

DATA STRUCTURE AND NON-LINEAR EFFECTS IN ADAPTIVE FILTERS

A. A. (Louis) Beex¹ and James R. Zeidler²

¹DSPRL, ECE-0111, Virginia Tech, Blacksburg, VA 24061-0111, USA
beex@vt.edu

²Communications and Information Systems, SPAWAR Systems Center, San Diego, CA 92152, USA
zeidler@spawar.navy.mil

Abstract: The non-linear effects that have been observed in adaptive filtering scenarios are explained from the point of view of the structure that underlies the desired data. While the model structure used by the conventional adaptive filter is a linear combination of tapped-delay line signals, that adaptive filter model does not generally correspond to the structure that best describes the desired data one is adapting to. The nonlinear effects in adaptive noise canceling, interference contaminated adaptive equalization, and adaptive linear prediction are explained here as being the result of forcing a filter model onto an essentially different data structure. The tapped delay line model can then only be compatible with the data if the filter weights become time-varying. If the adaptation captures the time-varying weight behavior, the adaptive filter performance can approach that associated with the data structure and thereby exceed the best performance associated with the corresponding conventional Wiener filter.

1. INTRODUCTION

Non-linear effects have been observed in several adaptive filter (AF) applications, usually at larger stepsizes and involving narrowband processes. Non-linear effects in adaptive noise canceling (ANC) [1], interference contaminated adaptive equalization (AEQ) [2], and adaptive linear prediction (ALP) [3] have been reported and analyzed previously. The origin of the nonlinear effects in ANC was studied recently [4]. The explanations depend on the existence of an underlying data structure that can produce performance well in excess of the performance associated with the optimal conventional tapped-delay line structure Wiener filter, which is time-invariant in the reported scenarios. When the tapped delay line structure is used in an AF context its weights can be time-varying. The fundamental mechanism, by which the weights adapt to the data structure, is through the error feedback, as the error reflects instantaneously the discrepancy between the structure of the data and its modeled behavior. The AF may capture some of the structure that underlies the data and exhibit improved performance.

2. NLMS ADAPTATION & MODELING

The nonlinear effects in adaptive filtering were observed when using the least-mean-square (LMS) algorithm and its normalized (NLMS) form. We focus here on using the NLMS algorithm, in which the error signal is computed by subtracting the desired signal estimate (the AF output, computed as a weighted linear combination of all its inputs) from the desired signal.

$$e_n = d_n - \hat{d}_n \quad (1)$$

$$\hat{d}_n = \mathbf{w}_n^H \mathbf{u}_n \quad (2)$$

The error signal is used to correct the weight vector.

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \bar{\mu} \frac{e_n^*}{\mathbf{u}_n^H \mathbf{u}_n} \mathbf{u}_n \quad (3)$$

While NLMS can be applied as given, its properties depend on the structure, or model, for the desired data. While we know the model that underlies the estimate provided by the adaptive filter, as given in (2), most often we do not know the structure that underlies the desired data. A common assumption is that the desired signal itself has the same linear structure as that used for the filter estimate, for some ideal set of constant weights, with perhaps some added white noise.

$$d_n = \mathbf{w}^o \mathbf{u}_n + \varepsilon_n \quad (4)$$

Under these assumptions the AF weights converge to the true weights, or a neighborhood about the true weights determined by the additive white noise, and the NLMS performance – in terms of MSE – approaches the variance of the additive white noise from above (depending on the excess MSE, which vanishes as the stepsize parameter $\bar{\mu}$ vanishes). The structure that underlies the data, as assumed in (4), is not necessarily what actually generated the data. If the model assumed in (2) is correct, but its order is too low, this forces part of the data to be modeled by the additive noise and/or time-varying weight behavior. If the model assumed in (2) is correct, but the additive noise in (4) is colored, it is possible to get a MSE smaller than the variance of the additive noise due to a biased parameterization. So far we have assumed that the structure underlying the data in (4) is time-invariant (TI). If the true weight vector in (4) is time-varying (TV), the MSE consists of a tracking error component in addition to the estimation error component.

Any TV behavior of the weight in (4) is immediately

reflected in the desired signal and, therefore, in the error signal. At large stepsizes this effects an instantaneous change in the a posteriori weight vector reflecting the discrepancy between the data structure (actual) and its tapped delay line filter model (assumed). We show that nonlinear effects can arise when forcing the tapped delay line adaptive filter model onto more general data structures.

3. 2-CHANNEL NLMS

The input vector \mathbf{u}_n used in (2) usually contains tapped-delay line values based on a single reference channel. For the purpose of revealing underlying data structures we assume a more general adaptive filter structure, where the input vector contains delay line values from the reference channel as well as from the auxiliary channel, in \mathbf{r}_n and \mathbf{x}_n respectively.

$$\mathbf{u}_n = \begin{bmatrix} \mathbf{x}_n \\ \mathbf{r}_n \end{bmatrix} \quad (5)$$

This definition renders (1)-(3) into a 2-channel NLMS adaptive filter, as reflected in Fig. 1.

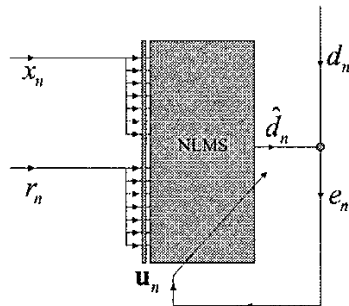


Fig. 1: Basic 2-Channel NLMS Filter.

In the conventional NLMS adaptive filter, the input vector does not contain the auxiliary vector partition directly, but some access to that data is provided in the form of the error signal (based on the desired data) which is used in the weight updates.

4. VARIOUS 2-CHANNEL SCENARIOS

We now look at the ANC, AEQ, and ALP scenarios from the 2-channel point of view. Depending on the particular scenario, the auxiliary channel is chosen to reflect the underlying structure of the data. Note that NLMS using (5) implies the data can be represented as in (4), using the 2-channel input. The latter has been verified by applying 2-channel NLMS and observing that both the weights and MSE performance approach those of the corresponding optimal TI Wiener filter.

4.1. ANC

In the conventional ANC scenario the desired signal estimate is derived from present and past values of the

reference signal. It is beneficial to also use past values of the desired signal [4]. Consequently, the appropriate 2-channel ANC scenario uses *past* values of the desired signal for the auxiliary channel, as reflected in Fig. 2.

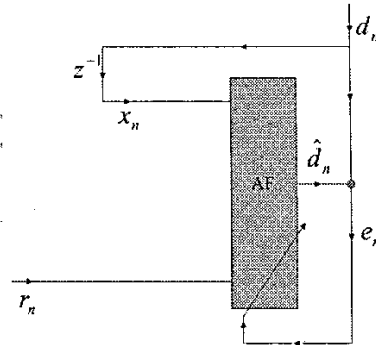


Fig. 2: 2-Channel ANC Scenario.

In the conventional ANC scenario, the reference process carries information about the desired process because of correlation. In the 2-channel case the past of the desired signal is similarly expected to carry additional information about its immediate future.

4.2. AEQ

The 2-channel AEQ scenario is depicted in Fig. 3.

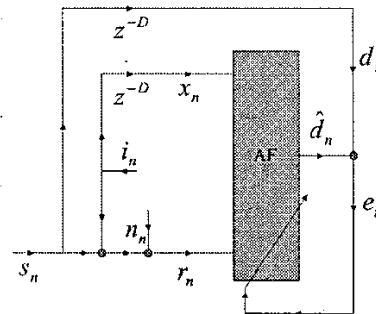


Fig. 3: 2-Channel AEQ Scenario.

In the conventional interference-contaminated AEQ scenario, the adaptive filter derives an estimate of the training sequence sample s_{n-D} from the reference $r_n = s_n + i_n + n_n$, an interference and measurement noise contaminated signal. Strong narrowband interference can be predicted quite well from its past values, so that having the interference channel as the auxiliary channel could be beneficial.

An approximation to this 2-channel structure could be implemented by using the reference input to provide an estimate for the interference signal.

4.3. ALP

In the conventional adaptive linear predictor the reference input consists of past values - delayed by one or more samples - of the process to be predicted. In the 2-channel version the auxiliary channel consists of the

immediate past values of the process to be predicted.

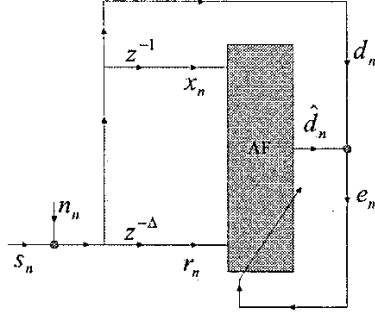


Fig. 4: 2-Channel ALP Scenario.

For white additive noise conditions the one-step predictor exhibits the best performance, but when the additive noise is correlated, it can be desirable to increase the delay in order to decorrelate the additive correlated noise component.

Recall that the main purpose of defining the above 2-channel scenarios is to reveal the structure underlying the desired signal in each case. In the ANC scenario, the auxiliary channel – the past of the desired signal – is actually available for use in a 2-channel approach. In the AEQ and ALP scenarios, the auxiliary channel – the interference signal and the immediate past of the signal to be predicted, respectively – is not measurable.

5. EQUIVALENT CONVENTIONAL FORM

As argued in Section 4, we start with the time-invariant 2-channel structure for the desired process.

$$\begin{aligned} d_n &= \mathbf{w}^o \mathbf{u}_n + \varepsilon_n \\ &= \begin{bmatrix} \mathbf{w}_x^H & \mathbf{w}_r^H \end{bmatrix} \begin{bmatrix} \mathbf{x}_n \\ \mathbf{r}_n \end{bmatrix} + \varepsilon_n \\ &= \hat{d}_n^o + \varepsilon_n \end{aligned} \quad (6)$$

When the conventional adaptive filter is used, with only the reference input channel, the underlying model to be estimated is of the following restricted form.

$$\begin{aligned} d_n &= \mathbf{w}_r^H \mathbf{r}_n + e_n \\ &= \hat{d}_n + e_n \end{aligned} \quad (7)$$

To see what happens when the data structure that underlies the desired process, as in (6), is forced onto the model used in conventional NLMS, as in (7), we use the device of linking sequences. The linking sequence $\rho_{n-l}^{(-m+l)}$ expresses the instantaneous connection between the elements x_{n-m} and r_{n-l} , of the auxiliary and reference channel respectively.

$$\rho_{n-l}^{(-m+l)} = \frac{x_{n-m}}{r_{n-l}} = \frac{x_{n-l+(-m+l)}}{r_{n-l}} \quad (8)$$

Consequently, the 2-channel structure in (6) can be rewritten in terms of the conventional model of (7). For example, supposing that the auxiliary channel has a single element x_n , this element can be rewritten in terms of the reference channel element r_{n-m} as $\rho_{n-m}^{(m)} r_{n-m}$. Using this substitution in (6) yields:

$$\begin{aligned} d_n &= \begin{bmatrix} \mathbf{w}_x^* \rho_{n-m}^{(m)} & \mathbf{w}_r^H \end{bmatrix} \begin{bmatrix} r_{n-m} \\ \mathbf{r}_n \end{bmatrix} + \varepsilon_n \\ &= \left(\mathbf{w}_x^* \rho_{n-m}^{(m)} \mathbf{1}_{m+1}^H + \mathbf{w}_r^H \right) \mathbf{r}_n + \varepsilon_n \end{aligned} \quad (9)$$

where $\mathbf{1}_{m+1}$ is an indicator vector of all zeros except for a 1 as the $(m+1)$ st element. Any choice for the latter location is equally valid, and any affine linear combination of such solutions is equally valid. The overall linear combination of these structures describes the manifold of solutions in which the a posteriori conventional NLMS weight vector resides.

Note that the structure in (9), rewritten in the form of the conventional structure, has the same – usually small – additive noise as the data structure in (6). However, even if the data structure in (6) has TI coefficients, its conventional model equivalent in (9) generally has time-varying coefficients, due to the presence of the linking sequence $\rho_{n-m}^{(m)}$. In the manifold a linear combination of

the linking sequences $\{\rho_{n-m}^{(m)}\}_{m=0}^{M-1}$ would be present, where M refers to the total number of elements in the reference vector. For any linear combination, the weight behavior in the equivalent conventional structure is determined by the behavior of the linking sequences. As we will show next, the specific nature of the linking sequences varies with the application scenarios given in Section 4.

5.1. ANC Linking Result

In the ANC application for which nonlinear effects have been reported the desired and reference processes are both narrowband AR(1), and centered at different frequencies. The linking sequences behave as follows.

$$\begin{aligned} \rho_{n-m}^{(m)} &= \frac{x_n}{r_{n-m}} = \frac{d_{n-1}}{r_{n-m}} \\ &= \frac{p_d d_{n-2} + \eta_{n-1}}{p_r r_{n-m-1} + v_{n-m}} \\ &= \frac{|p_d|}{|p_r|} e^{j(\omega_d - \omega_r)} \rho_{n-m-1}^{(m)} + v_{n-m} \end{aligned} \quad (10)$$

Note that for narrowband processes the pole radii are close to one, so that – away from the zero-crossings of the reference process – (10) implies that (9) contains a weight vector component that rotates with the difference

frequency of the desired and reference processes. In the manifold of solutions consecutive elements of the weight vector contain different linking sequences that are related as follows.

$$\begin{aligned}\rho_{n-m+1}^{(m-1)} &= \frac{x_n}{r_{n-m+1}} = \frac{d_{n-1}}{r_{n-m+1}} \\ &= \frac{d_{n-1}}{p_r r_{n-m} + v_{n-m+1}} \\ &= |p_r|^{-1} e^{-j\omega_r} \rho_{n-m}^{(m)} + v_{n-m+1}\end{aligned}\quad (11)$$

Note that neighboring weights are rotated (with random perturbation) based on the reference process center frequency. These weight behaviors correspond to the original description of the heterodyning operation [5].

5.2. AEQ Linking Result

In the AEQ application, the auxiliary channel consists of i_{n-D} and the reference process consists of consecutive elements centered about r_{n-D} . The linking sequences then behave as follows.

$$\begin{aligned}\rho_{n-D+l}^{(D-l)} &= \frac{x_n}{r_{n-D+l}} = \frac{i_{n-D}}{r_{n-D+l}} \\ &= \frac{p_i^{-l} i_{n-D+l} + \eta_{n-D}}{i_{n-D+l} + s_{n-D+l} + n_{n-D+l}} \\ &= |p_i|^{-l} e^{-j\omega_i l} + v_{n-D+l}\end{aligned}\quad (12)$$

Where the interference process dominates the noise and signal processes, the linking sequence produces in neighboring weights a rotation with magnitude nearly one and phase equal to the center frequency of the interference process. The time dependent behavior of the linking sequence lies in the additive noise term of the last line in (12), so that the TV aspect of the corresponding weight in (9) is a random drift.

5.3. ALP Linking Result

The ALP scenario linking sequences behave as:

$$\begin{aligned}\rho_{n-m}^{(m)} &= \frac{x_n}{r_{n-m}} = \frac{d_{n-1}}{d_{n-\Delta-m+1}} \\ &= \frac{p_d^{\Delta+m-1} d_{n-\Delta-m+1} + \eta_{n-1}}{d_{n-\Delta-m+1} + s_{n-\Delta-m+1} + n_{n-\Delta-m+1}} \\ &= |p_d|^{\Delta+m-1} e^{j\omega_d(\Delta+m-1)} + v_{n-m}\end{aligned}\quad (13)$$

In the manifold, the consecutive weight vector element portion that comes from projecting the interference signal onto the reference dimension undergoes a complex rotation, subject to additive noise. Only the random aspect produces time-varying weight behavior.

The nonlinear effects observed in ANC, AEQ, and ALP [1-4] are characterized by performance that is better than that for the conventional Wiener filter, and

by dependence on NLMS stepsize. We have shown that the conventional structure equivalent of the data exhibits weight behavior that is nearly deterministically time-varying (ANC) or subject to random drift (AEQ and ALP). NLMS adaptation implies the attempted tracking of the structure underlying the data and, depending on the measure of success in doing so, some fraction of the potential performance gain (determined by the structure of the desired data) may be realized.

6. CONCLUSIONS

The nonlinear effects that occur in adaptive filtering are explained by investigating what happens when the conventional tapped delay line model is forced upon the multi-channel structure that underlies the data. When the data structure does not correspond to a time-invariant version of the model used in the adaptive filter, the constrained structure becomes time-varying and an adaptive filter may be able to partially track that time-varying nature. The latter explains how an adaptive filter can outperform the corresponding time-invariant Wiener filter when both use the conventional tapped delay line model.

REFERENCES

- [1] M. Reuter and J. R. Zeidler, Nonlinear effects in LMS adaptive equalizers, *IEEE Trans. Signal Processing*, 47, 1570-1579, June 1999.
- [2] 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.
- [3] J. Han, J. R. Zeidler, and W. H. Ku, "Nonlinear Effects of the LMS Predictor for Chirped Input Signals," *EURASIP Journal of Applied Signal Processing, Special Issue on Nonlinear Signals and Applications, Part II*, January 2002.
- [4] A. A. (Louis) Beex and James R. Zeidler, "Nonlinear Effects in Adaptive Filters," in *Advances in LMS Adaptive Filters*, eds. S. Haykin and B. Widrow, John Wiley & Sons, 2002.
- [5] J. R. Glover, "Adaptive noise canceling applied to sinusoidal interference," *IEEE Trans on Acoustics, Speech and Signal Processing*, ASSP-25, 484-491, Dec. 1977.

Acknowledgement

The present work was funded by the Independent Laboratory Research Program at SPAWAR Systems Center, San Diego, and – in part – by the National Research Council, when the first author was a Senior Research Associate at SPAWAR Systems Center, San Diego, while on sabbatical from Virginia Tech.