

Non-Linear Effects in Interference Contaminated Adaptive Equalization

A. A. (Louis) Beex
DSPRL – ECE 0111
Virginia Tech
Blacksburg, VA 24061-0111, USA

& James R. Zeidler
Communications & Information Systems, 28505
SPAWAR Systems Center
San Diego, CA 92152, USA

ABSTRACT

The presence of non-linear effects in adaptive filtering was observed previously in a wide sense stationary equalization scenario with narrowband interference. The influence of a variety of parameters on the magnitude of such non-linear effects was also investigated. While the adaptive filter appears to exhibit behavior that one would expect in a wide sense stationary scenario, we show here that the adaptive filter is actually tracking a time-varying target solution. It is the latter property that is responsible for the performance improvement of the adaptive filter over the performance of the corresponding Wiener filter.

KEY WORDS: non-linear effects, NLMS, time-varying Wiener filter, multi-channel Wiener filter, multi-channel adaptive filter, adaptive equalization.

1. INTRODUCTION

In a wide-sense stationary equalization scenario, subject to narrowband interference, it was shown that a normalized least mean squares (NLMS) adaptive filter can sometimes provide better performance than the corresponding Wiener filter [1, 2]. A transfer function analysis was used to describe the nonlinear effects in terms of signal-to-noise ratio (SNR), signal-to-interference ratio (SIR), interference bandwidth, and NLMS stepsize. The latter was an important parameter that could take on large values, i.e. approaching unity.

It was shown [1,2] that the nonlinear effects originate from the effects of the past values of the error signal, which are fed back to the adaptive filter weight update algorithm. It was further shown that the performance was bounded by that of a nonlinear decision directed feedback equalizer (DFE), which directly utilized the past values of the estimates of the desired signal to update the filter coefficients. The mechanism by which the weight updates achieve the performance enhancements was not revealed.

In this paper we will analyze the dynamic behavior of the filter weights for the same interference contaminated equalizer applications considered earlier [1,2] and show that the improvements in performance result from time-varying dynamics in the weight updates. In addition we

illustrate that the performance of the interference contaminated adaptive equalizer is bounded by a two channel Wiener filter that utilizes the conventional filter channel, containing samples of the input process, and a second channel containing samples of the interference. This is an intuitively satisfying lower bound compared with the previous DFE results since the performance of the DFE is defined by the samples of the past estimates of the desired response which in turn consists of past values of both the training sequence and the interference. Since the training sequence consists of samples of an uncorrelated, i.i.d. process, there is no information in this component of the desired signal and the nonlinear response is defined by the past samples of the interference. An approximation to the two-channel Wiener filter considered here could be implemented using inputs from an array focused on the interfering signal.

It is first shown that this alternative two-channel Wiener filter provides the same performance as the ideal Wiener filter. Next, experimental evidence is provided to show that the performance improvement of the adaptive NLMS filter over the Wiener filter is the result of the time-varying nature of the adaptive filter. We then use the ideal version of the two-channel Wiener filter to derive a time-varying Wiener filter solution that forms the target for the adaptive NLMS filter. While the adaptive filter can realize at best a fraction of the performance improvement associated with the time-varying Wiener filter solution, it is sufficient to outperform the time-invariant Wiener filter for the wide-sense stationary scenario at hand.

2. ICAEQ SCENARIO

In interference contaminated adaptive equalization (ICAEQ), the wideband (white) QPSK (quadrature phase shift keying) signal q_n is contaminated by a narrowband interference i_n , a first order autoregressive (AR) process determined by a single pole at p (the pole location is at $0.9999\exp(j\pi/3)$ for all results shown). The interference-contaminated signal y_n is also contaminated by additive white Gaussian noise (AWGN) n_n to yield the observable process r_n . This case is reflected in Fig. 1.

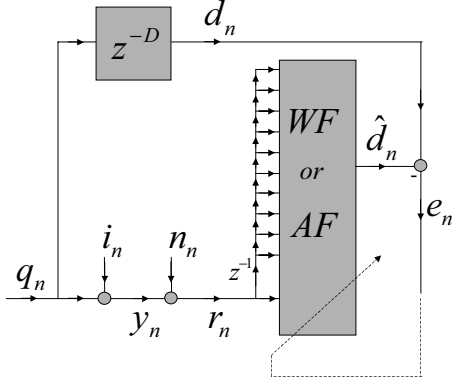


Fig. 1: Conventional ICAEQ Scenario.

The component processes of the measurable process are mutually independent.

$$r_n = q_n + i_n + n_n \quad (1)$$

The latter forms the input or reference process u_n to either a Wiener filter (WF) or an adaptive filter (AF). Fig. 1 reflects the ICAEQ scenario in the learning or training mode, i.e. where a known learning, or desired, sequence d_n is provided. For an SNR of 20 dB and an SIR of -20 dB significant nonlinear effects have been observed for NLMS stepsize $\bar{\mu} = 0.8$ [1, 2]. Note that the AWGN is weak relative to the signal, and that the signal is weak relative to the interference. Consequently, away from zero-crossings, the observed reference signal behaves largely like the interference signal.

3. 2-CHANNEL WF ALTERNATIVE

When the interference signal is the strongest component in the measurable reference signal, knowledge of the interference signal would produce vastly improved performance of the ICAEQ. This hypothetical situation is reflected in the ideal 2-channel version of the ICAEQ, as reflected in Fig. 2.

The input u_n , to the WF or AF, now consists of two components. Each of these components consists of a delay line fed by the reference signal and the interference signal respectively. The vector of inputs to the WF or AF is then defined as follows.

$$\mathbf{u}_n = \begin{bmatrix} \mathbf{i}_n \\ \mathbf{r}_n \end{bmatrix} \quad (2)$$

where the reference component contains $2D+1$ tap values, defined by the delay D at midpoint, and generally takes on the following form that is centered on r_{n-D} .

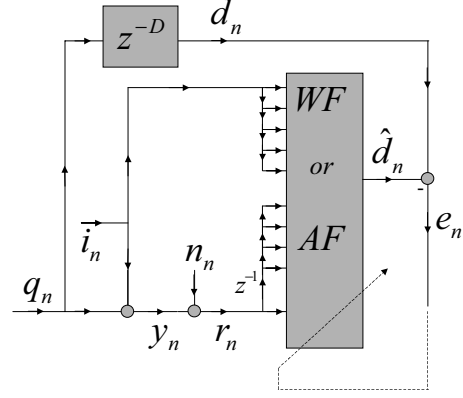


Fig. 2: Ideal 2-channel ICAEQ filtering.

$$\mathbf{r}_n = \begin{bmatrix} r_n \\ r_{n-1} \\ \vdots \\ r_{n-2D} \end{bmatrix} \quad (3)$$

The interference component can be thought of as being of a similar form, i.e. centered on the value i_{n-D} . Note from Figs. 1 and 2 that we desire to find estimates for q_{n-D} . From (1) we note that subtraction of i_{n-D} from r_{n-D} yields the QPSK signal contaminated only by weak AWGN, i.e. a very good estimate. Additional values of i_n do not yield further improvements in the estimate, so that it is actually sufficient to choose $\mathbf{i}_n = i_{n-D}$.

As the ideal situation reflected in Fig. 2 is not practically realizable – since i_n is not measurable – it only provides a bound on practically attainable performance. In Fig. 3 a realistic approximation to the ideal 2-channel ICAEQ case of Fig. 2 is presented.

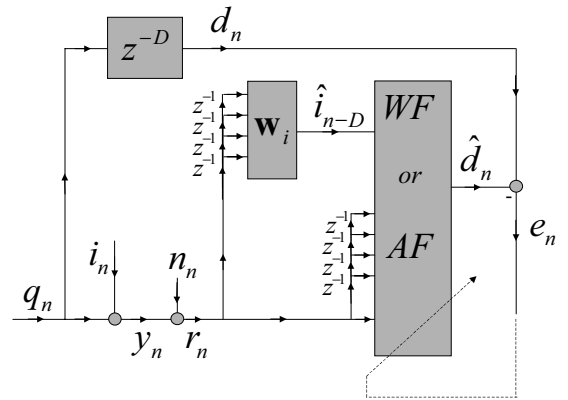


Fig. 3: Realistic 2-channel ICAEQ filter.

For the scenario under investigation, the reference signal behaves very much like the interference itself so that it seems logical to use the measurable reference input to derive a good estimate of the desired interference signal. In Fig. 3, therefore, we have replaced the unobservable interference channel input in Fig. 2 by an estimate of the interference, \hat{i}_{n-D} , based on the measurable r_n . Note that the FIR filter generating the estimate \hat{i}_{n-D} has as its input the reference input of the WF or AF.

4. WIENER FILTER SOLUTIONS

The Wiener filter approaches to each of the above cases can be captured by the following general formulation for a 2-channel WF, by using definitions appropriate to each.

In the ideal 2-channel ICAEQ case, the WF weights \mathbf{w}_{ir} operate on the input vector \mathbf{u}_n of (2) to generate the estimate $\hat{d}_{n-D} = \mathbf{w}_{ir}^H \mathbf{u}_n$. The design of this WF follows from the following Wiener-Hopf equation.

$$E \left\{ \begin{bmatrix} \mathbf{i}_n \\ \mathbf{r}_n \end{bmatrix} \begin{bmatrix} \mathbf{i}_n & \mathbf{r}_n \end{bmatrix}^H \right\} \mathbf{w}_{ir} = E \left\{ \begin{bmatrix} \mathbf{i}_n \\ \mathbf{r}_n \end{bmatrix} d_{n-D}^H \right\} \quad (4)$$

Recalling that the component processes q_n and n_n are white and zero-mean, that i_n is AR(1), and that all three are mutually independent, the component matrices needed in (4) are recognized as auto- or cross-covariances of ARMA (auto-regressive moving-average) processes. The latter can be evaluated using AR [3] and Sylvester matrix based techniques [4] respectively. The performance of the WF is then given by

$$MMSE_{w(N,M)} = E \left\{ |d_{n-D}^H|^2 \right\} - E \left\{ \begin{bmatrix} \mathbf{i}_n \\ \mathbf{r}_n \end{bmatrix} d_{n-D}^H \right\} \mathbf{w}_{ir}^H \quad (5)$$

The design and performance for the conventional ICAEQ scenario of Fig. 1 are obtained by deleting the partition corresponding to \mathbf{i}_n .

The design and performance for the realistic 2-channel ICAEQ case are obtained by replacing \mathbf{i}_n in (4) and (5) by \hat{i}_{n-D} , where the latter is obtained by designing the conventional WF for finding an estimate for i_{n-D} from the input $\mathbf{u}_n = \mathbf{r}_n$. The resulting WF is indicated as \mathbf{w}_i in Fig. 3. Note in the latter that the auto- and cross-covariances between \hat{i}_{n-D} and the component processes of r_n are again ARMA auto- and cross-covariances, so

that the design and performance can be evaluated from appropriate definitions in (4) and (5).

Another limit to performance is arrived at as follows. For a distortionless estimator response the weight operating on q_{n-D} , i.e. on r_{n-D} , must equal one. There is no information about q_{n-D} in $r_{n-l}, l \neq D$ because of the whiteness of the QPSK process. Consequently, $r_{n-l}, l \neq D$ must be used to estimate i_{n-D} so that it can be subtracted from r_{n-D} , thus rendering a good estimate of q_{n-D} . As i_n is an AR(1) process, its ideal estimate is derived from i_{n-1} and produces an estimation variance equal to that of its innovation process. Subtracting i_n from r_n , and adding the innovation variance to the white noise variance, leads to another indicator for the possible performance that we can expect. We will refer to the latter as the ideal interference predictor case.

Note that the scenarios reflected above are wide-sense stationary, so that all resulting WF solutions correspond to linear time-invariant (LTI) filters.

5. NLMS RESULTS

Both the conventional and two-channel adaptive filter (AF) use the NLMS algorithm, implemented as follows.

$$\begin{aligned} e_n &= d_{n-D} - \mathbf{w}_n^H \mathbf{u}_n \\ \mathbf{w}_{n+1} &= \mathbf{w}_n + \bar{\mu} \frac{e_n^*}{\mathbf{u}_n^H \mathbf{u}_n} \mathbf{u}_n \end{aligned} \quad (6)$$

The difference between the two cases lies in the definition that is used for the input vector; as in (3) for the conventional case and as in (2) for the two-channel case. In Fig. 4 we show the Wiener (WF) and adaptive filter (AF) performance results.

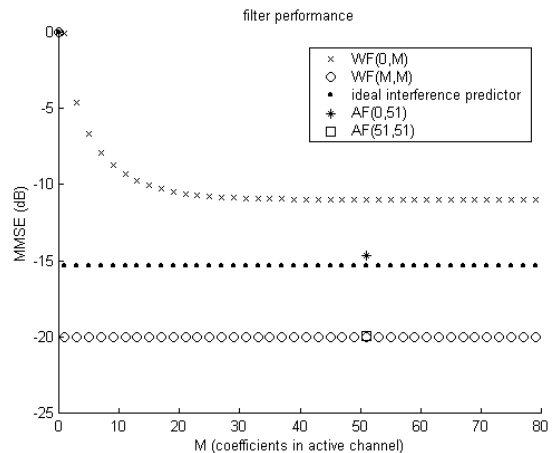


Fig. 4: Performance for Various Filters and Orders.

The NLMS AF result for $D=25$ (for $\bar{\mu} = 0.8$), in the conventional ICAEQ scenario, is indicated with *. WF performance for the conventional ICAEQ scenario is indicated by x's, ideal 2-channel ICAEQ WF performance is indicated by o's, and WF performance using the ideal interference predictor innovation variance is indicated by •'s. We note that the conventional NLMS AF outperforms the (optimal) conventional WF, thereby demonstrating the nonlinear effect in adaptive filtering. The AF performance comes close to the performance arrived at using the ideal interference predictor idea. The ideal 2-channel ICAEQ setup outperforms all other approaches and serves as the absolute upper bound on performance.

The performance of the realistic 2-channel ICAEQ turns out to be exactly that of the conventional WF (x's). Note that in both of the latter cases exactly the same (measurable) inputs are used to ultimately find the best, in MMSE (minimum mean squared error) sense, estimate of the desired signal.

Figure 5 shows the AF weights - for the conventional scenario - for 300 consecutive time indices, together with the time-invariant (TI) WF weights (o's), for two different process realizations (real part shown, imaginary part is similar).

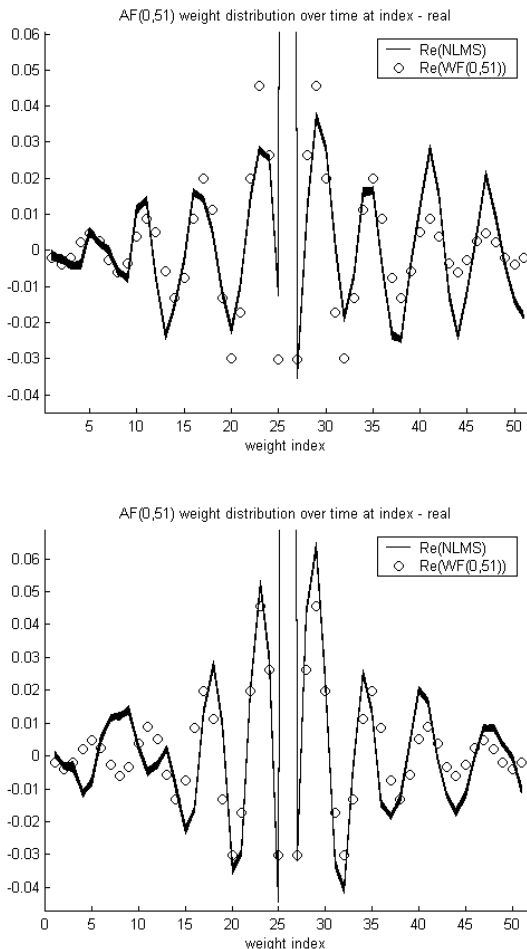


Fig. 5: NLMS Weight Behavior over 300 Samples.

The thickness of the lines reflects that the NLMS weights changed slightly over 300 iterations. Observing this behavior in either of the plots in Fig. 5 suggests that a TI weight solution exists. Observing the marked differences between the two plots in Fig. 5, however, suggests that the weight behavior is time-varying (TV). To further confirm this TV weight hypothesis, the experiment is repeated for 3000 consecutive time indices. The results of this experiment are shown in Fig. 6.

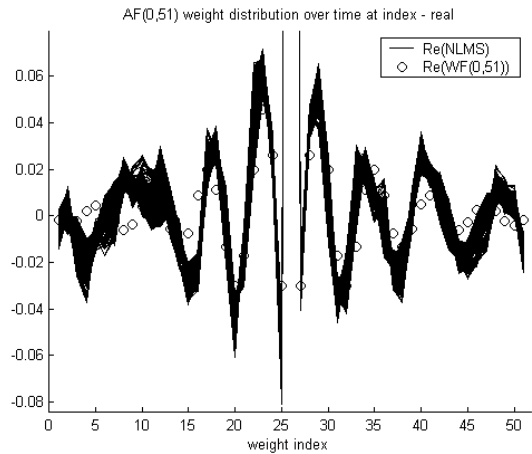


Fig. 6: NLMS Weight Behavior over 3000 Samples.

We observe that the weights keep changing and that the resulting manifold of solutions encompasses an increasingly significant portion of the conventional TI WF solution. The weight behavior versus time, over this longer time horizon, is shown in Fig. 7.

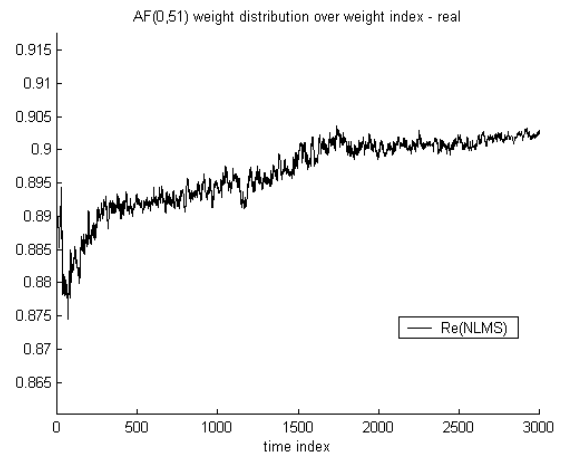


Fig. 7a: Midpoint Coefficient Behavior.

The midpoint coefficient in Fig. 7a operates on r_{n-D} , and would therefore be expected to have a value of one to yield a distortionless estimate. Of course it is generally possible to achieve a further reduction in MSE at the expense of some distortion (bias). In fact, even in the TI WF the solution reflects a compromise between signal distortion and noise contribution. Figure 7b shows the behavior of the weights other than the one at midpoint. All weights reflect slowly drifting behavior.

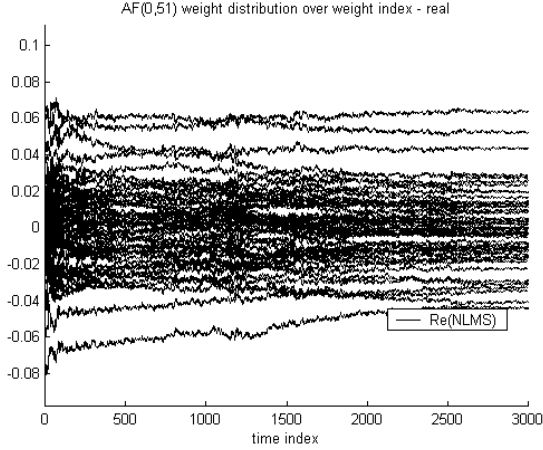


Fig. 7b: Behavior of Coefficients away from Midpoint.

These results lead to the hypothesis that the NLMS adaptive filter (using $\bar{\mu} = 0.8$) achieves its performance advantage over the conventional TI WF as a result of its time-varying weight behavior.

6. TV WF SOLUTION

The WF weight vector solution \mathbf{w}_{ir} , for the ideal 2-channel ICAEQ scenario, together with the corresponding AF weight vector (using $\bar{\mu} = 0.05$) produces the performance indicated in Fig. 4 by the intersecting \square and \circ . The corresponding weight vector itself is shown in Fig. 8. The thinness of the lines indicates that over 3000 samples there was no visible change in values, i.e. the AF behaves as expected when a TI weight vector solution exists that corresponds to the optimal performance.

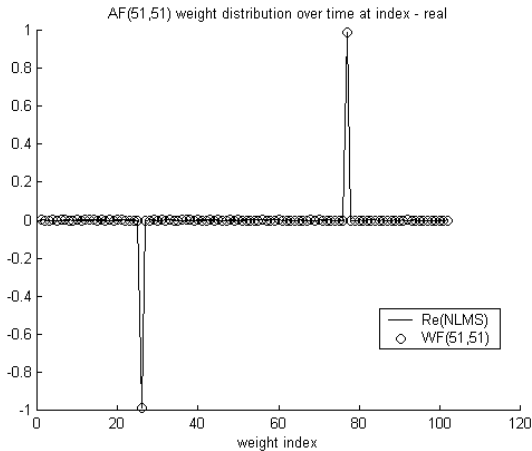


Fig. 8: Ideal 2-channel ICAEQ WF Weight Vector.

The result in Fig. 8 confirms that the only interference sample of interest is i_{n-D} , i.e. at the midpoint; its value is 0.99. We therefore have the following, very good, model for the desired signal q_{n-D} in terms of input values used in the ideal 2-channel ICAEQ case.

$$q_{n-D} = \begin{bmatrix} -0.99 & 0 & 0.99 & 0 \end{bmatrix} \begin{bmatrix} i_{n-D} \\ r_{n-D+1} \\ r_{n-D} \\ r_{n-D-1} \end{bmatrix} + \varepsilon_{n-D} \quad (7)$$

$$= \mathbf{w}_{ir}^H \mathbf{u}_n + \varepsilon_{n-D}$$

where ε_{n-D} represents a low level of noise. For the sake of clarity we have given the situation for $D=1$; the generalization is straightforward. The model in (7), being the most apt description for the desired signal, forms the target for the adaptive filter. We know this because the AF achieves its target for small stepsize, exhibits TI weight behavior, and achieves a level of performance equal to that of the ideal 2-channel WF.

We now derive the target weight vector for the conventional ICAEQ scenario. Thereto we define sequences linking the interference and reference signals.

$$\rho_{n-D+1}^{(-1)} = \frac{i_{n-D}}{r_{n-D+1}}$$

$$\rho_{n-D}^{(0)} = \frac{i_{n-D}}{r_{n-D}} \quad (8)$$

$$\rho_{n-D-1}^{(1)} = \frac{i_{n-D}}{r_{n-D-1}}$$

The linking sequences permit us to rewrite the desired signal in (7) in terms of reference vector values only. Using the first linking sequence in (8) yields:

$$q_{n-D} = -0.99 \rho_{n-D+1}^{(-1)} r_{n-D+1} + 0.99 r_{n-D} + \varepsilon_{n-D}$$

$$= 0.99 \begin{bmatrix} -\rho_{n-D+1}^{(-1)} & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{n-D+1} \\ r_{n-D} \\ r_{n-D-1} \end{bmatrix} + \varepsilon_{n-D} \quad (9a)$$

$$= \mathbf{w}_{0r}^{(1)H} \mathbf{u}_n + \varepsilon_{n-D}$$

Using the second linking sequence in (8) produces the following alternative:

$$q_{n-D} = -0.99 \rho_{n-D}^{(0)} r_{n-D} + 0.99 r_{n-D} + \varepsilon_{n-D}$$

$$= 0.99 \begin{bmatrix} 0 & 1 - \rho_{n-D}^{(0)} & 0 \end{bmatrix} \begin{bmatrix} r_{n-D+1} \\ r_{n-D} \\ r_{n-D-1} \end{bmatrix} + \varepsilon_{n-D} \quad (9b)$$

$$= \mathbf{w}_{0r}^{(2)H} \mathbf{u}_n + \varepsilon_{n-D}$$

The last linking sequence in (8) gives a third alternative way to express the target weight vector:

$$\begin{aligned}
q_{n-D} &= -0.99\rho_{n-D-1}^{(1)}r_{n-D-1} + 0.99r_{n-D} + \varepsilon_{n-D} \\
&= 0.99\begin{bmatrix} 0 & 1 & -\rho_{n-D-1}^{(1)} \end{bmatrix} \begin{bmatrix} r_{n-D+1} \\ r_{n-D} \\ r_{n-D-1} \end{bmatrix} + \varepsilon_{n-D} \quad (9c) \\
&= \mathbf{w}_{0r}^{(3)H} \mathbf{u}_n + \varepsilon_{n-D}
\end{aligned}$$

Any of the above alternatives is equally valid, and an affine linear combination of the right-hand sides produces the following general weight vector target:

$$\mathbf{w}_{0r,n} = 0.99 \begin{bmatrix} -\alpha_1 \rho_{n-D+1}^{(-1)*} \\ 1 - \alpha_2 \rho_{n-D}^{(0)*} \\ -\alpha_3 \rho_{n-D-1}^{(1)*} \end{bmatrix} \quad (10)$$

While the original weight vector target for the ideal 2-channel ICAEQ scenario, in (7), was TI, the corresponding weight vector target for the conventional ICAEQ case, in (10), is time-varying. The time dependence is indicated explicitly. The weight vectors in (7) and (10) produce the same desired signal.

Note from the first linking sequence in (8) that

$$\begin{aligned}
\rho_{n-D+1}^{(-1)} &= \frac{i_{n-D}}{i_{n-D+1} + q_{n-D+1} + n_{n-D+1}} \\
&= \frac{i_{n-D}}{pi_{n-D} + v_{n-D+1} + q_{n-D+1} + n_{n-D+1}} \quad (11a) \\
&\cong p^{-1} + \eta_{n-D+1}^{(-1)}
\end{aligned}$$

which represents a random variation about p^{-1} . For the second linking sequence we similarly find it to represent a random variation about 1.

$$\begin{aligned}
\rho_{n-D}^{(0)} &= \frac{i_{n-D}}{i_{n-D} + q_{n-D} + n_{n-D}} \\
&\cong 1 + \eta_{n-D}^{(0)}
\end{aligned} \quad (11b)$$

For the third linking sequence in (8) we find

$$\begin{aligned}
\rho_{n-D-1}^{(1)} &= \frac{i_{n-D}}{i_{n-D-1} + q_{n-D-1} + n_{n-D-1}} \\
&= \frac{pi_{n-D-1} + v_{n-D}}{i_{n-D-1} + q_{n-D-1} + n_{n-D-1}} \quad (11c) \\
&\cong p + \eta_{n-D-1}^{(1)}
\end{aligned}$$

which represents a random variation about p . Consequently, assuming a fixed linear combination, all elements of the weight vector target in (10) represent random variations about a constant at or very near one. The approximations are valid as long as the interference

signal is not too close to one of its zero crossings, so that the interference amplitude dominates the QPSK and AWGN amplitudes.

We thus find that the weight vector target elements behave like first-order Markov processes. Depending on the statistical characteristics that describe these processes we are tracking a nonstationary reference system when applying the conventional ICAEQ scenario in a wide sense stationary environment. The resulting MSE performance is thus a compromise between noise misadjustment and lag misadjustment, and therefore generally depends on NLMS stepsize $\bar{\mu}$.

7. CONCLUSION

We have shown that the nonlinear effects of adaptive filtering in the narrowband interference contaminated equalizer scenario are associated with slowly time-varying NLMS weights. Based on the ideal 2-channel ICAEQ scenario we find an expression for the time-varying target weight vector that the conventional NLMS adaptive filter tracks. The slowly time-varying nature corresponds to the underlying model for the data. The conventional NLMS adaptive filter can thus outperform the time-invariant Wiener filter of the same filter order and is bounded by the performance of the optimal 2-channel Wiener filter.

8. ACKNOWLEDGEMENT

The work of the first author was supported in part by his National Research Council Senior Research Associateship at SPAWAR Systems Center, San Diego, CA, during his Fall 2001 sabbatical semester from Virginia Tech.

REFERENCES

- [1] M. Reuter, K. Quirk, J. Zeidler, and L. Milstein, Nonlinear effects in LMS adaptive filters, *Proc. Symp. 2000 on Adaptive Systems for Signal Processing, Communications and Control*, 141-146, Lake Louise, Alberta, October 2000.
- [2] M. Reuter and J. R. Zeidler, Nonlinear effects in LMS adaptive equalizers, *IEEE Trans. Signal Processing*, 47, 1570-1579, June 1999.
- [3] J-P. Dugré, A. A. (Louis) Beex, and L. L. Scharf, Generating Covariance Sequences and the Calculation of Quantization and Rounding Error Variances in Digital Filters, *IEEE Trans. Acoustics, Speech, and Signal Processing*, 28, 102-104, February 1980.
- [4] A. A. (Louis) Beex, Efficient Generation of ARMA Cross-Covariance Sequences, *Proc. Int'l Conf. On Acoustics, Speech, and Signal Processing*, 327-330, Tampa FL, 26-29 March 1985.