NON-WIENER WEIGHT BEHAVIOR OF LMS TRANSVERSAL EQUALIZERS

Takeshi Ikuma[1], A. A. (Louis) Beex[1], and J. R. Zeidler[2]

[1]DSP Research Laboratory, Wireless @ Virginia Tech, Blacksburg, VA 24061-0111
[2]Dept. of ECE, Univ. of California at San Diego, La Jolla, CA 92093-0047

ABSTRACT

The least mean square (LMS) algorithm is widely assumed to operate around the corresponding Wiener filter solution. It has been observed that an exception to this popular perception occurs when the algorithm is used to adapt a transversal equalizer in the presence of additive narrowband interference. In the latter case, the steady-state LMS behavior does not correspond to the Wiener filter: its mean weights are different from the Wiener weights, and its mean squared error performance may be significantly better than the Wiener performance. Starting from the Butterweck expansion of the weight update equation, we derive a recursive approximation for the mean of the LMS weight vector in steady-state. The analytical approximation is good for all step-sizes where the expansion converges, as supported by the simulation results.

Index Term— adaptive equalization, iterative analysis, steady-state analysis, sinusoidal interference

1. INTRODUCTION

The least mean square (LMS) algorithm [1] is perhaps the most popular adaptive algorithm utilized today due to its simplicity and robustness. The LMS algorithm, in the vast majority of its applications, tends to the corresponding Wiener filter as the step-size is made smaller. Conversely, LMS performance worsens for larger step-sizes as the adapted weights vary more around the Wiener solution.

In adaptive transversal equalization, however, the LMS algorithm has been observed to contradict this widely expected behavior. This "non-Wiener" phenomenon occurs when the LMS adaptive transversal equalizer is utilized in an environment with narrowband interference. The mean of the large step-size LMS weights can be far removed from the expected Wiener solution.

This equalization-based interference mitigation technique was proposed by North et al. [2]. The latter also reported that the LMS equalizer with large step-size outperforms the fixed Wiener equalizer in terms of bit-error rate. That observation led to follow-up papers [3, 4] for estimating mean squared error performance and its bound, respectively. Towards an explanation of the phenomenon, Beex and Zeidler [5, 6] hypothesized that the LMS equalizer is tracking an underlying optimal time-varying Wiener filter. Lastly, Batra et al. [7] showed the slow convergence of this adaptation, even when it is operating with large step-sizes.

The potential for a shift in the mean of the LMS weights, away from the (time-invariant) Wiener weights, was noted earlier but not explicitly analyzed [6]. In this paper, we analytically derive the displacement of the mean LMS weights from the Wiener weights and provide an iterative expression leading to the mean of the LMS weights in steady state. The analysis starts with Butterweck's iterative expansion of the LMS update equation [8]. This expansion is attractive as each term can be seen as a linear time-invariant state update equation. We determine a recursive expression for the mean of the steady-state LMS weights.

The rest of this paper is organized as follows. Section 2 establishes the problem of equalization in an environment with narrowband interference and presents the alternative view of the problem that renders the derivation tractable. Section 3 presents the Butterweck iterative expansion and the iterative solution to the steady-state mean. Section 4 provides several observations on the analytical result, including a comparison to simulation results. Concluding remarks are given in Section 5.

2. LMS ADAPTIVE TRANSVERSAL EQUALIZER WITH NARROWBAND INTERFERENCE

Fig. 1 depicts the system diagram for the problem of adaptive equalization in an environment with narrowband interference. To isolate the system behavior due to the narrowband interference, we use a simplified version of the conventional equalization problem. The equalizer is fixed in the training mode. The channel is assumed to be ideal, i.e., it causes no intersymbol interference. Furthermore, the additive noise is assumed to be negligible relative to the signal and thus omitted from the analysis.

The transmitted signal x_n is complex valued and can be modeled as a white zero-mean wide-sense-stationary process with power σ_x^2. For the narrowband interference i_n, we consider the limiting case, a complex sinusoidal process:

$$i_n = \sigma_i \exp[j(\omega n + \phi)] \tag{1}$$

This process has power σ_i^2 and frequency ω. The phase ϕ is randomly drawn from $[0, 2\pi)$ but is fixed in each

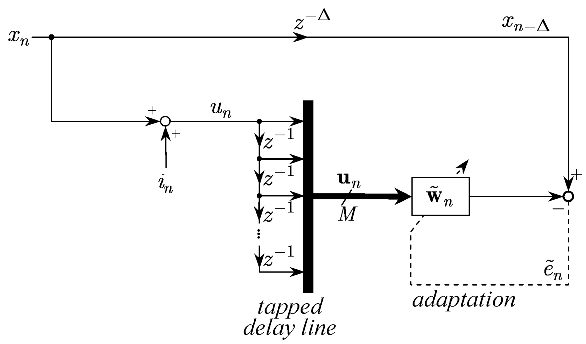

Fig. 1. Adaptive transversal equalization problem in a narrowband interference environment.

realization. The two processes, x_n and i_n, are uncorrelated.

The equalizer receives $u_n = x_n + i_n$ and forms the input vector $\mathbf{u}_n = [u_n \quad u_{n-1} \quad \cdots \quad u_{n-M+1}]^T$, where $(\bullet)^T$ is the transpose operator. We also denote the signal components of $\mathbf{u}_n$ as $\mathbf{x}_n$ and $\mathbf{i}_n$. The equalizer also gets the desired signal $x_{n-\Delta}$, a delayed version of the transmitted signal. The delay Δ must be chosen to be less than M.

The weight vector $\tilde{\mathbf{w}}_n$ is adapted using the LMS algorithm

$$\tilde{\mathbf{w}}_{n+1} = \tilde{\mathbf{w}}_n + \mu \mathbf{u}_n \tilde{e}_n^* \tag{2}$$

with step-size μ, driven by the error signal

$$\tilde{e}_n = x_{n-\Delta} - \tilde{\mathbf{w}}_n^H \mathbf{u}_n \tag{3}$$

Here, $(\bullet)^*$ is the complex conjugate operator, and $(\bullet)^H$ is the Hermitian (conjugate) transpose operator. Combining (2) and (3) yields

$$\tilde{\mathbf{w}}_{n+1} = \left(\mathbf{I} - \mu \mathbf{u}_n \mathbf{u}_n^H\right)\tilde{\mathbf{w}}_n + \mu \mathbf{u}_n x_{n-\Delta}^* \tag{4}$$

where $\mathbf{I}$ is the $(M \times M)$ identity matrix.

To make our iterative analysis tractable, we rewrite the LMS adaptation by redefining the adapted weights as

$$\mathbf{w}_n = \mathbf{p}_x - \tilde{\mathbf{w}}_n \tag{5}$$

where $\mathbf{p}_x$ is the $(\Delta + 1)$-st column of $\mathbf{I}$. The vector $\mathbf{p}_x$ arises from the fact that $E\{\mathbf{u}_n x_{n-\Delta}^*\} \triangleq \sigma_x^2 \mathbf{p}_x$. The redefined weights are updated by

$$\mathbf{w}_{n+1} = \left(\mathbf{I} - \mu \mathbf{u}_n \mathbf{u}_n^H\right)\mathbf{w}_n + \mu \mathbf{u}_n i_{n-\Delta}^* \tag{6}$$

In other words, the equalization problem is turned into the equivalent problem of estimating the interference component in $u_{n-\Delta}$.

3. EQUALIZER MEAN WEIGHT ANALYSIS

Our goal is to evaluate $\overline{\mathbf{w}} \triangleq E\{\mathbf{w}_n\}$, the steady-state mean of the weights updated by (6). Expanding $\mathbf{w}_n = \Sigma_{k=0}^{\infty} \mathbf{v}_{k,n}$, Butterweck [8] rewrites (6) in an iterative form, starting with the zero-order solution

$$\mathbf{v}_{0,n+1} = (\mathbf{I} - \mu \mathbf{R})\mathbf{v}_{0,n} + \mu \mathbf{u}_n i_{n-\Delta}^* \tag{7}$$

and the higher-order correction terms for $k > 0$

$$\mathbf{v}_{k,n+1} = (\mathbf{I} - \mu \mathbf{R})\mathbf{v}_{k,n} + \mu\left(\mathbf{R} - \mathbf{u}_n \mathbf{u}_n^H\right)\mathbf{v}_{k-1,n} \tag{8}$$

with $\mathbf{R}$ the input autocorrelation matrix. This matrix is well-defined for our problem and equals

$$\mathbf{R} = E\left\{\mathbf{u}_n \mathbf{u}_n^H\right\} = \sigma_x^2 \mathbf{I} + \sigma_i^2 \mathbf{e}\mathbf{e}^H \tag{9}$$

where $\mathbf{e} = [1 \quad e^{-j\omega} \quad \cdots \quad e^{-j\omega(M-1)}]^T$.

Defining $\overline{\mathbf{v}}_k \triangleq E\{\mathbf{v}_{k,n}\}$, the LMS weight expansion leads to $\overline{\mathbf{w}} = \Sigma_{k=0}^{\infty} \overline{\mathbf{v}}_k$. We evaluated the first few terms of $\overline{\mathbf{v}}_k$ and iteratively extended these results to the higher-order terms. The expected value of the zero-order solution in steady state is the Wiener solution to the problem:

$$\overline{\mathbf{v}}_0 = E\left\{\mu \sum_{p=0}^{\infty} (\mathbf{I} - \mu \mathbf{R})^p \, \mathbf{u}_{n-p} i_{n-\Delta-p}^*\right\} = \mathbf{R}^{-1}\mathbf{p} \tag{10}$$

where

$$\mathbf{p} \triangleq E\left\{\mathbf{u}_n i_{n-\Delta}^*\right\} = \sigma_i^2 \mathbf{e} e^{j\omega\Delta} \tag{11}$$

For $k > 0$, we have

$$\begin{aligned} \overline{\mathbf{v}}_k &= \mu \sum_{p=0}^{\infty} (\mathbf{I} - \mu \mathbf{R})^p \, E\left\{\left(\mathbf{R} - \mathbf{u}_{n-p}\mathbf{u}_{n-p}^H\right)\mathbf{v}_{k-1,n-p}\right\} \\ &= \mathbf{R}^{-1} E\left\{\left(\mathbf{R} - \mathbf{u}_n \mathbf{u}_n^H\right)\mathbf{v}_{k-1,n}\right\} \end{aligned} \tag{12}$$

Evaluating the expected value in (12) requires repeated expansion until no $\mathbf{v}_{k,n}$ term remains. We omit the detailed derivation here as it is too extensive for the space available. However, we will mention the two key assumptions that were asserted in that derivation. First, the $(\mathbf{R} - \mathbf{u}_n \mathbf{u}_n^H)$ term in (12) can be decomposed as follows

$$\begin{aligned} &\left(\mathbf{R} - \mathbf{u}_n \mathbf{u}_n^H\right) \\ &= \left(\mathbf{R}_\mathbf{x} - \mathbf{x}_n \mathbf{x}_n^H\right) + \left(\mathbf{R}_\mathbf{i} - \mathbf{i}_n \mathbf{i}_n^H\right) - \mathbf{x}_n \mathbf{i}_n^H - \mathbf{i}_n \mathbf{x}_n^H \end{aligned} \tag{13}$$

For the strong narrowband interference case, only one of the cross terms in each $(\mathbf{R} - \mathbf{u}_n \mathbf{u}_n^H)$ term in the expanded $\overline{\mathbf{v}}_k$ expression is a major contributor when evaluating the expected value in (12). The other terms are small and are eliminated in the derivation. Secondly, we assumed that

$$E\{\mathbf{x}_n \mathbf{x}_n^H \mathbf{Q} \mathbf{x}_m \mathbf{x}_m^H\} \approx E\{\mathbf{x}_n \mathbf{x}_n^H\}\mathbf{Q}E\{\mathbf{x}_m \mathbf{x}_m^H\} \tag{14}$$

for $n \neq m$ and $\mathbf{Q} = (\mathbf{I} - \mu \mathbf{R})^p$ with $p \geq 0$.

Using the above assumptions, we found that all odd $\overline{\mathbf{v}}_k$ are negligibly small. The first two even correction terms are shown here. The estimated mean of $\mathbf{v}_{2,n}$ is

$$\overline{\mathbf{v}}_2 \approx \mu \sigma_x^2 \mathbf{R}^{-1} \sum_{m=1}^{M-1} \mathbf{Z}^m \mathbf{R}^{-1} \mathbf{R}_{i,-m} (\mathbf{I} - \mu \mathbf{R})^{m-1} \mathbf{p} \tag{15}$$

and the mean of $\mathbf{v}_{4,n}$ is shown in (16) below. The matrix $\mathbf{Z}$

is a shift matrix, which has ones on the diagonal immediately below the main diagonal and zeros elsewhere, and introduced to express cross-correlation matrices:

$$E\left\{\mathbf{x}_n\mathbf{x}_{n-m}^H\right\} = \sigma_x^2\mathbf{Z}^m \quad m = 1 : M-1 \tag{17}$$

Also, the cross-correlation matrix $\mathbf{R}_{i,m}$ is defined as

$$\mathbf{R}_{i,m} \triangleq E\left\{\mathbf{i}_n\mathbf{i}_{n-m}^H\right\} = \sigma_i^2\mathbf{e}\mathbf{e}^H e^{j\omega m} \tag{18}$$

Based on (10), (15), and (16), we deduce the recursive estimate for the LMS mean weight vector to be

$$\bar{\mathbf{w}} \approx \mathbf{R}^{-1}\sum_{l=0}^{\infty}\mathbf{A}_l\mathbf{p} \tag{19}$$

where the matrix $\mathbf{A}_l$ is derived from $\bar{\mathbf{v}}_{2l}$ as

$$\mathbf{A}_l = \mu\sigma_x^2\sum_{m=1}^{M-1}\mathbf{Z}^m\mathbf{R}^{-1}\mathbf{A}_{l-1}\mathbf{R}_{i,-m}\left(\mathbf{I}-\mu\mathbf{R}\right)^{m-1} \tag{20}$$

for $l > 0$, with $\mathbf{A}_0 = \mathbf{I}$.

4. NUMERICAL ILLUSTRATIONS

We illustrate the weight behavior of the LMS algorithm using a fixed structure with $M = 7$ and $\Delta = 3$. Also, the complex sinusoidal process with fixed $\omega = 0.2\pi$ is used in this section. The mean weight behavior is studied as a function of the remaining two parameters: the step-size μ and the interference-to-signal ratio (ISR) σ_i^2 / σ_x^2.

First, the derived analytical mean of the LMS weights is compared against the simulated instantaneous weights (20,000 samples in steady state) on the complex plane, as shown in Fig. 2 together with the corresponding Wiener weights. The ISR is fixed to 20 dB, and the step-size is set to $\mu = \lambda_{\max}^{-1}$. Here, $\lambda_{\max} \triangleq \sigma_x^2 + M\sigma_i^2$ is the maximum eigenvalue of $\mathbf{R}$. The simulation employs a quadrature phase shift-keyed signal as x_n (i.e., a sample of x_n is randomly drawn from $\{\pm 1 \pm j\}$). The analysis and simulation agree well, forming a spiral, and clearly differ from the Wiener solution. Also, both the mean weights and the Wiener weights lie on the radial lines through $\mathbf{e}e^{j\omega\Delta}$. This condition is required for the filter to correctly estimate $i_{n-\Delta}$. Therefore, we concentrate on the magnitude of the weights for the rest of this section. Note that the magnitudes of the Wiener weights are equal.

Furthermore, we check the convergence behavior of the series in (19). Fig. 3 illustrates how the magnitude of each element of $\bar{\mathbf{w}}$ in Fig. 2 converges as more terms are added to the summation. All the values of $\bar{\mathbf{w}}$ illustrated in this section are estimated by computing the summation up to $l = L$, which satisfies $\left\|\mathbf{R}^{-1}\mathbf{A}_L\mathbf{p}\right\| < 10^{-6}$.

The second study examines $|\bar{\mathbf{w}}|$ as a function of μ, with

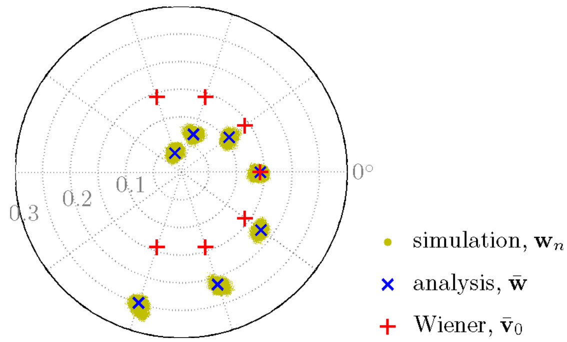

Fig. 2. LMS steady-state weight behavior (20,000 samples), analytical mean, and corresponding Wiener weights on $\mathbb{C}$ -plane.

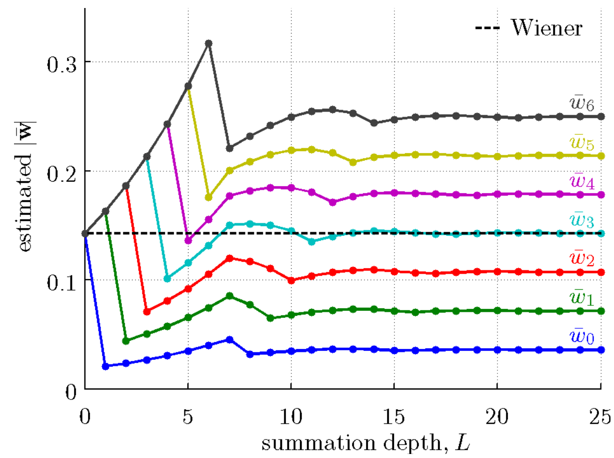

Fig. 3. Convergence behavior of (19).

the result shown in Fig. 4. The ISR is maintained at 20 dB, while the step-size is varied from 0 to $2\lambda_{\max}^{-1}$. The departure of the mean weights from the Wiener solution becomes prevalent around $\mu = 0.03\lambda_{\max}^{-1}$. Also, all the neighboring weight pairs are roughly equal distance apart for all μ.

The expression for the theoretical mean weight in (19) is found to diverge for large μ. While the exact condition for the divergence has not been determined, it stems from the sequence $\left\{\mathbf{v}_{k,n}\right\}$ diverging as $k \to \infty$ for all n in steady state. The LMS algorithm remains stable past the analysis breakdown point ($\mu \approx 1.28\lambda_{\max}^{-1}$) as illustrated by the simulation results in Fig. 4. All converged analysis results are in excellent agreement with the simulation. Lastly, the LMS algorithm diverges at $\mu \approx 1.93\lambda_{\max}^{-1}$.

The final study looks at the mean weight behavior as a function of the ISR as shown in Fig. 5. The step-size μ in this illustration is tied to the ISR as $\mu \triangleq \lambda_{\max}^{-1}\left(\sigma_x^2, \sigma_i^2\right)$. This result clearly illustrates that the non-Wiener mean weight behavior is caused by the narrowband interference. When the interference is weak the LMS mean weights behave as expected, that is, they follow the Wiener weights.

$$\bar{\mathbf{v}}_4 \approx \mu^2\sigma_x^4\mathbf{R}^{-1}\sum_{m=1}^{M-1}\mathbf{Z}^m\mathbf{R}^{-1}\sum_{l=1}^{M-1}\mathbf{Z}^l\mathbf{R}^{-1}\mathbf{R}_{i,-l}\left(\mathbf{I}-\mu\mathbf{R}\right)^{l-1}\mathbf{R}_{i,-m}\left(\mathbf{I}-\mu\mathbf{R}\right)^{m-1}\mathbf{p} \tag{16}$$

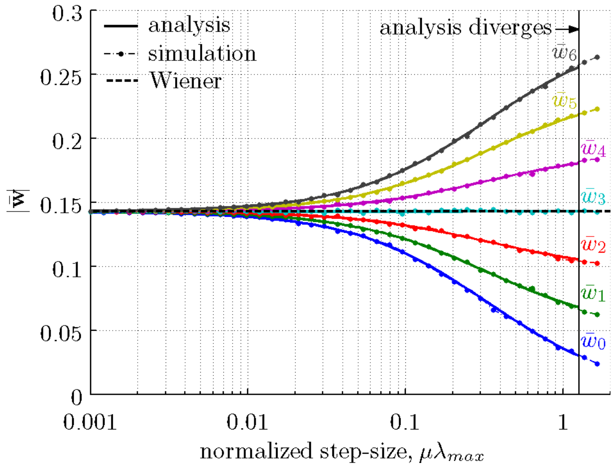

Fig. 4. Magnitudes of the mean LMS weights as functions of step-size.

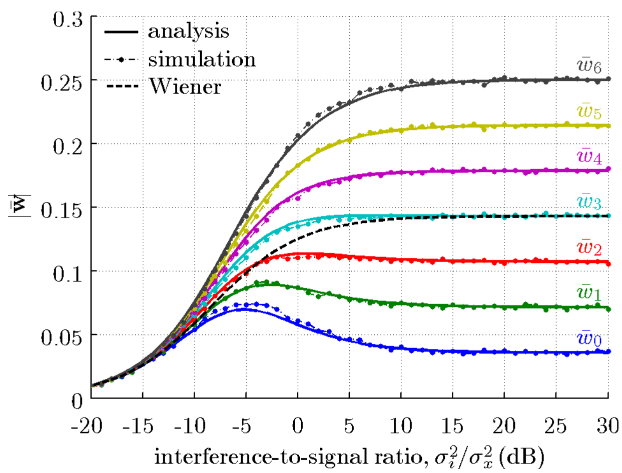

Fig. 5. Magnitude of mean weights as a function of ISR. Step-size is varied as a function of ISR.

As the interference becomes stronger, the LMS mean weights begin to move away from the Wiener weights to the spiral formation that is prevalently illustrated by Fig. 2. The analysis results show a slight deviation in the range about 0 dB ISR, presumably due to the assumptions made.

This study of the mean weight behavior for equalization in narrowband interference is expected to subsequently lead to new insights in analyzing the mean squared error (MSE). As observed in Fig. 6, the MSE of the equalizer with the mean $\bar{\mathbf{w}}$ as its fixed weights is larger than when the Wiener weights are used while the LMS equalizer operates around $\bar{\mathbf{w}}$ and produces better MSE than the Wiener weights. This suggests – as hypothesized by Beex and Zeidler [5, 6] – that the LMS instantaneous weight variation is responsible for the reduction, from the fixed $\bar{\mathbf{w}}$ MSE to the LMS MSE.

5. CONCLUSION

We have provided an analysis that shows that – and how – the mean LMS weights are different from the corresponding

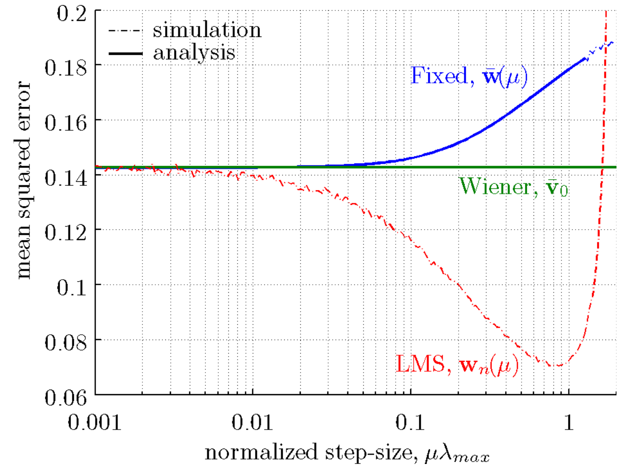

Fig. 6. Mean squared error as a function of normalized step-size for the LMS, fixed, and Wiener filters, corresponding to Fig. 4.

Wiener weights in an adaptive equalizer application for mitigating narrowband interference. The mean LMS weight values in steady state are approximated in a recursive form, based on the Butterweck expansion of the weight update equation. Excellent correspondence between the analytical and simulation results for the mean weight vector is observed over nearly the entire range of stable step-sizes.

6. REFERENCES

[1] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River, NJ: Prentice Hall, 2002.

[2] R. C. North, R. A. Axford, and J. R. Zeidler, "The performance of adaptive equalization for digital communication systems corrupted by interference," *27th Asilomar Conf.*, 1993, pp. 1548-1553.

[3] M. Reuter and J. R. Zeidler, "Nonlinear effects in LMS adaptive equalizers," *IEEE Trans. Signal Process.,* vol. 47, no. 6, pp. 1570-1579, 1999.

[4] K. J. Quirk, L. B. Milstein, and J. R. Zeidler, "A performance bound for the LMS estimator," *IEEE Trans. Inf. Theory,* vol. 46, no. 3, pp. 1150-1158, 2000.

[5] A. A. Beex and J. R. Zeidler, "Steady-state dynamic weight behavior in (N)LMS adaptive filters," in *Least-Mean-Square Adaptive Filters*, S. Haykin and B. Widrow, eds. Hoboken: John Wiley, 2003, pp. 335-443.

[6] A. A. Beex and J. R. Zeidler, "Non-linear effects in interference contaminated adaptive equalization," *IASTED Int'l Conf. SPPRA*, 2002, pp. 474-479.

[7] A. Batra, T. Ikuma, J. R. Zeidler, A. A. Beex, and J. G. Proakis, "Mitigation of unknown narrowband interference using instantaneous error updates," *37th Asilomar Conf.*, 2004, pp. 115-119.

[8] H. J. Butterweck, "Iterative analysis of the steady-state weight fluctuations in LMS-type adaptive filters," *IEEE Trans. Signal Process.,* vol. 47, no. 9, pp. 2558-2561, 1999.