where \( \lambda_{\text{max}} \) is the maximum eigenvalue of \( R \), which is given by

\[ \lambda_{\text{max}} = \sigma_x^2 + \sigma_n^2 + \sigma_i^2 M \]  \hfill (2-18)

Also, the cross-correlation between the input vector \( u_n \) and the desired signal \( x_{n-\Delta} \) evaluates to

\[ p_{ud} \triangleq E[u_n x_{n-\Delta}^*] = E[x_n^* x_{n-\Delta}] = \sigma_i^2 \mathbf{p}_\Delta \]  \hfill (2-19)

where the basis vector \( \mathbf{p}_\Delta \) is defined according to

\[ \mathbf{p}_k \triangleq \begin{bmatrix} 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \end{bmatrix} \in \mathbb{R}^M \]  \hfill (2-20)

As a final remark, we note that the pseudo-correlation matrices that complement the
correlation matrices in (2-12) through (2-14) vanish. That is,

\[ E[x_n x_{n-l}] = 0 \]  \hspace{1cm} (2-21)

\[ E[\mathbf{n}_n \mathbf{n}_{n-l}] = 0 \]  \hspace{1cm} (2-22)

\[ E[i_n i_{n-l}] = 0 \]  \hspace{1cm} (2-23)

These results follow directly from the properness of the input signal components as discussed earlier.

### 2.3. Adaptive Algorithms and Corresponding Wiener Filter

We consider two adaptive algorithms [8]: the LMS algorithm and the NLMS algorithm. The main focus of the analyses is on the LMS algorithm, but the results are afterwards extended to the NLMS algorithm. As depicted in Figure 2-2, both algorithms are driven by the error signal

\[ e_n = x_n - \mathbf{w}^H_n \mathbf{u}_n \]  \hspace{1cm} (2-24)

The actual error signals of the LMS and NLMS algorithms are different as their weights are adapted differently. We distinguish the outcomes of the two algorithms by using plain notation for the LMS algorithm and adding the superscript \(^{\text{NLMS}}\) to quantities pertaining to the NLMS algorithm. For example, the error signal \(e_n\) defined in (2-24) indicates the error of the LMS algorithm while \(e_n^{\text{NLMS}}\) denotes the NLMS error signal, which also uses (2-24) but with NLMS weights \(\mathbf{w}_n^{\text{NLMS}}\) instead of the LMS weights \(\mathbf{w}_n\). Similar superscript notation is also used later for the corresponding Wiener filter.

The weight vector \(\mathbf{w}_n\) is adapted using the LMS algorithm

\[ \mathbf{w}_{n+1} = \mathbf{w}_n + \mu \mathbf{u}_n e_n^* \]  \hspace{1cm} (2-25)

with step-size \(\mu\). As an alternate form, we combine (2-24) and (2-25) and obtain
\[
\hat{w}_{n+1} = \left( I - \mu u_n u_n^H \right) \hat{w}_n + \mu u_n x_{n-\Delta}^*
\]
(2-26)

where \( I \in \mathbb{R}^{M \times M} \) represents the identity matrix. This latter form is used in the analysis in Chapter 3. Similar to the LMS algorithm, the NLMS algorithm updates the weight vector according to

\[
\hat{w}_{n+1}^{(NLMS)} = \hat{w}_n^{(NLMS)} + \tilde{\mu} u_n \left( u_n^H u_n \right)^{-1} e_n^{(NLMS)*}
\]
(2-27)

with normalized step-size \( \tilde{\mu} \). The step-size \( \tilde{\mu} \) must be selected from \((0, 2)\) for stable operation.

The Wiener filter solution to this equalization problem can be computed from

\[
\hat{w}_w = R^{-1} p_{u_d}
\]
(2-28)

where the inverse of the input correlation matrix and the cross-correlation vector are defined in (2-17) and (2-19), respectively. Substituting these correlation expressions into (2-28), the Wiener filter can be expressed as

\[
\hat{w}_w = \eta \left( p_\Delta - \frac{\sigma_i^2}{\lambda_{\text{max}}} e^{j\omega_\Delta} \right)
\]
(2-29)

where

\[
\eta \triangleq \frac{\sigma_x^2}{\sigma_x^2 + \sigma_n^2}.
\]
(2-30)

This Wiener filter solution can be interpreted as a system with two filters in parallel, followed by a normalization gain \( \eta \). This interpretation is illustrated in Figure 2-3. The first filter with \( p_\Delta \) as weight vector selects \( u_{n-\Delta} \) from \( u_n \). Then the second filter with \( \frac{\sigma_i^2}{\lambda_{\text{max}}} e^{j\omega_\Delta} \) as weight vector estimates the interference component in \( u_{n-\Delta} \), so that this estimate can then be removed from \( u_{n-\Delta} \). The gain \( \eta \) normalizes the subtraction outcome so that its power is \( \sigma_x^2 \), i.e. equal to that of the desired signal. We will utilize this interpretation in Chapter 3.
Figure 2-3. Wiener filter interpretation as two filters in parallel with an output gain.

The mean squared error (MSE) for the Wiener filter is generically defined as [8 p.107]

\[
J_W \triangleq E\left[\left|e_n^{(Wiener)}\right|^2\right]_{\bar{w}_n = \bar{w}_w} = \sigma_d^2 - \bar{w}_w^H R \bar{w}_w \tag{2-31}
\]

where \( \sigma_d^2 \) is the variance of the desired signal. Using our desired signal \( x_{n-\Delta} \), (2-17), (2-19), and (2-30), the Wiener MSE for our problem is computed by

\[
J_W = \sigma_d^2 \left[1 - \eta \left(1 - \frac{\sigma_i^2}{\lambda_{\text{max}}}\right)\right] \tag{2-32}
\]

This MSE will be used as a reference to assess adaptive equalizer performance.

### 2.4. Observation of the Phenomenon

To illustrate the MSE performance benefit that can be obtained by the use of an adaptive equalizer, this section presents a series of simulation results for several scenarios. In Section 2.4.1, we concentrate on the equalization problem as defined by Figure 2-2 with a sinusoidal interference, which is the scope of the rest of the study. We then observe the system behavior under non-ideal channel or non-sinusoidal interference in Section 2.4.2. The latter section, including examples with AR(1) interference and a two-tap FIR filter channel, demonstrates how this phenomenon extends to further realistic situations. Also, all the results here are for NLMS equalizers; the LMS equalizer results will be shown later, together with the analysis results.
All cases involve quaternary phase shift-keying (QPSK) modulated signals as the communication signal. This modulation chooses its symbols from four possible values, \{1 + j, 1 − j, −1 + j, −1 − j\}, with appropriate scaling for the specified signal power. The transversal equalizer structure is fixed to \( M = 11 \) and \( \Delta = 5 \). The QPSK and noise powers are fixed at \( \sigma_z^2 = 1 \) and \( \sigma_n^2 = 0.0001 \), respectively, resulting in a signal-to-noise ratio (SNR), \( \sigma_z^2 / \sigma_n^2 \), of 40 dB. The noise \( \tilde{n}_n \) is modeled as a circularly symmetric complex Gaussian process. The interference power is described by the interference-to-signal ratio (ISR), \( \sigma_i^2 / \sigma_z^2 \). All the simulation results are captured in steady state, and the MSE is estimated by averaging 20,000 error signal samples in steady state.

### 2.4.1. Sinusoidal interference

Figure 2-4 shows the MSE performance as a function of the NLMS step-size parameter. The sinusoidal interference has ISR of 26 dB and \( \omega_i = 0.2\pi \text{ rad/S} \). All simulations are performed on the same received signal. The figure also includes the Wiener MSE (which is not a function of NLMS step-size) and the MSE of a fixed equalizer with its weights set to the final weight values of the NLMS equalizer simulation. This result is consistent with that of [14].

![Figure 2-4. MSEs of Wiener equalizer and NLMS adaptive equalizer (in steady state) with varying step-size and fixed equalizer using the final NLMS weights.](image)
The benefit of the adaptation is very clearly illustrated as the NLMS equalizer shows an almost 5-dB MSE performance gain over the time-invariant Wiener filter. Additional performance gain can be attained by increasing the filter order, as shown with the $M = 51$ case in [14] and later in Chapter 4. The fixed weight equalizer using the final value steady-state NLMS weights does not display the same MSE performance boost; on the contrary, it performs worse than the Wiener filter. This observation strongly suggests that the instantaneous, time-varying nature of the NLMS weights is the cause of the performance gain [17].

Next, the system behavior is more closely observed for the unit step-size ($\mu = 1$) case. First, the scatter plots of the equalizer output signal are shown in Figure 2-5. Although the QPSK symbol constellation is clearly observable in each of the equalizer output scatter plots, the NLMS equalizer (Figure 2-5a) produces by far the crispest scatter plot—with a curious sub-clustering within each symbol cluster. The Wiener equalizer (Figure 2-5b) and the equalizer with fixed NLMS weights (Figure 2-5c) show Gaussian-like distributions around each symbol. The inferiority of the fixed-NLMS-weight equalizer is observable around the symbol boundaries (it has more samples there than the Wiener equalizer).

![Figure 2-5. Scatter plots of equalizer outputs for each case in Figure 2-4. NLMS and fixed NLMS weight equalizers use unit step-size.](image)

Next, the instantaneous squared errors of the equalizers under study are shown in Figure 2-6. The squared error of the NLMS equalizer clearly has less variation than the fixed counterparts. Also, the NLMS equalizer does not outperform the Wiener filter on
a sample-by-sample basis. Furthermore, the smaller squared error variation in the NLMS equalizer is explained by the sub-clustering of the output signal observed in Figure 2-5. The apparent sub-clustering must be a direct consequence of some deterministic mechanism in the NLMS algorithm. With large step-size, the deterministic behavior supersedes the stochastic behavior of the NLMS algorithm. Since none of the deterministic solutions (i.e., the centers of the sub-clusters) aligns with the actual symbol location, the NLMS instantaneous error cannot become small and is dominated by the deterministic distances between the symbol location and the sub-cluster centers.

![Instantaneous squared errors associated with Figure 2-5 (250 samples shown).](image)

Lastly, the steady-state behaviors of the equalizer weights are examined. Figure 2-7 illustrates, on the complex plane, the steady-state weights of the NLMS equalizer with $\mu = 1$. The $k$-th cluster contains observations of $\tilde{w}_{k,n}$ over 20,000 samples. The Wiener equalizer weights are also indicated, using ‘x’s. The NLMS weights in steady-state appear to converge to some fixed non-Wiener solution. The NLMS weights reside on the same radial lines as the Wiener weights, but they exhibit a spiral shape—except for the $(\Delta + 1)$-st tap—instead of the circular shape of the Wiener weights. Finding the analytical solution to the observed displacement of the mean of the NLMS weight is the
Figure 2-7. The unit-step-size NLMS weights ('•') and Wiener weights ('x') associated with Figure 2-5 (numbers indicate the weight indices). Each cluster contains 20,000 observations of each weight and radial grid lines are separated by 0.2π rad.

main focus of Chapter 3.

Even when an individual weight is observed in either the time domain (Figure 2-8) or the frequency domain (Figure 2-9), the non-Wiener behavior appears to be some form of a weak low-pass process (weak compared to the fixed component) and—unlike the weight behavior in the adaptive noise cancellation case [17]—does not possess any deterministic quality.

Figure 2-8. Time-domain behavior of NLMS weight $\tilde{\omega}_{n,n}$ in Figure 2-7.
Chapter 2: Adaptive Equalizers and Their Non-Wiener Behaviors

Figure 2-9. Estimated PSD of NLMS weight $\tilde{w}_{3,n}$ in Figure 2-7. Welch’s method is applied to 20,000-sample data using 1,024-point Blackman window with 50% overlap. The single peak—at DC—is not shown here.

The simulation results shown thus far are based on very strong sinusoidal interference with 26-dB ISR. Figure 2-10 illustrates how the MSE performance varies as ISR is varied. Several different step-size configurations ($\hat{\mu} = 1, 0.6, 0.1,$ and $0.001$) are shown for the NLMS equalizer. The simulation is carried out with—other than the

Figure 2-10. MSE, as a function of ISR, of Wiener equalizer and NLMS adaptive equalizers with several different step-sizes.
interference magnitude—the same received signal. For large ISR above 3 dB, the largest step-size \( (\mu = 1) \) configuration produces the best NLMS performance, and the performance drops as the step-size is made smaller. All NLMS simulations but the smallest step-size \( (\mu = 0.001) \) case produced performance better than that of the Wiener equalizer for ISR above \(-2 \) dB. Below that ISR level, the NLMS MSE appears to be consistent with the conventional theory, that is, the small step-size configuration turns in the best MSE performance but—with (positive) excess MSE—is inferior to the Wiener MSE. It is worth noting that the NLMS non-Wiener effects are still beneficial for interference that is equally strong as (or even slightly weaker than) the communication signal. The step-size parameter of the NLMS algorithm needs to be reduced for optimal performance for a weaker interference.

Figure 2-11 looks at the MSE performance for a varying sinusoidal interference frequency while the ISR is fixed at 26 dB. The observed variation in the NLMS MSE curve is believed to be all statistical fluctuation (i.e., the locations of maxima and minima change from simulation to simulation). Neither Wiener nor NLMS equalizers appear to be affected by interference frequency variation.

![Figure 2-11. MSE of NLMS adaptive equalizer and Wiener equalizer as functions of NBI frequency.](image-url)
2.4.2. Other narrowband interference scenarios

So far, we have concentrated on the sinusoidal interference, which is the focus of the later analyses. We now extend the simulation to observe the existence of the non-Wiener effects under different circumstances. Three cases are presented: with autoregressive narrowband interference, with two sinusoidal interferers, and with a two-ray multi-path channel.

First, we observe the equalizer performance when a strong autoregressive narrowband interference is present; specifically, we use the autoregressive process of order one (AR(1) process) as the interference. In this example, all the interference parameters are consistent with those for Figure 2-4: an ISR of 26 dB and $\omega_i = 0.2\pi$ rad/S. The pole of the AR(1) process is located at $p e^{j\omega_i}$, and the power of the AR(1) innovation process is computed by

$$\sigma^2 = (1 - p^2)\sigma_i^2$$  \hspace{1cm} (2-33)

The auto-correlation function of this interference is derived as

$$n_{ii,l} = \sigma_i^2 (p e^{j\omega_i})^l$$  \hspace{1cm} (2-34)

Accordingly, the input correlation matrix (2-16) is modified, and the Wiener MSE (2-32) is computed based on the correlation matrix for this example.

Figure 2-12 illustrates the MSE behaviors of the Wiener equalizer and the NLMS equalizer for three step-size configurations ($\mu = 1, 0.1,$ and 0.001) as functions of the pole magnitude $p$ of the autoregressive interference of order one (AR(1) interference). The non-Wiener behavior is clearly present when the AR(1) process is extremely narrowband ($p > 0.99999$) and the performance cross-over occurs at $p \approx 0.999$. Above the cross-over point, the NLMS equalizer with larger step-size outperforms that with smaller step-size as well as the Wiener equalizer. On the other hand, as the AR(1) pole departs from the unit circle, the non-Wiener effect quickly disappears and becomes non-existent for $p < 0.9$. This example clearly illustrates that the non-Wiener effect of the adaptive equalizer occurs only when there is narrowband interference.
Next, we observe the existence of the non-Wiener effect when there are two equipower sinusoidal interferences. The two sinusoidal interferences are modeled together as

$$i_n = \frac{\sigma_i}{\sqrt{2}} [e^{j(\omega_1 n + \phi_1)} + e^{j(\omega_2 n + \phi_2)}]$$ \hfill (2-35)

Similar to the single sinusoidal interference case, the frequencies $\omega_1$ and $\omega_2$ are fixed while the phases $\phi_1$ and $\phi_2$ are randomly drawn from $[0, 2\pi)$ for each realization. The two phases are mutually uncorrelated. Evaluating the auto-correlation function of this interference yields

$$r_{ii,l} = \frac{\sigma_i^2}{2} (e^{j\omega_1} + e^{j\omega_2})$$ \hfill (2-36)

Accordingly, the input correlation matrix (2-16) is adjusted, and the Wiener MSE (2-32) is computed based on the correlation matrix for this example.

Figure 2-13 shows the MSE outcome of the simulation, in which the frequency of one sinusoid is varied while the other was fixed at $\omega_1 = 0.2\pi$ rad/S, as in the previous cases.
The overall interference power is maintained at $\sigma_i^2 = 400$. The non-Wiener benefit of the NLMS equalizer is prevalent for all $\omega_2$ values. The best MSE performance is observed when the second sinusoid frequency is slightly away from the first sinusoid, resulting in 8 dB gain in the MSE performance. When two sinusoids are well-separated, the performance gain is significantly smaller, a few dB, but exists across the entire range of interference frequency. The most interesting observation is that the MSE performance with only one sinusoidal interference (i.e., the $\omega_1 = \omega_2$ case) is inferior to that with two distinct, but close, sinusoids by as much as 3 dB.

![Figure 2-13. MSE of Wiener and NLMS adaptive equalizers with two equi-power sinusoidal interferences as a function of the frequency of the second sinusoid. The first sinusoid frequency is fixed at 0.2π.](image)

For the last example, we introduce a non-ideal channel for the communication signal. Thus far, the transmitted signal $x_n$ arrives at the equalizer output as is, uncontaminated, aside from the injection of the additive interference and noise. In this example, we introduce a multi-path effect to the channel so that the received communication signal, $\tilde{x}_n$ in Figure 2-1, is subject to intersymbol interference, and we no longer have $\tilde{x}_n = x_n$, which is assumed for the later analyses. Specifically, we employ a simple two-ray channel with the unit pulse response:
\[ h_n = \alpha \delta_n + \sqrt{1 - \alpha^2} \delta_{n-1} \] (2-37)

This channel maintains the total signal power for any \( \alpha \in [0, 1] \). Subsequently, the received communication signal becomes

\[ \tilde{x}_n = \sum_{k=0}^{\infty} h_k x_{k-n} = \alpha x_n + \sqrt{1 - \alpha^2} x_{n-1} \] (2-38)

This channel changes the correlation function (2-3) of the received communication signal to

\[ r_{\tilde{x}\tilde{x},l} = \begin{cases} \sigma_x^2, & l = 0 \\ \sigma_x^2 \alpha^2, & |l| = 1 \\ 0, & \text{otherwise} \end{cases} \] (2-39)

In this example, we go back to considering a single sinusoidal interference. The Wiener MSE (2-32) is computed with the adjusted input correlation matrix—(2-16) with the communication signal contribution changed according to (2-39).

Figure 2-14 demonstrates the non-Wiener effects in an NLMS equalizer for this two-ray channel as the signal power is split in a different ratio between the two paths. The sinusoidal interference is fixed to the 26-dB ISR (with respect to a perfectly combined communication signal). When there is a dominant multi-path effect, i.e., \( \alpha^2 \sim 0.5 \), most of the performance benefit due to NLMS non-Wiener effects is lost. By using smaller step-size, however, still a small gain—a fraction of a dB—is realized due to adaptation. A similar result was observed in the fractionally-spaced equalizer case [18 Fig.4].
In this chapter, we have firmly established the scope of this study—the adaptive equalization problem with narrowband interference. The included numerical examples clearly illustrate the non-Wiener phenomenon and its potential benefit to MSE performance in adaptive equalization. Both the analysis of the mean weight behavior (Chapter 3) and the analysis of MSE performance (Chapter 4) strictly follow the system descriptions defined in this chapter.
CHAPTER 3 ANALYSIS OF MEAN WEIGHT BEHAVIOR

The adapted weight behavior illustrated in Figure 2-7 is perhaps the most striking result revealed in the preliminary experimental observations in Section 2.3. The conventional theory indicates the mean of the NLMS weights to correspond to the corresponding Wiener filter solution [8], while study of the adaptive narrowband noise canceller exposed deterministic periodic weight behavior [17]. Neither one of these weight behaviors is apparent in Figure 2-7. The adapted weights are clearly away from the Wiener weights, and their only apparent deterministic behavior is the shift itself. The rest of the weight behavior appears random. As the main contribution of this work, this chapter presents the explanation, an analytical solution of the mean of the LMS weight vector in steady state under the large interference-to-signal ratio (ISR) assumption.

The analysis is carried out on the simplified equalization problem, which is depicted in Figure 2-2 with sinusoidal interference. To make the analysis tractable, we first transform the equalization problem into the equivalent interference estimation problem as described in Section 3.1. The derivation of the mean relies on Butterweck’s iterative expansion of the LMS weight update equation [32], which is covered in detail in Section 3.2. Then, Section 3.3 contains the derivation of the mean of the Butterweck partial weights (i.e., the expanded terms for the LMS weights), followed by the derivation of the expressions for the mean of the LMS equalizer weights in Section 3.4. The large ISR weight behavior is analyzed in Section 3.5. The analytical results are then applied to the NLMS equalizer in Section 3.6. Section 3.7 concludes the chapter with numerical illustrations. Simulation results are included to support the analytical results and show that the analytical result predicts the simulation results very well.
3.1. Transformation of Equalization Problem

In Section 2.3, we introduced the two-path interpretation of the Wiener filter (Figure 2-3) which naturally arose from the Wiener filter solution (2-29). Since the non-Wiener weight behavior occurs due to the LMS weights interacting with the narrowband interference, we transform the problem in order to isolate the LMS weight components that correspond to the interference estimation filter—i.e., the bottom path—of the two-path interpretation of the Wiener filter.

Hence, we define the transformed weights $\mathbf{w}_n$ to be related to the equalization weights $\mathbf{\bar{w}}_n$ as follows

$$\mathbf{\bar{w}}_n = \eta (\mathbf{p}_\Delta - \mathbf{w}_n) \quad (3-1)$$

By substituting (3-1) into (2-26), the redefined weights are updated by

$$\mathbf{w}_{n+1} = (\mathbf{I} - \mu \mathbf{u}_n \mathbf{u}_n^H) \mathbf{w}_n + \mu \mathbf{u}_n \mathbf{d}_n^* \quad (3-2)$$

with the new desired signal

$$\mathbf{d}_n \triangleq i_{n-\Delta} - \frac{\sigma_n^2}{\sigma_x^2} x_{n-\Delta} + \bar{n}_{n-\Delta} \cdot \quad (3-3)$$

This transformed desired signal can be seen as a noisy narrowband interference signal, in which the transmitted communication signal $x_n$ acts as an additional noise component. The LMS algorithm in (3-2) adapts $\mathbf{w}_n$ so that $\mathbf{w}_n^H \mathbf{u}_n$ is an estimate of $d_n$. All the analysis and discussion in this chapter center on $\mathbf{w}_n$.

To find the Wiener solution to this problem, we first determine the cross-correlation vector $E[\mathbf{u}_n \mathbf{d}_n^*]$. 
\[
E[u_n d_n^*] = E \left[ (x_n + \hat{n}_n + i_n) \left( -\frac{\sigma_n^2}{\sigma_x^2} x_{n-\Delta} + \hat{n}_{n-\Delta} + i_{n-\Delta} \right) \right]
\]
\[
= -\frac{\sigma_n^2}{\sigma_x^2} E[x_n x_{n-\Delta} + E[\hat{n}_n \hat{n}_{n-\Delta}] + E[i_n i_{n-\Delta}]
\]
\[
= -\frac{\sigma_n^2}{\sigma_x^2} \sigma_\Delta^2 + \sigma_\Delta^2 + E[i_n i_{n-\Delta}]
\]
\[
= E[i_n i_{n-\Delta}]
\]

(3-4)

In the above computation of \( E[u_n d_n^*] \), it is interesting to point out that the terms with \( x_n \) and \( \hat{n}_n \) effectively cancel out, leaving the cross-correlation vector to depend only on \( i_n \). The non-interference, white components in (3-3) are eliminated in (3-4), showing that the problem has been converted to an interference estimation problem. The Wiener filter for this equivalent problem is determined to be

\[
w_w = R^{-1} E[u_n d_n^*]
\]
\[
= \frac{\sigma_i^2}{\lambda_{\max}} e^{j\omega\Delta}
\]

(3-5)

which is, as expected, the interference estimation portion of the original Wiener filter in (2-29). This filter solution is computed using \( R^{-1} \) in (2-17) and \( E[u_n d_n^*] \) in (3-4), and agrees with the interference estimation path of the two-path interpretation in Figure 2-3.

What is implied in (3-4) is that the transformed weights see the sum of the two white input components \( -x_n \) and \( \hat{n}_n \) as one entity. Thus, we introduce the combined white process

\[
n_n \triangleq x_n + \hat{n}_n
\]

(3-6)

with variance \( \sigma_n^2 \triangleq \sigma_x^2 + \sigma_\hat{n}^2 \). Following (2-12) and (2-13), the correlation matrix of this process is defined as

\[
E[n_n n_{n-l}^H] = \begin{cases} 
\sigma_n^2 \mathbf{Z}, & \text{if } 0 \leq l < M \\
\sigma_n^2 \mathbf{N}, & \text{if } -M < l < 0 \\
\mathbf{0}, & \text{otherwise}
\end{cases}
\]

(3-7)
Using this new notation, the input vector is alternatively written as \( u_n = i_n + n_n \), and the alternative expression for the autocorrelation matrix (2-16) of the input becomes

\[
R = \sigma_i^2 ee^H + \sigma_n^2 I
\]  
(3-8)

Moreover, we also notate the white components of the desired signal (3-3) as

\[
n_{d,n} \triangleq -\frac{\sigma_n^2}{\sigma_x^2} x_n + \tilde{n}_n
\]  
(3-9)

Using this notation, the composition of the desired signal can be expressed as \( d_n = i_{n-\Delta} + n_{d,n-\Delta} \). Lastly, \( n_{d,n} \) is uncorrelated with \( n_n \), which can be shown as follows

\[
E\left[n_n n_{d,n-l}^*\right] = E\left[(x_n + \tilde{n}_n)\left(-\frac{\sigma_n^2}{\sigma_x^2} x_{n-l}^* + \tilde{n}_{n-l}^*\right)\right]
= -\frac{\sigma_n^2}{\sigma_x^2} E\left[x_n x_{n-l}^*\right] + E\left[\tilde{n}_n \tilde{n}_{n-l}^*\right]  
(3-10)
= -\frac{\sigma_n^2}{\sigma_x^2} \delta_{l} + \sigma_n^2 \delta_{l}
= 0
\]

Consequently, \( n_{d,n} \) is also uncorrelated with \( u_n \).

### 3.2. Butterweck LMS Weight Expansion

Expanding \( w_n = \sum_{k=0}^{\infty} v_{k,n} \), Butterweck [32] rewrites the LMS weight update equation (3-2) for each of the partial weights \( v_{k,n} \):

\[
v_{k,n+1} = \begin{cases} 
Pv_{0,n} + \mu u_n d_{n}^*, & k = 0 \\
Pv_{k,n} + \mu (R - u_n u_n^H) v_{k-1,n}, & k > 0 \end{cases}
\]  
(3-11)

where \( P \triangleq I - \mu R \). The partial weight \( v_{0,n} \) is referred to as the zero-order solution and the others are called the higher-order correction terms. The partial weight update equations in (3-11) can be viewed as state update equations for a cascade of identical multiple-input multiple-output linear time-invariant systems as depicted in Figure 3-1.
The matrix $P$ acts as the system matrix for each system. The input to the chain is $u_n d_n^*$, while the inputs to the higher-order stages are the product of $(R - u_n u_n^H)$ and the current output of the previous stage, $v_{k-1,n}$.

![Diagram](image)

Figure 3-1. Systems view of the Butterweck expansion of the LMS weight update equation.

This system-based perspective provides several properties useful for the ensuing analysis. First, the system matrix $P$ has one eigenvalue of $1 - \mu \lambda_{\text{max}}$ and $(M - 1)$ repeated eigenvalues of $1 - \mu \sigma_n^2$. For stable operation, the LMS step-size must be chosen so that all these eigenvalues (or equivalently the poles of the system) are located inside the unit circle, i.e., the eigenvalues must fall in between $-1$ and $1$; consequently, $0 < \mu < 2\lambda_{\text{max}}^{-1}$ is the necessary condition for stability. The upper bound results from the fact that the unique eigenvalue reaches the $-1$ limit quicker than the repeated eigenvalues because $\lambda_{\text{max}} \geq \sigma_n^2$.

Secondly, since all the systems are linear and time-invariant, the steady-state solutions are well-defined and as follows.
\[
v_{k,n} = \begin{cases} 
\mu \sum_{p=1}^{\infty} P^{p-1} u_{n-p} q_{n-p}, & k = 0 \\
\mu \sum_{p=1}^{\infty} P^{p-1} (R - u_{n-p} u_{n-p}^H) v_{k-1,n-p}, & k > 0 
\end{cases}
\quad (3-12)
\]

The system is assumed to be operating perpetually in steady-state; therefore, the initial conditions—invoking \( v_{k,0} \) terms—are omitted from (3-12) (as stable system operation will lead to the influence of the initial condition having vanished in steady-state), and its summations are extended to infinity. The state transition matrix \( P^n \) can be expressed as follows (its derivation is included in Appendix A).

\[
P^n = (1 - \mu \sigma_n^2)^n \left[ I + e e^H \frac{1}{M} \left( \left( \frac{1 - \mu \lambda_{\text{max}}}{1 - \mu \sigma_n^2} \right)^n - 1 \right) \right]
\quad (3-13)
\]

Lastly, when a large step-size, i.e., \( \mu \sim \lambda_{\text{max}}^{-1} \), is utilized in the adaptation, the time constant of these systems is almost exclusively determined by the poles at \( 1 - \mu \sigma_n^2 \)—or equivalently the \( (1 - \mu \sigma_n^2)^n \) term in (3-13)—because the eigenvalue \( 1 - \mu \lambda_{\text{max}} \) is located close to the origin. The time constant is found to be

\[
\kappa = -\ln^{-1} \left[ 1 - \mu \sigma_n^2 \right] \approx \frac{1}{\mu \sigma_n^2}
\quad (3-14)
\]

where the final term in (3-14) is by first-order Taylor series approximation. By definition, the time constant indicates the time needed for the unit pulse response of the stable system to decay to \( e^{-1} \) of its initial value. It also gives an idea of the memory depth of the system: the effective number of past input samples contributing to produce the current output sample. By substituting \( \mu = \lambda_{\text{max}}^{-1} \) we find that

\[
\kappa \bigg|_{\mu = \lambda_{\text{max}}^{-1}} = \frac{\sigma_i^2}{\sigma_n^2} M
\quad (3-15)
\]

In other words, the numerical convergence of the summations in (3-12) (i.e., the time it takes for each \( v_{k,n} \) to reach its steady state) is directly proportional to both the ISR—
assuming SNR large—and the filter length.

This slowly converging nature of the state transition matrix $P^n$ leads to a useful approximation. Because $P^n$ does not converge until $p > \kappa \gg M$ for large ISR (i.e., $\sigma_i^2 \gg \sigma_n^2$), the following approximation is valid:

$$E\left\{ n_n^H \sum_{p=1}^{\infty} P^{p-1} n_{n-p} n_{n-p}^H \right\} \approx E\left\{ n_n^H \right\} \sum_{p=1}^{\infty} P^{p-1} E\left\{ n_{n-p} n_{n-p}^H \right\} \quad (3-16)$$

This results from three facts. First, for $p = 1$ and $p \geq M$,

$$E\left\{ n_n^H P^{p-1} n_{n-p} n_{n-p}^H \right\} = E\left\{ n_n^H \right\} P^{p-1} E\left\{ n_{n-p} n_{n-p}^H \right\} \quad (3-17)$$

due to the whiteness of $n_n$. Secondly, that according to (3-15), $P^n$ does not converge until $p \gg M$ for large ISR (i.e., $\sigma_i^2 \gg \sigma_n^2$). Lastly, both expected values on the right hand side of (3-17) are sufficiently large for $p \geq M$. Consequently, the uncorrelated terms for $p \geq M$ dominate the correlated terms (those for $p = 2$ to $M-1$), resulting in (3-16).

A corollary to (3-16) is that the mean of the Butterweck partial weight vectors $v_k \triangleq E[v_{k,n}]$ for all $k$ is essentially uncorrelated to the square of the white component of the input, $n_n$, as defined in (2-12) and (2-13):

$$E[n_n^H v_{k,n}] \approx E[n_n^H] E[v_{k,n}] \quad (3-18)$$

To show the validity of this approximation, expanding $v_{k,n}$ using (3-12) yields

$$E[n_n^H v_{k,n}] = \begin{cases} \mu E\left[ n_n^H \sum_{p=1}^{\infty} P^{p-1} u_{n-p} d_{n-p}^* \right], & k = 0 \\ \mu E\left[ n_n^H \sum_{p=1}^{\infty} P^{p-1} (R - u_{n-p} u_{n-p}^H) v_{k-1,n-p} \right], & k > 0 \end{cases} \quad (3-19)$$

This expansion reveals the infinite-summation form that is reminiscent of (3-16). The only $v_{k,n}$ constituent terms that are correlated to the $n_n^H$ term are terms containing $n_n$, i.e., $u_{n-p}$, $d_{n-p}$, and $v_{k-1,n-p}$, all are delayed by the summation index $p$, similar to
(3-16). Because $n_n$ is white, only the first $M - 1$ terms of the summation can be correlated to $n_n$. Consequently, applying the approximation in (3-16), $n_n n_n^H$ and $v_{k,n}$ are essentially uncorrelated, resulting in (3-18), as long as the expected value of the terms inside of the summation is sufficiently large, i.e., for partial weights with sufficiently significant magnitude. The approximation in (3-18) can be applied to related forms such as

$$E[ n_n n_n^H v_{k,n-q} ] \approx E[ n_n n_n^H ] E[ v_{k,n-q} ], \quad q \geq 0$$

and

$$E[ v_{k,n-q} n_n^H p^{q-1} n_n - v_{k,n-q} ] \approx E[ v_{k,n-q} ] E[ n_n^H p^{q-1} n_n - v_{k,n-q} ], \quad q > 0$$

The delays introduced in (3-20) and (3-21) do not change the approximation as long as the partial weight vector is the oldest term (which is guaranteed by the $q > 0$ condition). We will use these approximations, (3-18), (3-20), and (3-21), in the next section.

### 3.3. Mean of Butterweck Partial Weights

Defining the expected value of the Butterweck partial weights as $\bar{v}_k \triangleq E[v_{k,n}]$, the mean LMS weight can be found by $\bar{w} = \sum_{k=0}^{\infty} \bar{v}_k$. We first establish the basic cases by evaluating $\bar{v}_0$ and $\bar{v}_1$, showing that the latter mean weight vanishes for large ISR. Then, the expected value of the recursive expression in (3-12) is analyzed.

Taking the expected value of (3-12) yields

$$\bar{v}_k = \begin{cases} \mu \sum_{p=1}^{\infty} p^{p-1} E[ u_{n-p} d_{n-p}^c ], & k = 0 \\ \mu \sum_{p=1}^{\infty} p^{p-1} E[ (R - u_{n-p} u_{n-p}^H) v_{k-1,n-p} ], & k > 0 \end{cases}$$

(3-22)

Because all the signals involved are wide-sense-stationary processes, the expected values in (3-22) are invariant to a uniform time-shift of its signals. We also have
\[ \mu \sum_{p=1}^{\infty} P^{p-1} = \mu (I - P)^{-1} = R^{-1} \] (3-23)

Consequently, (3-22) simplifies to

\[ \bar{v}_k = \begin{cases} R^{-1}E[u_n d_n^*], & k = 0 \\ R^{-1}E[(R - u_n u_n^H) v_{k-1,n}], & k > 0 \end{cases} \] (3-24)

It is apparent that \( \bar{v}_0 \) in (3-24) is the Wiener solution (3-5) of the transformed interference estimation problem. Hence,

\[ \bar{v}_0 = w_w \] (3-25)

In a typical application of the LMS algorithm, we expect the weights to vary around this Wiener solution; in other words, the higher-order partial weights are zero-mean. In our problem, however, we are anticipating non-zero means for the higher-order correction terms to explain the observed shift in the mean weights away from the Wiener solution.

We now take the expression for \( (k > 0) \) in (3-24) and will observe from a later result that only two terms, \( n_n i_n^H \) and \( i_n n_n^H \), in \( (R - u_n u_n^H) \) influence the expression for the mean. The higher-order expression of (3-24) expands to

\[ \bar{v}_k = R^{-1}E[(R - n_n n_n^H - i_n i_n^H) v_{k-1,n}] - R^{-1}E[(n_n i_n^H + i_n n_n^H) v_{k-1,n}] \] (3-26)

Evaluating the first right-hand-side term in (3-26), we expand it further to

\[ E[(R - n_n n_n^H - i_n i_n^H) v_{k-1,n}] = R \bar{v}_{k-1} - \sigma_n^2 \bar{v}_{k-1} - \sigma_i^2 ee^H \bar{v}_{k-1} \] (3-27)

The expected value of the second term follows the approximation in (3-18). The expected value of the last term follows from the sinusoidal process being deterministic in each realization as shown earlier in (2-11). Finally, recognizing (3-8) in (3-27) yields
\[
E\left[ (R - n_n^H n_n^H - i_n^H i_n^H) v_{k-1,n} \right] \approx \left[ R - (\sigma_n^2 I + \sigma_i^2 e e^H) \right] \bar{v}_{k-1} = 0
\] (3-28)

Hence, the first right-hand-side term in (3-26) vanishes, which results in (3-26) being reduced to

\[
\bar{v}_k \approx -R^{-1}E\left[ (n_n^H i_n^H + i_n^H n_n^H) v_{k-1,n} \right]
\] (3-29)

From (3-29), we next proceed to evaluate \( \bar{v}_1 \) and a recursive expression for \( \bar{v}_k \) for \( k > 1 \).

For \( k = 1 \), substituting \( v_{0,n} \) in (3-12) into (3-29) yields

\[
\bar{v}_1 = -R^{-1}E\left[ (n_n^H i_n^H + i_n^H n_n^H) \mu \sum_{p=1}^{\infty} P^{p-1} u_{n-p}^* d_{n-p}^\ast \right]
\] (3-30)

First, by \( n_{d,n-\Delta} \) being uncorrelated to both \( u_n \) and \( n_n \), (3-30) simplifies to

\[
\bar{v}_1 = -R^{-1}E\left[ (n_n^H i_n^H + i_n^H n_n^H) \mu \sum_{p=1}^{\infty} P^{p-1} u_{n-p}^* i_{n-p-\Delta}^* \right]
\] (3-31)

Expanding \( u_n = i_n + n_n \), (3-30) becomes

\[
\bar{v}_1 = -R^{-1}E\left[ n_n^H i_n^H \mu \sum_{p=1}^{\infty} P^{p-1} i_{n-p}^* i_{n-p-\Delta}^* \right]^{(b)} - R^{-1}E\left[ n_n^H \mu \sum_{p=1}^{\infty} P^{p-1} n_{n-p}^* i_{n-p-\Delta}^* \right]^{(c)}
\]

\[
-R^{-1}E\left[ i_n^H n_n^H \mu \sum_{p=1}^{\infty} P^{p-1} i_{n-p}^* i_{n-p-\Delta}^* \right]^{(b)} - R^{-1}E\left[ i_n^H \mu \sum_{p=1}^{\infty} P^{p-1} n_{n-p}^* i_{n-p-\Delta}^* \right]^{(a)}
\] (3-32)

Only the last term, labeled (a), evaluates to a non-zero value. The terms with the label (b) vanish because \( n_n \) is zero-mean and is uncorrelated to \( i_n \). The (c) term vanishes by the properness of \( n_n \) and \( i_n \). Removing all the vanishing terms results in

\[
\bar{v}_1 = -R^{-1} \mu \sum_{p=1}^{\infty} E\left[ i_n^H n_n^H P^{p-1} n_{n-p}^* i_{n-p-\Delta}^* \right].
\] (3-33)
This term can be rewritten as

\[
\bar{v}_1 = -R^{-1}\mu \sum_{p=1}^{\infty} E\left[i_{n_{n-p}}^*\right] tr \left\{ E\left[ P^{p-1} n_{n-p} n_n^H \right] \right\} \\
= -R^{-1}\mu \sum_{p=1}^{\infty} E\left[i_{n_{n-p}}^*\right] tr \left\{ P^{p-1} E\left[ n_{n-p} n_n^H \right] \right\} \\
= -R^{-1}\mu \sigma_n^2 e^{j\omega_\Delta} \sigma_n^2 \mu \sum_{p=1}^{M-1} e^{j\omega_\Delta p} tr \left\{ P^{p-1} N^p \right\} \tag{3-34}
\]

Using (3-13), the trace expression in (3-34) results in

\[
tr \left\{ P^{p-1} N^p \right\} = (1 - \mu \sigma_n^2)^{p-1} tr \left\{ N^p + ee^H N^p \frac{1}{M} \left[ \left( \frac{1 - \mu \lambda_{\max}}{1 - \mu \sigma_n^2} \right)^{p-1} - 1 \right] \right\} \tag{3-35}
\]

Noting that \( tr\{N^p\} = 0 \) for all \( p > 0 \) and \( tr\{ee^H N^p\} = e^H N^p e = (M - p) e^{-j\omega p} \), (3-35) can be further manipulated

\[
tr \left\{ P^{p-1} N^p \right\} = (1 - \mu \sigma_n^2)^{p-1} \left[ \frac{1 - \mu \lambda_{\max}}{1 - \mu \sigma_n^2} \right]^{p-1} - 1 \frac{M - p}{M} e^{-j\omega p} \\
= \left[ (1 - \mu \lambda_{\max})^{p-1} - (1 - \mu \sigma_n^2)^{p-1} \right] \frac{M - p}{M} e^{-j\omega p} \tag{3-36}
\]

With (3-36), the term inside the summation of (3-34) can be shown to be bounded. Substituting (3-36) into the summation portion of (3-34) yields

\[
e^{j\omega p} tr \left\{ P^{p-1} N^p \right\} = \left[ (1 - \mu \lambda_{\max})^{p-1} - (1 - \mu \sigma_n^2)^{p-1} \right] \frac{M - p}{M} \tag{3-37}
\]

Note that this quantity is real-valued. Evaluating over the stable step-size range, i.e., \( 0 < \mu < 2\lambda_{\max}^{-1} \) and for all relevant \( p > 0 \), the following (loose but sufficient) bounds to each term of (3-37) are determined:

\[
-1 < 1 - \mu \lambda_{\max}^{-1} < 1 \tag{3-38}
\]
Using (3-38) through (3-40), we find the bounds to (3-37) to be

\[-2 < e^{j\omega_p} tr \left\{ P^{p-1} N^p \right\} < \frac{2\sigma_n^2}{\lambda_{\text{max}}} \]  

(3-41)

Consequently, the upper bound to \( \bar{v}_1 \) in (3-34) is determined using the lower bound found in (3-41):

\[ \bar{v}_1 < w_w \sigma_n^2 \mu \sum_{p=1}^{M-1} 2 = w_w \sigma_n^2 \mu (M - 1) \]  

(3-42)

The Wiener solution \( w_w \) is substituted into (3-42) for \( R^{-1}\sigma_i^2 e^{j\omega_i \Delta} \). This bound is monotonically increasing as a function of \( \mu \); hence, we can further bound it by substituting \( \mu = 2\lambda_{\text{max}}^{-1} \):

\[ \bar{v}_1 < w_w \frac{4(M - 1)\sigma_n^2}{\lambda_{\text{max}}} = w_w \frac{4(M - 1)\sigma_n^2}{\sigma_i^2 M + \sigma_n^2} \]  

(3-43)

Finally, assuming the \( \sigma_i^2 \gg \sigma_n^2 \) condition, we conclude that \( \bar{v}_1 \rightarrow 0 \).

For the higher order \( \bar{v}_k \) terms for \( k > 1 \), the key to establishing the recursive expression for the mean of the partial weight vector is to show that \( n_i^H i_n^H + i_n^H n_n^H \) and \( v_{k-1,n} \) are essentially uncorrelated. Replacing \( v_{k-1,n} \) in (3-29) for \( k > 1 \), from (3-12) yields

\[ \bar{v}_k \approx -R^{-1}E \left[ \left( n_i^H i_n^H + i_n^H n_n^H \right) \mu \sum_{p=1}^{\infty} P^{p-1} \left( R - u_{n-p}^H u_{n-p}^H \right) v_{k-1,n-p} \right] \]  

(3-44)

By following the steps from (3-26) to (3-29) to reduce \( \left( R - u_{n-p}^H u_{n-p}^H \right) \), (3-44) becomes...
\[
\bar{v}_k \approx R^{-1} E \left[ (n_i i_n^H + i_n n_n^H) \mu \sum_{p=1}^{\infty} P^{p-1} (n_{n-p} i_n^H + i_{n-p} n_n^H) v_{k-2,n-p} \right]
\]  
(3-45)

Expanding (3-45) yields:

\[
\bar{v}_k \approx R^{-1} E \left[ \mu \sum_{p=1}^{\infty} n_i i_n^H P^{p-1} n_{n-p} i_{n-p} v_{k-2,n-p} \right] \quad \cdots (b)
\]

\[
+ R^{-1} E \left[ \mu \sum_{p=1}^{\infty} i_n n_n^H P^{p-1} i_{n-p} n_{n-p} v_{k-2,n-p} \right] \quad \cdots (a)
\]

\[
+ R^{-1} E \left[ \mu \sum_{p=1}^{\infty} i_n n_n^H P^{p-1} n_{n-p} i_{n-p} v_{k-2,n-p} \right] \quad \cdots (a)
\]

\[
+ R^{-1} E \left[ \mu \sum_{p=1}^{\infty} i_n n_n^H P^{p-1} i_{n-p} n_{n-p} v_{k-2,n-p} \right] \quad \cdots (b)
\]

(3-46)

The (a) terms result in a non-zero expected value because \( n_i \) and \( n_{n-p}^H \) (or the conjugates of them) are correlated for \( 1 \leq p < M \). On the other hand, by the properness of \( i_n \) and \( n_n \), the signals in the (b) terms result in pseudo-correlations equal to zero. Hence, for the (b) terms to be non-zero, its signal vectors must correlate with the partial weight vector \( v_{k-2,n-p} \). However, \( v_{k-2,n-p} \) is constructed using noise vectors older than \( n_{n-p-1} \); therefore, the correlation of the noise vector in (3-46), e.g., \( n_i \) and \( n_{n-p} \), and the previous noise contribution to \( v_{k-2,n-p} \) is much smaller than the correlated noise vectors in the (a) terms. Hence, the (a) terms dominate the (b) terms, resulting in

\[
\bar{v}_k \approx R^{-1} \mu \sum_{p=1}^{\infty} E \left[ n_i i_n^H P^{p-1} i_{n-p} n_{n-p}^H v_{k-2,n-p} \right] \quad \cdots (3-47)
\]

From (3-47), the terms inside of the expected value operators are regrouped, and the expected values involving interference terms are separated from the sinusoidal process, which is deterministic in each realization as shown earlier in (2-11):
\[
\bar{v}_k = R^{-1} \mu \sum_{p=1}^{\infty} E\left[ n_n n_{n-p}^H v_{k-2,n-p} \right] E\left[ i_n^H P^{p-1} i_{n-p} \right] + R^{-1} \mu \sum_{p=1}^{\infty} E\left[ i_n^H i_{n-p} \right] E\left[ v_{k-2,n-p} n_n^H P^{p-1} n_{n-p} \right]
\] (3-48)

The expected value of the partial weights can be separated by employing the approximation in (3-20) for the \( E\left[ n_n n_{n-p}^H v_{k-2,n-p} \right] \) term and the approximation in (3-21) for \( E\left[ v_{k-2,n-p} n_n^H P^{p-1} n_{n-p} \right] \). Employing these approximations to (3-48) (with minor matrix manipulation) yields

\[
\bar{v}_k = R^{-1} \mu \sum_{p=1}^{\infty} E\left[ n_n n_{n-p}^H \right] \bar{v}_{k-2} tr\left\{ P^{p-1} E\left[ i_{n-p} i_n^H \right] \right\} + R^{-1} \mu \sum_{p=1}^{\infty} E\left[ i_n i_{n-p} \right] \bar{v}_{k-2} tr\left\{ P^{p-1} E\left[ n_{n-p} n_n^H \right] \right\}
\] (3-49)

Using (2-14) for the \( E\{i_n i_{n-p}^H\} \) terms and (3-7) for the \( E\{n_n n_{n-p}^H\} \) terms, substituting for the expected value expressions in (3-49) results in

\[
\bar{v}_k = R^{-1} \sigma^2_n \sigma^2_i \mu \sum_{p=1}^{M-1} Z^p \bar{v}_{k-2} e^{H} P^{p-1} e^{-j\omega_p} + R^{-1} \sigma^2_n \sigma^2_i \mu \sum_{p=1}^{M-1} e e^{H} \bar{v}_{k-2} tr\left\{ P^{p-1} N^p \right\} e^{j\omega_p}
\] (3-50)

The upper limit for the summations in (3-50) has became finite as \( Z^p = N^p = O \) for \( p \geq M \). Of the two terms appearing in (3-50) only the first term is found to be significant. The last term vanishes as follows. Using the upper bound defined in (3-41) for \( e^{j\omega_p} tr\{P^{p-1}N^p\} \), the last term of (3-50) is bounded above

\[
R^{-1} e e^{H} \bar{v}_{k-2} \sigma^2_n \sigma^2_i \mu \sum_{p=1}^{M-1} tr\left\{ P^{p-1} N^p \right\} e^{j\omega_p} < R^{-1} e e^{H} \bar{v}_{k-2} \sigma^2_n \sigma^2_i \mu \frac{2\sigma^2_n}{\lambda_{\max}} (M - 1)
\] (3-51)

Since
\[
\frac{\sigma_t^2 (M - 1)}{\lambda_{\text{max}}} < 1 \tag{3-52}
\]

(3-51) can be further bounded by

\[
\mathbf{R}^{-1}\mathbf{e}\mathbf{e}^H \mathbf{v}_{k-2} \sigma_n^2 \sigma_t^2 \mu \sum_{p=1}^{M-1} \text{tr} \left\{ \mathbf{P}^{p-1} \mathbf{N}^p \right\} e^{j\omega p} < \mathbf{R}^{-1}\mathbf{e}\mathbf{e}^H \mathbf{v}_{k-2} 2 \sigma_n^2 \sigma_t^2 \mu \tag{3-53}
\]

Substituting (2-17), \( \mathbf{R}^{-1}\mathbf{e}\mathbf{e}^H \) in (3-53) simplifies as follows

\[
\mathbf{R}^{-1}\mathbf{e}\mathbf{e}^H = \frac{1}{\sigma_n^2} \left( \mathbf{I} - \frac{\sigma_t^2}{\lambda_{\text{max}}} \mathbf{e}\mathbf{e}^H \right) \mathbf{e}\mathbf{e}^H = \frac{1}{\sigma_n^2} \mathbf{e} \left( 1 - \frac{\sigma_t^2 M}{\lambda_{\text{max}}} \right) \mathbf{e}^H = \frac{1}{\lambda_{\text{max}}} \mathbf{e}\mathbf{e}^H \tag{3-54}
\]

Substituting (3-54) into the right hand side of (3-53) yields

\[
\mathbf{R}^{-1}\mathbf{e}\mathbf{e}^H \mathbf{v}_{k-2} \sigma_n^2 \sigma_t^2 \mu \sum_{p=1}^{M-1} \text{tr} \left\{ \mathbf{P}^{p-1} \mathbf{N}^p \right\} e^{j\omega p} < \mathbf{v}_{k-2} 2 \mu \sigma_n^2 \frac{\sigma_n^2}{\lambda_{\text{max}}} \tag{3-55}
\]

\[
= \mathbf{v}_{k-2} 2 \mu \sigma_n^2 \left\{ 1 + \frac{\sigma_t^2 M}{\sigma_n^2} \right\}
\]

The denominator in the final expression contains the ISR, \( \sigma_t^2 / \sigma_n^2 \). Therefore, as the ISR increases, the denominator increases at the same rate, resulting in the bound going to zero. Hence, this term also vanishes under the large ISR condition. Hence, (3-50) becomes

\[
\mathbf{v}_k = \mathbf{R}^{-1} \sigma_n^2 \sigma_t^2 \mu \sum_{p=1}^{M-1} Z^p \mathbf{v}_{k-2} \mathbf{e}\mathbf{e}^H \mathbf{P}^{p-1} \mathbf{e} e^{-j\omega p} \tag{3-56}
\]

We introduce the following results for \( p > 0 \), derived from (3-13):

\[
\mathbf{e}\mathbf{e}^H \mathbf{P}^{p-1} \mathbf{e} = M \left( 1 - \mu \lambda_{\text{max}} \right)^{p-1} \tag{3-57}
\]
Substituting (3-57) in (3-56) yields
\[
\bar{v}_k \simeq R^{-1}\sigma_\eta^2\sigma_i^2 M \mu \sum_{p=1}^{M-1} Z^p \bar{v}_{k-2} e^{-j\omega_p} (1 - \mu \lambda_{\text{max}})^{p-1}
\] (3-58)

For a more compact notation, we introduce
\[
Q \triangleq \sigma_\eta^2\sigma_i^2 M \sum_{p=1}^{M-1} Z^p e^{-j\omega_p} \gamma^{p-1}
\] (3-59)

with \(\gamma \triangleq 1 - \mu \lambda_{\text{max}}\) so that (3-58) becomes
\[
\bar{v}_k = \mu R^{-1}Q\bar{v}_{k-2}.
\] (3-60)

The matrix \(Q \in \mathbb{C}^{M \times M}\) is a lower-triangular Toeplitz matrix with all diagonal elements equal to zero. The \(m\)-th element of the first column of \(Q\) equals
\[
q_{m,1} = \begin{cases} 
0, & m = 1 \\
\sigma_\eta^2\sigma_i^2 M \gamma^{(m-2)} e^{-j(m-1)\omega}, & 2 \leq m \leq M 
\end{cases}
\] (3-61)

The recursive expression in (3-60) for even \(k\) has its base case in (3-25). Meanwhile, all \(\bar{v}_k\) for odd \(k\) vanish as ISR increases, along the lines of their base case \(\bar{v}_1\), which was found to be bounded in (3-43) and consequently vanished.

### 3.4. Mean of LMS Weights

The derived Butterweck partial weight vectors in the previous section are combined to form the mean for the LMS weight vector in Section 3.4.1. This section also provides the further analysis on the resultant mean vector to separate the magnitude and phase components. Furthermore, in Section 3.4.2, the mean expression for the each weight is obtained from the mean vector.

#### 3.4.1. Mean of LMS weight vector and its magnitude-phase decomposition

Assuming large ISR, only the \(\bar{v}_k\) for even \(k\) contribute to the mean of the steady-
state LMS weight $\bar{w}$. Unwrapping the recursion in (3-60), the $2l$-th even term is computed from

$$\bar{v}_{2l} = (\mu R^{-1}Q)^l \bar{v}_0$$

(3-62)

Hence, the mean of the steady-state LMS weight vector is found as

$$\bar{w} \approx \sum_{l=0}^{\infty} \bar{v}_{2l}$$

$$= \sum_{l=0}^{\infty} (\mu R^{-1}Q)^l \bar{v}_0$$

(3-63)

Hence, the closed-form solution for the mean LMS weight vector is then

$$\bar{w} = (I - \mu R^{-1}Q)^{-1} \bar{v}_0$$

(3-64)

As $\mu \to 0$, $(I - \mu R^{-1}Q)^{-1} \to I$ and subsequently, as expected, (3-64) tends to the Wiener solution as step-size vanishes. However, the large-step-size behavior is not apparent from this formulation.

From (3-64), the magnitude and phase components can be separated (also applicable to the partial weights). The phase of the mean weights only depends on the interference frequency $\omega_i$ and on the desired signal delay $\Delta$ by

$$\angle \bar{w} = e^{j\omega_i \Delta}$$

(3-65)

as it is a required condition for the filter to correctly estimate $i_{n-\Delta}$. The magnitude of the mean of the weights, however, is found to be invariant to $\omega_i$ and $\Delta$. In other words, the magnitude vector of the mean LMS weights for any arbitrary $\omega_i$ can be computed from

$$|\bar{w}| = \bar{w}|_{\omega_i=0}$$

(3-66)

For $\omega_i = 0$, $R^{-1}$, defined in (2-17) becomes
\[ \Xi \triangleq R^{-1}|_{\omega_i=0} = \frac{1}{\sigma_i^2} \left( I - \frac{\sigma_i^2}{\lambda_{\max}} 11^H \right) \] (3-67)

where \( \mathbf{1} \in \mathbb{R}^M \) is a vector of all ones, i.e., \( \mathbf{e}_{\omega_i=0} \). Likewise, \( Q \) in (3-59) becomes

\[ \Theta \triangleq Q|_{\omega_i=0} = \sigma_i^2 \sigma_i^2 M \sum_{p=1}^{M-1} \mathbf{Z}^p \gamma^p \] (3-68)

Consequently, the magnitude expression for \( \overline{v}_{2l} \)—which is defined in (3-25) for \( l = 0 \) and in (3-63) for \( l > 0 \)—is expressed as

\[ |\overline{v}_{2l}| = \begin{cases} \frac{\sigma_i^2}{\lambda_{\max}} \mathbf{1}, & l = 0 \\ \frac{\sigma_i^2}{\lambda_{\max}} (\mu \Xi \Theta)^i \mathbf{1}, & l > 0 \end{cases} \] (3-69)

Finally, the magnitude vector of the mean of the LMS weights becomes

\[ |\overline{w}| = \frac{\sigma_i^2}{\lambda_{\max}} (I - \mu \Xi \Theta)^{-1} \mathbf{1}. \] (3-70)

A detailed derivation is given in Appendix B.

3.4.2. Expression for individual weights

Additionally, we now derive an inverse-free expression for (3-70), which leads to the final expression for the individual weights. We first expand (3-70) by substituting (3-67) and (3-68):

\[ |\overline{w}| = \frac{\sigma_i^2}{\lambda_{\max}} \left( I - \mu \sigma_i^2 M \frac{1}{\lambda_{\max}} \mathbf{11}^T \right) \sum_{p=1}^{M-1} \mathbf{Z}^p \gamma^p \right)^{-1} \mathbf{1} \]

\[ = \frac{\sigma_i^2}{\lambda_{\max}} \left( I - \mu \sigma_i^2 M \sum_{p=1}^{M-1} \mathbf{Z}^p \gamma^p \right) \frac{1}{\lambda_{\max}} \sum_{p=1}^{M-1} \mathbf{Z}^p \gamma^p \right)^{-1} \mathbf{1} \] (3-71)
In preparation to applying the matrix inversion lemma to the inverse term in (3-71), we define the following.

\[
L^{-1} \triangleq I - \mu \sigma_i^2 M \sum_{p=1}^{M-1} Z^p \gamma^{p-1}. \tag{3-72}
\]

Since the right-hand side matrix in (3-72) is unit lower triangular and Toeplitz, its inverse $L$ exists, and $L$ is also a unit lower triangular Toeplitz matrix. The $m$-th element of the first column $L$ is found in a closed-form solution as (see Appendix C for the derivation)

\[
l_{m,1} = \begin{cases} 
1, & m = 1 \\
\mu \sigma_i^2 M \left(1 - \mu \sigma_n^2 \right)^{m-2}, & 1 < m \leq M 
\end{cases} \tag{3-73}
\]

Because of its Toeplitz structure, (3-73) completely defines the matrix $L$. We also define

\[
\theta^T \triangleq 1^T \sum_{p=1}^{M-1} Z^p \gamma^{p-1}. \tag{3-74}
\]

of which the elements are found to be

\[
\theta_m \triangleq \{ \theta^T \}_{m} \\
= \left\{ \sum_{p=1}^{M-m} \gamma^{p-1} \right\}_m \\
= \frac{1 - \gamma^{M-m}}{\mu \lambda_{\max}}, \quad |\gamma| < 1 \tag{3-75}
\]

The denominator in (3-75) results from the definition, $\gamma \triangleq 1 - \mu \lambda_{\max}$. The expression in (3-75) evaluates to $M - m$ when $\gamma = 1$. Substituting (3-72) and (3-74) into (3-71) yields

\[
|\bar{w}| = \frac{\sigma_i^2}{\lambda_{\max}} \left( L^{-1} + \mu \sigma_i^2 M \frac{\sigma_i^2}{\lambda_{\max}} 1^T \right)^{-1} 1. \tag{3-76}
\]

Now, we apply the Woodbury identity [34 Appendix A] to (3-76) and get
\[ |\bar{w}| = \frac{\sigma_i^2}{\lambda_{\text{max}}} \left( L - \frac{L_1 \theta^T L}{\left( \mu \sigma_i^2 M \frac{\sigma_i^2}{\lambda_{\text{max}}} \right)^{-1} + \theta^T L_1} \right) 1. \] (3-77)

Further algebraic manipulation results in

\[ |\bar{w}| = L_1 \frac{\sigma_i^2}{\lambda_{\text{max}}} \frac{1}{1 + \mu \sigma_i^2 M \theta^T L_1 \frac{\sigma_i^2}{\lambda_{\text{max}}}}. \] (3-78)

Since \( L \) is a Toeplitz lower triangular matrix, we find the \( m \)-th element of the vector \( L_1 \), which we denote as \( f_m \), by

\[ f_m \triangleq \{L_1\}_m = \sum_{k=1}^{m} l_{k,1}. \] (3-79)

Substituting (3-73) into (3-79) yields

\[ f_m = \begin{cases} 1, & m = 1 \\ 1 + \mu \sigma_i^2 M \sum_{k=0}^{m-2} (1 - \mu \sigma_n^2)^k, & 1 < m \leq M \\ \frac{\lambda_{\text{max}} - \sigma_i^2 M \left( 1 - \mu \sigma_n^2 \right)^{m-1}}{\sigma_n^2}, & 1 < m \leq M \end{cases} \] (3-80)

We denote the denominator of (3-78) as

\[ L \triangleq 1 + \mu \sigma_i^2 M \frac{\sigma_i^2}{\lambda_{\text{max}}} \sum_{m=1}^{M} \theta_m f_m \] (3-81)

Finally, combining (3-65), (3-78), and (3-80), \( \bar{w}_m \), the \( (m + 1) \)-st element of \( \bar{w} \), is given by
Chapter 3: Analysis of Mean Weight Behavior

\[
\bar{w}_m = \frac{\sigma_i^2}{\lambda_{\text{max}} L} f_m e^{j\omega_i (\Delta - m)}
\]

\[
= \begin{cases} 
\frac{\sigma_i^2}{\lambda_{\text{max}} L} e^{j\omega_i \Delta}, & m = 0 \\
\frac{\sigma_i^2}{\sigma_n^2 L} \left[ 1 - \frac{\sigma_i^2 M}{\lambda_{\text{max}}} (1 - \mu \sigma_n^2)^m \right] e^{j\omega_i (\Delta - m)}, & 1 < m < M 
\end{cases}
\]  

Care must be taken when computing (3-82) as the weight index in (3-82) is zero-base (to be used in a standard convolution formula) while the index in (3-75) is in one-base notation (more natural for indexing vector elements).

3.5. LMS Weight Behavior with Large ISR

The weight expression obtained in (3-82) is still complex and does not provide instant insight to the equalizer weight behavior. We now evaluate the limiting behavior of the equalizer as the ISR is driven large. First, we observe the limiting behavior of the Wiener weight magnitude (which is used in (3-82)):

\[
\lim_{\sigma_i^2/\sigma_n^2 \to \infty} \frac{\lim_{\sigma_i^2/\sigma_n^2 \to \infty} \frac{\sigma_i^2}{\lambda_{\text{max}}}}{1 - \frac{\sigma_i^2 M}{\sigma_n^2 + \sigma_i^2 M}} = \frac{1}{M}
\]  

(3-83)

Next, the limit of the \( f_m \) term is evaluated from its summation form—i.e., the first right hand side term in (3-80)—because, for the stable step-size range, we have \( (1 - \mu \sigma_n^2)^k \to 1 \) as the ISR is increased.

\[
\lim_{\sigma_i^2/\sigma_n^2 \to \infty} f_m = \begin{cases} 
1, & m = 1 \\
\lim_{\sigma_i^2/\sigma_n^2 \to \infty} \left[ 1 + \mu \sigma_i^2 M \sum_{k=0}^{m-2} (1 - \mu \sigma_n^2)^k \right], & 1 < m \leq M. 
\end{cases}
\]  

(3-84)

\[
= 1 + \mu \sigma_i^2 M (m - 1)
\]

The last term in (3-82) remaining is \( L \). The mathematical solution for the limit of the
summation term in (3-81) has not been derived; however, its tendency as the ISR increase is observed in simulation and is hypothesized as

\[
\sum_{m=1}^{M} \theta_m f_m \xrightarrow{\sigma_i^2/\sigma_i^2 \to \infty} \frac{M-1}{2} \quad (3-85)
\]

Using (3-85) and (3-83) we have

\[
\lim_{\sigma_i^2/\sigma_i^2 \to \infty} L = 1 + \mu \sigma_i^2 \frac{M-1}{2} \quad (3-86)
\]

The limiting weight behavior is found by taking the limit of (3-82) as the ISR is increased and substituting (3-83), (3-84), and (3-86) into the limit expression.

\[
\lim_{\sigma_i^2/\sigma_i^2 \to \infty} \bar{w}_m = \frac{2(1 + m \mu \sigma_i^2)}{M[2 + (M-1)\mu \sigma_i^2]} e^{j\omega_i(\Delta-m)} \quad (3-87)
\]

Finally, evaluating (3-87) for \(\mu = 1/\sigma_i^2\), we get

\[
\lim_{\sigma_i^2/\sigma_i^2 \to \infty} \bar{w}_m |_{\mu = \sigma_i^{-2}} = \frac{2(m+1)}{M(M+1)} e^{j\omega_i(\Delta-m)} \quad (3-88)
\]

This solution corresponds to the spiral weight formation in Figure 2-7. The weight magnitude is an increasing linear function of the weight index. That is, the newest sample \(u_{n-M+1}\) in the input delay line gets the smallest weight, and conversely the oldest \(u_{n-M+1}\) is weighted the most. We hypothesize that this weight behavior is for the LMS algorithm to take advantage of the input sample that has been in the system for a while through some form of time-varying mechanism. The exact nature of the time-varying behavior is not captured in the analysis of the mean as the statistical analysis neglects any instantaneous behavior of the algorithm. The investigation of the instantaneous weight perturbation is left as an open extension to this study.
3.6. Application to NLMS Algorithm

As a final remark regarding the mean weight analysis result, the derived mean weight expression can be applied to the NLMS algorithm (see Section 2.3) by a simple adjustment in the LMS step-size definition. First, the transformation of the problem applied in Section 3.1 to the LMS algorithm directly applies to the NLMS algorithm. Using the weight relationship in (3-1), the NLMS weight update equation (2-27) with the error signal (2-24) transforms to

$$w_{n+1}^{(NLMS)} = (I - \hat{\mu}_n (u_n^H u_n)^{-1} u_n^H) w_n^{(NLMS)} + \hat{\mu}_n (u_n^H u_n)^{-1} d_n$$

(3-89)

with the same desired signal as in (3-3). Equivalently, we can express the NLMS algorithm as the LMS algorithm (3-2) with a variable step-size such that

$$\mu_n = \hat{\mu} (u_n^H u_n)^{-1}$$

(3-90)

Approximating the inverse term in (3-90) by

$$u_n^H u_n \approx M(\sigma_n^2 + \sigma_i^2)$$

(3-91)

the expected value for the NLMS weight vector is then

$$\bar{w}_{NLMS} = \left(I - \frac{\hat{\mu}}{M(\sigma_n^2 + \sigma_i^2)}R^{-1}Q\right)^{-1} w_w$$

(3-92)

following the LMS expression in (3-64). The expression for the individual weights can also be determined, by substituting (3-90) into (3-82):

$$\bar{w}_m^{(NLMS)} = \begin{cases} \frac{\sigma_i^2}{\lambda_{max} L} e^{j\omega_i \Delta}, & m = 0 \\ \frac{\sigma_i^2}{\sigma_n^2 L} \left[1 - \frac{\sigma_i^2 M}{\lambda_{max}} \left(1 - \frac{\hat{\mu}\sigma_n^2}{M(\sigma_n^2 + \sigma_i^2)}\right)^m\right] e^{j\omega_i(\Delta - m)}, & 1 < m < M \end{cases}$$

(3-93)
3.7. Numerical Illustrations

This section illustrates the weight behavior of the LMS algorithm using a fixed structure with $M = 7$ and $\Delta = 3$. All results are shown for the transformed weights as defined in (3-1). Also, the complex sinusoidal process with fixed $\omega_i = 0.2$ is used in this section. The signal-to-noise ratio (SNR), $\sigma_w^2 / \sigma_n^2$, is fixed at 25 dB throughout the section as the SNR does not contribute to the weight behavior. The behavior of the weights is illustrated as a function of the remaining two parameters: the step-size $\mu$ and the interference-to-signal ratio (ISR) $\sigma_i^2 / \sigma_w^2$.

Each example in this section is accompanied by simulations to illustrate the agreement between the analytical solution and the corresponding experimental results. All the simulations in this section employ, as $x_n$, an eight-point quadrature amplitude modulated (8-QAM) signal with minimum distance between its signal points [1 p.277]. In each simulation, every 8-QAM symbol is drawn randomly from its 8 possible signal points with equal probabilities. The power of the 8-QAM signal is fixed at $\sigma_x^2 = 4.732$ in all examples. The noise $\hat{n}_n$ in the simulation is modeled as a circularly symmetric complex Gaussian process. Finally, the mean of the weights is estimated from the simulated weights by averaging 10,000 samples in steady state.

First, the derived analytical mean of the LMS weights is compared against the simulated instantaneous weights on the complex plane, as shown in Figure 3-2 together with the corresponding Wiener weights. The ISR is fixed in this example to 20 dB. The step-size is set to $\mu = \lambda_{\max}^{-1} = 3.015 \times 10^{-4}$. The analysis and simulation results are in good agreement, forming a spiral, and are very distinct from the Wiener solution. Also, both the mean of the LMS weights and the Wiener weights lie on the radial lines through $e^{j\omega\Delta}$ as expected from the analysis. For the remainder of the section, we concentrate on the magnitude of the weights. Note that the Wiener solution has equal amplitude for all weights.
Figure 3-2. LMS steady-state weight behavior (10,000 samples), analytical mean, and corresponding Wiener weights on the complex plane.

Figure 3-3 illustrates the behavior of $|\bar{w}|$ as a function of $\mu$ as the step-size is varied from 0 to $2\lambda_{\text{max}}^{-1}$ while the ISR is maintained at 20 dB. Figure 3-3a captures the large step-size behavior, as the step-size is shown on a linear scale. Excellent agreement between the theory and simulation is observed until LMS algorithm divergence in simulation, which occurs at $\mu \approx 1.93\lambda_{\text{max}}^{-1}$. Although wider weight fluctuation is observed as the step-size approaches the point of divergence, the fluctuation occurs around the expected weight values. Furthermore, Figure 3-3b emphasizes the small step-size behavior for the same configuration. The departure of the mean weights from the Wiener solution (which aligns with $w_3$) becomes prevalent around $\mu = 0.03\lambda_{\text{max}}^{-1}$. Also, all the neighboring weight pairs are roughly equal distance apart for all $\mu$.
Figure 3-3. Magnitudes of the mean LMS weights as a function of step-size (a) on a linear scale, with the Wiener solution, and (b) on a log scale.

Figure 3-4 shows the behavior of the mean of the weights as a function of ISR. The step-size $\mu$ in this example is determined as a function of ISR as follows:

$$\mu = \frac{1}{M \left( \sigma_x^2 + \sigma_{\hat{n}}^2 + \sigma_t^2 \right)} \quad (3-94)$$

As the ISR is reduced, the LMS algorithm diverges at a much smaller step-size than the step-size used in the first example, $\lambda^{-1}_{\text{max}} = (\sigma_x^2 + \sigma_{\hat{n}}^2 + M \sigma_t^2)^{-1}$. Hence, (3-94) is used to set the step-size more conservatively.
Figure 3-4. Magnitude of mean LMS weights as a function of ISR. Step-size is varied as a function of ISR.

The result in Figure 3-4 clearly illustrates that the non-Wiener behavior of the mean of the weight vector is caused by the narrowband interference. When the interference is weak, the mean of the LMS weights behaves as expected, that is, they follow the Wiener weights. As the interference becomes stronger, the mean of the LMS weights begins to move away from the Wiener weights to the spiral formation that is prevalently illustrated in Figure 3-2. The theoretical model deviates from the simulation results over the range of ISR from −15 dB to 5 dB due to the theory depending on the large ISR assumption. However, even over the latter range, the mean of the weights in simulation is much closer to the high-ISR predicted mean of the weight vector than it is to the Wiener weight vector, i.e., the prediction is relatively robust to violating the high-ISR assumption.

Lastly, Figure 3-5 shows the magnitude behavior of the mean of the NLMS weights as a function of ISR. The NLMS step-size is set to \( \tilde{\mu} = 1 \), which puts the NLMS adaptation at a comparable step-size to that of the LMS algorithm as defined in (3-94) to generate Figure 3-4. The discrepancy between theory and simulation is greater than that observed in Figure 3-4, due to the additional assumption in (3-91). The NLMS algorithm keeps the means of its weights closer to the Wiener solution longer than the LMS
algorithm (and theory) for smaller ISR. On the other hand, when the non-Wiener effect kicks in, the NLMS weights reach the maximal non-Wiener weight spread at lower ISR than the LMS weights.

![Figure 3-5](image)

**Figure 3-5.** Magnitude of mean NLMS weights \((\bar{\rho} = 1)\) as a function of ISR.

### 3.8. Summary

This chapter has presented the analysis of the steady-state mean of the LMS equalizer weights when the equalizer is under the influence of narrowband interference. The analysis utilized the Butterweck expansion of the LMS weights, and an expression for the mean of the weights was derived. The expression derived here for the mean of the weights was shown to agree with simulation results over a wide range of ISR, even though the derivation relied on the large-ISR assumption. In the coming chapter we will use this expression for the mean of the weight vector to assess the mean squared error performance of the equalizer.
CHAPTER 4  ANALYSIS OF MEAN SQUARED ERROR BEHAVIOR

This chapter contains discussions of the mean squared error (MSE) performance of the LMS adaptive equalizer when it is subject to narrowband interference. Specifically, a new MSE estimate is derived based on the so-called transfer function approximation of the LMS algorithm [15] using the expression for the mean of the weights derived in Chapter 3.

The method of computing the MSE using the transfer function approximation itself is not new as it was originally proposed by Reuter and Zeidler [14]. However, we provide a more rigorous and explicit derivation, as well as the reasoning for appropriateness of transfer-function based MSE estimation (Section 4.1). Also, a new MSE expression is proposed for the equalization problem (Section 4.2) complemented with a comparative discussion of the new estimator relative to the Reuter-Zeidler MSE estimator and the Wiener MSE (Section 4.3). Lastly, the new MSE estimator is extended to apply to the NLMS equalizer following [14] (Section 4.4).

Numerical examples are provided in Section 4.5. The new estimator is shown to closely follow the simulation results under the assumed conditions and to consistently outperform the original MSE estimator by Reuter and Zeidler. In addition, breakdown of both MSE estimators is demonstrated to occur when the interference is weak (but still noticeable) compared to the signal.

4.1. Estimation of Mean Square Error Based on Transfer Function Approximation of the LMS Algorithm

This section presents the general steady-state MSE analysis of the LMS algorithm
with narrowband inputs, using the transfer-function approximation of the algorithm. The first half of the section presents the generalized transfer function approximation framework that renders steady-state analysis possible. The latter half of the section constructs the MSE estimate from the transfer function formulation in the spectral domain. It is important to remember that this method of estimating the MSE only provides valid estimates for an input signal with a strong narrowband component and for LMS step-sizes and scenarios in general that produce small steady-state weight fluctuations.

The idea behind the transfer function approximation is to exploit the implicit relationship between the desired signal $d_n$ and the error signal $e_n$ with proper weight initialization. The approach can be applied to steady-state analysis by removing the dependence of the weight evolution on its initial condition and assessing the relationship between two sample instances in steady state.

For some arbitrary time instances $m$ and $n > m$, the relationship between the weight vectors $\mathbf{w}_m$ and $\mathbf{w}_n$ is found by recursively substituting (2-25) into itself, resulting in

$$\mathbf{w}_n = \mathbf{w}_m + \mu \sum_{k=1}^{n-m} \mathbf{u}_{n-k} e^*_n e_{n-k} \tag{4-1}$$

In previous presentations of the transfer function approximation [14, 15], $m$ is set to zero to describe the system with respect to the initial LMS weight vector $\mathbf{w}_0$. However, for the analysis of the steady-state MSE of the LMS filter, setting $m = 0$ has little relevance. The analysis needs to be carried out strictly in steady state so that the LMS error signal $e_n$ can be assumed to be stationary.

To analyze the steady-state behavior of the LMS algorithm, the reference point $m$ is set to a time sample index occurring after the LMS algorithm has reached its steady state. Consequently, (4-1) represents the relationship of the weight vectors in steady state. Here, we are assuming that the steady-state LMS weights, as random processes, are wide sense stationary, inheriting that property from the input process. Once the reference point $m$ is fixed, the weight vector $\mathbf{w}_m$ is then modeled as a random vector
Chapter 4: Analysis of Mean Squared Error Behavior

that is fixed in each realization but changes across realizations.

Substituting (4-1) into the error signal, defined in (2-24), yields

\[ e_n = (\hat{d}_n - \hat{w}_n^H u_n) - \mu \sum_{k=1}^{n-m} e_{n-k} u_n^H u_n \]  

This equation can be interpreted as a time-varying difference equation as it can be reformulated (after additional algebraic manipulation) as

\[ \sum_{k=0}^{n-m} g_{k,n} e_{n-k} = e_{m,n} \]  

where

\[ g_{k,n} = \begin{cases} 
1, & k = 0 \\
\mu u_n^H u_n, & k > 0 
\end{cases} \]  

(4-4)

and

\[ e_{m,n} \triangleq d_n - \hat{w}_n^H u_n. \]  

(4-5)

Here, the signal \( g_{k,n} \) acts as the feedback coefficients, and the error \( e_{m,n} \) is due to the filter with fixed weight vector \( \hat{w}_n \). Difference equation (4-3) represents a time-varying infinite-order recursive linear system with \( e_{m,n} \) as input and the LMS error signal \( e_n \) as output. Clarkson and White set \( \hat{w}_m \big|_{m=0} \) as input and defined the transfer function from the desired signal \( d_n \) to the LMS error signal \( e_n \) [15]. Reuter and Zeidler, on the other hand, initialized \( \hat{w}_m \big|_{m=0} \) to the corresponding Wiener filter weight vector \( \hat{w}_w \) so that the transfer function was intended to describe the relationship between the Wiener filter error signal \( (d_n - \hat{w}_w^H u_n) \) and the LMS error signal \( e_n \) [14].

Using the ergodicity assumption on the input process, Clarkson and White asserted that the autocorrelation function \( r_{uu,l} \) can be used as a good time-invariant approximation for \( g_{k,n}, k > 0 \) [15]:
Although the validity of this approximation for a general family of stochastic signals is challenged by Butterweck [29], it is a valid approximation when the input signal is predominantly a deterministic stochastic process, i.e. when it consists of a strong narrowband process or a sum of such processes.

The approximation (4-6) turns the difference equation (4-3) into its constant coefficient variant:

$$g_{k,n} = \mu u_{n-k-H}^H u_n \approx \mu M_{uu,k} \triangleq g_k$$  (4-6)

$$\sum_{k=0}^{n-m} g_k e_{n-k} \approx e_{m,n}$$  (4-7)

The system represented by (4-7) is both linear and time-invariant, describing the transformation from \(e_{m,n}\) to \(e_n\). As our analysis is for the steady-state behavior, we assume that the system has been operating in steady state ever since negative infinity. Then, it follows that we can modify (4-7) to:

$$\sum_{k=0}^{\infty} g_k e_{n-k} \approx e_{m,n}$$  (4-8)

Given wide sense stationary \(u_n\) and \(d_n\) and assuming that the system is in steady state at time \(m\), both \(e_n\) and \(e_{m,n}\) are wide sense stationary and their power spectra exist. Moreover, the power spectrum of \(e_n\), \(S_{ee}(\omega)\), and that of \(e_{m,n}\), \(S_{e_m e_m}(\omega)\), can be related by

$$S_{ee}(\omega) \approx \frac{1}{|G(\omega)|^2} S_{e_m e_m}(\omega)$$  (4-9)

where

$$G(\omega) = \sum_{k=0}^{\infty} g_k e^{-j\omega k}$$  (4-10)

Substituting the definition of \(g_k\) (4-6) for \(k > 0\) and \(g_0 = 1\) from (4-4) into (4-10)
yields

$$G(\omega) = 1 + \mu \sum_{k=1}^{\infty} r_{uu,k} e^{-j\omega k}. \quad (4-11)$$

Consequently, the MSE of the LMS algorithm is found from

$$J \equiv E[|e_n|^2] = E[e_n^2] = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{ee}(\omega) e^{j\omega l} d\omega \bigg|_{l=0}$$

$$\approx \frac{1}{2\pi} \int_{-\pi}^{\pi} |G(\omega)|^2 d\omega \quad (4-12)$$

To determine the power spectrum $S_{ee}^{m,n}(\omega)$, the autocorrelation function of $e_{m,n}$ needs to be determined first. From (4-5), we find

$$E[e_{m,n}^* e_{m,n-l}^*] = E[(d_n - \tilde{\mathbf{w}}_m^H \mathbf{u}_n)(d_{n-l} - \tilde{\mathbf{w}}_m^H \mathbf{u}_{n-l})^*]$$

$$= E[d_n d_{n-l}^*] - E[d_n^* \mathbf{u}_n^H \tilde{\mathbf{w}}_m] - E[\tilde{\mathbf{w}}_m^H \mathbf{u}_n d_{n-l}^*] + E[\tilde{\mathbf{w}}_m^H \mathbf{u}_n^H \tilde{\mathbf{w}}_m] \quad (4-13)$$

The expression for $S_{ee}^{m,n}(\omega)$ is found by taking the discrete-time Fourier transform of (4-13). We denote the auto- and cross-spectra involving signals $d_n$ and $u_n$ as follows.

$$S_{dd}(\omega) \triangleq \sum_{l=-\infty}^{\infty} E[d_n d_{n-l}^*] e^{-j\omega l} \quad (4-14)$$

$$S_{du}(\omega) \triangleq \sum_{l=-\infty}^{\infty} E[d_n^* u_{n-l}] e^{-j\omega l} \quad (4-15)$$

$$S_{ud}(\omega) \triangleq S_{du}^*(\omega) \quad (4-16)$$

$$S_{uu}(\omega) \triangleq \sum_{l=-\infty}^{\infty} E[u_n u_{n-l}^*] e^{-j\omega l} \quad (4-17)$$
Chapter 4: Analysis of Mean Squared Error Behavior

To analyze the terms with $\tilde{w}_m$ in (4-13), we assume that the steady-state time instances $m$ and $n$ are sufficiently separated, i.e., $n \gg m$, so that $d_n$ and $u_n$ are uncorrelated to $\tilde{w}_m$ (except for periodic components if there are any). Using this assumption, the $E[d_n u_n^H \tilde{w}_m]$ term in (4-13) expands to the scalar-summation format as follows

$$E[d_n u_n^H \tilde{w}_m] = E \left[ d_n \sum_{k=0}^{M-1} u_{n-l-k}^* \tilde{w}_{m,k} \right] = \sum_{k=0}^{M-1} E[d_n u_{n-l-k}^*] E[\tilde{w}_{m,k}]$$  (4-18)

where $E[\tilde{w}_{m,k}]$ is the ensemble-average of the $(k+1)$-st element of $\tilde{w}_m$. With the final expression in (4-18), we can use the standard spectral input-output relationship of linear systems [35 p.332] to find the representation of this term in the spectral domain:

$$\sum_{l=-\infty}^{\infty} E[d_n u_n^H \tilde{w}_m] e^{-j\omega l} = \tilde{W}(\omega) S_{du}(\omega)$$  (4-19)

where

$$\tilde{W}(\omega) \triangleq \sum_{k=0}^{M-1} E[\tilde{w}_{k,m}] e^{-j\omega k}.$$  (4-20)

Since we assumed that the LMS algorithm is in steady-state at time $m$, (4-20) is the frequency response of the fixed filter with as coefficients the mean of the weights in steady-state.

The spectrum of the $E[\tilde{w}_m^H u_n^* d_{n-l}^*]$ term of (4-13), is simply the conjugate of (4-19). Using (4-16), we have

$$\sum_{l=-\infty}^{\infty} E[\tilde{w}_m^H u_n^* d_{n-l}^*] e^{-j\omega l} = \tilde{W}^*(\omega) S_{ud}(\omega)$$  (4-21)

Lastly, again using the uncorrelatedness assumption between the weights $\tilde{w}_{m,k}$ and the signal $u_n$, the last term $E[\tilde{w}_m^H u_n^H \tilde{w}_m]$ of (4-13) can be notated in summation form as


\[
E \left[ \tilde{w}_m^H u_n^H \tilde{w}_{m-n} \right] = E \left[ \sum_{p=0}^{M-1} \sum_{q=0}^{M-1} \tilde{w}_{m,p}^* u_{n-p}^* u_{n-l-q}^* \tilde{w}_{m,q} \right]
= \sum_{p=0}^{M-1} \sum_{q=0}^{M-1} E \left[ \tilde{w}_{m,p}^* \tilde{w}_{m,q}^* \right] E \left[ u_{n-p}^* u_{n-l-q} \right]
= \sum_{p=0}^{M-1} \sum_{q=0}^{M-1} E \left[ \tilde{w}_{m,q}^* \tilde{w}_{m,p}^* \right] r_{uu,l-p+q}
\]

(4-22)

Accordingly, the spectrum of this term is computed from (4-22) as follows.

\[
\sum_{l=-\infty}^{\infty} E \left[ \tilde{w}_m^H u_n^H u_{n-l} \tilde{w}_m \right] e^{-j\omega l} = \sum_{l=-\infty}^{\infty} \sum_{p=0}^{M-1} \sum_{q=0}^{M-1} E \left[ \tilde{w}_{m,q}^* \tilde{w}_{m,p}^* \right] r_{uu,l-p+q} e^{-j\omega l}
= \sum_{p=0}^{M-1} \sum_{q=0}^{M-1} E \left[ \tilde{w}_{m,q}^* \tilde{w}_{m,p}^* \right] \sum_{n=-\infty}^{\infty} r_{uu,l-p+q} e^{-j\omega l}
= S_{uu} (\omega) E \left[ \sum_{p=0}^{M-1} \tilde{w}_{m,p} e^{j\omega p} \sum_{q=0}^{M-1} \tilde{w}_{m,q} e^{-j\omega q} \right]
\triangleq S_{uu} (\omega) E \left[ \tilde{W}_m (\omega) \right]^2
\]

(4-23)

The expected value \( E[\tilde{w}_{m,q}^* \tilde{w}_{m,p}^*] \) is the second moment of two weights, and by the definition of covariance,

\[
E \left[ \tilde{w}_{m,q}^* \tilde{w}_{m,p}^* \right] = Cov \left( \tilde{w}_{m,q}^* , \tilde{w}_{m,p}^* \right) + E \left[ \tilde{w}_{m,q}^* \right] E \left[ \tilde{w}_{m,p}^* \right]
\]

(4-24)

where \( Cov \left( X,Y \right) \) is the covariance of random variables \( X \) and \( Y \). Consequently, the covariance of the steady-state weights influences \( E[|\tilde{W}_m (\omega)|^2] \) in (4-23). In other words, the amount of adapted weight fluctuation in steady state changes \( E[|\tilde{W}_m (\omega)|^2] \).

Assuming that \( Cov \left( \tilde{w}_{m,q}, \tilde{w}_{m,p} \right) \approx 0 \), i.e., the LMS adaptation results in variations of the steady-state weights that are small with respect to the magnitudes of their mean, we approximate

\[
E \left[ |\tilde{W}_m (\omega)|^2 \right] \approx |\tilde{W}_m (\omega)|^2
\]

(4-25)

Substituting (4-25) into (4-23) yields
Accordingly, taking the discrete-time Fourier transform of (4-13) and substituting (4-14), (4-19), (4-21), and (4-26) into the transformed expression, we find

\[
S_{e_m e_m}(\omega) \approx S_{dd}(\omega) - \hat{W}(\omega) S_{d u}^*(\omega) - \hat{W}^*(\omega) S_{d u}(\omega) + |\hat{W}(\omega)|^2 S_{u u}(\omega) \tag{4-27}
\]

This spectral-domain approximation implies that the resulting LMS MSE estimate is not reliable if the weight fluctuation due to adaptation is significant, commonly occurring when using a large step-size. This includes narrowband applications of the LMS algorithm, in which large step-size facilitates variation of the weights by design, such as shown in the narrowband adaptive noise cancellation application [17]. On the other hand, it appears to be a suitable method to estimate MSE for the equalizer scenario, as we observed in Figure 2-7 that the adapted weights form rather tight clusters around their mean under strong narrowband interference.

To summarize, the MSE of the LMS algorithm can be evaluated from (4-12), together with (4-9) and (4-27). The key approximations in (4-6) and (4-25) must be valid for this MSE estimator to produce good estimates. Lastly, this derivation, namely (4-20), explicitly indicates that the steady-state mean of the LMS weights is needed in order to assess the MSE of the LMS filter.

4.2. Estimated LMS MSE for the Equalization Problem

In this section, we apply the MSE formula (4-12) to the simplified equalization problem (Figure 2-2) using our expression for the mean of the LMS weights from Chapter 3. We start off establishing the error spectrum \(S_{e_m e_m}(\omega)\) (4-27) for the equalization problem. According to the signal definitions in Section 2.2, spectra involving the input and desired signals are found as follows:

\[
S_{dd}(\omega) = \sigma_x^2 \tag{4-28}
\]

\[
S_{d u}(\omega) = \sigma_x^2 e^{-j\omega \Delta} \tag{4-29}
\]
\[ S_{uu}(\omega) = (\sigma_x^2 + \sigma_n^2) + \sigma_i^2 \delta(\omega - \omega_i) \]  

Substituting (4-28) through (4-30) into (4-27), we get

\[ S_{e_me_m}(\omega) \approx \sigma_x^2 - \sigma_x^2 e^{j\omega \Delta} \bar{W}(\omega) - \sigma_x^2 e^{-j\omega \Delta} \bar{W}^*(\omega) + |\bar{W}(\omega)|^2 \left( \sigma_x^2 + \sigma_n^2 + \sigma_i^2 \delta(\omega - \omega_i) \right) \]

\[ = \sigma_x^2 \left( 1 - e^{j\omega \Delta} \bar{W}(\omega) - e^{-j\omega \Delta} \bar{W}^*(\omega) + |\bar{W}(\omega)|^2 \right) + \left[ \sigma_n^2 + \sigma_i^2 \delta(\omega - \omega_i) \right] |\bar{W}(\omega)|^2 \]

In other words, \( e_{m,n} \) can be generated by the two-channel structure depicted in Figure 4-1, which is consistent with the equalization problem in Figure 2-2.

![Figure 4-1](image-url)

**Figure 4-1.** Two-channel generation of the fixed filter error signal.

To obtain the magnitude responses of the two filters in Figure 4-1, we first evaluate the frequency response \( \bar{W}(\omega) \) by first taking the expected value of (3-1):

\[ E[\bar{w}] = \eta(p_\Delta - \bar{w}) \]  

This introduces the expression \( \bar{w} \) for the mean of the weight vector in the transformed problem of Chapter 3. Then, representing the magnitude-phase decomposition of \( \bar{w}_n \) — which is defined in (3-82) — as \( |\bar{w}_n| e^{j\omega_i(n-\Delta)} \), the unit-pulse response of the mean-weight fixed equalizer is denoted by

\[ E[\bar{w}_n] = \eta \left( \delta_{n-\Delta} - |\bar{w}_n| e^{-j\omega_i(n-\Delta)} \right) \]  

(4-33)

Here, \( |\bar{w}_n| = 0 \) for \( n < 0 \) and \( n \geq M \). The discrete-time Fourier transform of (4-33) evaluates to
\[ \hat{W}(\omega) = \eta \left[ e^{-j\omega\Delta} - \sum_{m=0}^{M-1} |\bar{w}_m| e^{-j(\hat{\omega}m + \omega\Delta)} \right]. \] (4-34)

where

\[ \hat{\omega} \triangleq \omega - \omega_i. \] (4-35)

Now, using (4-34), we can evaluate the two magnitude responses in (4-31) as follows.

\[
\left| \hat{W}(\omega) \right|^2 = \left| \eta \left[ e^{-j\omega\Delta} - \sum_{n=0}^{M-1} |\bar{w}_n| e^{-j(\hat{\omega}n + \omega\Delta)} \right] \right|^2
= \eta^2 - 2\eta^2 \sum_{n=0}^{M-1} |\bar{w}_n| \cos \{ \hat{\omega}(n - \Delta) \} + \eta^2 \sum_{n=0}^{M-1} |\bar{w}_n|^2
+ 2\eta^2 \sum_{n=1}^{M-1} \sum_{m=0}^{n-1} |\bar{w}_n||\bar{w}_m| \cos \{ \hat{\omega}(n - m) \} \] (4-36)

\[
\left| e^{-j\omega\Delta} - \hat{W}(\omega) \right|^2 = \left| (1 - \eta) e^{-j\omega\Delta} + \eta \sum_{n=0}^{M-1} |\bar{w}_n| e^{-j(\hat{\omega}n + \omega\Delta)} \right|^2
= (1 - \eta)^2 + 2\eta (1 - \eta) \sum_{n=0}^{M-1} \sum_{n=1}^{M-1} |\bar{w}_n||\bar{w}_m| \cos \{ \hat{\omega}(n - m) \}
+ 2\eta^2 \sum_{n=1}^{M-1} \sum_{m=0}^{n-1} |\bar{w}_n||\bar{w}_m| \cos \{ \hat{\omega}(n - m) \} \] (4-37)

The expressions involving \( \eta \) in (4-38) are subject to simplification. Using the definition
in (2-30), the following identities hold:

\[(1 - \eta)^2 \sigma_x^2 + \eta^2 \sigma_n^2 = \eta \sigma_n^2 \]  
\[\eta (1 - \eta) \sigma_x^2 - \eta^2 \sigma_n^2 = 0 \]  
\[\eta^2 (\sigma_x^2 + \sigma_n^2) = \eta \sigma_x^2 \]

Using (4-39) through (4-41), (4-38) simplifies to

\[S_{\text{err}}(\omega) = \eta \sigma_n^2 + \eta^2 \sigma_i^2 \delta(\omega - \omega_i) - 2[\eta^2 \sigma_i^2 \delta(\omega - \omega_i)] \sum_{n=0}^{M-1} |\bar{w}_n| \cos \{\tilde{\omega}(n - \Delta)\} \]
\[+ \left[ \eta \sigma_x^2 + \eta^2 \sigma_i^2 \delta(\omega - \omega_i) \right] \left( \sum_{n=0}^{M-1} |\bar{w}_n|^2 + 2 \sum_{n=1}^{M-1} \sum_{m=0}^{n-1} |\bar{w}_n||\bar{w}_m| \cos \{\tilde{\omega}(n - m)\} \right) \]  
\[(4-42)\]

Next, we evaluate \(G(\omega)\) for this problem; substituting the input autocorrelation function in (2-6) into (4-7) produces

\[G(\omega) = 1 + \mu M \sum_{k=1}^{\infty} \left[ (\sigma_x^2 + \sigma_n^2) \delta_k + \sigma_i^2 e^{j\omega k} \right] e^{-j\omega k} \]
\[= 1 + \mu \sigma_x^2 M \sum_{k=1}^{\infty} e^{-j(\omega - \omega_i)k} \]
\[= 1 + \mu \sigma_x^2 M \frac{e^{-j(\omega - \omega_i)}}{1 - e^{-j(\omega - \omega_i)}} \]
\[= 1 - \frac{(1 - \mu \sigma_x^2 M) e^{-j(\omega - \omega_i)}}{1 - e^{-j(\omega - \omega_i)}} \]  
\[(4-43)\]

To make the notation more compact, we use (4-35) and introduce

\[\alpha \triangleq 1 - \mu \sigma_n^2 M \]  
\[(4-44)\]

Subsequently, (4-43) becomes

\[G(\omega) = \frac{1 - \alpha e^{-j\tilde{\omega}}}{1 - e^{-j\omega}} \]  
\[(4-45)\]
From (4-45), we further compute

\[ |G(\omega)|^2 = \frac{1 + \alpha^2 - 2\alpha \cos \tilde{\omega}}{2(1 - \cos \tilde{\omega})}. \]  

This system has a pole on the unit circle at \( e^{i\omega_0} \); therefore, the inverse of this system, as used in (4-9), has a zero at \( e^{i\omega_0} \) and completely rejects \( i_n \). Also, we observe that (4-46) tends to one as the step-size or the interference power tends to zero, indicating the loss of the non-Wiener effects. Interestingly, what is not modeled, however, is the loss of the non-Wiener effects when the signal power increases, i.e., the effect as a function of ISR.

The MSE estimate can be obtained from substituting (4-42) and (4-46) into (4-12) (with all the terms in (4-42) involving \( i_n \) dropping out) and changing integration variable from \( \omega \) to \( \tilde{\omega} \):

\[
J \approx \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[ \eta \sigma_n^2 + \sqrt{\sigma_n^2 + \sigma_x^2} \left| \sum_{n=0}^{M-1} |\tilde{w}_n|^2 + \sum_{n=1}^{M-1} \sum_{m=0}^{n-1} |\tilde{w}_n||\tilde{w}_m| \cos \{ \tilde{\omega}(n - m) \} \right] \right] d\tilde{\omega} \\
= \eta \left[ \sigma_n^2 + \sigma_x^2 \sum_{n=0}^{M-1} |\tilde{w}_n|^2 \right] \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{2(1 - \cos \tilde{\omega})}{1 + \alpha^2 - 2\alpha \cos \tilde{\omega}} d\omega \\
\quad + 2\eta \sigma_x^2 \sum_{n=1}^{M-1} \sum_{m=0}^{n-1} |\tilde{w}_n||\tilde{w}_m| \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{2(1 - \cos \tilde{\omega}) \cos \{ \tilde{\omega}(n - m) \}}{1 + \alpha^2 - 2\alpha \cos \tilde{\omega}} d\tilde{\omega} \]  

The two integrals in (4-47) may be solved by using the following standard integrals [36 p.385-386]:

\[
\int_0^\pi \frac{\cos k\omega}{1 + \alpha^2 - 2\alpha \cos \omega} d\omega = \frac{\pi \alpha^k}{1 - \alpha^2}, \quad \alpha < 1, k \geq 0 \]  

and

\[
\int_0^\pi \frac{\cos \omega \cos k\omega}{1 - 2\alpha \cos \omega + \alpha^2} d\omega = \frac{\pi}{2} \frac{1 + \alpha^2}{1 - \alpha^2} \alpha^{k-1}, \quad \alpha < 1, k \geq 0 \]  

Note that these integrands are both even functions. Using (4-48) twice with \( k = 0 \) and 1,
the first integral in (4-47) evaluates to

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1 - \cos \hat{\omega}}{1 + \alpha^2 - 2\alpha \cos \hat{\omega}} d\hat{\omega} = \frac{1}{\pi} \left( \frac{\pi}{1 - \alpha^2} - \frac{\pi\alpha}{1 - \alpha^2} \right)$$

$$= \frac{1}{1 + \alpha}$$

(4-50)

Likewise, using (4-48) and (4-49), both with \( k = n - m \) (a strictly positive quantity by (4-47)), the second integral in (4-47) evaluates to

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{2(1 - \cos \hat{\omega}) \cos \{ \hat{\omega} (n - m) \}}{1 + \alpha^2 - 2\alpha \cos \hat{\omega}} d\hat{\omega} = \frac{2}{\pi} \left( \frac{\pi\alpha^{n-m}}{1 - \alpha^2} - \frac{\pi}{1 - \alpha^2} \frac{\alpha^{n-m}}{2} \frac{1 + \alpha^2}{\alpha^{n-m-1}} \right)$$

$$= -\frac{1 - \alpha}{1 + \alpha} \alpha^{n-m-1}$$

(4-51)

Finally, substituting these two results into (4-47) yields the MSE estimate

$$J \approx \eta \left[ \sigma_n^2 + \sigma_x^2 \sum_{n=0}^{M-1} |\bar{w}_n|^2 \right] \frac{2}{1 + \alpha} - 2\eta \sigma_x^2 \sum_{n=1}^{M-1} \sum_{m=0}^{n-1} |\bar{w}_n| \| \bar{w}_m \| \alpha^{n-m-1} \frac{1 - \alpha}{1 + \alpha}$$

$$= \frac{2\eta}{1 + \alpha} \left[ \sigma_n^2 + \sigma_x^2 \left( \sum_{n=0}^{M-1} |\bar{w}_n|^2 - (1 - \alpha) \sum_{n=1}^{M-1} \sum_{m=0}^{n-1} |\bar{w}_n| \| \bar{w}_m \| \alpha^{n-m-1} \right) \right]$$

(4-52)

This solution may be expressed in vector-matrix form as follows

$$J \approx \frac{2\eta}{1 + \alpha} \left[ \sigma_n^2 + \sigma_x^2 \left\{ \| \bar{w} \|^2 - (1 - \alpha) \| \bar{w} \|^T A \| \bar{w} \| \right\} \right]$$

(4-53)

where \( \| \cdot \| \) computes the Euclidian norm of the vector, and

$$A \triangleq \begin{bmatrix} 0 & \cdots & 0 \\ 1 & 0 \\ \alpha & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \alpha^{M-2} & \cdots & \alpha & 1 & 0 \end{bmatrix}$$

(4-54)
4.3. Observations on the Analytical MSE Expression

We now compare our result to the expressions for the Wiener MSE and for the Reuter-Zeidler MSE estimate [14]. The Wiener filter for the equalization problem is defined previously in (2-29), and its MSE in (2-32). The three MSE expressions are restated here for convenience:

Ikuma—

\[ J \approx \frac{2\eta}{1 + \alpha} \left[ \sigma_n^2 + \sigma_x^2 \left\{ ||\bar{w}||^2 - (1 - \alpha) |\bar{w}|^T A |\bar{w}| \right\} \right] \]  

(4-53)

Wiener—

\[ J_W = \eta \left( \sigma_n^2 + \sigma_x^2 \frac{\sigma_i^2}{\lambda_{\text{max}}} \right) \]  

(4-55)

Reuter-Zeidler—

\[ J_{RZ} = \frac{2\eta}{1 + \alpha} \left[ \sigma_n^2 + \sigma_x^2 \left( \frac{\sigma_i^2}{\lambda_{\text{max}}} \right)^2 \frac{1 - \alpha^M}{1 - \alpha} \right] \]  

(4-56)

The expressions for Wiener and Reuter-Zeidler are manipulated from their original forms—(2-32) and [14 (41)], respectively—so that they are in a form comparable to the one in (4-53). These three MSE expressions appear to be very similar, and this section includes two comparative remarks on them.

First, we can readily identify the dissimilar portions of (4-53) and (4-56):

\[ ||\bar{w}||^2 - (1 - \alpha) |\bar{w}|^T A |\bar{w}| \sim \left( \frac{\sigma_i^2}{\lambda_{\text{max}}} \right)^2 \frac{1 - \alpha^M}{1 - \alpha} \]  

(4-57)

It can be shown that the two sides of (4-57) will actually be equal if we replace \( \bar{w} \) with \( w_0 \); in other words, (4-56) can be derived from (4-53) by setting the mean weight to be the Wiener weights. From (2-29), we can recognize that the constant magnitude vector of the Wiener filter is

\[ |\bar{w}_w| = \frac{\sigma_i^2}{\lambda_{\text{max}}} \]  

(4-58)

Substituting (4-58) into the left hand side of (4-57) as |\bar{w}| yields
\[
|\bar{w}_w|^2 - (1 - \alpha)|\bar{w}_w|^T A|\bar{w}_w| = \left( \frac{\sigma^2_i}{\lambda_{\text{max}}} \right)^2 \left[ M - (1 - \alpha)1^T A1 \right]
\]  

(4-59)

where \(1 \in \mathbb{R}^M\) is a vector of all ones. Referring back to (4-52) for the scalar formulation of the \(1^T A 1\) term, we get

\[
|\bar{w}_w|^2 - (1 - \alpha)|\bar{w}_w|^T A|\bar{w}_w| \\
= \sum_{n=0}^{M-1} |\bar{w}_{w,n}|^2 - (1 - \alpha) \sum_{n=1}^{M-1} \sum_{m=0}^{n-1} |\bar{w}_{w,n}| |\bar{w}_{w,m}| \alpha^{n-m-1} \\
= \left( \frac{\sigma^2_i}{\lambda_{\text{max}}} \right)^2 M - (1 - \alpha) \left( \frac{\sigma^2_i}{\lambda_{\text{max}}} \right)^2 \sum_{n=1}^{M-1} \sum_{m=0}^{n-1} \alpha^{n-m-1} \\
= \left( \frac{\sigma^2_i}{\lambda_{\text{max}}} \right)^2 \left[ M - (1 - \alpha) \sum_{n=1}^{M-1} \sum_{m=0}^{n-1} \alpha^{n-m-1} \right] 
\]  

(4-60)

Applying the solution to the geometric series twice to the double summation, we can further simplify (4-60).

\[
|\bar{w}|^2 - (1 - \alpha)|\bar{w}|^T A|\bar{w}| = \left( \frac{\sigma^2_i}{\lambda_{\text{max}}} \right)^2 \left[ M - (1 - \alpha) \sum_{n=1}^{M-1} \alpha^{n-1} \frac{1 - \alpha^{-n}}{1 - \alpha^{-1}} \right] \\
= \left( \frac{\sigma^2_i}{\lambda_{\text{max}}} \right)^2 \left[ M + \sum_{n=1}^{M-1} (\alpha^n - 1) \right] \\
= \left( \frac{\sigma^2_i}{\lambda_{\text{max}}} \right)^2 \left[ M + \left( \frac{1 - \alpha^M}{1 - \alpha} - 1 \right) - (M - 1) \right] \\
= \left( \frac{\sigma^2_i}{\lambda_{\text{max}}} \right)^2 \frac{1 - \alpha^M}{1 - \alpha}
\]  

(4-61)

The final expression is the right hand side of (4-57); thus, (4-53) with the Wiener weights equals (4-56).

Secondly, as \(\mu \to 0\) (or equivalently \(\alpha \to 1\)), we expect the LMS algorithm to approach the Wiener filter in terms of the mean of the weights and in terms of MSE performance. In Section 3.4, the convergence of \(\bar{w}\) to \(w_w\) has been shown; moreover, substituting \(w_w\) into (4-32) yields \(\bar{w}_w\) in (2-29). Since (4-53) becomes (4-56) when the mean of the weights equals the Wiener weights, it suffices to show that (4-56) tends to
(4-55) as $\mu \to 0$ for both estimators. We first note the following mathematical result

$$\lim_{\alpha \to 1} \frac{1-\alpha^n}{1-\alpha} = m. \quad (4-62)$$

Using (4-62), we can show the limiting behavior of (4-56).

$$\lim_{\alpha \to 1} J_{RZ} = \eta \left[ \sigma_n^2 + \frac{\sigma_i^2 \sigma_i^2}{\lambda_{\text{max}}} \frac{M}{\lambda_{\text{max}}} \right] \quad (4-63)$$

Comparing to the Wiener MSE in (4-55), there is an excess term, $\sigma_i^2 M / \lambda_{\text{max}}$, still remaining in (4-63). This term tends to one as the interference becomes strong with respect to the other two components of the input signal, i.e., $\sigma_i^2 \gg (\sigma_i^2 + \sigma_n^2)$. Likewise, as the interference fades away, i.e., $\sigma_i^2 \to 0$, (4-63) also asymptotically equals the Wiener MSE. However, anywhere in between, in which the presence of the interference is noticed but not dominating the other components, this excess term is expected to cause some error in the estimate. This term is caused by the approximation in (4-6) which was based on modeling the LMS algorithm as a time-invariant system to render the transfer function approach tractable. In Section 4.5, we observe how this error impacts the MSE estimates.

4.4. Application to the NLMS Adaptive Equalizer

We briefly note here that the new MSE estimate in (4-53) can be applied to estimate the performance of the NLMS equalizer as well. This modification follows exactly how Reuter and Zeidler applied theirs to the NLMS algorithm [14].

Performing the same iterative expansion for the NLMS algorithm as we have done for the LMS algorithm in (4-1), results in a difference equation for the NLMS algorithm

$$\sum_{k=0}^{n-m} \hat{g}_{k,n} e_{n-k} = e_{m,n} \quad (4-64)$$

with the time-varying coefficients defined as
Chapter 4: Analysis of Mean Squared Error Behavior

\[ \tilde{g}_{k,n} = \begin{cases} 1, & k = 0 \\ \hat{\mu} (u_{n-k}^H u_{n-k})^{-1} u_{n-k}^H u_n, & k > 0 \end{cases} \]  
\tag{4-65}

Applying the Clarkson-White approximation (4-6) to both dot products in (4-65) yields the time-invariant approximation of the coefficients

\[ \tilde{g}_{k,n} = \frac{\hat{\mu}}{\sigma_u^2 M} r_{uu,k}, \quad k > 0. \]  
\tag{4-66}

where \( \sigma_u^2 \triangleq \sigma_x^2 + \sigma_n^2 + \sigma_i^2 \) is the variance of the input signal. Since the only change in (4-66) from (4-6) is essentially the step-size, the rest of the derivation for the LMS algorithm applies directly to the NLMS algorithm, and the NLMS MSE estimate is thus taking the same form as in (4-53), that is,

\[ \tilde{J} \approx \frac{2\eta}{(1 + \hat{\alpha})} \left[ \sigma_n^2 + \sigma_x^2 \left( \| \tilde{\mathbf{w}} \|^2 - (1 - \tilde{\alpha}) \left| \tilde{\mathbf{w}}^T \tilde{\mathbf{A}} \tilde{\mathbf{w}} \right| \right) \right] \]  
\tag{4-67}

but with

\[ \tilde{\alpha} \triangleq 1 - \hat{\mu} \frac{\sigma^2 M}{\sigma_u^2} \]  
\tag{4-68}

and

\[ \tilde{\mathbf{A}} \triangleq \begin{bmatrix} 0 & \cdots & 0 \\ 1 & 0 \\ \tilde{\alpha} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \tilde{\alpha}^{M-2} & \cdots & \tilde{\alpha} & 1 & 0 \end{bmatrix} \]  
\tag{4-69}

4.5. Numerical Illustrations

The main purpose of this section is to show the improvement of the new MSE estimator based on the mean of the LMS weights relative to the Reuter-Zeidler MSE
estimator. We will also illustrate when these MSE estimates are valid and when they are not. Moreover, for every example, the MSE estimates from computer simulation are provided to serve as a reference. There are four control parameters of interest to us: step-size $\mu$, the number of equalizer taps $M$, the interference-to-signal ratio (ISR) $\sigma_i^2 / \sigma_x^2$, and the signal-to-noise ratio (SNR) $\sigma_n^2 / \sigma_x^2$. The results are distinguished by their subscripts: $I$ for the new estimator, $RZ$ for the Reuter-Zeidler estimator, $W$ for the Wiener filter of the same structure, and $\text{sim}$ for the simulation results.

The computer simulation is set up as follows. For the transmitted signal $x_n$, we use an eight-point quadrature amplitude modulated (8-QAM) signal with minimum distance between its signal points [1, p.277]. Each $x_n$ sample that is received by the equalizer is one of the 8 possible 8-QAM symbols on the complex plane. Every 8-QAM symbol is generated randomly with equal probabilities among all symbols. The power of the 8-QAM signal is fixed at $\sigma_x^2 = 2$ in all examples. The noise $\hat{n}_n$ in the simulation is modeled as a circularly symmetric complex Gaussian process. The frequency of the interference $i_n$ is fixed at $\omega_i = 0.2\pi$ radians, as this parameter does not affect the MSE outcome. Likewise, the delay for the desired signal is fixed at $\Delta = 0$ for the same reason. The equalizer is configured to operate in error-free mode all the time. Finally, the weight vector elements for every simulation trial are initialized to their mean according to the expected mean of the weight vector (3-64) and each simulation runs for 200,000 samples. The MSE estimates are computed by time-averaging the last 100,000 samples. The first 100,000 samples are excluded from averaging to account for the expected error in the estimates for the mean of the weights, which is known to occur in the moderate ISR case, as illustrated in Figure 3-4.

Figure 4-2 shows the relative MSE with respect to the Wiener MSE as a function of the LMS step-size. The step-size is indicated with respect to the maximum eigenvalue $\lambda_{\text{max}}$ of the input correlation matrix. This result is obtained for a 51-tap equalizer with 25-dB SNR and 20-dB ISR. It is apparent from the figure that the LMS MSE is smaller than the Wiener MSE—i.e., the relative MSE is less than 0 dB—up to $\mu = 1.69\lambda_{\text{max}}^{-1}$, and that the minimum MSE is observed at $\mu = 0.47\lambda_{\text{max}}^{-1}$, which is much larger than a typically used step-size. The new estimator $J_I$ is for the most part in agreement with
the simulation (it starts to deviate for very large step-size) while the $J_{RZ}$ estimator predicts a smaller MSE than what is actually attainable. The estimation error in the $J_I$ estimator grows as the step-size increases due to the violation of the assumption in (4-25) (i.e., small weight fluctuation in steady state). As the new model MSE diverges from the simulation result at very large step-size, it tends to the Reuter-Zeidler model MSE.

![Figure 4-2. Relative MSE as a function of step-size ($M = 51$, ISR = 20 dB, SNR = 25 dB).](image)

One of the claims made in earlier work is that the transfer function approach is only valid for a large filter order so that the approximation in (4-6) gets a sufficient amount of time-averaging [14]. This condition is found to not be necessary when the input contains a dominant sinusoidal interference, as shown in Figure 4-3. These relative MSE results are for an equalizer with only 3 taps, which is the smallest filter order needed to observe this non-Wiener behavior. The new estimator tracks the simulation results as closely as for the $M = 51$ case in Figure 4-2, although for $M = 3$ the deviation at very large step-size starts earlier ($\mu = 1.25\lambda^{-1}_{\text{max}}$). The Reuter-Zeidler model, on the other hand, shows a noticeable deviation from simulation, and its predicted minimum MSE ($\mu = \lambda^{-1}_{\text{max}}$) does not correspond to that of the simulation ($\mu = 0.5\lambda^{-1}_{\text{max}}$).
Figure 4-3. Relative MSE as a function of step-size \((M = 3, \text{ISR} = 20 \text{ dB}, \text{SNR} = 25 \text{ dB})\).

Using the same ISR and SNR as in the above two examples, the MSE as a function of the equalizer length is shown in Figure 4-4a using the optimal step-size, which is as shown in Figure 4-4b. The optimal step-size is numerically determined using the golden section search with parabolic interpolation [37 p.270] with the termination tolerance of the step-size set to \(10^{-12}\). The optimal step-size for the simulation was determined by repeating the simulation with the same input signal. Its (near) optimality is verified with other input signal for reasonably sized step-sizes. This method is used for the examples hereafter when the optimal step-size is determined.
As the equalizer length is increased, the LMS equalizer initially shows more drastic MSE improvement than the corresponding Wiener equalizer, and the Wiener filter asymptotically catches up as the equalizer length is kept increasing. The maximum MSE spread between the LMS and Wiener equalizers is found to be $-5.9$ dB at $M = 34$. Our MSE model follows the simulation results very closely for both MSE and optimal step-size, while the Reuter-Zeidler model appears to provide more of a lower bound for the MSE. Moreover, in Figure 4-4b, the optimal step-size of the new model follows the dip in the simulation optimal step-size as the filter length is shortened below 10.
Figure 4-5a illustrates the MSE as a function of SNR while the ISR is fixed at 25 dB. The equalizer length of 51 is used in this example. The optimal step-size is determined for each point of evaluation as shown in Figure 4-5b. This result shows that the MSE difference between the LMS equalizer and the Wiener equalizer widens up to –11 dB as the additive noise vanishes. The Reuter-Zeidler estimator exhibits more deviation from the simulation at higher SNR, and the use of the mean of the weights in our model improves the MSE estimate to the point where it closely tracks the simulation results.

Figure 4-5. MSE as a function of SNR (a) using its optimal step-size (b) ($M = 51$, ISR = 20 dB).
The estimated optimal step-sizes in Figure 4-5b correspond to their associated MSE results in Figure 4-5a as well as to our intuition. When the additive noise level is large—i.e. SNR is small—the benefit of the LMS adaptation vanishes; there is nothing that LMS can do better than a fixed filter when the desired information in the input is covered up by strong white noise that is not correlated to the desired signal. Consequently, it is then better to adapt the LMS algorithm with a smaller step-size. The use of a smaller step-size leads to the MSE of the LMS equalizer roughly equaling that of the Wiener equalizer. On the other hand, the benefit of the LMS adaptation shows the strongest when the input signal is comprised of only the desired signal and the narrowband interference. Under the latter circumstances, the LMS step-size can be maximized according to the input ISR.

The remaining parameter that we have yet to assess is the ISR, which—of the four parameters—is the one our MSE estimator is most sensitive to. In Section 4.1, we noted that the transfer-function approximation of the LMS algorithm is only valid when the narrowband process is the dominant component in the input signal $u_n$ so that (4-7) is a valid approximation. As we reduce the ISR, the required dominance of $i_n$ over $x_n$ in $u_n$ no longer holds, and both MSE estimators break down as illustrated in Figure 4-6.
As expected, when the interference is weak (i.e., ISR below –14 dB), the LMS algorithm follows the Wiener filter behavior in both MSE in Figure 4-6a and in near-zero optimal step-size in Figure 4-6b. The full performance benefit of the LMS adaptation is achieved above 15-dB ISR, in other words, with strong interference. For ISR between –14 dB and 15 dB, the system is in a transition state where the LMS performance does not reach the predicted performance level estimated by the models, while performing clearly better than the Wiener filter. In both transition and weak-interference states, both MSE models, $J_I$ and $J_{RZ}$, break down—their predicted MSE
and optimal step-size do not correspond to the simulation results—because of the strong narrowband component assumption in the transfer function approximation being violated in those situations. On the other hand, if the interference is strong, these models agree with simulation (in case of the new model, $J_I$, almost exactly) as illustrated in Figure 4-2 to Figure 4-5.

The predicted optimal step-sizes in Figure 4-6b are clearly far from the actual optimal step-size determined from simulation, and it is logical to observe that the MSE estimates based on those predicted optimal step-sizes are not likely to be accurate. Figure 4-7 shows the relative MSE, with respect to the Wiener MSE, of the results in Figure 4-6a—$J_I(\mu_I)$ and $J_{sim}(\mu_{sim})$—and the relative MSE of the new estimator using the optimal step-size found in simulation, $J_I(\mu_{sim})$, and the simulated relative MSE using the optimal step-size determined by the estimator, $J_{sim}(\mu_I)$. Using the actual optimal step-sizes, $J_I$ improves its estimates over the ISR range of $-25$ dB to $5$ dB even though they still do not match the MSE of the simulation. On the other hand, using the predicted optimal step-sizes causes the simulation to diverge below $-6$-dB ISR as indicated by the disappearance of the $J_{sim}(\mu_I)$ curve.

![Figure 4-7. Relative LMS MSE as a function of ISR with optimal step-sizes in Figure 4-6b ($M = 51$, SNR = 25 dB).](image-url)
Chapter 4: Analysis of Mean Squared Error Behavior

The final LMS equalizer example shows the relative MSE surface as a function of both ISR and step-size. Figure 4-8a shows the relative MSE surface observed in simulation, while the relative MSE surface modeled by the new estimator is shown in Figure 4-8b. First, we discuss the simulation results in Figure 4-8a. The loss of the non-Wiener effect is visible as the convex shape for the ISR above −5 dB quickly turns into a monotonically increasing function of step-size for ISR below −12 dB. In both modes of

![Relative LMS MSE surfaces as functions of ISR and normalized step-size: (a) simulation and (b) new model (M = 51, SNR = 25 dB, relative MSE above 10 dB clipped).](image-url)
operation (i.e., with and without the non-Wiener effect), the relative MSE tends to 0 dB as the step-size tends to zero for all values of ISR. Lastly, for moderate ISR around 5 dB, the divergence (i.e., very large step-size) characteristic of the LMS algorithm appears to be different from elsewhere; its surface is not smooth, which indicates that the MSE is sensitive to specifics in each realization.

Now, we focus on the relative MSE produced by the estimator. The predicted relative MSE surface closely matches that of the simulation for ISR above 10 dB, and—in that case—for most values of step-size. Also, the small step-size behavior is modeled properly for an even wider ISR range (above −10 dB) before it starts deviating from the Wiener MSE. On the other hand, the above results clearly show the breakdown in the proposed model, taking us back to the two deficiencies of the model: insensitivity to the weight fluctuation due to (4-25), and the modeling error for small-step-size, low-ISR scenarios as shown in (4-63). The former is more prevalent in Figure 4-8 than the latter. As the interference weakens, the LMS weights begin to fluctuate more for the same step-size, affected by the white processes which become more influential, which in turn causes the MSE to increase. Because the effect of the weight fluctuation is neglected in (4-25), the proposed model does not reflect its effects on the MSE estimate. Qualitatively, this explains the low MSE observed for high step-size, low ISR scenarios, and the constant MSE as a function of step-size at −30-dB ISR. The latter shortcoming, the modeling error for the small-step-size, is observed in the absence of MSE deviation from 0 dB for small-step-size low-ISR conditions.

Lastly, Figure 4-9 presents the NLMS equalizer example—which is equivalently configured to the last LMS example in Figure 4-8—showing the relative MSE surfaces as functions of both ISR and step-size for simulation (a) and the new model (b). The theoretical results in Figure 4-9b are no different than those in Figure 4-8b because the latter result is shown with the same step-size scaling as used in (4-66). The simulation results, however, have noticeable differences for the very large step-sizes. The rough surface observed around 5-dB ISR in Figure 4-8a has vanished in Figure 4-9a. This behavioral change between the two algorithms stems from the controlled step-size of the NLMS algorithm; that is, the stability of the NLMS algorithm is guaranteed as long
Chapter 4: Analysis of Mean Squared Error Behavior

Figure 4-9. Relative NLMS MSE surfaces as functions of ISR and step-size: (a) simulation and (b) new model ($M = 51$, SNR = 25 dB, relative MSE above 10 dB clipped).

as its step-size is bounded between 0 and 2. The LMS algorithm, on the other hand, does not have that guarantee, and over the medium-ISR rough area in Figure 4-8a, the stability condition of the LMS algorithm is in a different mode than for other ISR ranges.

4.6. Summary

This chapter covered the method of estimating the mean squared error for generic
LMS applications that involve narrowband signals. This method is applied to the adaptive equalization problem, and we have shown that the resulting MSE estimator accurately predicts the performance of the equalizer when dominant interference is present. By means of numerical experiments the breakdown of the estimator was demonstrated when the interference component is weakened, i.e. when two of the main conditions—assumed for its derivation—become violated.
5.1. Conclusions

This study focused on the behaviors of adaptive transversal equalizers operating in the presence of narrowband interference. Specifically, the research interest lies in understanding the behavior of the equalizers when LMS or NLMS adaptive algorithms are used in the presence of a single sinusoidal interferer. Such adaptive systems exhibit three notable deviations from what conventional adaptive filter theory expects. First, the mean-squared-error (MSE) performance of these adaptive equalizers—over a wide range of adaptation step-sizes—exceeds the performance of the optimal fixed equalizer, i.e., the Wiener filter, of the same structure. Second, the optimal step-sizes that minimize the steady-state MSE are much larger than the small step-size that is typically utilized in conventional applications in steady state. Third, the mean of the adaptive equalizer weights at those large step-sizes is located away from the corresponding Wiener filter weights.

The main contribution of this work is the discovery of the shift in the mean of the adapted weights from the corresponding Wiener filter weights, and the analysis of that discovered displacement under the assumption of large interference-to-signal ratio. The mean of the LMS weight vector in steady state is analytically derived based on the Butterweck expansion of the weight update equation. Excellent correspondence between the analytical and simulation results for the mean of the LMS weight vector is observed over nearly the entire range of stable step-sizes. The analytical expression, while not exact, captures the transition from Wiener to non-Wiener weight behavior as the ISR increases. A straightforward modification of the expression for the mean of the
LMS weight vector gives a good estimate of the mean of the weight values of the NLMS algorithm.

As a second contribution, the MSE performance of the LMS and NLMS equalizers is studied. This portion of the work is based on the original development by Reuter and Zeidler [14]. We established a generalized framework for the transfer function approximation of the LMS algorithm and provided a rigorous and explicit analysis of the LMS MSE estimation based on the transfer function approximation. The analysis resulted in a solid explanation for the use of the mean of the adaptive filter weights in the MSE estimator, instead of the previously assumed Wiener weight initialization [14]. Subsequently, using the new analytical expression for the mean of the weight vector, the new MSE estimator is a vast improvement over the Reuter-Zeidler MSE estimator, under the assumption of an interference dominated environment. Using the new MSE estimator, the unnecessary assumption of needing large order filters has been dispelled in the case of strong sinusoidal interference.

Moreover, analyzing the transfer-function based method of the MSE estimation revealed a critical assumption, namely that the LMS weight fluctuation must be assumed to be negligibly small in steady-state. This assumption was not brought to the fore earlier [14] since the assumption turned out to hold when the narrowband interference is dominating the other input components. The breakdown of the MSE estimators when the interference is not strong is mainly due to the violation of the small weight fluctuation assumption in steady state.

5.2. Future Directions of the Research

To extend the research topic presented in this dissertation, there are three different tangential paths to be pursued: analysis of other system properties of the current system model, analysis of similar equalization environments, and analysis of different adaptive algorithms for the same system. We conclude this dissertation by briefly covering each potential extension.

While this work analyzed two of the most important characteristics of adaptive transversal equalizers in the presence of narrowband interference, there are many
others that can be studied in depth. Here, we present three potential topics of interest, noting that the list is not an exhaustive one. First, analyzing the LMS weight fluctuation for arbitrary step-size can further enhance the current MSE estimator by filling a missing link. For this analysis, the Butterweck expansion may be suitable if similar uncorrelatedness of partial weights is observed for the second-moment analysis. The second characteristic is the transient behavior of the adaptive transversal equalizer. These equalizers are experimentally observed to converge very slowly, much slower than suggested by the average time constant [8]. Lastly, the instantaneous fluctuation of the weights can be studied to search for deterministic or linked weight behavior along the lines of the periodic weight behavior observed in the narrowband adaptive noise canceller [17].

Next, similar analyses of the behavior of the mean of the LMS weights and the corresponding MSE performance estimate may be performed for other operating environments. Each of the three that were illustrated in Section 2.4.2 poses an interesting extension. First is to consider the AR(1) interference. Reuter and Zeidler [14] already looked at the MSE performance for this type of interference, but without considering the shift in the mean of the weight vector. The second case, i.e., two sinusoidal interferers with close frequencies, may be the most fascinating of the three examples. Having two interferers can result in better MSE performance than having just one interferer. Lastly, the diminishing non-Wiener effect due to a non-ideal channel could be analyzed. This problem relates to having colored noise in the equivalent interference estimation problem.

Finally, the equalizer may be adapted by other adaptive algorithms, and the presence of the non-Wiener effect, or lack thereof, can be analyzed. One candidate is the recursive least squares algorithm, which was shown to operate in non-Wiener fashion in the adaptive noise cancelling application [38], albeit in a muted way.
APPENDIX A DERIVATION OF (3-12)

To show the expanded form of the state transition matrix $P^n$ in (3-12), we start off by substituting (2-16) into:

$$P^n = (I - \mu R)^n$$
$$= [I - \mu (\sigma_n^2 I + \sigma_i^2 ee^H)]^n$$
$$= [(1 - \mu \sigma_n^2)I - \mu \sigma_i^2 ee^H]^n \quad (A-1)$$

Using the binomial theorem [39 p.10], (A-1) becomes

$$P^n = \sum_{k=0}^{n} \binom{n}{k} (1 - \mu \sigma_n^2)^{n-k} (-\mu \sigma_i^2 ee^H)^k$$
$$= (1 - \mu \sigma_n^2)^n \left[ I + \sum_{k=1}^{n} \binom{n}{k} \left( -\mu \sigma_i^2 \right)^k \left( ee^H \right)^k \right] \quad (A-2)$$

Because $ee^H = M$ from its definition, we get

$$P^n = (1 - \mu \sigma_n^2)^n \left[ I + ee^H \sum_{k=1}^{n} \binom{n}{k} \left( -\mu \sigma_i^2 \right)^k \left( M \right)^{k-1} \right]$$
$$= (1 - \mu \sigma_n^2)^n \left[ I + ee^H \frac{1}{M} \sum_{k=1}^{n} \binom{n}{k} \left( -\mu \sigma_i^2 \frac{M}{1 - \mu \sigma_n^2} \right)^k \right] \quad (A-3)$$
$$= (1 - \mu \sigma_n^2)^n \left[ I + ee^H \frac{1}{M} \left( \sum_{k=0}^{n} \binom{n}{k} \left( -\mu \sigma_i^2 \frac{M}{1 - \mu \sigma_n^2} \right)^k - 1 \right) \right]$$

The final summation term in (A-3) can be reverted to polynomial form by again
Appendix A: Derivation of (3-12)

applying the binomial theorem:

\[ P^n = \left(1 - \mu \sigma_n^2\right)^n \left[ I + ee^H \frac{1}{M} \left\{ \left(1 - \frac{\mu \sigma_i^2 M}{1 - \mu \sigma_n^2} \right)^n - 1 \right\} \right] \tag{A-4} \]

From (A-4), we find (3-12) by using (2-18) followed by some algebra.
APPENDIX B  EXPRESSION FOR MEAN OF WEIGHT MAGNITUDE

The derivation of the expression for the magnitude of the mean of the LMS weights is approached in two steps outlined as follows. We start with the recursive solutions of the even-order partial weights $\mathbf{v}_{2l}$ —as found in (3-60) and (3-25)—and show that every $\mathbf{v}_{2l}$ can be expressed in the form

$$\mathbf{v}_{2l} = \mathbf{V}_l e^{j\omega\Delta} \quad \text{(B-1)}$$

where

$$\mathbf{V}_l = \text{diag} \{ |\mathbf{v}_{2l,0}|, |\mathbf{v}_{2l,1}|, \ldots, |\mathbf{v}_{2l,M-1}| \} \quad \text{(B-2)}$$

is the $\mathbb{R}^{M \times M}$ diagonal matrix containing the magnitude of each element in $\mathbf{v}_{2l}$. Then, the expression derived for the real diagonal matrix $\mathbf{V}_l$ is columnized to show that the vector of magnitudes of the even-order partial weights is as defined in (3-69).

We prove the proposition in (B-1) by induction. The $l = 0$ base case is straightforward from (3-25) and the Wiener solution in (3-5):

$$\mathbf{V}_0 = \frac{\sigma^2}{\lambda_{\text{max}}} \mathbf{I} \quad \text{(B-3)}$$

For the induction step, we show that $\mathbf{V}_l$ is real and diagonal given that $\mathbf{V}_{l-1}$ is real and diagonal. Substituting (B-1) into (3-60) yields

$$\mathbf{V}_l e^{j\omega\Delta} = (\mu \mathbf{R}^{-1} \mathbf{Q}) \mathbf{V}_{l-1} e^{j\omega\Delta} \quad \text{(B-4)}$$
Expanding $\mathbf{R}^{-1}$ and $\mathbf{Q}$ according to (2-17) and (3-59), respectively, (B-4) becomes

$$
\mathbf{V}_i \mathbf{e}^{j\omega \Delta} = \mu \sigma_i^2 M \left( \mathbf{I} - \frac{\sigma_i^2}{\lambda_{\text{max}}} \mathbf{e} \mathbf{e}^H \right) \sum_{p=1}^{M-1} \mathbf{Z}^p e^{-j\omega \gamma_p} \mathbf{V}_{l-1} \mathbf{e}^{j\omega \Delta} \tag{B-5}
$$

Introducing the algebraic result

$$
\mathbf{Z}^p \mathbf{V}_{l-1} \mathbf{e}^{-j\omega \gamma_p} = \mathbf{Z}^p \mathbf{V}_{l-1} \mathbf{N}^p \mathbf{e}, \tag{B-6}
$$

we get

$$
\mathbf{V}_i \mathbf{e}^{j\omega \Delta} = \mu \sigma_i^2 M \left( \mathbf{I} - \frac{\sigma_i^2}{\lambda_{\text{max}}} \mathbf{e} \mathbf{e}^H \right) \sum_{p=1}^{M-1} \gamma_p^{-1} \mathbf{Z}^p \mathbf{V}_{l-1} \mathbf{N}^p \mathbf{e} \mathbf{e}^{j\omega \Delta} \tag{B-7}
$$

Because $\mathbf{V}_i$ is real and diagonal, and the $\mathbf{e}$ vectors are the only complex quantities, we have

$$
\mathbf{e}^H \mathbf{Z}^p \mathbf{V}_{l-1} \mathbf{N}^p \mathbf{e} = \sum_{m=0}^{M-p-1} \mathbf{e}^{j\omega (m+p)} |\mathbf{\bar{v}}_{2(l-1),m}| \mathbf{e}^{-j\omega (m+p)}
$$

$$
= \sum_{m=0}^{M-p-1} |\mathbf{\bar{v}}_{2(l-1),m}| \tag{B-8}
$$

$$
= \sum_{m=0}^{M-p-1} 1 |\mathbf{\bar{v}}_{2(l-1),m}| 1
$$

$$
= \mathbf{1}^H \mathbf{Z}^p \mathbf{V}_{l-1} \mathbf{N}^p \mathbf{1}
$$

This quantity is a real scalar. Hence,

$$
\mathbf{V}_i = \mu \sigma_i^2 M \sum_{p=1}^{M-1} \gamma_p^{-1} \left[ \mathbf{Z}^p \mathbf{V}_{l-1} \mathbf{N}^p - \frac{\sigma_i^2}{\lambda_{\text{max}}} (\mathbf{1}^H \mathbf{Z}^p \mathbf{V}_{l-1} \mathbf{N}^p \mathbf{1}) \mathbf{I} \right]. \tag{B-9}
$$
This matrix is real (since it is a function of exclusively real entities) and is diagonal (as $\mathbf{Z}^p \mathbf{V}_{l-1} \mathbf{N}^p$ and $\mathbf{I}$ are both diagonal matrices). Hence, we conclude the first step of the derivation since the base case is true and the inductive step is true.

Now, we form the magnitude vector $|\mathbf{v}_l|$ of the partial weights from the diagonal elements of $\mathbf{V}_l$, i.e.,

$$|\mathbf{v}_l| \triangleq \begin{bmatrix} |\mathbf{v}_{2l,0}| & |\mathbf{v}_{2l,1}| & \cdots & |\mathbf{v}_{2l,M-1}| \end{bmatrix}^T$$  \hspace{1cm} (B-10)

Noting that $\mathbf{Z}^p \mathbf{V}_{l-1} \mathbf{N}^p$ shifts the diagonal elements of $\mathbf{V}_{l-1}$ down diagonally by $p$ elements, its equivalent operation with respect to $|\mathbf{v}_l|$ is $\mathbf{Z}^p |\mathbf{v}_{l-1}|$, shifting the elements in $|\mathbf{v}_{l-1}|$ down by $p$. We also observe that

$$1^H \mathbf{Z}^p \mathbf{V}_{l-1} \mathbf{N}^p \mathbf{1} = \sum_{q=p}^{M-1} b_{l-1,q} = 1^T \mathbf{Z}^p |\mathbf{v}_{l-1}|$$  \hspace{1cm} (B-11)

Transforming (B-9) using the above two results yields

$$|\mathbf{v}_l| = \mu \sigma_i^2 M \sum_{p=1}^{M-1} \gamma^{p-1} \left[ \mathbf{Z}^p |\mathbf{v}_{l-1}| - \frac{\sigma_i^2}{\lambda_{\text{max}}} 11^H \mathbf{Z}^p |\mathbf{v}_{l-1}| \right]$$

$$= \mu \sigma_i^2 M \sum_{p=1}^{M-1} \gamma^{p-1} \left[ \mathbf{I} - \frac{\sigma_i^2}{\lambda_{\text{max}}} 11^H \right] \mathbf{Z}^p |\mathbf{v}_{l-1}|$$

$$= \mu \frac{1}{\sigma_w^2} \left[ \mathbf{I} - \frac{\sigma_i^2}{\lambda_{\text{max}}} 11^H \right] \sigma^2_w \sigma_i^2 M \sum_{p=1}^{M-1} \gamma^{p-1} \mathbf{Z}^p |\mathbf{v}_{l-1}|$$  \hspace{1cm} (B-12)

All the ingredients for the matrices $\Xi$ and $\Theta$ are found in (B-12) as defined in (3-67) and (3-68), respectively. Substituting $\Xi$ and $\Theta$ into (B-12) we get

$$|\mathbf{v}_l| = \mu \Xi \Theta |\mathbf{v}_{l-1}|.$$  \hspace{1cm} (B-13)

Obtaining (3-69) and (3-70) from (B-13) follows the same steps as shown from (3-60) to (3-64).
APPENDIX C  MATRIX INVERSE OF (3-72)

Because $L^{-1}$ in (3-72) is a lower triangular matrix, the first column $l_1$ of $L$ can be found by solving the system of equations:

$$L^{-1}l_1 = p_0$$

The basis vector $p_0$ has as first element a one and the rest are zeros, as defined according to (2-20). Subsequently, the $m$-th element of the first column of $L$ is found by

$$l_{m,1} = \begin{cases} 1, & m = 1 \\ \mu \sigma_i^2 M \sum_{k=1}^{m-1} \gamma^{m-k-1} l_{k,1}, & 1 < m \leq M \end{cases}$$

The recursion in (C-2) is completely unwrapped for the first few terms to identify the pattern:

$$l_{1,1} = 1$$
$$l_{2,1} = \mu \sigma_i^2 M$$
$$l_{3,1} = \mu \sigma_i^2 M (\gamma + \mu \sigma_i^2 M)$$
$$l_{4,1} = \mu \sigma_i^2 M \left[ \gamma^2 + \gamma \mu \sigma_i^2 M + \mu \sigma_i^2 M (\gamma + \mu \sigma_i^2 M) \right]$$
$$l_{5,1} = \mu \sigma_i^2 M \left[ \gamma^3 + \gamma^2 \mu \sigma_i^2 M + \gamma \mu \sigma_i^2 M (\gamma + \mu \sigma_i^2 M) + \mu \sigma_i^2 M \left\{ \gamma^2 + \gamma \mu \sigma_i^2 M + \mu \sigma_i^2 M (\gamma + \mu \sigma_i^2 M) \right\} \right]$$

Simplifying each unwrapped term in (C-3) yields
\[ l_{i,1} = 1 \]
\[ l_{2,1} = \mu \sigma_i^2 M \]
\[ l_{3,1} = \mu \sigma_i^2 M \left( \gamma + \mu \sigma_i^2 M \right) \]
\[ l_{4,1} = \mu \sigma_i^2 M \left( \gamma + \mu \sigma_i^2 M \right)^2 \]
\[ l_{5,1} = \mu \sigma_i^2 M \left( \gamma + \mu \sigma_i^2 M \right)^3 \]

We can readily identify the geometric series starting with the \( l_{2,1} \) term. Accordingly, the \( m \)-th term is then defined as

\[ l_{m,1} = \mu \sigma_i^2 M \left( \gamma + \mu \sigma_i^2 M \right)^{m-2}, \quad 1 < m \leq M \]  

(C-5)

By the definition of \( \gamma \triangleq 1 - \mu \lambda_{\text{max}} \),

\[ l_{m,1} = \mu \sigma_i^2 M \left( 1 - \mu \sigma_i^2 \right)^{m-2}, \quad 1 < m \leq M \]  

(C-6)

Hence, the elements of the first column of \( L \) are fully defined by incorporating the \( l_{1,1} = 1 \) case, and results in (3-73).
References

References


Takeshi Ikuma received his dual Bachelor of Science degrees in Electrical Engineering and Computer Engineering in May 1999 from Virginia Tech. He remained at Virginia Tech for graduate study and completed the Master of Science degree in Electrical Engineering in August 2001 under the supervision of Dr. A. A. (Louis) Beex (Thesis: “Model-based identification of POTS local loops for DSL connectivity prediction”).

His research interests are in the areas of digital signal processing and stochastic signal processing, with special emphasis in adaptive filters. During his tenure at Virginia Tech, he has compiled the following publications.


