

**EFFECTS OF ILL-CONDITIONED DATA
ON LEAST SQUARES ADAPTIVE FILTERS**

**Gary A. Ybarra
and
S.T. Alexander**

**Center for Communications and Signal Processing
Electrical and Computer Engineering Department
North Carolina State University**

CCSP-TR-88/6

January 1988

ABSTRACT

An ill-conditioned least squares (LS) problem has a solution which may be highly sensitive to small perturbations in the data. This paper presents sensitivity results for the LS weight vector of an all-zero adaptive equalizer in an ill-conditioned signal environment. The ill-conditioned data results from severe amplitude distortion introduced by the data channel. Specifically, a theoretical upper bound for the relative change in the magnitude of the weight vector due to additive noise is derived as a function of the channel characteristics and equalizer order, and simulation results are presented which indicate the derived bound to be applicable over a wide range of equalizer orders.

1. INTRODUCTION

Many applications of adaptive filtering such as channel equalization [1], AR spectrum estimation [2], and line tracking [3] must operate in environments which are extremely ill-conditioned. Ill-conditioned data is marked by a large eigenvalue spread in the correlation matrix and this implies a high degree of signal correlation. In LMS filtering, the smallest eigenvalue is responsible for the duration of the adaptation period, and if λ_{\min} is nearly zero, as occurs frequently in many applications, LMS type algorithms may require hundreds of thousands of iterations to converge even for stationary data.

However, the adaptive parameter convergence rate in exact least squares (LS) algorithms is less affected by signal correlation properties, making LS methods attractive in applications requiring rapid adaptation in poorly conditioned signal environments. Adaptive LS filters typically converge shortly after the filter has filled with data, provided the noise power is relatively small compared to the signal power. Very ill-conditioned data can cause two distinctly different effects on the least squares solution: (1) numerical error accumulation and (2) increased solution sensitivity. Concerning error accumulation, an ill-

conditioned LS problem can cause the accumulation of round-off errors leading to a computed numerical solution which deviates significantly from the theoretical, infinite precision solution. In fact, ill-conditioning is a prime contributor to rapid divergence of fast, adaptive, recursive L.S. squares filters [4].

Concerning solution sensitivity, an ill-conditioned LS problem also has a solution which is highly sensitive to small changes or perturbations in the data [5]. Additive noise, inaccuracies in signal measurement devices and quantization are examples of small data perturbations. Another example is identification of systems whose inputs are inaccessible. If the inaccessible input is active during identification, then the data used by the filter is perturbed impulse response data, and bias will be present in the estimated system parameters.

This paper examines the sensitivity to additive noise of an adaptive LS transversal equalizer operating in an ill-conditioned signal environment. Additionally, this introductory paper is limited to an investigation of the steady state, or converged LS filter solution. The main result is the derivation of a simple upper bound for the relative change in the magnitude of the weight vector as a function of the equalizer order and channel properties. Finally, simulation results are presented which compare this bound to actual values. Results in this paper show that the bound is very nearly attained for a wide range of equalizer orders, provided the noise level is relatively low.

2. THE EQUALIZATION PROBLEM

A data channel with transversal equalizer in the training mode is shown in Figure 1. The presence of $d(n)$ and the DELAY device in the lower branch of Figure 1 signifies $d(n)$ is present during the training phase at the receiver. For simplicity in this initial examination, the data $d(n)$ is assumed to be a pseudo-random sequence of ± 1 's. The channel is modeled as a causal, linear phase M th order FIR so that only amplitude distortion is introduced. The

The symmetric matrix $C^T C$ is the statistical auto-correlation matrix of the equalizer input $x(n)$ provided the data sequence $d(n)$ is white and $v(n) = 0$. The vector $C^T \pi$ is the statistical cross-correlation vector of $x(n)$ with $d(n-D)$ and is simply the channel impulse response sequence in reverse order, embedded in a null vector. Hence (3) is identical to the statistical normal equations given by

$$\mathbf{R}w = \mathbf{p} \quad (4)$$

where \mathbf{R} is the $N \times N$ auto-correlation matrix of $x(n)$ and \mathbf{p} is the $N \times 1$ cross-correlation vector of $x(n)$ with $d(n-D)$. The value of the delay D determines the location of the unit entry in π and hence the center of the channel impulse response in \mathbf{p} , and is usually chosen so that the unit entry in π is near the center of the vector.

In recursive least squares algorithms (RLS), a set of sample normal equations is solved at each time iteration to adapt the equalizer tap values. As time progresses, and if the data forgetting factor λ is set to unity as in the prewindowed case, the sample normal equations evolve into the statistical normal equations [2]. In this paper, it is assumed that enough data has been processed so that the sample normal equations may be replaced by the statistical normal equations.

The autocorrelation values for $x(n)$ are defined by

$$R(m) = E \{x(n)x(n+m)\} \quad (5a)$$

where $x(n)$ is given by (1) and $E\{\cdot\}$ denotes expected value. With the uncorrelated additive noise $v(n)$ present, substituting (1) into (5a) yields

$$R_v(0) = R(0) + \sigma_v^2, \quad R_v(m) = R(m) \quad m \neq 0 \quad (5b)$$

where the subscript v has been attached to the autocorrelation values of $x(n)$ with the noise present. From (5b) it is seen that the noise alters only the main diagonal of \mathbf{R} in (4), producing a "noisy" set of equations given by

$$\left(\mathbf{R} + \sigma_v^2 \mathbf{I}\right) \mathbf{w}_v = \mathbf{p}, \quad (6)$$

where \mathbf{w}_v is the solution to the noisy problem. The following section next quantifies the relative change from \mathbf{w} to \mathbf{w}_v as a function of data conditioning.

3. EFFECTS OF ILL-CONDITIONED DATA

Ill-conditioning in the equalization problem occurs as a result of severe amplitude distortion imposed by the data channel. Phase distortion produced by the channel has no impact on the conditioning [6]. The degree of ill-conditioning present is quantified by the condition number $K(\mathbf{R})$ of the correlation matrix given by [5]

$$K(\mathbf{R}) = \|\mathbf{R}\| \|\mathbf{R}^{-1}\| = \frac{\lambda_{\max}}{\lambda_{\min}} \quad (7)$$

where the Euclidean norm has been used for the matrix norms, and is defined for a symmetric matrix \mathbf{A} as the maximum eigenvalue of \mathbf{A} . Let λ_{\max} and λ_{\min} be the maximum and minimum eigenvalues of the correlation matrix \mathbf{R} . These are bounded by the maximum and minimum, respectively, of the channel power transfer function [6]. Thus, if the channel power transfer function contains a zero on the unit circle, λ_{\min} may attain a very small value causing $K(\mathbf{R})$ from (7) to become extremely large. It should be noted that with the all-zero channel model chosen, a large condition number is the result of a small λ_{\min} and not a large λ_{\max} . The relative change in the LS solution to

$$\left(\mathbf{R} + \epsilon \Delta \mathbf{R}\right) \mathbf{w}_v = \mathbf{p} \quad (8)$$

due to small ϵ has an upper bound [5] given by

$$\delta_{\mathbf{w}} = \frac{\|\mathbf{w}_v - \mathbf{w}\|}{\|\mathbf{w}\|} \leq K(\mathbf{R}) \frac{\|\epsilon \Delta \mathbf{R}\|}{\|\mathbf{R}\|} + O(\epsilon^2) \quad (9)$$

where from (6) $\Delta \mathbf{R}$ may be taken as the identity matrix and ϵ the noise variance σ_v^2 . Substituting (7) into (9) and neglecting the higher order terms, then gives

$$\delta_{\mathbf{w}} \leq \|\mathbf{R}\| \|\mathbf{R}^{-1}\| \frac{\|\sigma_v^2 \mathbf{I}\|}{\|\mathbf{R}\|} = \|\mathbf{R}^{-1}\| \|\sigma_v^2 \mathbf{I}\|. \quad (10)$$

But since $\|\mathbf{R}^{-1}\| = 1/\lambda_{\min}$ and $\|\sigma_v^2 \mathbf{I}\| = \sigma_v^2$, then the following simple bound results from (10):

$$\delta_{\mathbf{w}} \leq \frac{\sigma_v^2}{\lambda_{\min}}. \quad (11)$$

It remains to determine the relationship between λ_{\min} and the channel power transfer function. Consider the normalized channel spectrum shown in Figure 2, where P_{\min} is the minimum value of the channel spectrum. Let there be $k - 1$ vanishing derivatives of $P(\omega)$ at $\omega = \omega_0$, where the k th derivative is non-zero and has a value of β . Then, as shown by Parter [7], the minimum eigenvalue of the Toeplitz matrix \mathbf{R} has the asymptotic expansion

$$\lambda_{\min} = P_{\min} + \frac{A \beta}{k! N^k} + O\left(\frac{1}{N^k}\right), \quad (12)$$

where A is a constant dependent on k . In general, the next term of the expansion is unknown. If $P_{\min} = 0$, then λ_{\min} asymptotically approaches zero as the equalizer length is increased. Thus, a severely ill-conditioned problem arises when the channel has a spectral null with multiple vanishing derivatives at the null and a lengthy equalizer is applied.

The rate of growth of the upper bound (11) is obtained by substituting (12) into (11). Assuming $P_{\min} = 0$, and retaining only the second term of the expansion gives

$$\delta_{\mathbf{w}} \leq \frac{\sigma_v^2 k! N^k}{A \beta}. \quad (13)$$

Thus, the number of vanishing derivatives of $P(\omega)$ at $\omega = \omega_0$ dominates the rate of growth of $\delta_{\mathbf{w}}$ as a function of the equalizer order.

4. EXPERIMENTAL RESULTS

In order to test the analytical upper bound for $\delta_{\mathbf{w}}$, three different channel models were examined, each with a single zero at $\omega_0 = \pi/3$. Channel 1 has an impulse response

sequence $(1/\sqrt{3}, -1/\sqrt{3}, 1/\sqrt{3})$. Its corresponding power spectrum has a single vanishing derivative at ω_0 and is illustrated in Figure 3. Channel 2 was obtained by convolving the Channel 1 impulse response with itself and again normalizing to unit energy. The resulting power spectrum has three vanishing derivatives at ω_0 . Finally, Channel 3 was obtained by convolving the Channel 2 impulse response with itself yielding a channel spectrum with seven vanishing derivatives at ω_0 . Plots comparing the actual values for δ_w with the bound (11) for channels 2 and 3 are shown in Figure 4. The actual values were computed by taking the solutions to (4) and (6), and substituting into the definition (9) for δ_w for $5 \leq N \leq 50$. The upper bound (11) was computed using an eigenvalue-eigenvector decomposition for \mathbf{R} . Since the random data $d(n)$ has unit power, the signal-to-noise ratio (SNR) at the input to the equalizer is

$$\text{SNR} = 1/\sigma_v^2. \quad (14)$$

In all cases, the bound for δ_w had a rate of growth predicted by (13), e.g. N^4 for Channel 2 and N^8 for Channel 3. The value of the noise variance for Channel 2 in Figure 4a is $\sigma_v^2 = 5 \cdot 10^{-4}$. As can be seen, δ_w is approximately three orders of magnitude larger than the additive noise for $N > 35$. The rate of growth of the actual δ_w and its bound are nearly the same over the range of equalizer orders shown. This is in contrast to Figure 4b which shows the actual value to be leveling off for $N > 25$, while the bound continues to grow as N^8 . The leveling off of the actual δ_w is a result of improved conditioning caused by the noise. As N is increased, the condition number $K(\mathbf{R})$ grows as N^8 . However, the small noise variance $\sigma_v^2 = 10^{-3}$ added to the main diagonal of \mathbf{R} significantly improves the problem condition. For small values of N , the improvement in conditioning due to the noise is negligible, and therefore the upper bound and actual values for δ_w grow at nearly the same rate.

In terms of the physical problem of equalization, these results have the following application. If the channel contains a spectral null with several vanishing derivatives such as Channel 3, then the equalizer tries to produce a large gain over the weak spectral region. However, if broadband noise above a certain level is added by the channel, the spectral gap is filled enough so that the large number of vanishing derivatives no longer dominates the conditioning of the problem determined by (12). As an example, figure 4c shows that if the noise level is reduced to $\sigma_v^2 = 10^{-6}$, then the actual δ_w very nearly attains the upper bound over the range of equalizer orders $10 \leq N \leq 40$. No leveling in the actual value of δ_w occurs because the noise does not significantly improve the problem condition.

To demonstrate how the results of the steady state analysis may be extended to adaptive LS filter applications, a 31-tap fast transversal equalizer was simulated with Channel 2. Trajectories of the main tap value w_{15} for both noisy ($\sigma_v^2 = 5 \cdot 10^{-4}$) and no-noise data are shown in Figure 5. The data in Figure 4a predicts $\delta_w = 11\%$ which suggests that an approximate 11% relative change can be expected from the no-noise to noisy steady state tap values. After 5000 iterations, the relative change in the coefficient vector was found to be $\delta_w = 9.4\%$ and the relative change in w_{15} was 7%. Interestingly, it is also evident from Figure 5 that the severe ill-condition of the data has significantly slowed the convergence rate of the FTF algorithm. Regular RLS using the matrix inverse lemma converged even more slowly.

5. Conclusion

This paper has presented a simple bound for the additive noise-induced change in the magnitude of the coefficient vector of a transversal equalizer operating in an ill-conditioned signal environment. An asymptotic expansion for the minimum eigenvalue of a Toeplitz matrix has allowed a bound for the noise-induced change in the equalizer to be predicted as

a function of the equalizer length and noise variance. Simulation results have shown that the bound is accurate for highly ill-conditioned channels operating with relatively low-level additive noise. It has also been shown that the ill-conditioned problem arises as a consequence of vanishing derivatives of the channel power transfer function at its zeroes.

References

- [1] S. Qureshi, "Adaptive Equalization," *Proc. of the IEEE*, Oct. 1985.
- [2] S.T. Alexander, *Adaptive Signal Processing*, Springer-Verlag, New York, 1986.
- [3] W.S. Hodgkiss Jr. and J.A. Presley Jr., "Adaptive Tracking of Multiple Sinusoids Whose Power Levels are Widely Separated," *IEEE Trans. on ASSP*, vol. ASSP-29, No. 3, pp. 710-721, June 1981.
- [4] J. Cioffi, "Limited-Precision Effects in Adaptive Filtering," *IEEE Trans. on CAS*, vol. CAS-34, pp. 821-833, July 1987.
- [5] G.H. Golub and C.F. Van Loan, "Matrix Computations," *John Hopkins Press*, Baltimore, Md. 1983.
- [6] F. Milinazzo, C. Zala, and I. Barrodale, "On the Rate of Growth on Condition Numbers for Convolution Matrices," *IEEE Trans. on ASSP*, vol. ASSP 35, no. 4, pp.471-475, April 1987.
- [7] S.V. Parter, "On the Extreme Eigenvalues of Truncated Toeplitz Matrices," *Bulletin of Amer. Math. Soc.*, vol. 67, pp. 191-196, 1961.

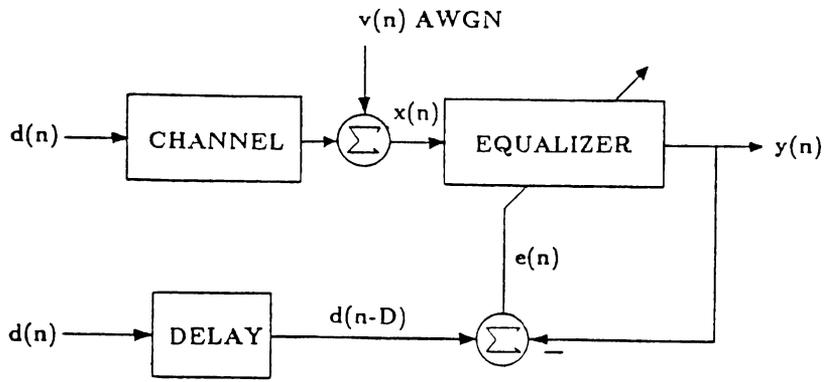


Figure 1 Transversal Equalizer in Training Mode.

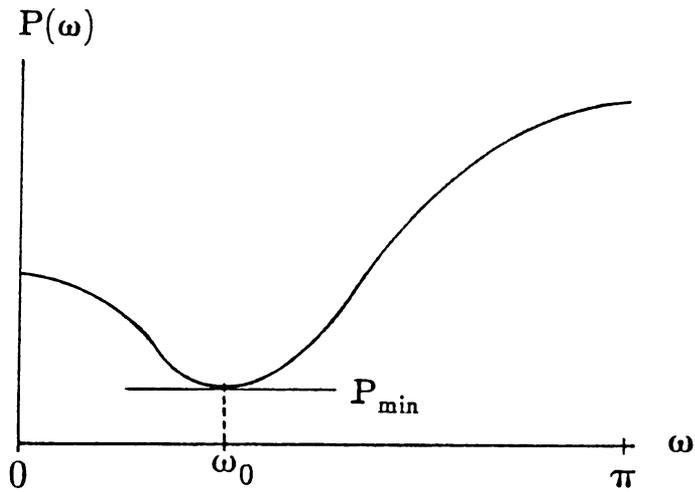


Fig. 2 Normalized Channel Power Transfer Function.

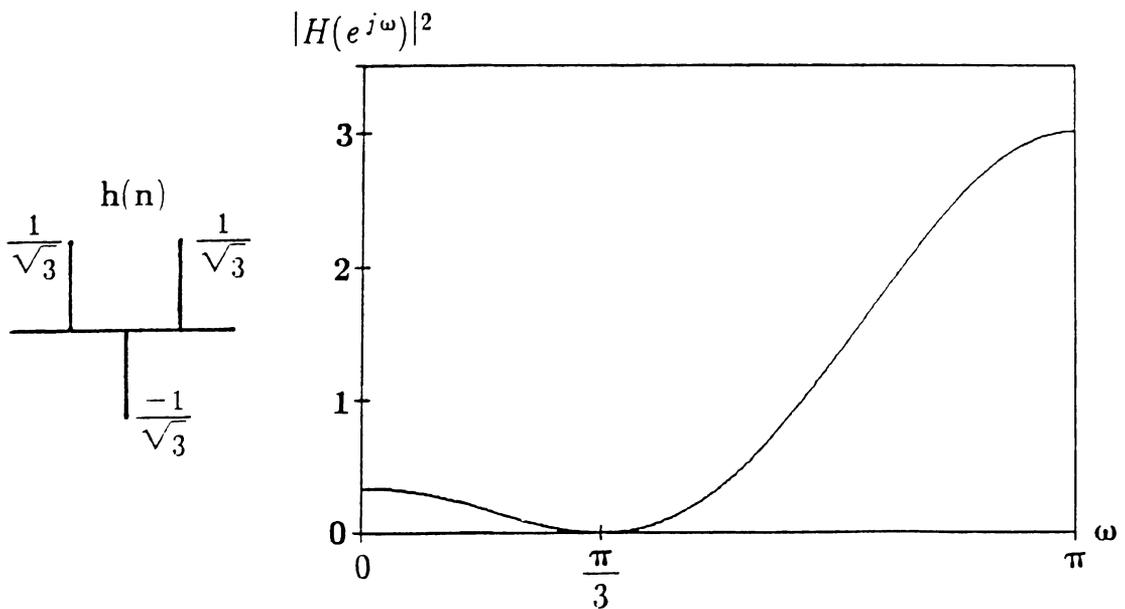


Figure 3 Impulse Response and Power Transfer Function of Channel 1.

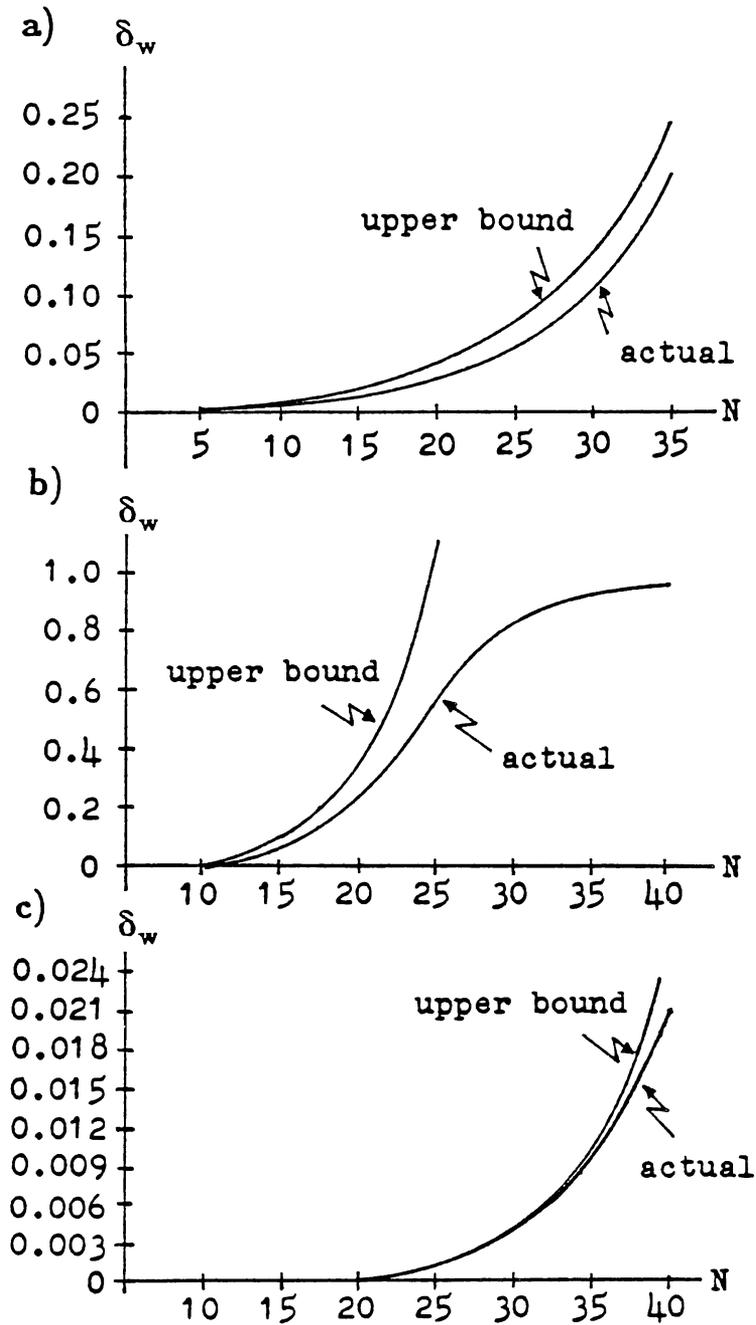


Figure 4 Relative Change in Equalizer Tap Values and Theoretical Upper Bound as a Function of Equalizer Order.
a) Channel 2 with SNR = 33 dB
b) Channel 3 with SNR = 30 dB
c) Channel 3 with SNR = 60 dB.

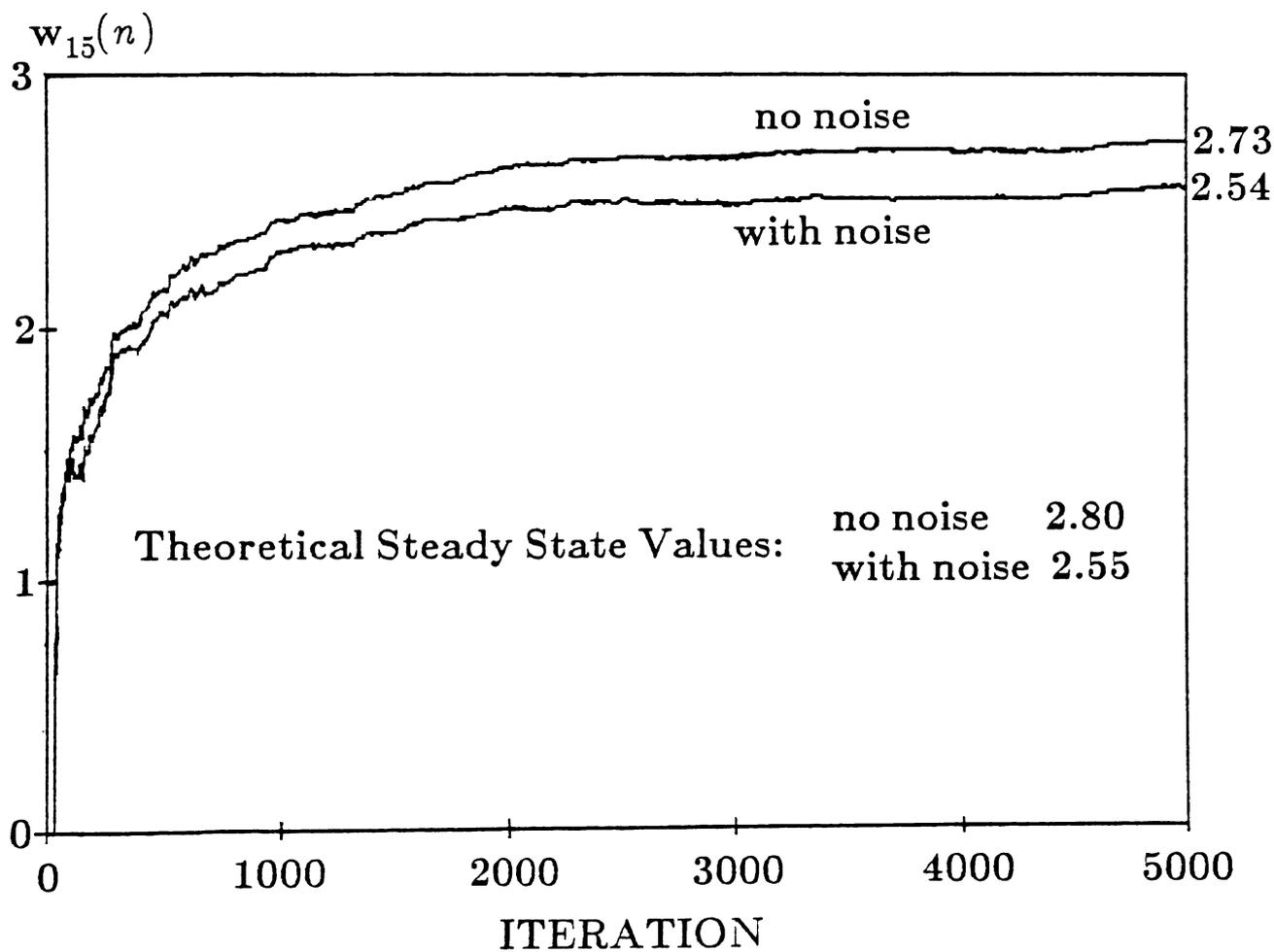


Figure 5 Main Tap Trajectories for a 31 Tap Fast Transversal Equalizer Operating with Channel 2 and SNR = 33 dB.