ADAPTIVE COMBINATION OF SECOND ORDER VOLTERRA FILTERS WITH NLMS AND SIGN-NLMS ALGORITHMS FOR NONLINEAR ACOUSTIC ECHO CANCELLATION

Cristian CONTAN, Mircea VAIDA, Tudor PALADE, Marina TOPA
Technical University of Cluj-Napoca, Romania, cristian.contan@bel.utcluj.ro

Abstract: In this paper, starting from a robust statistics (RS) adaptive approach presented in a previous work entitled the combined NLMS-Sign (CNLMS-S) adaptive filter, an automatic combination technique with similar performances is proposed. Thus, in order to obtain better performances in acoustic echo cancellation (AEC) setups than with the normalized least-mean square (NLMS) algorithm, in the CNLMS-S case the decision between the two algorithms (NLMS and Sign) is based on a set error threshold. The error threshold can be empirically determined or known a priori if the signal-to-noise ratio (SNR) value from the loudspeaker-enclosure-microphone (LEM) setup is available or if the local noise levels can be determined from the silences. Here, to overcome this shortcoming, an adaptive combination of the two algorithms involved in RS is highlighted, providing similar results regarding convergence and final misadjustment. Also, the need of the error threshold set by the user is removed, the combination being controlled only by a step-size parameter, independent on the LEM, constrained only by the stability range. The proposed method is compared to the CNLMS-S in nonlinear LEM setups using measured linear and quadratic Volterra kernels, tracking the behavior of the echo-return loss enhancement (ERLE) characteristic. As input sequences, audio signals with different PDFs are used and WGN is added as local noise. Simulation results justify the efficiency of the proposed method, both in convergence and steady-state error against the CNLMS-S and, implicitly the NLMS and the Sign-NLMS algorithms.

Keywords: *Acoustic echo cancellation, adaptive filters, Volterra models, the NLMS algorithm.*

I. INTRODUCTION

Adaptive algorithms are largely used in various audio applications, such as linear or nonlinear acoustic echo cancellation (AEC), active noise control (ANC) or double talk detection (DTD) [1 - 3] due to their simplicity in design and robustness. However, the efficiency of these applications is limited by a tradeoff between the speed of convergence and the steady-state error of the adaptive processes. As a straightforward consequence, the principal effort in designing an adaptive filter is to minimize the convergence time and the steady state error in the same time. Adaptive filter research is constantly searching for the optimum value of a cost function, usually square error-dependent, such as the benchmark normalized least mean square (NLMS) filter [4, 5]. The convergence of the later is improved by using parametric or nonparametric variable step-sizes [3] and exponentially weighted step sizes [6]. The search for faster adaptation processes conducted to the development of alternative cost functions and adaptations: the simplified LMS [4], the mean fourth error [7], the modified normalized least mean fourth (NLMF) [8] or the convex combination of the adaptive filters [9].

Thus, in the same manner of speeding the convergence of adaptive AEC structures, the starting point of this paper is the so-called combined NLMS-Sign (CNLMS-S) adaptive filter that uses a robust statistics algorithm discussed in [10, 11]. Its cost function is piecewise-defined: when the absolute value of the current error is less than a threshold, the cost function equals the square error; otherwise it takes the scaled value of the absolute error, using in each case the norm of the regressor. The procedure provides noticeable performances, but the drawback of the algorithm is the existence of an error threshold that should be measured, mathematically determined or empirically selected by the user. Thereby, our proposal is to eliminate the threshold by adaptively combining the outputs of an NLMS and a sign-NLMS adaptive filter, deploying a convex combination approach as in [9]. With the adaptive combination we aim each time at the progress of the best individual updating algorithm, being always focused on the smaller error. The CNLMS-S and the proposed adaptive combination of NLMS and Sign-NLMS (ACNLMS-S) are tested and compared in a nonlinear AEC application using second-order Volterra structures [12].

The paper is organized as follows. First, in section II, the linear and nonlinear AEC signal model is defined for its proper approximation level of the acoustic enclosure along with the standard individual updating algorithms: NLMS and Sign-NLMS. Next, in section III the CNLMS-S and ACNLMS-S methods are presented, highlighting the advantages and key points of the ACNLMS-S in a more productive practical adaptation of the Volterra kernels against the CNLMS-S, without the need of the error threshold. Section IV, containing simulation results, enables the efficiency of the ACNLMS-S for nonstationary regressors with different PDFs and WGN as local noise. Finally, conclusions are drawn in section V.

Manuscript received March 23, 2015; revised June 3, 2015

II. SIGNAL MODEL FOR ACOUSTIC ECHO CANCELLATION (AEC)

In mobile communications or teleconferencing applications that involve hands-free devices, an important issue is the cancellation of the undesired echoes that are being fed back to the far-end speaker. Acoustic echoes arise in the Loudspeaker-Enclosure-Microphone (LEM) configuration from the acoustic coupling between the loudspeaker and the microphone [13].

2.1 Linear AEC model

The general scheme for AEC is presented in Figure 1. First, let us neglect the blocks drawn with dotted lines, which are the nonlinear components in the AEC setup. The far-end signal $x[k]$ is aired by a loudspeaker and it is affected by multipath propagation of sound waves. A microphone picks up $d[k]$ that is composed of the distorted far-end signal and a local noise $n[k]$. A replica of the LEM is held by an adaptive filter (linear component). Filtering the far-end signal $x[k]$ with the LEM replica results in $\hat{y}[k]$ which is then subtracted from $d[k]$, thus resembling the residual error signal $e[k]$. The residual error is used to adaptively update the filter taps. When the adaptive filter has converged, for a fair adaptation, the error signal $e[k]$ should resemble the local noise $n[k]$ (in this case, in the absence of local speech).

The achieved echo reduction of an AEC system is evaluated here in terms of echo-return loss enhancement (ERLE):

$$\mathrm{ERLE}[k] := 10\log_{10}\frac{E\{d^2[k]\}}{E\{e^2[k]\}} \quad [\mathrm{dB}], \tag{1}$$

where $E\{\cdot\}$ denotes statistical expectation.

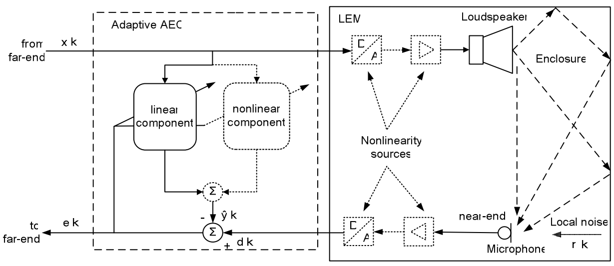

Figure 1. Acoustic echo cancellation setup

2.2 Nonlinear AEC model

However, in practical situations, nonlinear AEC models represent a more accurate tool that takes into consideration the nonlinear behavior of the available acoustic hardware. An illustration of a nonlinear acoustic echo path can be observed in Fig. 1, where sources of nonlinear distortions – digital to analog converters, amplifiers, analog to digital converter – are drawn with dotted lines. We are mostly interested in nonlinearities with memory, generated by the small loudspeaker driven at high volume and memoryless nonlinearities due to overdriven amplifiers [2, 14] that corrupt the microphone signal. Certain polynomial models were used to incorporate the type of nonlinearities from the LEM setup like Wiener and Volterra models as in [15] and cascaded adaptive filtering involving memoryless nonlinearity in [16].

In this work, we consider the second-order Volterra structure (2VF) as a sufficient tool to enclose the nonlinear distortions found in the typical LEM. As suggested in Fig. 1, it is composed by both the linear and nonlinear component and its output follows the expression:

$$\begin{aligned}\hat{y}[k] &= y_{2\mathrm{VF}}[k] = \\ &= \sum_{m_1=0}^{M_1-1}\hat{h}_{1,m_1}[k]x[k-m_1] + \sum_{m_1=0}^{M_2-1}\sum_{m_2=m_1}^{M_2-1}\hat{h}_{2,m_1,m_2}[k]x[k-m_1]x[k-m_2],\end{aligned} \tag{2}$$

where $y_{2\mathrm{VF}}[k]$ is the output of the 2VF, $\hat{h}_{1,m_1}[k]$ and $\hat{h}_{2,m_1,m_2}[k]$ represent the first and the second-order Volterra kernels of length M_1, respectively M_2. The Volterra kernels employed in (2) present a general symmetry which consists in using only terms with non-decreasing indices ($m_2 \geq m_1$). For simplicity, equation (2) can be rewritten in vector notation as:

$$y_{2\mathrm{VF}}[k] = \hat{\mathbf{h}}_1^T[k]\mathbf{x}_1[k] + \hat{\mathbf{h}}_2^T[k]\mathbf{x}_2[k] \tag{3}$$

using the associated vector definitions:

$$\begin{aligned}&\mathbf{x}_1[k] = (x[k],\ x[k-1],\ \ldots,\ x[k-M_1+1])^T;\\ &\hat{\mathbf{h}}_1[k] = (\hat{h}_{1,0}[k],\ \hat{h}_{1,1}[k],\ \ldots,\ \hat{h}_{1,M_1-1}[k])^T;\\ &\mathbf{x}_2[k] = (x^2[k],\ x[k]x[k-1],\ \ldots,\ x[k]x[k-M_2+1],\\ &\qquad x^2[k-1],\ x[k-1]x[k-2],\ \ldots,\ x^2[k-M_2+1])^T;\\ &\hat{\mathbf{h}}_2[k] = (\hat{h}_{2,0,0}[k],\ \hat{h}_{2,0,1}[k],\ \ldots,\ \hat{h}_{2,0,M_2-1}[k],\\ &\qquad \hat{h}_{2,1,1}[k],\ \hat{h}_{2,1,2}[k],\ \ldots,\ \hat{h}_{2,M_2-1,M_2-1}[k])^T,\end{aligned} \tag{4}$$

where $()^T$ denotes the transposing operation.

The first and the second-order Volterra kernels are obtained by adaptive processes, with the initiative to minimize the residual error signal $e[k]$ defined as:

$$e[k] = d[k] - y_{2\mathrm{VF}}[k]. \tag{5}$$

The nonlinear behavior of the LEM is modeled using the following representation of the microphone signal $d[k]$:

$$d[k] = \underbrace{\mathbf{h}_1^T[k]\mathbf{x}_1[k]}_{linear} + \alpha\underbrace{\mathbf{h}_2^T[k]\mathbf{x}_2[k]}_{nonlinear} + \beta\underbrace{n[k]}_{local\ noise}, \tag{6}$$

where $\mathbf{h}_1[k]$ and $\mathbf{h}_2[k]$ are the linear and the quadratic kernels written in vector notation as in (4). These kernels were obtained from measurements conducted in a low reverberant room with low-cost acoustic components [17]. The memory lengths of the kernels are equal to 320 and 64×64 taps, to include all the coefficients with significant nonzero values. The linear-to-nonlinear ratio ($\mathrm{LNLR} = 10\log_{10}\frac{\mathrm{E}\{(\mathbf{h}_1^T[k]\mathbf{x}_1[k])^2\}}{\alpha^2\mathrm{E}\{(\mathbf{h}_2^T[k]\mathbf{x}_2[k])^2\}}$ [dB]) and the signal-to-noise ratio ($\mathrm{SNR} = 10\log_{10}\frac{\mathrm{E}\{(\mathbf{h}_1^T[k]\mathbf{x}_1[k] + \alpha\mathbf{h}_2^T[k]\mathbf{x}_2[k])^2\}}{\beta^2\mathrm{E}\{n^2[k]\}}$ [dB]) quantities are kept constant throughout simulations by choosing proper values for parameters α and β.

With the polynomial description of the LEM and the Volterra model corresponding to the summarized nonlinearities, in the following, the necessary adaptive algorithms are briefly presented, ranging from the traditional ones to the proposed combined technique. The next pair of algorithms includes the individual ones that compose the RS technique and are implemented in the 2VF CNLMS-S adaptation from [11]. Their corresponding cost functions and kernels updating equations are briefly discussed.

2.3 Volterra kernel update with NLMS

The conventional derivation of the NLMS algorithm [18] considers the cost function to be the square error, and the step size is divided by the norm of the regressor. An alternate approach is to leave the step size parameter unchanged but the cost function is the square of the error divided by the norm of the regressor. Both approaches result in the same update process. In this work the second option of the cost function is used. The update of the second-order Volterra filter kernels that uses the NLMS adaptation consists in the minimization of a cost function defined as:

$$\hat{J}[k]^{(\mathrm{NLMS})} = \frac{e^2[k]}{2\left\|\mathbf{x}_i[k]\right\|^2}, \tag{7}$$

where $i = \{1, 2\}$. The update equations of the two Volterra kernels using the NLMS algorithm are presented in (8):

$$\begin{cases} \hat{\mathbf{h}}_1[k+1]^{(\mathrm{NLMS})} = \hat{\mathbf{h}}_1[k]^{(\mathrm{NLMS})} + \dfrac{\mu_1}{\left\|\mathbf{x}_1[k]\right\|^2 + \phi} e[k]\mathbf{x}_1[k], \\ \hat{\mathbf{h}}_2[k+1]^{(\mathrm{NLMS})} = \hat{\mathbf{h}}_2[k]^{(\mathrm{NLMS})} + \dfrac{\mu_2}{\left\|\mathbf{x}_2[k]\right\|^2 + \phi} e[k]\mathbf{x}_2[k], \end{cases} \tag{8}$$

where φ is a positive constant introduced to prevent division by zero or a very small value. The step-size parameters μ_1 and μ_2 control the convergence rate and steady-state error of the adaptive filter. To ensure stability of the system error norm in the mean, the step size parameters should be selected from the range (0, 2). While these values are constant, they are the leading cause of the compromise between convergence and final misalignment in adaptive structures that involve the NLMS update. Thus, different algorithms and adaptive techniques have been studied to lower the existing tradeoff of constant step sizes.

2.4 Volterra kernel update with sign-NLMS

The cost function of the sign-NLMS [4] uses the absolute value of the error signal as:

$$\hat{J}[k]^{(\mathrm{SIGN})} = \frac{|e[k]|}{\left\|\mathbf{x}_i[k]\right\|^2} - c, \tag{9}$$

where c is a positive constant. As in the NLMS case, starting from the steepest-descent recursion where the cost function is defined in this case by (9), the resulting Volterra kernel update equations are written:

$$\begin{cases} \hat{\mathbf{h}}_1[k+1]^{(SIGN)} = \hat{\mathbf{h}}_1[k]^{(SIGN)} + \dfrac{\mu_1}{\left\|\mathbf{x}_1[k]\right\|^2 + \phi} sign(e[k])\mathbf{x}_1[k], \\ \hat{\mathbf{h}}_2[k+1]^{(SIGN)} = \hat{\mathbf{h}}_2[k]^{(SIGN)} + \dfrac{\mu_2}{\left\|\mathbf{x}_2[k]\right\|^2 + \phi} sign(e[k])\mathbf{x}_2[k]. \end{cases} \tag{10}$$

Note that in the update equations the sign-NLMS algorithm uses a normalized step size, in comparison to the sign-LMS algorithm.

The two algorithms show raised convergence (SIGN) and low steady-state error (NLMS) but not simultaneously. Thereby, alternatives are desired to cover both convergence and low final misadjustment making use of the mentioned features of the individual algorithms.

III. COMBINED NLMS AND SIGN-NLMS ALGORITHMS

3.1 The combined NLMS-sign (CNLMS-S) Algorithm

In the previous section, the general forms of two adaptive processes were presented. The parabolic characteristic of (7) has high return values for large errors, while the linear characteristic of (9) is better suited in this situation, because it may be insensitive for large outliers of the error signal. The combination of both characteristics results in a robust statistic algorithm that was proposed in [10] and used on Volterra AEC structures in [11]. The combined cost function $J_{CNLMS\text{-}S}[k]$ of CNLMS-S in (11) is slightly modified from the one published in the prior art [10]. $J_{CNLMS\text{-}S}[k]$ is still defined as a piecewise function, where the user should select an error threshold $e_{\lim}$. According to this threshold, the sequent cost functions associated with their proposed methods are minimized:

$$\hat{J}[k]^{(CNLMS-S)} = \begin{cases} \hat{J}[k]^{(NLMS)}, & |e[k]| < e_{\lim} \\ \hat{J}[k]^{(SIGN)}, & |e[k]| \geq e_{\lim}. \end{cases} \tag{11}$$

As provided in [11] in the derivation process of the cost function, the update equations in the second-order Volterra case are:

$$\begin{cases} \hat{\mathbf{h}}_1[k+1]^{(CNLMS-S)} = \hat{\mathbf{h}}_1[k]^{(CNLMS-S)} + \dfrac{\mu_1}{\left\|\mathbf{x}_1[k]\right\|^2 + \phi} f(e[k], e_{\lim})\mathbf{x}_1[k], \\ \hat{\mathbf{h}}_2[k+1]^{(CNLMS-S)} = \hat{\mathbf{h}}_2[k]^{(CNLMS-S)} + \dfrac{\mu_2}{\left\|\mathbf{x}_2[k]\right\| + \phi} f(e[k], e_{\lim})\mathbf{x}_2[k], \end{cases} \tag{12}$$

where $f(\xi, \xi_{\lim})$ can be seen as an activation function and it is defined as:

$$f(\xi, \xi_{\lim}) = \begin{cases} \xi & \text{if } |\xi| < \xi_{\lim} \\ sign(\xi), & \text{if } |\xi| \geq \xi_{\lim}. \end{cases} \tag{13}$$

In *Figure* 2 the implementation of the CNLMS-S is depicted. The Digital Delay Line is used to align the input samples $x[k]$ into a vector that forms the regressor $\boldsymbol{x}_1[k]$, and to compute the vector $\boldsymbol{x}_2[k]$, the regressor of the nonlinear component of the Volterra filter. The regressors $\boldsymbol{x}_1[k]$ and $\boldsymbol{x}_2[k]$ are fed to the linear and nonlinear components. The output of each component is added and their sum returns the current output of the Volterra structure, $y_{2\mathrm{VF}}[k]$. The $y_{2\mathrm{VF}}[k]$ and $d[k]$ signals are used to compute the current error sample $e[k]$, which is the subject of further processing by the activation function f. There is no explicit formula for the threshold $e_{\lim}$. The selection of its value is subjective and may give inappropriate results. Its optimum value e_{opt} is related to the SNR of the system which is directly not accessible or unknown or when the local noise can be measured during

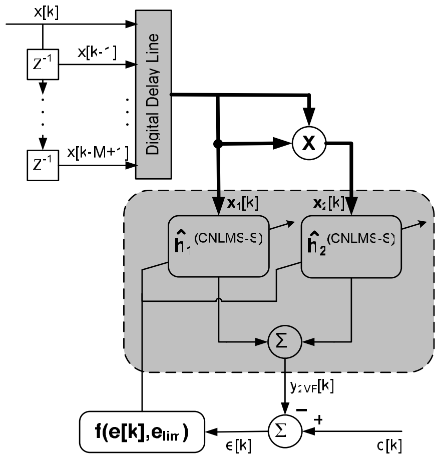

Figure 2. Second-order Volterra filter structure with CNLMS-S adaptation

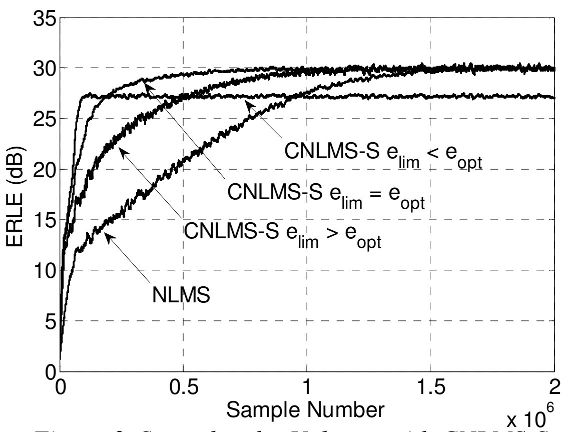

Figure 3. Second-order Volterra with CNLMS-S adaptation with different values of e_{lim}

silences. Here e_{lim} is empirically selected as in [11, 19]. In the following we will briefly show that the value of e_{lim} can affect the outcome of the adaptation process.

Let us consider a basic nonlinear AEC setup where the following quantities are applied: the SNR is 30 dB, the LNLR is 10 dB, the test signals have Gaussian distributions, the step-size parameters are μ_1 =0.002, μ_2 = 0.001. The ERLE of four system identification processes are depicted in *Figure 3*. One characteristic is obtained from a second-order Volterra structure with NLMS, which is considered as a benchmark. Three characteristics were obtained using 2VF with the CNLMS-S adaptation technique in the next situations: (i) e_{lim} is equal to e_{opt}; (ii) e_{lim} is less than the optimum value e_{opt}; (iii) e_{lim} is larger than e_{opt} [11]. Let us interpret the characteristics from Figure 3. In the first case, the optimum value of e_{lim} is set to e_{opt} = 0.05. A better ERLE convergence is achieved with the CNLMS-S adaptation than with the standard NLMS. Also, the steady-state ERLE is similar for both; the CNLMS-S method does not affect the stabilization value of the error. The steady-state ERLE is 30 dB, equal to the SNR value.

In the second case, the characteristic is obtained for e_{lim} = 0.005, a decision threshold way smaller than the optimum value e_{opt}. In this case, an expansion of the error range is gained, where the sign-NLMS adaptation procedure is applied. This will unreasonably update the Volterra kernels with the NLMS algorithm for error samples smaller than the optimum value. Analyzing the performances of the CNLMS-S adaptation method in comparison with the NLMS Volterra, the CNLMS-S version offers a better convergence than in the previous case (e_{lim} = e_{opt}). But it does not provide the expected steady-state ERLE, the adaptive filter stabilizes below the established SNR value.

In the third case, the decision threshold is chosen e_{lim} = 0.1. The steady-state ERLE for the CNLMS-S scenario returns to the imposed SNR amount and the obtained convergence rate is also higher than in the NLMS second-order Volterra case. However, the convergence speed is slightly reduced from the CNLMS-S from the first case. The developed convergence makes this threshold also valid but it misses out some high error values for the applicability of the sign-NLMS technique. This offers the observed convergence rate difference in contrast to the optimum threshold scenario. As observed, the selection process of the error threshold once again causes a compromise between convergence rate and steady-state ERLE. Although, the behavior of the method is significantly improved for threshold values close to the optimum, yet obtaining the optimum threshold is a challenging task in practical on-line AEC applications.

3.2 The adaptive combination of NLMS and sign (ACNLMS-S) Algorithm

As mentioned, a combination technique is regarded to eliminate the need of the error threshold and still keep the improved convergence rate and steady-state misadjustment of the CNLMS-S. To achieve this, the LMS-based convex combination approach [20] is applied on the nonlinear NLMS and Sign algorithms included in the CNLMS-S approach. The block structure of the proposed adaptive combination of NLMS and SIGN algorithms (ACNLMS-S) is illustrated in Figure 4. First, the outputs of the two individual block structures (NLMS and SIGN) are defined:

$$\begin{aligned}\hat{y}[k]^{(\text{NLMS})} &= \hat{y}_1[k]^{(\text{NLMS})} + \hat{y}_2[k]^{(\text{NLMS})} = \\ &\quad \hat{\mathbf{h}}_1^T[k]^{(\text{NLMS})}\mathbf{x}_1[k] + \hat{\mathbf{h}}_2^T[k]^{(\text{NLMS})}\mathbf{x}_2[k], \\ \hat{y}[k]^{(\text{SIGN})} &= \hat{y}_1[k]^{(\text{SIGN})} + \hat{y}_2[k]^{(\text{SIGN})} = \\ &\quad \hat{\mathbf{h}}_1^T[k]^{(\text{SIGN})}\mathbf{x}_1[k] + \hat{\mathbf{h}}_2^T[k]^{(\text{SIGN})}\mathbf{x}_2[k],\end{aligned} \tag{14}$$

together with their related residual error signals:

$$\begin{aligned}e[k]^{(\text{NLMS})} &= d[k] - \hat{y}[k]^{(\text{NLMS})}, \\ e[k]^{(\text{SIGN})} &= d[k] - \hat{y}[k]^{(\text{SIGN})}.\end{aligned} \tag{15}$$

Next, the two outputs of the Volterra structures are combined using a mixing parameter $\lambda[k]$ as follows, as in [21]:

$$\hat{y}_{2\text{VF}}[k] = \lambda[k]\hat{y}[k]^{(\text{SIGN})} + \left(1-\lambda[k]\right)\hat{y}[k]^{(\text{NLMS})}, \tag{16}$$

$$\lambda[k] = \text{sgm}(a[k]) = \frac{1}{1+e^{-a[k]}}, \tag{17}$$

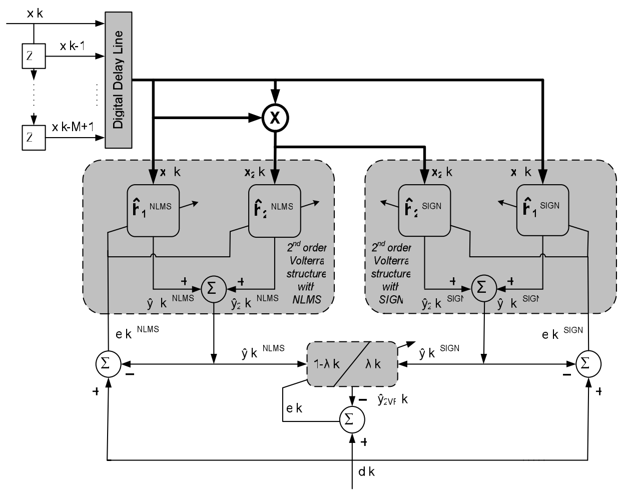

Figure 4. Second-order Volterra filter with ACNLMS-S adaptation

$$a[k+1]=a[k]+\mu\left(e[k]^{(\mathrm{NLMS})}-e[k]^{(\mathrm{SIGN})}\right)e[k]\lambda[k]\left(1-\lambda[k]\right), \quad (18)$$

where $e[k]=d[k]-\hat{y}_{2\mathrm{VF}}[k]$ and μ is the step-size parameter of the adaptive combination.

On account of the difference sign between the two independent residual errors $e[k]^{(\mathrm{NLMS})}-e[k]^{(\mathrm{SIGN})}$, the global output of the method will equal that with smaller associated error. Therefore, if $e[k]^{(\mathrm{NLMS})}>e[k]^{(\mathrm{SIGN})}$, the value of $a[k]$ will constantly grow, making $\lambda[k]$ to prone towards 1. For the nonce, $\hat{y}_{2\mathrm{VF}}[k]$ becomes approximately $\hat{y}[k]^{(\mathrm{SIGN})}$, with the smaller residual error. On the contrary, if $e[k]^{(\mathrm{NLMS})}<e[k]^{(\mathrm{SIGN})}$, $a[k]$ decays, becoming eventually negative, thus providing $\lambda[k]$ close to 0. The global outcome $\hat{y}_{2\mathrm{VF}}[k]$ will reach $\hat{y}[k]^{(\mathrm{NLMS})}$ due to the lower individual error. As a remark, when the two errors are equal, the global output maintains its previous value. Herewith, an automatic piecewise adaptation is accomplished without the necessity of the comparison to an error threshold, only comparing the two independent errors conjunctively.

IV. SIMULATION RESULTS

On account of the presented adaptation techniques, the performances of the ACNLMS-S technique will be compared to those of the CNLMS-S for a LEM setup, designed as in (6). The LNLR equals 10 dB and the SNR is set to 30 dB. Gradually, each individual second-order adaptive model is tested (NLMS and SIGN) in comparison to a linear NLMS algorithm. Nonstationary input signals are used, while AWGN is used as local noise. The step-size values are chosen $\mu_1=\mu_{\mathrm{lin}}=0.002$ to update the linear kernel and the reference linear NLMS filter, $\mu_2=0.001$ for the quadratic kernel and the threshold, estimated as in [19], is found to be optimum $e_{\mathrm{lim}}=0.01$ in all cases due to the constant SNR. The ERLE curves are averaged over 22000 samples, using a sliding window. As for the step-size value of the combination, $\mu=1$ is used.

Hence, in the next four figures, a recorded song fragment is used as source sequence. On the top of Figure 5, the independent nonlinear approaches that form the ACNLMS-S method are tested in terms of convergence and steady-state error while at the bottom, the microphone signal is depicted. Here, the mentioned tradeoff is clearly emphasized from the evolution of the two ERLE characteristics. The linear NLMS reaches only 10 dB ERLE due to the constant LNLR value. Figure 6 illustrates the convergence and misadjustment improvement of the CNLMS-S technique ($e_{\mathrm{lim}}=0.01$) from [11] compared to the ERLE behavior of the 2VF NLMS filter. It can be seen that the CNLMS-S is a fair solution to minimize the tradeoff, at the cost of the error threshold estimation. To remove the need of the threshold, a combination of the two algorithms from Figure 5 is proposed that only requires a convergence step-size μ to be set by the user, usually chosen positive, smaller than 1. For the ERLE characteristics from Figure 7, the ACNLMS-S method was applied to improve the performances of the particular filters compared in Figure 5. The ERLE characteristic of the proposed method follows the behavior of the Sign algorithm in convergence and that of the NLMS when reaching saturation, acting each time as the most suitable model. This process can be also examined from the development of the mixing parameter $\lambda[k]$ from bottom of Figure 7. It takes values close to 1 in convergence, yielding $\hat{y}_{2\mathrm{VF}}[k]\approx\hat{y}[k]^{(\mathrm{SIGN})}$. These values slowly drop to approximately 0 in the steady-state phase, giving $\hat{y}_{2\mathrm{VF}}[k]\approx\hat{y}[k]^{(\mathrm{NLMS})}$. In Figure 8 one can compare the evolutions of the CNLMS-S and ACNLMS-S approaches in terms of convergence rate and steady-state misadjustment. As remarked, the progressions of the two procedures are similar, here with slightly better convergence from the CNLMS-S, at the cost of obtaining the optimum e_{lim}.

The same ERLE charts for speech and instrumental music as source signals are depicted from Figure 9 to Figure 12, respectively form Figure 13 to Figure 16. One can again observe from the ERLE curves and also from the development of the mixing parameter $\lambda[k]$ that the ACNLMS-S removes the need of a threshold, gaining in the same time similar performances in ERLE evolution as the CNLMS-S approach in all three nonstationary input sequences. While both strategies (CNLMS-S and ACNLMS-S) improve the performances of the individual filters, the 2VF NLMS and 2VF Sign, only the ACNLMS-S approach eliminates the need of the error threshold.

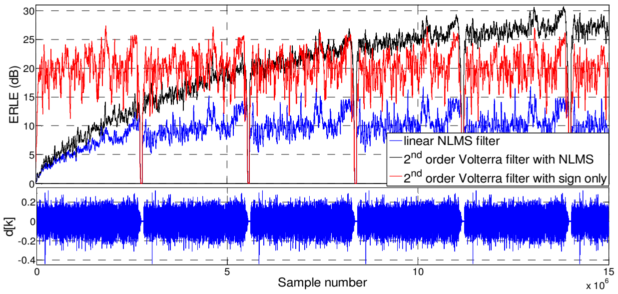

Figure 5. Evolution of ERLE for 2VF Sign, 2VF NLMS, linear NLMS (top) and the microphone signal (distorted song fragment plus noise) (bottom)

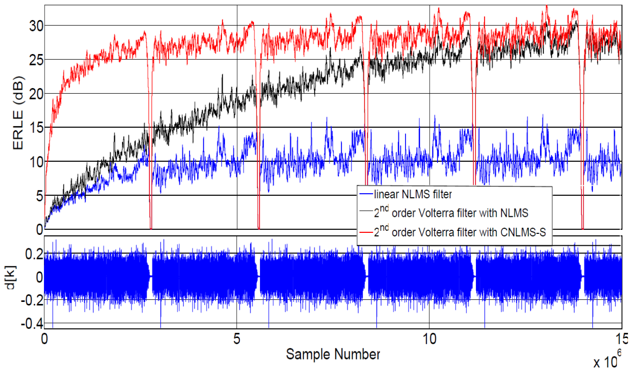

Figure 6. Evolution of ERLE for 2VF CNLMS-S, 2VF NLMS, linear NLMS and the microphone signal (distorted song fragment plus noise)

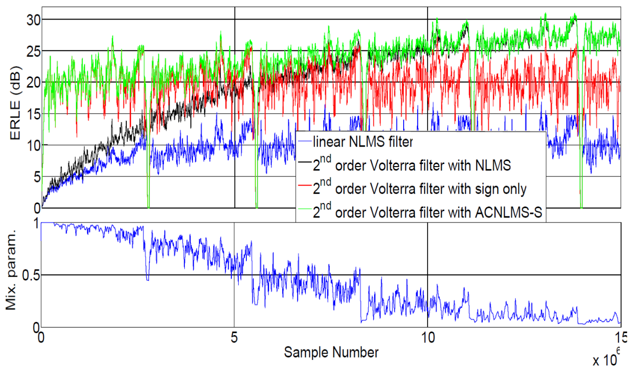

Figure 7. Evolution of ERLE for 2VF ACNLMS-S, 2VF Sign, 2VF NLMS, linear NLMS and the mixing parameter

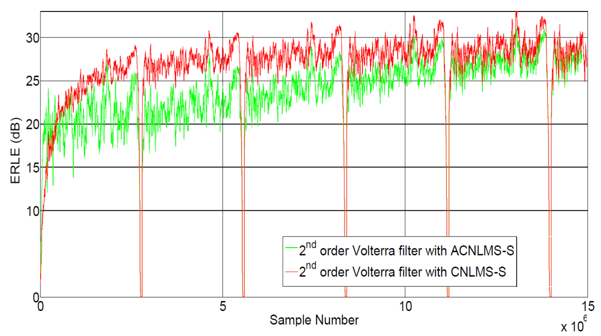

Figure 8. Comparison between 2VF ACNLMS-S and 2VF CNLMS-S for distorted song fragment plus noise as microphone signal

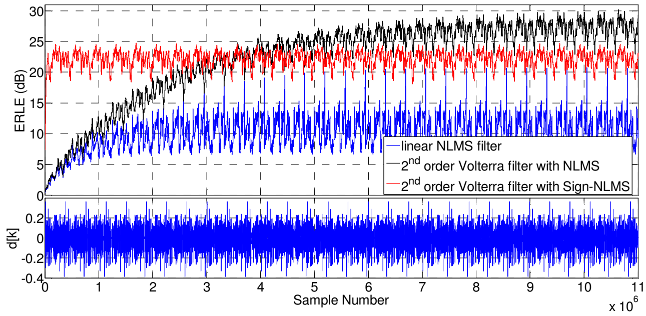

Figure 9. Evolution of ERLE for 2VF Sign, 2VF NLMS, linear NLMS and the microphone signal (distorted speech plus noise)

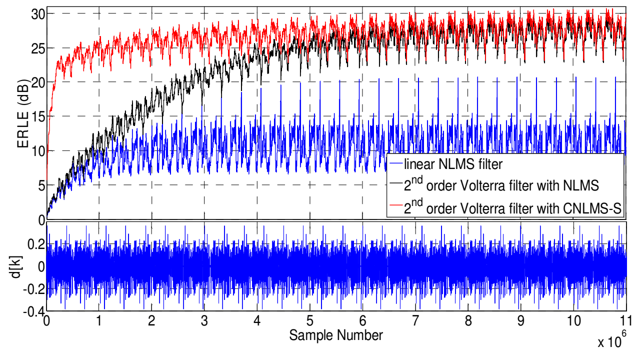

Figure 10. Evolution of ERLE for 2VF CNLMS-S, 2VF NLMS, linear NLMS and the microphone signal (distorted speech plus noise)

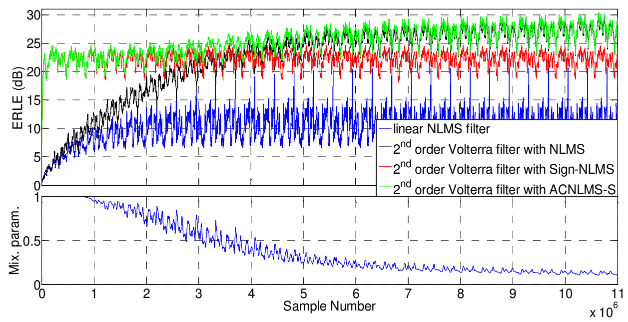

Figure 11. Evolution of ERLE for 2VF ACNLMS-S, 2VF Sign, 2VF NLMS, linear NLMS and the mixing parameter

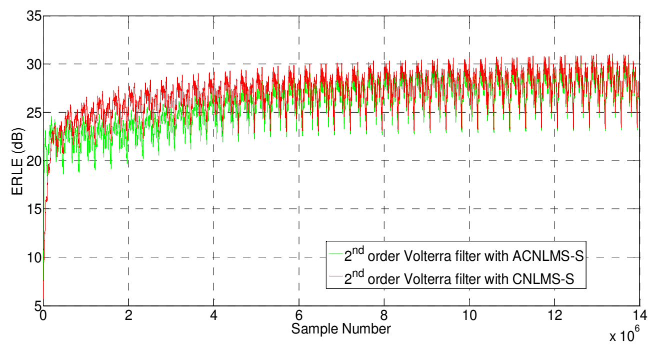

Figure 12. Comparison between 2VF ACNLMS-S and 2VF CNLMS-S for distorted speech plus noise as microphone signal

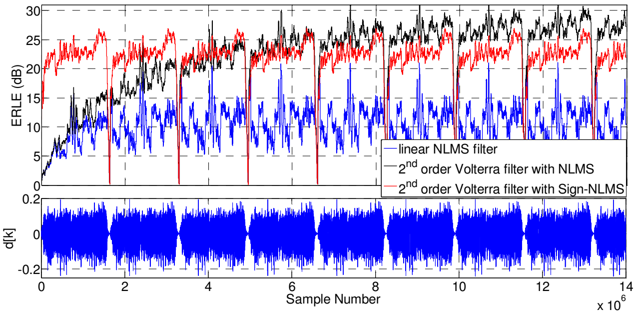

Figure 13. Evolution of ERLE for 2VF Sign, 2VF NLMS, linear NLMS and the microphone signal (distorted instrumental music plus noise)

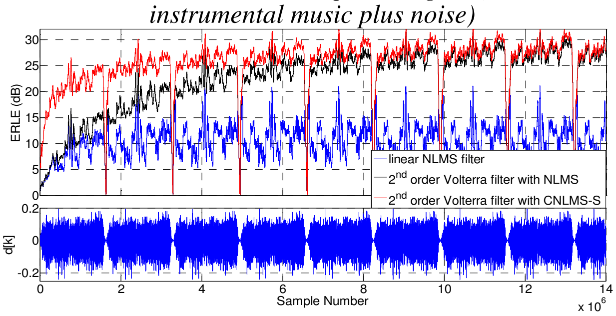

Figure 14. Evolution of ERLE for 2VF CNLMS-S, 2VF NLMS, linear NLMS and the microphone signal (distorted instrumental music plus noise)

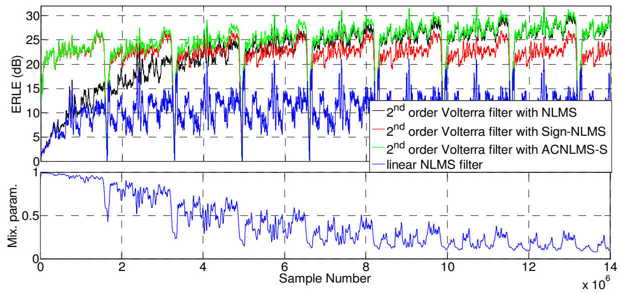

Figure 15. Evolution of ERLE for 2VF ACNLMS-S, 2VF Sign, 2VF NLMS, linear NLMS and the mixing parameter

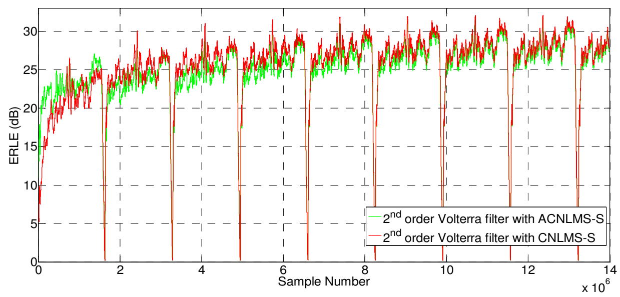

Figure 16. Comparison between 2VF ACNLMS-S and 2VF CNLMS-S for distorted instrumental music plus noise as microphone signal

V. CONCLUSIONS

In this paper, a robust second-order Volterra adaptive procedure was insight that aimed at improving the convergence rate and final misadjustment of conventional acoustic echo cancellation (AEC) methods and also those of a more recent one inspired from the robust-statistics (RS) appliance called the combined NLMS-Sign (CNLMS-S). In the CNLMS-S event, the adaptation of the kernels was carried out in a piecewise manner, relying on an optimum error threshold. Thus, according to the actual error value compared to the aforementioned threshold, the Sign NLMS algorithm was chosen in convergence, while the NLMS was desired in steady-state. The disadvantage of the CNLMS-S was the definition of its optimum adaptation limit, wherefore additional information was needed that in most practical applications is not available. However, to eliminate this issue, an adaptive combination of the two individual algorithms that define the RS tool was proposed (ACNLMS-S), while maintaining the performances of the CNLMS-S. The ACNLMS-S method was tested and compared to the other established structures in terms of the error power in nonlinear AEC scenarios involving measured Volterra kernels and different audio signals as input. In these circumstances, the efficiency of the proposed method was obvious in implementation but also in performance.

ACKNOWLEDGMENT

This paper was supported by the Post-Doctoral Programme POSDRU/159/1.5/S/137516, project co-funded from European Social Fund through the Human Resources Sectorial Operational Program 2007-2013. The authors would also like to thank Prof. Dr.-Ing. Walter Kellermann and Dr.-Ing. Marcus Zeller from the University of Erlangen-Nuremberg, Germany, for providing measured Volterra kernels.

REFERENCES

[1] E. Hänsler, G. Schmidt, *Topics in Acoustic Echo and Noise Control,* Springer, 2010.
[2] R. Martin, U. Heute, C. Antweiler, *Advances in Digital Speech Transmission*, Wiley, 2008.
[3] C. Paleologu, J.Benesty, S. Ciochina, *Sparse Adaptive Filters for Echo Cancellation*, Morgan & Claypool Publishers, UK, 2010.
[4] B. Farhang-Boroujeny, *Adaptive Filters Theory and Applications*, Wiley, Oct. 1998.
[5] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, New Jersey, 1996.
[6] S. Makino, Y. Kaneda, N. Koizumi, "Exponentially weighted stepsize NLMS adaptive filter based on the statistics of a room impulse response," *Speech and Audio Processing, IEEE Transactions on* , vol. 1, no. 1, pp. 101-108, Jan 1993.
[7] E. Walach, B. Widrow, "The Least Mean Fourth (LMF) Adaptive Algorithm and Its Family", *IEEE Transactions on Information Theory*, Vol. IT-30, No. 2, March 1984.
[8] C. Contan, B. S. Kirei, M. D. Topa, "Modified NLMF adaptation of Volterra filters used for nonlinear acoustic echo cancellation", *Signal Processing*, vol. 93, pp. 1152-1161, 2013.
[9] J. Arenas-Garcia, A. R. Figueiras-Vidal, A. H. Sayed, "Mean-square performance of a convex combination of two adaptive filters," *Signal Processing, IEEE Transactions on* , Vol. 54, No. 3, pp. 1078- 1090, March 2006.
[10] L. R. Vega, H. Rey, J. Benesty, S. Tressens, "A New Robust Variable Step-Size NLMS Algorithm," *Signal Processing, IEEE Transactions on*, vol. 56, no. 5, pp. 1878-1893, May 2008.
[11] C. Contan, M. D. Topa, B. S. Kirei, I. M. Kovacs, "A Derived Robust Statistics Approach For Adaptive Volterra Filters Applied In Nonlinear Acoustic Echo Cancellation Scenarios," *Acta Technica Napocensis – Electronics and Telecommunications*, ISSN 1221-6542, Volume 54, Number 1, pp. 1-8, 2013.
[12] M. Zeller, W. Kellermann, "Advances in Identification and Compensation of Nonlinear Systems by Adaptive Volterra Models", *Proc. 44th Asilomar Conference on Signals, Systems,*

and Computers, Pacific Grove (CA), USA, Nov. 2010.

[13] F. A. Everest, *The Master Handbook of Acoustics*, fourth edition, McGraw-Hill, 2001.

[14] B. S. Nollet, D. L. Jones, ″Nonlinear echo cancellation for hands-free speakerphones″, *Proc. NSIP'97*, Michigan USA, September 8-10, 1997.

[15] T. Ogunfunmi, *Adaptive Nonlinear System Identification: The Volterra and Wiener Model Approaches*, Springer, 2007.

[16] A. Stenger, W. Kellermann, R. Rabenstein, "Adaptation of Acoustic Echo Cancellers Incorporating a Memoryless Nonlinearity," *IEEE Workshop on Acoustic Echo and Noise Control (IWAENC)*, Pocono Manor, PA, USA, Sept. 1999.

[17] L. A. Azpicueta-Ruiz, M. Zeller, A. R. Figueiras-Vidal, J. Arenas-Garcia, W. Kellermann, "Adaptive Combination of Volterra Kernels and Its Application to Nonlinear Acoustic Echo Cancellation," *Audio, Speech, and Language Processing, IEEE Transactions on* , vol.19, no.1, pp. 97-110, Jan. 2011.

[18] C. Contan, *Acoustic Echo Cancellation in Nonlinear Systems*, PhD Thesis, Technical University of Cluj-Napoca, Dec. 2012.

[19] C. Contan, M. Topa, "Variable Step Size Adaptive Nonlinear Echo Canceller," *Electronics and Telecommunications (ISETC), 10th International Symposium on,* pp. 267 – 270, Nov. 2012.

[20] J. Arenas-Garcia, A. Figueiras-Vidal, A. H. Sayed, "Mean Square Performance of a Convex Combination of Two Adaptive Filters", *IEEE Transactions on Signal Processing*, Vol. 54, No.3, pp. 1078-1090, 2006.

[21] L. A. Azpicueta-Ruiz, A. R. Figueiras-Vidal, and J. Arenas-Garcia, "A normalized adaptation scheme for the convex combination of two adaptive filters," *in Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Process*., pp. 3301–3304, Las Vegas, Nov. 2008.