Eurécom
Department of Multimedia Communication
2229, route des Crêtes
B.P. 193
06904 Sophia-Antipolis
FRANCE

Research Report RR-10-240

Acoustic echo cancellation in non-linear and noisy environments

May 25th, 2010

Moctar I. Mossi, Christelle Yemdji Tchassi, Nicholas W. D. Evans and Christophe Beaugeant

Tel : (+33) 4 93 00 81 00
Fax : (+33) 4 93 00 82 00
Email : {mossi, yemdji, evans}@eurecom.fr, christophe.beaugeant@infineon.com

[1]Eurécom's research is partially supported by its industrial members: BMW Group Research & Technology - BMW Group Company, Bouygues Télécom, Cisco Systems, France Télécom, Hitachi Europe, SFR, Sharp, STMicroelectronics, Swisscom, Thales.

Acoustic echo cancellation in non-linear and noisy environments

Moctar I. Mossi, Christelle Yemdji Tchassi, Nicholas W. D. Evans and Christophe Beaugeant

Abstract

This paper presents a new assessment of four approaches to acoustic echo cancellation in the presence of non-linear echo. The comparison is performed with algorithms that are configured to give equivalent performance under linear-only echo conditions, thereby giving a more meaningful assessment. We also compare the effect of non-linear echo to that of noise and show that, whilst performance differs in non-linear environments, there are negligible differences in noisy environments. The computationally efficient FBLMS algorithm is shown to perform as well as other algorithms in noisy environments but performs poorly under non-linear echo conditions. We also show how the correlation between non-linearities and the speech signal can corrupt the echo path estimate and leads to small attenuation of non-linear components at the expense of reduced attenuation for linear echo. Finally we discuss the merits of modelling non-linearities as a linear environment with a time varying echo path.

Index Terms

echo cancellation, non-linear distortion, noise, AEC, LMS, NLMS, APA, FBLMS.

1 Introduction

The general problem of acoustic echo cancellation (AEC) has been actively researched for many years and adaptive filtering has proved to be the most popular solution in the communications environment. This has led to the development of many different algorithms to optimize adaptive filtering in the presence of additive noise and the inevitable echo path variations that generally decrease filter performance. In noisy environments adaptive filters are configured so that convergence (and/or divergence) occur gradually, and thus there is an inbuilt robustness to sudden increases in noise level e.g. [1–3]. Fast adaptive filters ensure robustness to echo path variations. Additional challenges come from the miniaturization of components which can lead to non-linear echo artefacts. Many efforts to improve adaptive filters for non-linear environments have been reported and two dominant solutions have emerged. The first is based on the Volterra filter [4] and the second involves post-filtering [5] in combination with AEC adaptive filtering. The Volterra solution is generally slow to convergence and is highly computationally complex. Post-filters are less complex but rely on the performance of linear adaptive filters that are still disturbed by non-linear echo. Thus linear adaptive filtering is still popular and it is of interest to assess adaptive filters in such environments, as reported in our previous work [6].

In [6] the performance of adaptive filters was assessed by comparing the degradation in echo return loss enhancement $ERLE$ with (i) linear echo and (ii) linear and non-linear echo. However, this assessment perhaps does not best reflect the true robustness of each algorithm to non-linear echo, since the comparison in [6] was made independently of the $ERLE$ under linear conditions. The contribution in this paper is thus new experimental work which compares the same approaches to AEC that were assessed in [6] but here with identical adaptation step sizes and with tuned regularization factors that give each approach equivalent $ERLE$ performance under *linear* echo conditions. This form of assessment aims to give a more meaningful and fairer comparison of each approach under *non-linear* echo conditions.

A second contribution in this paper is a new comparison of the effects of non-linear echo to additive background noise. The aim here is to judge the relative importance of non-linear echo and associated degradation in $ERLE$ to that caused by noise. Since the effects of non-linear echo and noise are uncorrelated they are accordingly assessed independently. Therefore, the experimental results that are presented in this paper can be used to predict performance in other environments if the level of each perturbation is known. This can help to decide which approach to AEC is most applicable in which environment.

The remainder of the paper is organized as follows. In Section 2 a general echo cancellation system is described with the four different approaches to AEC that are investigated in this paper. In Section 3 we introduce the non-linear echo and noise model which is used in our assessments. Experimental work and results are

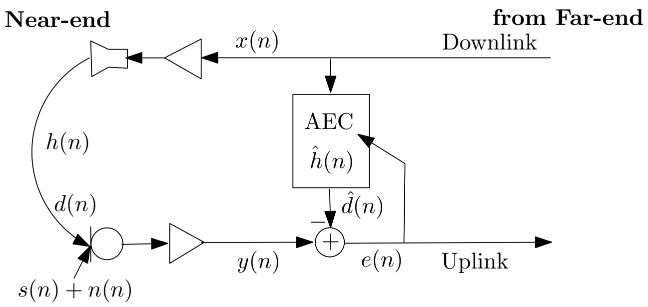


Figure 1: System/echo model illustrating the acoustical coupling between the loudspeaker and microphone and a general approach to adaptive AEC.

presented in Section 4. The implications of the work are discussed in Section 5 and our conclusions are presented in Section 6.

2 Acoustic echo cancellation

We here introduce a typical system/echo model and a general framework for AEC with adaptive filtering. Also described are the four 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 down-link (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 additive background noise $n(n)$ the uplink signal potentially includes an additional echo component $d(n)$, which is a result of the acoustical coupling between the loudspeaker and the microphone. It is generally modelled with a linear convolution, $d(n) = x(n) * h(n)$, where $h(n)$ is the impulse response which characterizes the acoustical coupling. AEC may thus be implemented by estimating $h(n)$ with a filter $\hat{h}(n)$ in order to give an estimate of the coupled echo signal $\hat{d}(n) = x(n) * \hat{h}(n)$. The echo is attenuated simply by subtracting $\hat{d}(n)$ from the uplink signal. Since the acoustical coupling is generally time varying $\hat{h}(n)$ is usually an adaptive filter. Near-end speech disturbs the adaptive filter and so $\hat{h}(n)$ is usually updated in echo-only periods, i.e. when $s(n) = 0$. Noise can also disturb the adaptive filter but, if we also suppose that the noise is negligible, i.e. $n(n) = 0$, then $y(n) = d(n)$ and thus the resulting error signal, $e(n)$ is simply 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.

AEC rarely operates under such ideal conditions, however, and thus it is interesting to study the robustness under more realistic conditions. i.e. with near-end speech, non-linear echo and additive background noise. As adaptation is simply paused during intervals of near-end speech, only disturbances from non-linear echo and background noise are considered here. Each of the approaches to AEC that are considered are described below.

2.2 Linear adaptive filtering algorithms

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

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

where $\hat{\mathbf{h}}(n)$ is the vector of filter taps at time n and where $\Delta\mathbf{h}(n)$ is the gradient used to update the filter. Everywhere in this paper boldface denotes vectors whereas boldface capitals denote matrices. The gradient is different for each algorithm but, in all cases, should ensure that $\hat{\mathbf{h}}(n)$ converges to the optimal Wiener solution $\mathbf{h}_{opt}$ after sufficient iterations. Only the barest of details for each approach considered are given below as full details can be found in the open literature [7].

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), **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 $\frac{\mu}{\|\mathbf{x}(n)\|^2}\mathbf{x}(n)e(n)$.

Affine Projection Algorithm (APA): The update $\Delta\mathbf{h}(n)$ is here given by $\mu\mathbf{X}(n)[\mathbf{X}^T(n)\mathbf{X}(n) + \epsilon\mathbf{I}_N]^{-1}\mathbf{e}(n)$ 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, $\mathbf{I}_N$ is the identity matrix and $\mathbf{e}(n)$ is now a vector. For all experiments represented here we use $N = 2$.

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 $\mu\sum_{m=0}^{B-1} e(nB+m)\mathbf{x}(nB+m)$ where n is now a block index, m is the block sample index and B is the block length. We use $B = 256$.

3 Non-linear model

In this paper we aim to assess the effects of non-linear echo. So that direct comparisons between (i) linear echo and (ii) linear and non-linear echo may be made with otherwise identical conditions, for all experiments reported here, non-linearities are added artificially. In practice, non-linearities are introduced by component imperfections, i.e. from the miniaturization of components, and can be divided into two groups: those which arise in the downlink path and those which arise in the uplink path. Previous work [4] has shown that non-linearities coming from the loudspeaker and amplifiers in the downlink path dominate those in the uplink path due to the fact that microphone and uplink amplifiers generally operate on lower-level signals. As in [6] a third-order polynomial model is used here to simulate non-linearities and is an approximation to the Volterra model. The output of the loudspeaker is given by $x_{nl}(n) = x(n) + \alpha x^2(n) + \beta x^3(n)$, where $x(n)$ is the far-end signal and $\alpha x^2(n) + \beta x^3(n)$ are the non-linear components introduced by the downlink loudspeaker and amplifiers. The parameters α and β are used to control the relative levels of second and third order non-linear distortions. A full description of this setup is given in our previous article [6].

4 Experimental work

We report different tests on each of the adaptive filters and compare the effects of non-linearities and white noise. The assessment is based on $ERLE$, convergence time and system distance. A 60-second speech signal is used as the far-end signal $x(n)$ and is sufficient to ensure the convergence of each algorithm. In all cases $ERLE$ measurements relate to intervals in which the algorithms are deemed to have converged. Non-linear artefacts are introduced into the down-link signal according to the model described in Section 3. The loudspeaker output is composed of the original speech signal $x(n)$ and a non-linear component $\alpha x^2(n) + \beta x^3(n)$ which are both convolved with the echo path $h(n)$. This leads to a linear echo component $x(n) * h(n)$ and a non-linear echo component $[\alpha x^2(n) + \beta x^3(n)] * h(n)$. Then, a linear echo to non-linear echo ratio ($SNeR$) is computed as in [8]:

$$SNeR = \frac{1}{K} \sum_{i=1}^{K} SNeR_{seg}(i), \tag{2}$$

where $SNR_{seg}(i)$ is given by:

$$SNeR_{seg}(i) = 10log_{10} \frac{\sum_{m=0}^{M-1} d_i^2(n)}{\sum_{m=0}^{M-1} d_{nl,i}^2(n)} \tag{3}$$

and where $d_i(n)$ and $d_{nl,i}(n)$ are the linear and non-linear echo components respectively in the i^{th} segment of analysed signals. The $SNR_{seg}(i)$ is computed using a window of $32ms$ ($M = 256$ for a sampling rate of $8kHz$) according to speech stationarity. The $SNeR$ is used to generate a noisy signal with linear echo, where the mean SNR is equal to $SNeR$. In so doing we have two linear echo signals that are equally disturbed, one with non-linear echo, and another with additive noise. The weighting factors α and β are in the range of $[0, 1]$ as in [6]. This permits us to artificially increase the level of the non-linear echo component (and noise) by increasing α and/or β. We compare the behaviour of each adaptive filter, with both both non-linear echo and noise, when they are configured with the same step size μ, and to obtain approximately the same level of $ERLE$. This is achieved

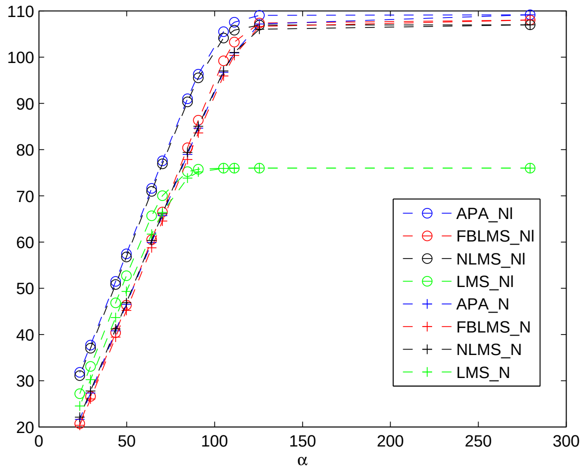


Figure 2: Maximum $ERLE$ (in dB) achieved after convergence as a function of $SNR/SNeR$ (also in dB). Here the SNR or $SNeR$ corresponds to added noise or to non-linear echo as indicated. Profiles are illustrated for both perturbations and for each of the four approaches to AEC. APA, FBLMS and NLMS are all configured to give equivalent performance under linear echo conditions.

by varying the regularization factor in each case. The APA, FBLMS, and NLMS algorithms obtain $ERLEs$ of $\sim 110dB$ in linear echo conditions. LMS does not perform sufficiently well and gives an $ERLE$ of $\sim 80dB$.

4.1 Echo Return Loss Enhancement

Figure 2 shows the maximum $ERLE$ achieved by each algorithm in non-linear and noisy environments. The maximum $ERLE$ is the mean $ERLE$ obtained during a 10 second period where each algorithm has converged. Figure 2 shows the maximum $ERLE$ on the vertical axis and the SNR ($SNeR = SNR$) on the horizontal axis. We can observe that whatever the perturbation (non-linear echo or noise) performance decreases for all adaptive filters. In non-linear environments APA and NLMS algorithms show similar behaviour; decreases of approximately $80dB$ are observed between the linear conditions (right side of Figure 2) and non-linear echo conditions (left side of Figure 2). This shows the sensitivity of linear adaptive filters as, in this range, the non-linearities are inaudible. The FBLMS algorithm is the most affected. Performance decreases by about 90dB over the same range and for $SNeRs$ less than $75dB$ performance is worse than that for the standard LMS algorithm. This is explained in [6] as the effect of block-by-block processing which is more susceptible to non-linear effects than a sample-by-sample process. We see that the LMS algorithm is the most robust of all adaptive filters considered; it has the least degradation in performance as the SNR or $SNeR$

decreases. This is due to its poor ERLE performance which is so low that the algorithm cannot even be configured to give equivalent performance to the other algorithms under linear echo conditions.

In noisy environments the performance of APA, NLMS and FBLMS algorithms decreases by approximately the same amount. For the APA and NLMS algorithms, and when the $SNR < 100dB$, the difference between the $ERLE$ in non-linear and noisy environments is about $10dB$ with better performance in non-linear environments than noisy environments. The FBLMS algorithm seems to show the smallest differences between non-linear and noisy environments. This can again be explained by the averaging effect of block-by-block approaches. In the case of noise the perturbation is effectively averaged over the block and thus has a reduced impact on performance. This is not the case with non-linear echo, which is correlated with the input signal. The result is that noise perturbations have less of an effect than they do for the other approaches and that noise and non-linear echo have an equivalent effect on the performance of the FBLMS algorithm.

The difference between the effects of non-linearities and those of noise are explained by two hypotheses:

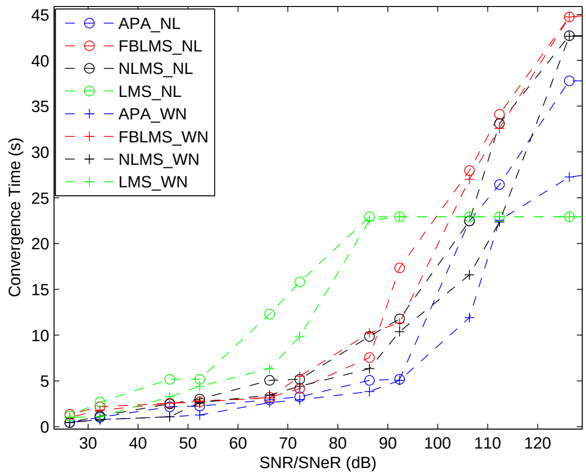
Frequency range of the noise: The convolution of non-linearities with the echo path reduces the effective frequency range. In general the echo path has energy concentrated at low frequencies. Consequently the convolution will attenuate components at higher frequencies that are introduce by non-linear distortions. The echo paths used for these tests are estimated in real conditions with noise so some filter harmonics that are not excited by the linear signal can be introduced by the artificially added non-linear distortion. The noise has a flat spectrum so it will perturb more the adaptive filter than will the non-linearities.

Non-linearities are correlated with the far-end signal: Since non-linearities are correlated with the input signal, this can result in the adaptive filter under estimating the linear part but slightly attenuating the non-linearities. This is less so the case for the noisy environments as there is no correlation between the noise and the far-end signal.

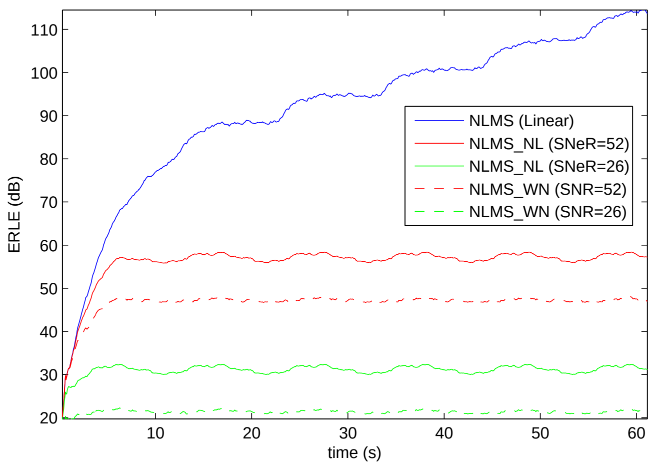
4.2 Convergence Time

The convergence time for each algorithm is computed as given in [6], and is defined as the time needed for the adaptive filter to reach 95% of its maximum $ERLE$ value. Convergence times are determined using the same speech signals as used previously and are estimated for both conditions: linear echo with non-linear echo, and linear echo with noise. Figure 3(a) shows the convergence time in seconds against $SNR/SNeR$ for each of the four algorithms and both perturbations.

We see that, with the exception of the LMS algorithm, all profiles have a similar trend even though differences in convergence time are in the order of $25s$ at $110dB$. In addition, for each algorithm, convergence times are greater for non-linear perturbations than they are for noise. The LMS algorithm is the slowest



(a)



(b)

Figure 3: Convergence performance with non-linear (NL) and white noise (WN) perturbations for (a) APA, FBLMS, NLMS and LMS algorithms plotted as convergence time against SNR, and (b) the NLMS algorithm plotted as $ERLE$ against time.

to converge where the $SNeR$ or SNR is low but the fastest where they are high ($> 100dB$). This is explained by the fact that the $ERLE$ obtained is lower: about $80dB$ compared to $110dB$ for all other algorithms in linear echo conditions (right side of Figure 2). We remark that, in all cases, the more the perturbations increase the lower the convergence time, since the $ERLE$ obtained is lower.

The plots in Figure 3(a) show the absolute convergence time in seconds but do not give an impression of the dynamic performance and, as already discussed, nor do they reflect the $ERLE$ that is eventually achieved. They are thus potentially misleading and for this reason we present in Figure 3(b) a plot of $ERLE$ against time, here for the NLMS algorithm only to better illustrate the dynamic and absolute performance. Figure 3(b) shows the $ERLE$ against time with linear echo only and added non-linear echo or noise at 52 and $26dB$.

These plots show that higher levels of perturbation result in lower levels of $ERLE$. In the case of linear echo (top profile) convergence is slow and is not even reached during the $60s$ illustrated. Crucially, though, the $ERLE$ is much higher than it is for non-linear and noise perturbations. However, in these cases the algorithm converges faster, but to a lower level (i.e. $\sim 55dB$ for non-linear echo with an $SNeR$ of $52dB$ and $\sim 20dB$ at $26dB$ $SNeR$, cf. $\sim 45dB$ for noise with an SNR of $52dB$ and $\sim 25dB$ at $26dB$ SNR). Hence consideration of the convergence time or maximum obtained $ERLE$ are not sufficient on their own to properly appreciate the performance of each approach. Similar profiles were obtained for the other adaptive filters and show an identical trend to that shown here for the NLMS algorithm albeit to different levels of $ERLE$. Finally, since all algorithms are shown to converge reasonably quickly in noise and non-linear environments it is of questionable advantage to focus effort on more computationally efficient algorithms; efforts are better directed toward the development of more robust algorithms. Indeed, more stable and straight forward algorithms, such as NLMS, are arguably of more interest for mobile terminal applications than their less stable and more computationally demanding alternatives.

4.3 Linear echo estimation

The assessment of performance with linear echo is commonly measured according to the system distance which is measured as $10log_{10}[|h(n) - \hat{h}(n)|^2/|h(n)|^2]$. It is less appropriate in the case of non-linear echo as the system distance shows only how well the linear echo path is estimated by the adaptive filter. In linear echo environments, the system distance indicates how effective is the echo cancellation. In the case of non-linear echo, the system distance indicates only how well the linear component is estimated but does not necessarily reflect the level of echo attenuation actually achieved. Figure 4 shows the behaviour of the NLMS system distance as a function of time. Whilst there are differences in exact values of system distance, the order of the profiles and general trends are indicative of performance for all the other filters. In general, the better the system distance, the better the $ERLE$. However, upon comparison of Figures 3(b) and Figure 4 we observe an apparent disparity. Figure 3(b) shows that performance with non-linear echo is generally better than that under additive noise with the same SNR, whereas Figure 4 shows almost no differences. This is due to the fact that system distance is only equivalent to $ERLE$ under the condition of total linearity. The $ERLE$ reflects the global performance according to the residual error, whereas the system distance reflects the accuracy of $\hat{h}(n)$. Equivalent values of system distance show

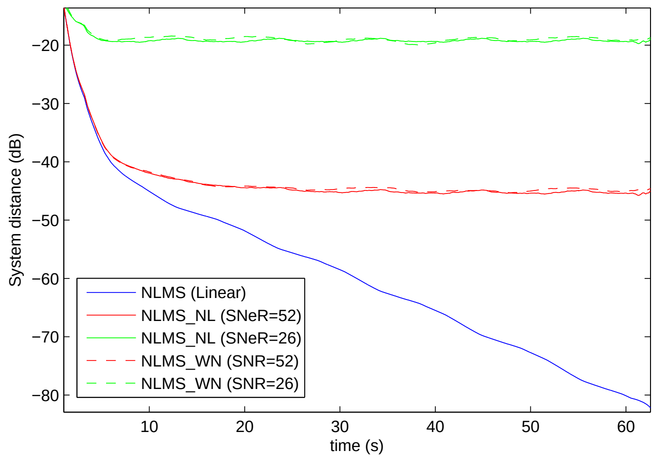


Figure 4: Plots of system distance (in dB) against time (in seconds) for the NLMS algorithm. Profiles are illustrated for linear echo and also for linear echo with either non-linear echo or added noise at two different levels.

that linear echo can be attenuated equally well with either non-linear echo or noise perturbations. The differences in the $ERLE$, however, show that non-linear echo perturbations are better attenuated than noise. This is due to the fact that in non-linear environments some of the non-linearities are indeed effectively attenuated by the adaptive filter even if the residual error is still higher than in the linear situation. This is due to the fact that adaptive filters aim to reduce the correlation (increase the orthogonality) between the error and the input signal. Since non-linear echo is correlated with the input signal it can also be attenuated, albeit only slightly. This is not the case with additive noise. This does not imply that adaptive filters are better in non-linear environments than they are in noisy environments as the adaptive filter does not aim to reduce the noise, but rather the echo signal which includes the non-linear component. In the next section we try to illustrate the implications of correlation, the relation to convolution and the potential of modelling non-linear environments as time-varying systems with the assumption of a time invariant echo path (or an echo path which varies more slowly than the speech signal.).

5 Discussion

The general equation for the LMS filter in a linear environment is given by:

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \mu\mathbf{x}(n)e(n), \tag{4}$$

where the error is equal to $d(n) - \hat{d}(n)$ and can be written as:

$$e(n) = \mathbf{h}^T\mathbf{x}(n) - \hat{\mathbf{h}}^T(n)\mathbf{x}(n) \tag{5}$$

If we denote the error in non-linear environments as $e_{nl}(n)$, then we can write:

$$e_{nl}(n) = \mathbf{h}^T[\mathbf{x}(n) + g(\mathbf{x}(n))] - \hat{\mathbf{h}}^T(n)\mathbf{x}(n), \quad (6)$$

where the subscript nl denotes non-linear and $g(x)$ is a function responsible for generating the non-linear echo components. If we suppose that the perturbation can be considered as an echo path variation then we can write the new time varying system as:

$$\mathbf{h}_{t.nl}(n) = \mathbf{h} + \Delta\mathbf{h}_{t.nl}(n), \quad (7)$$

where t denotes time variation, and where $\Delta\mathbf{h}_{t.nl}(n)$ is the variation around the true, static time invariant echo path **h** caused by non-linear distortions. We can then write the time varying non-linear component of $\mathbf{h}_{t.nl}$ as:

$$\mathbf{h}^T_{t.nl}(n)\mathbf{x}(n) = \mathbf{h}^T[\mathbf{x}(n) + g(\mathbf{x}(n))], \quad (8)$$

and thus:

$$\Delta\mathbf{h}^T_{t.nl}(n)\mathbf{x}(n) = \mathbf{h}^T g(\mathbf{x}(n)). \quad (9)$$

We note that Equation 9 is similar to the Wiener expression and can thus be written using the short term Fourier transformation as in [1] (Chapter 5):

$$\begin{aligned} \Delta\mathbf{H}_{t.nl}(f) &= \frac{\gamma_{h*g(x(n)),x(n)}(f)}{\gamma_{x(n)}(f)} \\ &= \mathbf{H}(f) \times \frac{\gamma_{G(X(f)),X(f)}(f)}{\gamma_{X(f)}(f)} \\ &= \mathbf{H}(f) \times c(f), \end{aligned} \quad (10)$$

where $c(f)$ is given by:

$$c(f) = \frac{\gamma_{G(X(f)),X(f)}(f)}{\gamma_{X(f)}(f)} \quad (11)$$

where $\gamma_{G(X(f)),X(f)}$ is the cross power spectral density of the non-linear signal component, $g(x(n))$ and the original signal, $x(n)$. $\gamma_{X(f)}$ is the power spectral density of $x(n)$. The ratio $c(f)$ is time varying and indicates the variability around the static echo path. For an intuitive explanation of Equation 10 let us assume that $g(x(n)) = x(n)$. As can easily be shown this leads to $\Delta\mathbf{h}_{t.nl}(n) = \mathbf{h}$. Consequently the AEC will converge to $2 \times \mathbf{h}$. This will not perturb the filter as $c(f)$ is a constant equal to 1.

Equation 10 shows that the time varying frequency components of $\Delta\mathbf{H}_{t.nl}(f)$ are obtained by the multiplication of $\mathbf{H}(f)$ and $c(f)$. Therefore they have the same frequency range. This explains why we observe a poorer estimation of the linear echo path in noisy environments than in non-linear environments, as described in our first hypothesis in Section 4.1. Since the adaptive filter aims to track the time varying echo path, it is natural that non-linear echo is slightly attenuated but that this leads to an under estimate of the linear echo and hence reduced attenuation of linear echo components. Another point highlighted by Equation 10 is that, the more the perturbation is correlated with the far-end signal, the higher is $c(f)$. As $c(f)$ is not constant in time and frequency this will lead to more variability in time and introduce more perturbation. The consideration of the environment as a problem of time varying filter tracking has been reported previously [9, 10]. The modelling of perturbations as a time varying system has the potential to give a better parametrization of the adaptive filter and is the subject of our on-going work.

6 Conclusions

This paper presents a new comparison of the effects of non-linearities and noise on four adaptive filters. Experimental results show that APA and NLMS have comparable behaviour in non-linear environments whereas FBLMS is badly affected. In noisy environments, however, there is little difference between each approach and, being less computationally demanding than the other approaches, FBLMS is an appealing solution in this case. We also show that, as the level of perturbations increase, performance decreases in both non-linear and noisy environments. Nevertheless, the echo canceller seems to be more robust to non-linearities than noise with a similar SNR (with the exception of the FBLMS algorithm). We show that the linear component of the echo path is under estimated but is as accurate in the case of non-linear echo as it is in noisy environments, again with a similar SNR. In addition, as the non-linear component is correlated with the far-end signal a fraction of non-linearities are effectively attenuated. Noise, in contrast, neither correlated, nor attenuated.

Finally we show how non-linear echo cancellation can be addressed as a problem of time varying filter estimation and that this approach has potential to bring improvements in non-linear environments. Given the correlation between the input speech signal and non-linear echo, this model illustrates why echo cancellers are less perturbed by non-linear echo than they are by additive noise. The model also introduces an alternative approach to cope with non-linear echo and is the subject of our on-going work.

References

[1] E. Hansler and G. Schmidt, *Acoustic Echo and Noise Control: A Practical Approach*. Address: Wiley-IEEE Press, 2004.

[2] M. H. Costa and J. C. M. Bermudez, "A noise resilient variable step-size LMS algorithm," *Signal Processing*, vol. 88, pp. 733–748, 2008.

[3] Y. Zhang, J. A. Chambers, W. Wang, P. Kendrick, and T. J. Cox, "A new variable step-size LMS algorithm with robustness to nonstationary noise," in *ICASSP 2007*, Honolulu, US , April 4 2007, pp. 257 – 260 .

[4] A. Stenger, L. Trautmann, and R. Rabenstein, “Nonlinear acoustic echo cancellation with 2nd order adaptive Volterra filters,” in *ICASSP 1999*, Phoenix, AZ, USA, March 15-19. 1999, pp. 877–880.

[5] K. Shi, X. Ma and G. T. Zhou, “A residual echo suppression technique for systems with non-linear acoustic echo paths,” in *ICASSP 2008*, Las Vegas, Nevada, USA, April 1-4. 2008, pp. 257–260.

[6] M. I. Mossi, N. W. D. Evans, and C. Beaugeant, “An assessment of linear adaptive filter performance with non-linear distortions,” to *in Proc. ICASSP 2010*, Dallas, Texas, March 14-19. 2010.

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

[8] M. Vondrasek, P. Pollak, “Methods for Speech SNR Estimation: Evaluation Tool and Analysis of VAD Dependency,” *RADIOENGINEERING -PRAGUE-*, vol. 1, pp. 11–25, Jan. 1999.

[9] A.K. Kohli, D.K. Mehra, “Tracking of time-varying channels using two-step LMS-type adaptive algorithm,” *Signal Processing, IEEE Transactions on* , vol. 54, pp. 2606 –2615, June. 2007.

[10] B. Farhang-Boroujeny, S. Gazor,“Performance of LMS-based adaptive filters in tracking a time-varying plant ,” *Signal Processing, IEEE Transactions on* , vol. 44, pp. 2868 –2871, August. 2002.