COEFFICIENT-DEPENDENT STEP-SIZE FOR ADAPTIVE SECOND-ORDER VOLTERRA FILTERS

Fabian Kuech and Walter Kellermann

Telecommunications Laboratory,
University of Erlangen-Nuremberg
Cauerstr. 7, D-91058 Erlangen, Germany
{kuech,wk}@LNT.de

ABSTRACT

In this contribution we propose a coefficient-dependent and time-variant step-size for the least mean square (LMS) algorithm applied to adaptive second-order Volterra filters. The optimum step-size is derived by introducing a novel optimality criterion which is given by the minimum mean squared error between the coefficient error of an adaptive Volterra filter coefficient and the respective LMS update term of that coefficient. As the optimum step-size includes statistical terms that are in general not accessible, we also present models for estimating these quantities for the application in nonlinear acoustic echo cancellation.

1. INTRODUCTION

The LMS algorithm represents a popular approach in linear and nonlinear adaptive filtering [1],[2], not at least in the context of acoustic echo cancellation. The general set-up of the acoustic echo cancellation problem is shown in Fig. 1. The

PSfrag replacements

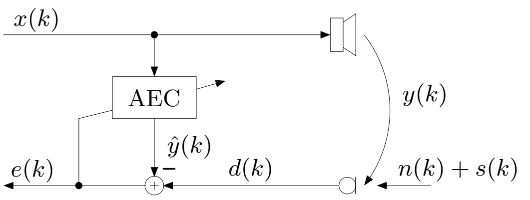

Figure 1: General set-up of the acoustic echo cancellation problem.

acoustic echo canceler (AEC) seeks to minimize the power of the error signal $e(n)$ by subtracting an estimate of the echo signal $\hat{y}(n)$ from the microphone signal $d(n)$. Here, we consider situations, where the loudspeaker systems introduce nonnegligible nonlinear distortions, e.g., caused by low-cost loudspeakers driven at high volume. Thus, the AEC has to be based on a nonlinear model to achieve an acceptable level of echo attenuation. A common approach to modeling the nonlinear behavior of loudspeakers is given by second-order Volterra filters [3] which are, therefore, considered here.

This paper is organized as follows. In Section 2 we derive an optimum step-size for the LMS algorithm applied to adaptive second-order Volterra filters based on a novel optimality criterion. In order to discuss its properties, we introduce a factorized version of the optimum step-size in Section 3. For an actual realization of the optimum step-size, we additionally propose estimates for certain statistical terms that are required but not measurable. It turns out that there is a strong link between the proposed approximated step-size and the proportionate normalized LMS (PNLMS) for second-order Volterra filters [4]. Finally, simulation results are presented in Section 4 in order to evaluate the performance of both, the optimum step-size and its approximated version compared to the normalized LMS (NLMS).

2. DERIVATION OF THE OPTIMUM STEP-SIZE

In the following we assume that the unknown echo path, i.e., the cascade of nonlinear loudspeaker and room impulse response, can be modeled by a finite-length second-order Volterra filter and, thus, the echo signal $y(k)$ can be expressed by

$$y(k) = y_1(k) + y_2(k), \tag{1}$$

where $y_1(k)$ represents the output of a linear filter

$$y_1(k) = \sum_{l=0}^{N_1-1} c_l^{(1)}(k)x(k-l), \tag{2}$$

and $y_2(k)$ is the output of a homogeneous quadratic Volterra kernel, i.e.,

$$y_2(k) = \sum_{l_1=0}^{N_2-1}\sum_{l_2=l_1}^{N_2-1} c_{l_1,l_2}^{(2)}(k)x(k-l_1)x(k-l_2). \tag{3}$$

Here, $c_l^{(1)}(k)$ and $c_{l_1,l_2}^{(2)}(k)$ denote the coefficients of the linear and quadratic Volterra kernel, respectively. The microphone signal

$$d(k) = y(k) + n(k) + s(k) \tag{4}$$

is composed of the echo signal $y(k)$, the noise signal $n(k)$ accounting for background noise, and the speech signal of a near-end talker $s(k)$. In the following we assume that $x(k)$, $n(k)$, and $s(k)$ are zero-mean, mutually statistically independent processes. Denoting the adaptive kernel coefficients of the echo canceler by $h_l^{(1)}(k)$ and $h_{l_1,l_2}^{(2)}(k)$, the output of the AEC reads

$$\hat{y}(k) = \hat{y}_1(k) + \hat{y}_1(k), \tag{5}$$

with

$$\begin{aligned} \hat{y}_1(k) &= \sum_{l=0}^{N_1-1} h_l^{(1)}(k)x(k-l), & (6)\\ \hat{y}_2(k) &= \sum_{l_1=0}^{N_2-1}\sum_{l_2=l_1}^{N_2-1} h_{l_1,l_2}^{(2)}(k)x(k-l_1)x(k-l_2). & (7)\end{aligned}$$

The coefficient errors of the linear and quadratic kernel coefficients are defined as

$$\begin{aligned} m_l^{(1)}(k) &= c_l^{(1)}(k) - h_l^{(1)}(k), & (8)\\ m_{l_1,l_2}^{(2)}(k) &= c_{l_1,l_2}^{(2)}(k) - h_{l_1,l_2}^{(2)}(k), & (9)\end{aligned}$$

respectively. The residual echos resulting from the misadjustment of the linear and the quadratic kernel coefficients of the AEC, respectively, yield

$$\varepsilon_1(k) = \sum_{l=0}^{N_1-1} m_l^{(1)}(k) x(k-l), \tag{10}$$

$$\varepsilon_2(k) = \sum_{l_1=0}^{N_2-1} \sum_{l_2=l_1}^{N_2-1} m_{l_1,l_2}^{(2)}(k) x(k-l_1) x(k-l_2). \tag{11}$$

The overall residual echo $\varepsilon(k) = y(k) - \hat{y}(k)$ can then be written as

$$\varepsilon(k) = \varepsilon_1(k) + \varepsilon_2(k). \tag{12}$$

The error signal $e(k) = d(k) - y(k)$ is then given by

$$e(k) = \varepsilon(k) + n(k) + s(k). \tag{13}$$

The following considerations are based on the so-called independence theory [1], i.e., we assume that the input signal $x(k)$ is a zero-mean, independent identically distributed (iid) process, and we assume that the coefficients of the adaptive Volterra filter are statistically independent of the input $x(k)$, implying that the coefficient errors $m_l^{(1)}(k)$ and $m_{l_1,l_2}^{(2)}(k)$ are also statistically independent of $x(k)$.

2.1 Optimum step-size for the linear kernel

The update of the linear kernel coefficients $h_l^{(1)}(k)$ applying the LMS algorithm [1] is given by

$$h_l^{(1)}(k+1) = h_l^{(1)}(k) + \mu_l^{(1)}(k) e(k) x(k-l), \tag{14}$$

where $\mu_l^{(1)}(k)$ denotes a positive step-size parameter to control the adaptation. However, the optimum update term for $h_l^{(1)}(k+1)$ would in fact be the respective coefficient error $m_l^{(1)}(k)$ if $c_l^{(1)}(k)$ is time-invariant. Thus, it is very intuitive to define a novel optimality criterion for the determination of the optimum step-size $\mu_l^{(1)}(k)$, namely, the mean squared error between $m_l^{(1)}(k)$ and the LMS update term for $h_l^{(1)}(k+1)$:

$$J_1 = E\left\{ \left[m_l^{(1)}(k) - \mu_l^{(1)}(k) e(k) x(k-l) \right]^2 \right\}. \tag{15}$$

Here, $E\{\cdot\}$ denotes expectation. The step-size $\mu_l^{(1)}(k)$ is then chosen to minimize the cost function J_1, yielding

$$\mu_{l,\mathrm{opt}}^{(1)}(k) = \frac{E\left\{ m_l^{(1)}(k) e(k) x(k-l) \right\}}{E\{e^2(k) x^2(k-l)\}}. \tag{16}$$

Note that the step-size according to (16) also assures convergence w.r.t. the mean squared coefficient error, i.e.,

$$E\left\{ \left[m_l^{(1)}(k+1) \right]^2 \right\} - E\left\{ \left[m_l^{(1)}(k) \right]^2 \right\} < 0. \tag{17}$$

It can be easily verified that (17) is fulfilled if $\mu_l^{(1)}(k)$ is chosen according to

$$0 < \mu_l^{(1)}(k) < 2\mu_{l,\mathrm{opt}}^{(1)}(k). \tag{18}$$

As $x(k)$ is a zero-mean iid process, and regarding the definitions (10)–(13), the numerator of (16) can be written as

$$E\left\{ m_l^{(1)}(k) e(k) x(k-l) \right\} = E\left\{ \left[m_l^{(1)}(k) \right]^2 \right\} E\left\{ x^2(k-l) \right\}. \tag{19}$$

Accounting for the mutual independence of $x(k)$, $n(k)$, and $s(k)$, the denominator of (16) can be expressed by

$$\begin{aligned} &E\left\{ e^2(k) x^2(k-l) \right\} = \\ &\quad E\left\{ \varepsilon^2(k) x^2(k-l) \right\} + E\left\{ n^2(k) + s^2(k) \right\} E\left\{ x^2(k-l) \right\}. \end{aligned} \tag{20}$$

For simplifying (16) we introduce the approximation

$$E\left\{ \varepsilon^2(k) x^2(k-l) \right\} \approx E\left\{ \varepsilon^2(k) \right\} E\left\{ x^2(k-l) \right\}. \tag{21}$$

It can be easily verified that in general (21) represents a valid approximation if $N_i \gg 1$, $i \in \{1, 2\}$. Inserting (21) in (20) and regarding (19) leads to a simplification of the optimum step-size (16) according to

$$\mu_{l,\mathrm{opt}}^{(1)}(k) = \frac{E\left\{ \left[m_l^{(1)}(k) \right]^2 \right\}}{E\{\varepsilon^2(k) + n^2(k) + s^2(k)\}}. \tag{22}$$

A detailed discussion of the optimum step-size and its realization is given in Section 3.

2.2 Optimum step-size for the quadratic kernel

Next, we consider the derivation of the corresponding optimum step-size for the quadratic kernel coefficients. The update equation for the quadratic kernel coefficients applying the LMS algorithm [2] is given by

$$h_{l_1,l_2}^{(2)}(k+1) = h_{l_1,l_2}^{(2)}(k) + \mu_{l_1,l_2}^{(2)}(k) e(k) x(k-l_1) x(k-l_2). \tag{23}$$

Analogously to Section 2.1, we obtain the optimum step-size for the quadratic kernel coefficients by minimizing

$$J_2 = E\left\{ \left[m_{l_1,l_2}^{(2)}(k) - \mu_{l_1,l_2}^{(2)}(k) e(k) x(k-l_1) x(k-l_2) \right]^2 \right\} \tag{24}$$

with respect to $\mu_{l_1,l_2}^{(2)}(k)$, yielding

$$\mu_{l_1,l_2,\mathrm{opt}}^{(2)}(k) = \frac{E\left\{ m_{l_1,l_2}^{(2)}(k) e(k) x(k-l_1) x(k-l_2) \right\}}{E\{e^2(k) x^2(k-l_1) x^2(k-l_2)\}}. \tag{25}$$

It can also be shown that convergence of the mean squared coefficient error can be assured for

$$0 < \mu_{l_1,l_2}^{(2)}(k) < 2\mu_{l_1,l_2,\mathrm{opt}}^{(2)}(k). \tag{26}$$

In order to obtain a simplification of the optimum step-size (25) that corresponds to (22), we have to analyze the statistical properties of the coefficient errors $m_{l_1,l_2}^{(2)}(k)$ for the considered application in acoustic echo cancellation. The results presented in [5] suggest that for white input $x(k)$ the coefficient errors of the linear kernel $m_l^{(1)}(k)$ are mutually orthogonal for different l. Furthermore, we note that the nonlinear acoustic echo path consists of the cascade of a nonlinear part (loudspeaker) and a linear part (room impulse response), where the linear component dominates the overall echo path. As reported in [3], the characteristics of the quadratic kernel of the corresponding Volterra filter representation of the echo path are then mainly determined by the characteristics of the impulse response of the linear kernel. Thus, it is reasonable to assume that the coefficients lying on the main diagonal of the quadratic kernel are also mutually orthogonal, i.e.,

$$E\left\{ m_{i,i}^{(2)}(k) m_{j,j}^{(2)}(k) \right\} = 0, \quad \text{for } i \neq j. \tag{27}$$

With the definitions (10)–(13) and regarding (27), we obtain

$$E\Big\{m^{(2)}_{l_1,l_2}(k)e(k)x(k-l_1)x(k-l_2)\Big\} = E\Big\{\Big[m^{(2)}_{l_1,l_2}(k)\Big]^2\Big\}\,E\big\{x^2(k-l_1)x^2(k-l_2)\big\}\,. \quad (28)$$

In order to simplify (25), we introduce an approximation similar to (21), i.e.,

$$E\big\{\varepsilon^2(k)x^2(k-l_1)x^2(k-l_2)\big\} \approx E\big\{\varepsilon^2(k)\big\}\,E\big\{x^2(k-l_1)x^2(k-l_2)\big\} \quad (29)$$

It can be shown that the approximation (29) is in general valid for $N_i \gg 1$, $i \in \{1,2\}$. The desired simplification of the optimum step-size for the quadratic kernel is finally obtained by regarding (28) and (29) for computing (25):

$$\mu^{(2)}_{l_1,l_2,\mathrm{opt}}(k) = \frac{E\Big\{\Big[m^{(2)}_{l_1,l_2}(k)\Big]^2\Big\}}{E\{\varepsilon^2(k)+n^2(k)+s^2(k)\}}. \quad (30)$$

Comparing (22) and (30), the analogy between the optimum step-sizes of the linear and quadratic kernel coefficients becomes obvious. However, it should be emphasized that for the derivation of (22) no assumptions with respect to the properties of the echo path are required, whereas (30) has been derived with explicit reference to acoustic echo cancellation.

3. DISCUSSION AND REALIZATION OF THE OPTIMUM STEP-SIZE

For a better understanding of the optimum step-size, we introduce the auxiliary step-size factors

$$\mu_{\mathrm{dt}}(k) = \frac{E\{\varepsilon^2(k)+n^2(k)\}}{E\{\varepsilon^2(k)+n^2(k)+s^2(k)\}}, \quad (31)$$

$$\mu_{\mathrm{bn}}(k) = \frac{E\{\varepsilon^2(k)\}}{E\{\varepsilon^2(k)+n^2(k)\}} \quad (32)$$

$$\mu_{\varepsilon_i}(k) = \frac{E\{\varepsilon_i^2(k)\}}{E\{\varepsilon^2(k)\}}, \quad i \in \{1,2\}. \quad (33)$$

Note that $\mu_{\mathrm{dt}}(k)$ and $\mu_{\mathrm{bn}}(k)$ represent kernel-independent step-size factors, whereas $\mu_{\varepsilon_1}(k)$ and $\mu_{\varepsilon_2}(k)$ are kernel-dependent but coefficient-independent step-size factors. Additionally, we define

$$\alpha^{(1)}_l(k) = \frac{E\Big\{\Big[m^{(1)}_l(k)\Big]^2\Big\}}{\sum\limits_{n=0}^{N_1-1} E\Big\{\Big[m^{(1)}_n(k)\Big]^2\Big\}\,E\{x^2(k-n)\}}, \quad (34)$$

where the denominator in (34) equals $E\{\varepsilon_1^2(k)\}$, and

$$\alpha^{(2)}_{l_1,l_2}(k) = \frac{E\Big\{\Big[m^{(2)}_{l_1,l_2}(k)\Big]^2\Big\}}{\sum\limits_{n_1=0}^{N_2-1}\sum\limits_{n_2=n_1}^{N_2-1} E\Big\{\Big[m^{(2)}_{n_1,n_2}(k)\Big]^2\Big\}\,E\{x^2(k-n_1)x^2(k-n_2)\}} \quad (35)$$

which represent coefficient-dependent step-size parameters. Note that the denominator in (35) corresponds to $E\{\varepsilon_2^2(k)\}$, if (27) is applied. The above definitions of step-size factors are used to factorize the optimum step-sizes according to

$$\mu^{(1)}_{l,\mathrm{opt}}(k) = \mu_{\mathrm{dt}}(k)\,\mu_{\mathrm{bn}}(k)\,\mu_{\varepsilon_1}(k)\,\alpha^{(1)}_l(k), \quad (36)$$

$$\mu^{(2)}_{l_1,l_2,\mathrm{opt}}(k) = \mu_{\mathrm{dt}}(k)\,\mu_{\mathrm{bn}}(k)\,\mu_{\varepsilon_2}(k)\,\alpha^{(2)}_{l_1,l_2}(k). \quad (37)$$

The influence of the different step-size parameters on the control of the adaptation of the Volterra filter coefficients is discussed in the following.

From the definition of $\mu_{\mathrm{dt}}(k)$ we notice that it accounts for double-talk (dt) situations, i.e., for $s(k) \neq 0$. In the echo cancellation context it is reasonable to realize $\mu_{\mathrm{dt}}(k)$ as an on/off switch in combination with a double-talk detector, i.e., $\mu_{\mathrm{dt}}(k) = 0$ if a near-end talker is active in order to avoid divergence of the adaptive filter coefficients, and $\mu_{\mathrm{dt}}(k) = 1$ otherwise. The step-size factor $\mu_{\mathrm{bn}}(k)$ controls the adaptation of the AEC with respect to the distortion introduced by the background noise (bn) $n(k)$. Several methods for the estimation of the combination $\mu_{\mathrm{c}}(k) = \mu_{\mathrm{dt}}(k)\mu_{\mathrm{bn}}(k)$ have been presented in [6] for linear adaptive filters. With some modifications, these methods can also be applied to nonlinear AEC and, thus, they are not discussed in more detail here.

For an interpretation of $\mu_{\varepsilon_i}(k)$, $i \in \{1,2\}$, we note that the error introduced by a misadjusted linear kernel acts as a distortion for the adaptation of the quadratic kernel and vice versa. Hence, the step-size factors $\mu_{\varepsilon_i}(k)$ can be interpreted as an adaptation control with respect to distortions caused by the different misadjusted Volterra kernels. As follows from (33), the determination of $\mu_{\varepsilon_i}(k)$ requires knowledge of $\varepsilon_1(k)$ and $\varepsilon_2(k)$ (or at least of the ratio of the second-order moments of $\varepsilon_i(k)$ and $\varepsilon(k)$) which is in general not accessible. Therefore, we introduce a model for estimating the respective second-order moments. More precisely we assume that the second-order moment of the residual echo of the linear kernel, i.e., $\varepsilon_1(k)$, is proportionate to the output of the adaptive linear kernel, i.e., $\hat{y}_1(k)$ and, analogously, the second-order moment of $\varepsilon_2(k)$ is assumed to be proportionate to $\hat{y}_2(k)$:

$$E\{\varepsilon_i^2(k)\} \approx \gamma(k)\left[\delta_i + \beta_i(k)\overline{|\hat{y}_i(k)|}\right], \quad (38)$$

where $\overline{|\hat{y}_i(k)|}$ denotes a smoothed version of the magnitude of $\hat{y}_i(k)$ and δ_i is a small constant which is required especially in the beginning of the adaptation, where $\hat{y}_i(k) = 0$ if the Volterra coefficients are initialized with zero. The approximation of $\mu_{\varepsilon_i}(k)$ is obtained by introducing (38) in (33), and assuming that the residual echo of the linear and quadratic kernel are orthogonal, yielding

$$E\{\varepsilon^2(k)\} = E\{\varepsilon_1^2(k)+\varepsilon_2^2(k)\}\,. \quad (39)$$

Note that (39) holds if the probability density function of the amplitude of $x(k)$ is an even function, as then $E\{x^3(k)\} = 0$. The proportionality factor $\gamma(k)$ does not have to be known explicitly, as (33) can be reduced with respect to $\gamma(k)$.

The coefficient-variable step-size parameters $\alpha^{(1)}_l(k)$ and $\alpha^{(2)}_{l_1,l_2}(k)$ can finally be used to speed-up the adaptation of coefficients that cause large coefficient errors. However, the coefficient errors are not known and, therefore, we have to use models for estimating the respective second-order moments. A common assumption is that large coefficient magnitudes also cause large error magnitudes (see, e.g., [7]). Consequently, we assume that the second-order moment of a certain coefficient error is proportionate to the magnitude of the corresponding adaptive coefficient:

$$E\Big\{\Big[m^{(1)}_l(k)\Big]^2\Big\} \approx \gamma_1(k)\left[\rho_1 + \lambda_1(k)\left|h^{(1)}_l(k)\right|\right], \quad (40)$$

$$E\Big\{\Big[m^{(2)}_{l_1,l_2}(k)\Big]^2\Big\} \approx \gamma_2(k)\left[\rho_2 + \lambda_2(k)\left|h^{(2)}_{l_1,l_2}(k)\right|\right] \quad (41)$$

where the proportionality factors $\gamma_1(k)$ and $\gamma_2(k)$ do not have to be specified explicitly. Furthermore, we replace the expectation operator with respect to $x(k)$ in (35) and (34) by the corresponding instantaneous values.

Interestingly, the approximations (40) and (41) directly lead to the concept of the PNLMS for second-order Volterra filters [4], if $\lambda_1(k)$ and $\lambda_2(k)$ are chosen according to

$$\lambda_1^{-1}(k) = \sum_{l=0}^{N_1-1} \left| h_l^{(1)}(k) \right|, \tag{42}$$

$$\lambda_2^{-1}(k) = \sum_{l_1=0}^{N_2-1} \sum_{l_2=l_1}^{N_2-1} \left| h_{l_1,l_2}^{(2)}(k) \right|. \tag{43}$$

It should be mentioned that the NLMS algorithm represents a special case for the step-size presented here. Assuming that the magnitudes of all coefficient errors of both, the linear and quadratic kernel are equal, i.e.,

$$\left| m_l^{(1)}(k) \right| = \left| m_{l_1,l_2}^{(2)}(k) \right| = \overline{m}(k), \quad \forall\, l, l_1, l_2, \tag{44}$$

the product terms $\mu_{\varepsilon_1}(k)\alpha_l^{(1)}(k)$ and $\mu_{\varepsilon_2}(k)\alpha_{l_1,l_2}^{(2)}(k)$ are replaced by a kernel-independent normalization factor $1/C(k)$,

$$C(k) = \sum_{l=0}^{N_1-1} x^2(k-l) + \sum_{l_1=0}^{N_2-1} \sum_{l_2=l_1}^{N_2-1} x^2(k-l_1)x^2(k-l_2), \tag{45}$$

resulting in the NLMS algorithm for Volterra filters [2]. The product term $\mu_{\mathrm{NLMS}}(k) = \mu_{\mathrm{dt}}(k)\mu_{\mathrm{bn}}(k)$ represents the step-size control with respect to $s(k)$ and $n(k)$ according to [6].

4. SIMULATION RESULTS

To evaluate the performance of the proposed optimum step-size, we present simulation results obtained for an acoustic echo cancellation application. In the first experiment the input has been colored noise with a power density spectrum (pds) corresponding to the long-term pds of speech and the echo path has been modeled by a second-order Volterra filter with memory lengths of 400 taps and 80 taps for the linear and quadratic kernel, respectively. The same memory lengths have been chosen for the adaptive filter, i.e., $N_1 = 400$ and $N_2 = 80$. As double-talk detection algorithms are not in the scope of this paper, we set $s(k) = 0$ in the following, implying $\mu_{\mathrm{dt}} = 1$. An SNR of 30 dB has been preset with respect to $n(k)$. As we are mainly interested in the improvement resulting from the kernel-dependent and the coefficient-dependent step-size parameters, a fixed value $\mu_{\mathrm{bn}} = 0.5$ is used. The echo return loss enhancement (ERLE) that has been achieved using the optimum step-size parameters $\mu_{\varepsilon_i}(k)$, $\alpha_l^{(1)}(k)$, and $\alpha_{l_1,l_2}^{(2)}(k)$ (assuming that the coefficient errors are known) is shown in Fig. 2 together with the result obtained for an NLMS algorithm with step-size $\mu_{\mathrm{NLMS}} = \mu_{\mathrm{bn}} = 0.5$. Furthermore, the ERLE graph resulting from the approximation of the optimum step-size using (38)-(41) are shown there. The model parameters for (38) have been $\delta_i = 0.001$ and $\beta_i = 1$. Following [4], $\rho_1 = 1/2N_1^2$ and $\lambda_1(k)$ according to (42) have been chosen for the linear kernel. For the quadratic kernel $\rho_2 = 1/(N_2^2 + N_2)$ and $\lambda_2(k)$ according to (43) have been used. As can be noticed from Fig. 2, both the optimum step-size and its approximation clearly outperform the NLMS algorithm in terms of convergence speed. The optimum step-size is significantly superior to the approximated version as it does not rely on model assumptions. The ERLE graphs presented in Fig. 2 (bottom) base on recorded speech data from a low-cost loudspeaker in an enclosure with low reverberation, where the same parameters have been applied as in the previous experimental set-up. Note that only the results obtained for the approximated step-sizes and the NLMS algorithm are shown here, as the physical echo path is not known and, thus, the optimum step-sizes (22) and (30) cannot be determined. Again, a remarkable increase in convergence speed is achieved by the proposed coefficient-dependent step-size compared to the NLMS algorithm.

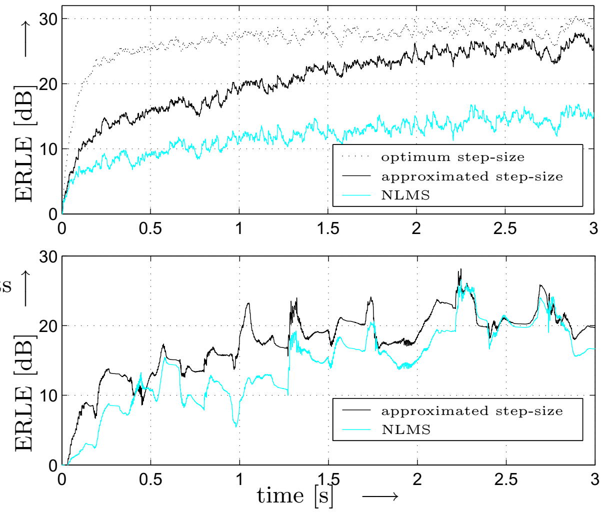

Figure 2: ERLE for LMS with optimum step-size and approximated step-size compared to NLMS for stationary colored noise input (top), and recorded speech data (bottom).

5. CONCLUSION

We presented a novel step-size control for the LMS algorithm applied to adaptive second-order Volterra filters. The derived optimum step-size has then been approximated, as a direct implementation is not possible due to non-accessible statistical terms included in the definition of the optimum-step-size. Simulation results with respect to nonlinear acoustic echo cancellation using adaptive second-order Volterra filters have shown that the proposed step-size control for the LMS algorithm leads to a significant improvement of convergence speed compared to an NLMS algorithm.

REFERENCES

[1] S. Haykin, *Adaptive Filter Theory*, New Jersey: Prentice Hall, 2001.

[2] V. J. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*, New York: Wiley, 2000.

[3] M. J. Reed and M. O. Hawksford, "Practical modelling of nonlinear audio systems using Volterra series," *Journal of the Audio Engineering Society* Preprint 4264, 1996.

[4] F. Kuech and W. Kellermann, "Proportionate NLMS algorithm for second-order Volterra filters and its application to nonlinear echo cancellation," in *Proc. Workshop on Acoustic Echo and Noise Control*, Kyoto, Sept. 2003.

[5] W. Kellermann, "Some properties of echo path impulse responses of microphone arrays and consequences for acoustic echo cancellation," in *Proc. Workshop on Acoustic Echo and Noise Control*, Røros, June 1995.

[6] A. Mader, H. Puder, and G. U. Schmidt, "Step-size control for acoustic echo cancellation filters - an overview," *Signal Processing*, vol. 80, pp. 1697–1719, Sept. 2000.

[7] M. Rupp and J. Cezanne, "Robustness conditions of the LMS algorithm with time-variant matrix step-size," *Signal Processing*, vol. 80, pp. 1787–1794, Sept. 2000.