AN ASSESSMENT OF LINEAR ADAPTIVE FILTER PERFORMANCE WITH NONLINEAR DISTORTIONS

Moctar I. Mossi[*] *and Nicholas W. D. Evans*

EURECOM
Sophia-Antipolis, France
{mossi, evans}@eurecom.fr

Christophe Beaugeant

Infineon Technologies
Sophia-Antipolis, France
christophe.beaugeant@infineon.com

ABSTRACT

Acoustic echo cancellers are generally based on the assumption of a linear echo path between the transducers. However the small loudspeakers that are commonly used in todays terminals can introduce nonlinear distortions that reduce the performance of echo cancellation. In order to evaluate the degradation in performance, this paper assesses the behaviour of five linear echo cancellers in the presence of nonlinearities and presents the first thorough comparison of their robustness. Even if the performance of all the echo cancellers degrades as expected, some algorithms are shown to be more robust than others: fast converging algorithms and block signal processing are more perturbed in nonlinear environments.

Index Terms— echo cancellation, nonlinear distortion, AEC, LMS, NLMS, TDLMS, APA, FBLMS, Volterra model.

1. INTRODUCTION

This paper addresses the well known problem of acoustic echo. Typically, the round-trip delay of mobile and IP networks exceeds 200 ms. Any acoustic feedback between the loudspeaker and the microphone of a terminal can be particularly disturbing for the far-end user who can be disturbed by hearing his/her own delayed voice. Consequently many different approaches to Acoustic Echo Cancellation (AEC) have been proposed over recent years. Common to much of this work is the assumed linearity of electronic components and the acoustic echo path between the loudspeaker and microphone. Under such conditions AEC algorithms generally perform well. However, the miniaturization of transducers and enclosures introduces nonlinear distortions which are known to degrade the performance of linear AEC algorithms [1, 2]. Researchers have thus sought to develop effective solutions to nonlinear AEC.

One common solution to nonlinear AEC is based on Volterra filters [1, 3, 4]. Volterra solutions lead to improved nonlinear echo cancellation but tend to come at the expense of higher complexity, slow convergence and often lead to sub-optimal, local minima MSE solutions [5]. An alternative approach involves the use of linear adaptive filters followed by a postfilter to attenuate residual nonlinear echo. These solutions tend to be less complex than Volterra-based solutions but rely more heavily upon efficient linear AEC in the presence of nonlinearities [6, 7].

Both approaches thus rely to some extent on effective linear AEC performance in the presence of nonlinearities. It is though, perhaps surprisingly, difficult to find a thorough comparison of the robustness of linear AEC algorithms to nonlinear distortion (one notable exception being [8]) and thus herein lie the contributions of this paper. We present an assessment of five popular, standard linear AEC algorithms under the presence of artificially generated but realistic nonlinear distortions. Contrary to the findings of [8] our work shows that advanced AEC algorithms such as the Affine Projection Algorithm (APA) do indeed outperform the more conventional approaches in linear environments but attain only comparable performance in highly nonlinear environments. We also show that in the presence of nonlinearities block processing algorithms are more affected. In addition we present new experimental work which assesses the performance of each algorithm under varying degrees of nonlinear distortion and highlight conditions where the more conventional algorithms might nonetheless be of benefit. The work should be of particular relevance to further work in nonlinear AEC in guiding the choice of linear filter used with postfilters and the adaptation of the linear component of Volterra filters.

The remainder of the paper is organised as follows. A general system/echo model is introduced in Section 2 before the five standard linear AEC algorithms are briefly described. In Section 3 we introduce the nonlinear model which was used to synthesize nonlinear distortions for our experimental work which is presented in Section 4. Finally our conclusions are presented in Section 5.

2. ACOUSTIC ECHO CANCELLATION

In this section we introduce a typical system/echo model and a general framework for AEC with adaptive filtering. Also described are the five approaches to AEC that are investigated in this paper.

2.1. System/echo model

A general system/echo model, which was used for all experiments reported in this paper, is illustrated in Figure 1. The terminal receives a downlink (or loudspeaker) signal, $x(n)$, from a far-end speaker, and transmits an uplink (or microphone) signal $y(n)$. In addition to near-end speech $s(n)$ and noise $n(n)$ the uplink signal potentially includes an additional echo component $d(n)$, a result of the acoustical coupling between the loudspeaker and the microphone.

The acoustical coupling is generally modelled with a linear convolution, $d(n) = x(n) * h_{opt}(n)$, where $h_{opt}(n)$ is the impulse response which characterises the acoustical coupling. AEC may thus be implemented by estimating $h_{opt}(n)$ with a filter $h(n)$ in order to estimate the coupled echo signal $\hat{d}(n) = x(n) * h(n)$. The echo is attenuated simply by subtracting $\hat{d}(n)$ from the uplink signal. Since the acoustical coupling is time varying $h(n)$ is usually an adaptive filter. Near-end speech disturbs the adaptive filter and so $h(n)$ is usually updated in echo-only periods, i.e. $s(n) = 0$. In this work it is also supposed that the background noise is negligable, i.e. where

[*]M. I. Mossi is supported by Infineon Technologies

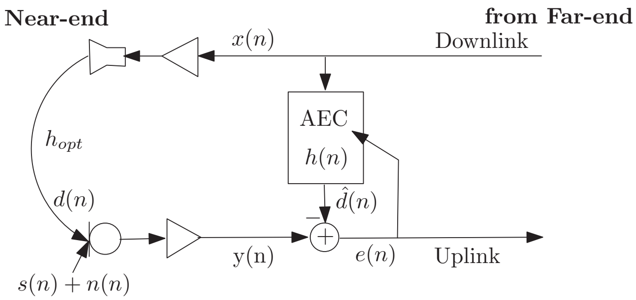


Fig. 1. System/echo model illustrating the acoustical coupling between the loudspeaker and microphone and a general approach to adaptive AEC.

$n(n) = 0$. Under such conditions $y(n) = d(n)$ and thus the resulting error signal, $e(n)$ is the difference between the echo signal and its estimate, i.e. $e(n) = d(n) - \hat{d}(n)$. The error $e(n)$ is used to update the filter $h(n)$ whose goal is to drive $e(n)$ to zero.

Since the linear filter can thus influence nonlinear filtering performance, it is of interest to study the robustness of the linear filter to nonlinearities. This is even more important to postfiltering approaches, given the inherent dependency between a conventional linear adaptive filter, used to attenuate the linear echo, and the postfilter, which is used to attenuate residual (nonlinear) echo. In this paper we present some new experimental work which assesses the performance of five different, standard algorithms, each of which is described below.

2.2. Linear adaptive filter algorithms

The adaptive filters considered in this paper are updated according to a general adaptation recursion given by:

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \Delta\mathbf{h}(n), \qquad (1)$$

where $\mathbf{h}(n)$ is the vector of filter taps at time n, and where $\Delta\mathbf{h}(n)$ is the gradient used to update the filter. It is different for each algorithm and should ensure that $\mathbf{h}$ converges to $\mathbf{h}_{opt}$ after sufficient iterations. In the following we identify the five commonly used adaptive AEC filters that are investigated in this paper. Only the barest of details are given as full details can be found in the open literature [9].

Least Mean Square (LMS): The LMS filter update $\Delta\mathbf{h}(n)$ is equal to $\mu\mathbf{x}(n)e(n)$, where μ is a scalar or step size which aims to control the rate of adaptation (and hence convergence/divergence), $\mathbf{x(n)} = [x(n), x(n-1), ..., x(n-L+1)]^T$ is the input vector of the filter and L is the filter length (256 for all algorithms used here).

Normalized-LMS (NLMS): The NLMS algorithm uses a normalized step size μ. Here the update $\Delta\mathbf{h}(n)$ is equal to $\frac{\mu}{\|\mathbf{x}(n)\|^2}\mathbf{x}(n)e(n)$.

Transform Domain-LMS (TDLMS): We use a Discrete Cosine Transform-LMS (DCTLMS), with an update $\Delta\mathbf{h}(n)$ equal to $\mu\bar{\mathbf{x}}(n)e(n)$, where $\bar{\mathbf{x}}(n) = \mathbf{x}(n)\mathbf{T}$. $\mathbf{T}$ is the Discrete Cosine Transform (DCT) matrix.

Affine Projection Algorithm (APA): The update $\Delta\mathbf{h}(n)$ is here given by $\mu\mathbf{X}(n)[\mathbf{X}^T(n)\mathbf{X}(n)+\epsilon I_N]^{-1}\mathbf{e}(n)$ where $\mathbf{X}(n) = [\mathbf{x}(n)\mathbf{x}(n-1)...\mathbf{x}(n-N+1)]$, an $L \times N$ matrix. L is the length of the filter, N is the order of the APA, I_N is the identity matrix and $\mathbf{e}(n)$ is now a vector. In this paper we use only $N = 2$ (APA2) (higher order APA filters were investigated with similar results to those presented in this article).

Frequency Block-LMS (FBLMS): FBLMS is an implementation of a block-by-block LMS using fast convolution. In the time domain the update $\Delta\mathbf{h}(n)$ is given by $\mu\sum_{m=0}^{m=B-1} e(nB+m)\mathbf{x}(nB+m)$ where n is now a block index, m is the block sample index and B is the block length. We use $B = 256$.

3. NONLINEAR MODEL

It is the objective of this paper to report the first thorough assessment of standard linear AEC robustness to nonlinearities. Since this requires comparisons of performance both with and without nonlinearities under otherwise identical conditions it is necessary that nonlinear distortions be generated artificially. It is these aspects of the test setup which are described here. All other aspects of the test setup are described in Section 4.

In general nonlinearities are introduced by the uplink and downlink amplifiers, by the loudspeaker, the microphone, resonance from the mobile terminal housing and the acoustic echo path. However, since the loudspeaker signal is usually of high level, especially in handsfree mode, it is commonly assumed that nonlinearities from the downlink amplifier and loudspeaker dominate and that, consequently, all other sources are negligible [3, 10]. Under this assumption the acoustic path may then be considered as linear.

As in [3, 5] both downlink nonlinearities may be adequately modelled using a Volterra model [3]. As in the work of [10] the full Volterra model of amplifier and loudspeaker nonlinearities may be approximated by a cascade of memoryless saturation characteristics. We take into account only the second and third order nonlinearities as they are generally assumed to be the most dominant components [2, 3]. As in [7, 8] for all experimental work reported here nonlinearities are generated according to:

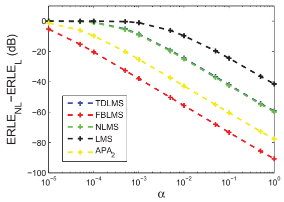
$$x_{nl}(n) = x(n) + \alpha x^2(n) + \beta x^3(n), \qquad (2)$$

where $x_{nl}(n)$ is the nonlinear output of the loudspeaker. α and β are the respectively second and third order weighting components and lie in the range of $(\alpha, \beta) = [0, 1]$. It is worth mentionning that the couple $(\alpha, \beta) = (0, 0)$ corresponds to the linear case. This range of parameters was deemed to be representative of realistic nonlinearities measured through laboratory tests of several popular, current mobile phones. It also agrees with those in the general literature, e.g. [11]. The loudspeaker signal $x_{nl}(n)$ is then convolved with an impulse response $h_{opt}(n)$ to simulate the linear echo path between the loudspeaker and the microphone.

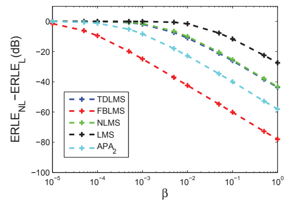
4. EXPERIMENTAL WORK

Each algorithm is assessed in terms of Echo Return Loss Enhancement (ERLE), i.e. the reduction in energy (in dB) of $d(n)$ achieved by echo reduction. It is also assessed in terms of convergence time which we define as the time needed for the ERLE to reach 95% of its maximum.

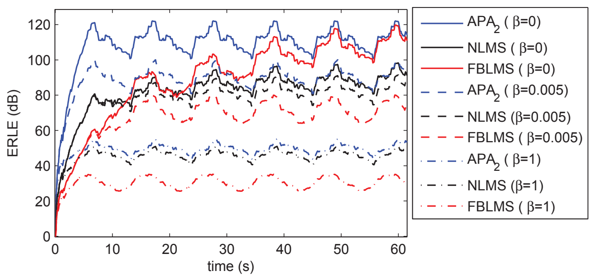
In order to illustrate our experimental setup we describe here one particular experiment extracted from a larger setup described below. A 10 second long speech signal is concatenated 6 times to produce test signal $x(n)$ of sufficient duration to ensure the convergence of each algorithm. $x(n)$ is used to synthesize downlink amplifier and loudspeaker nonlinearities according to Equation 2. Since we assume a linear echo path the nonlinear signals $x_{nl}(n)$ are subsequently convolved with a 256-tap filter h_{opt}, to simulate the microphone signal $d(n)$. Each of the five AEC algorithms are then applied to $d(n)$ according to the general scheme of Figure 1, using



a) Difference in ERLE ($\beta = 0$)

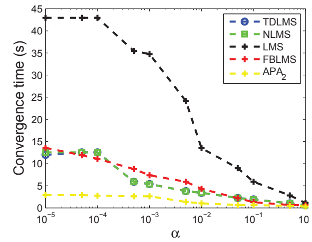


b) Difference in ERLE ($\alpha = 0$)

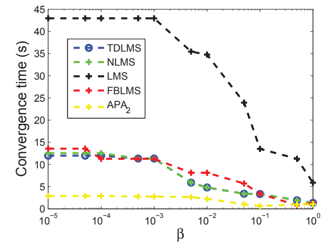


c) ERLE over time of NLMS, FBLMS and APA ($\alpha = 0$)

Fig. 2. ERLE test results to compare the performance in linear and nonlinear environments.



a)($\beta = 0$)



b)($\alpha = 0$)

Fig. 3. Convergence time decreasing in presence of nonlinearities

$x(n)$ as the reference signal. The typical set-up described here is extracted from a larger test setup, using different impulse responses h_{opt} (measured experimentally using a mobile terminal in an office room) and different input signals (4 speakers, 2 languages). The larger test set-up leads to identical conclusions as presented below.

The behavior of all algorithms is dependent on the step size μ. We have chosen suitable values of μ for each algorithm based on thorough empirical optimization in order to achieve maximum ERLE after convergence, leading to $\mu = 1$ for APA and NLMS, $\mu = 0.5$ for FBLMS and $\mu = 0.15$ for LMS. Additional experiments (not reported here) show that for different values of μ, the influence of nonlinear distortion is similar to the effects described here. We have also checked that the influence of μ on the performance of the AEC is similar in linear and nonlinear environments.

4.1. Echo Return Loss Enhancement (ERLE)

Figures 2(a) and 2(b) show the difference between the ERLE (vertical axis) in a linear environment $(\alpha, \beta) = (0, 0)$ and in a nonlinear environment $(\alpha, \beta) \neq (0, 0)$ after convergence for each of the five different algorithms. We define the ERLE value after convergence as the mean of the ERLE on the 6^{th} period of our test sequence. Figure 2(a) (resp. Figure 2(b)) illustrates the influence on the ERLE for different values of α (horizontal axis) when $\beta = 0$ (resp. β when $\alpha = 0$). Results where both $\beta \neq 0$ and $\alpha \neq 0$ are similar to profiles depicted in these two figures, An idea of performance for such test cases can be accordingly extrapolated from these two curves.

The general trend of these curves shows that the difference in ERLE of all the algorithms increases when the nonlinearity increases. Also evident is the greater influence of the second order weighting factor α than the third order factor β. This can be explained easily considering the model of Equation 2: for a given normalized $x(n)$ ($\|x(n)\| \leq 1$), $x^2(n) > x^3(n)$ so that the second order α has a stronger influence than the third order β.

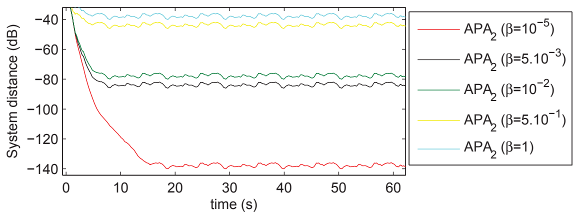
Generally, for small values of α and β the ERLE difference is close to zero, indicating a low degradation in echo cancellation performance due to small nonlinearities. For the LMS, when $\alpha \leq 10^{-3}$ and $\beta \leq 10^{-2}$ and for NLMS and TDLMS when $\alpha \leq 10^{-4}$ and $\beta \leq 10^{-3}$, the ERLE is almost unaffected by the nonlinearities. This is shown by the flatness of the curves in these ranges. The most affected echo cancellers are the APA and FBLMS, where the difference in ERLE decreases even for small values of α and β.

To better illustrate the behavior of the ERLE over time, Figure 2 (c) gives the ERLE for the APA2, NLMS and FBLMS for different values of β ($\alpha = 0$). Similar curves are obtained by considering different values of α. TDLMS gives almost identical behavior to the NLMS and so its profile is not presented. These curves show clearly how nonlinearities reduce the maximum ERLE reached by each algorithm. As already mentioned the FBLMS is the most affected and its ERLE is lower than that of NLMS when $\beta > 10^{-3}$. Even if the APA filter is disturbed significantly by nonlinearities, it still reaches a better ERLE than other algorithms after convergence. From these experiments, a first conclusion is that the faster an algorithm converges the more it is affected by nonlinearities. The APA, for instance, is known to convergence quickly compared to the NLMS but its performance drastically decreases when nonlinearities increase.

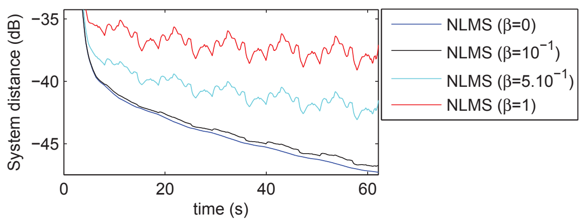
FBLMS, however, is severely affected even though it does not converge quickly in linear environments. This behavior is explained by the block-by-block processing nature of FBLMS. According to Equation 2, small input signals $x(n)$ lead to small nonlinearities. As a result, even for high values of α and β, a sample-based algorithm will be, for certain periods of low $x(n)$, equivalent to a linear environment and thus, during such periods, it will be relatively less disturbed by nonlinearities. Considering block-based processing such as FBLMS, a whole frame of low level $x(n)$ is needed to have the same effect. As a result, block-based algorithms are more disturbed by the same level of nonlinear distortion.

4.2. Convergence Time

Figures 3(a) and 3(b) show the influence of the nonlinear weighting factor on the convergence time. These results clearly show that in nonlinear environments all the algorithms converge faster than in linear environments. Such unexpected results are explained by the fact that the algorithms converge in practice to a lower ERLE level; this ERLE level is in fact reached faster simply because it is lower. Looking, for instance, at the profile for LMS, its convergence time decreases from 45s to less that 5s for α varying between 0 and 1, but at the same time the ERLE achieved by LMS collapses by 30 dB. It is nevertheless an important result that echo cancellers operating in nonlinear environments provide less echo reduction but their maximum level of echo reduction is reached relatively quickly. Accord-



a) APA



b) NLMS

Fig. 4. System distance over time ($\alpha = 0, \beta$)

ingly, fast converging algorithms such as APA can be of less interest in nonlinear environments as the argument to use such algorithms due to their reduced convergence time may no longer hold.

4.3. Estimation of Linear Echo Path

Plotted in Figures 4(a) and 4(b) is the evolution of the system distance $\left|\frac{h_{opt}-h(n)}{h_{opt}}\right|$, over time in dB. We present here the results with only APA2 and NLMS. The system distance allows us to judge the accuracy of the AEC in estimating the linear component of the echo signal. We can first observe that APA2 results in better estimation in the presence of low level nonlinearities, but less accurate estimation when nonlinearities increase. The NLMS shows slower convergence than APA2 but its estimate is closer to the linear case until the level of nonlinearities exceeds $\beta = 10^{-1}$. This shows that the estimation of the linear component of the echo is more robust when using NLMS in highly nonlinear environments. The behavior of the NLMS is similar to that of TDLMS and LMS (results not shown here). The FBLMS system distance is also more affected as was the case for the ERLE. One could easily assume that the linear echo canceller aims at estimating the linear component of $d(n)$, but this assumption is not supported by these results. Indeed the system distance increases when the nonlinearities increases. This means that, in practice, echo cancellers do not converge to a reliable estimate of the linear component of the echo path h_{opt}. This is of particular interest as many algorithms assume that a nonlinear system can be accurately modelled by a cascade of a linear echo canceller and post cancellation of the residual nonlinear echo [6]. Even in [3] a poor estimation of the linear component of the echo will influence the performance of the whole system.

5. CONCLUSIONS

This paper reports an assessment of linear AEC performance in nonlinear environments modelled by a Volterra approximation. We compare the performance of five common standard algorithms. Experimental results show that APA achieves similar performance to NLMS in highly nonlinear environments. The FBLMS performance collapses even for relatively small nonlinearities. We also show that, in presence of nonlinearities, the linear component of the echo is not well estimated by conventional approaches to AEC. This leads us to question the common application of linear AEC to cancel the linear component in nonlinear environments.

Thus the experimental results reported here show that performance varies greatly across the different algorithms investigated. The study highlights the need for further work to confirm these results on a wider array of AEC approaches to confirm the interpretation proposed in this article, i.e. the low robustness of fast converging algorithms and block-based processing facing nonlinearities. More generally, assessing the performance of linear AEC is an important step to provide effective nonlinear AEC systems. Such an investigation has, perhaps surprisingly, not been published previously and thus this article sheds new light on the robustness of linear echo cancellers to nonlinear distortion.

6. REFERENCES

[1] A. Stenger, L. Trautmann, and R. Rabenstein, "Nonlinear acoustic echo cancellation with 2nd order adaptive volterra filters," *ICASSP*, vol. 2, pp. 877 – 880, Mar 1999.

[2] A.N. Birkett and R.A. Goubran, "Limitations of handsfree acoustic echo cancellers due to nonlinear loudspeaker distortion and enclosure vibration effects," *IEEE ASSP Workshop*, pp. 103 – 106, Oct 1995.

[3] A. Guerin, G. Faucon, and R. Le Bouquin-Jeannes, "Nonlinear acoustic echo cancellation based on volterra filters," *IEEE Trans. on Speech and Audio Proc.*, vol. 11, pp. 672 – 683, Nov 2003.

[4] D. Zhou, V. DeBrunner, Y. Zhai, and M. Yeary, "Efficient adaptive nonlinear echo cancellation, using sub-band implementation of the adaptive volterra filter," *ICASSP*, vol. 5, 2006.

[5] A. Fermo, A. Carini, and G.L. Sicuranza, "Nonlinear acoustic echo cancellation using adaptive orthogonalized power filters-analysis of different low complexity nonlinear filters for acoustic echo cancellation," *IWISPA*, pp. 261–266, June 2000.

[6] O. Hoshuyama and A. Sugiyama, "An acoustic echo suppressor based on a frequency-domain model of highly nonlinear residual echo," *ICASSP*, vol. 5, May 2006.

[7] K. Shi, X. Ma, and G. T. Zhou, "A residual echo suppression technique for systems with nonlinear acoustic echo paths," *ICASSP*, pp. 257 – 260, Apr 2008.

[8] R. Niemisto and T. Makela, "On performance of linear adaptive filtering algorithms in acoustic echo control in presence of distorting loudspeakers," *IWAENC*, pp. 79–82, Sept 2003.

[9] S. Haykin, *Adaptive Filter Theory* 4^{th} *Ed*, Prentice Hall, 2001.

[10] F. Kuech, A. Mitnacht, and W. Kellermann, "Nonlinear acoustic echo cancellation using adaptive orthogonalized power filters," *ICASSP*, vol. 3, pp. 105–108, Mar 2005.

[11] W. Frank, "An efficient approximation to the quadratic volterra filter and its application in real-time loudspeaker linearization," *Signal Processing*, vol. 45, pp. 97–113, Jul 1995.