A NOVEL MULTIDELAY ADAPTIVE ALGORITHM FOR VOLTERRA FILTERS IN DIAGONAL COORDINATE REPRESENTATION

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 present a novel algorithm for the efficient computation of the output of Volterra filters in diagonal coordinate representation (DCR), while allowing for different memory length for each kernel. This is achieved by extending partitioned block filtering methods for fast convolution in the discrete Fourier transform (DFT) domain to Volterra filters. It is shown that the DCR is particularly favorable for modeling the cascade of a Volterra filter followed by a linear filter, as required for applications such as nonlinear acoustic echo cancellation. To obtain a corresponding adaptive structure of the proposed approach, we introduce a generalization of a known DFT-domain adaptive algorithm for linear systems to Volterra filters.

1. INTRODUCTION

Volterra filters are known to be capable to model a large variety of real world nonlinear systems and thus, there is a wide range of applications of adaptive Volterra filters [1]. In the following we will refer to acoustic echo cancellation as a specific application. The general set-up of the acoustic echo cancellation problem is shown in Fig. 1. The acoustic echo canceler (AEC) seeks to minimize

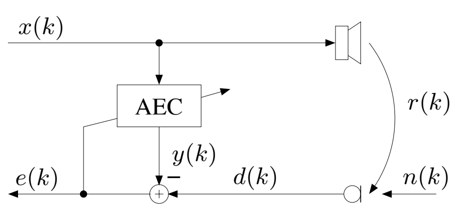

Fig. 1. General set-up of the acoustic echo cancellation problem.

the power of the error signal $e(n)$ by subtracting an estimate of the echo signal $y(n)$ from the microphone signal $d(n)$. Standard approaches for the cancellation of acoustic echos rely on the assumption that the echo path to be identified can be modeled by a linear filter. However, in some practical situations loudspeaker systems introduce nonnegligible nonlinear distortions, e.g., caused by low-cost loudspeakers driven at high volume. With this nonlinear distortion, the performance of a linear acoustic echo canceler degrades substantially and thereby, greatly impairs quality of voice communication. Thus, nonlinear models have to be considered. A common approach to modeling the nonlinear behavior of loudspeakers is given by second-order Volterra filters, where the required memory length of the linear kernel is larger than that of the quadratic kernel [2].

In this paper we first introduce an efficient DFT-domain algorithm for the computation of the output of a P-th order Volterra filter in DCR allowing for different memory length for each kernel. It is shown that the DCR is especially efficient for modeling the cascade of a Volterra filter followed by a linear filter, as required for the compensation of nonlinear acoustic echos described above. Then, an adaptive realization of this approach is introduced that can be considered as a generalization of the well-known DFT-domain adaptive algorithm for linear systems [3] to P-th order Volterra filters. The performance of the adaptive algorithm for the second-order case is then evaluated for the acoustic echo cancellation application using real measured data.

2. VOLTERRA FILTERS IN DIAGONAL COORDINATE REPRESENTATION (DCR)

The input/output relation of a finite length P-th order Volterra filter is given by

$$y(k) = \sum_{p=1}^{P} y_p(k) \tag{1}$$

where the output $y_p(k)$ of the p-th order Volterra kernel reads

$$y_p(k) = \sum_{n_{p,1}=0}^{N_p-1} \cdots \sum_{n_{p,p}=n_{p,p-1}}^{N_p-1} h_{\mathbf{n}_p} \prod_{i=1}^{p} x(k-n_{p,i}). \tag{2}$$

$\mathbf{n}_p = [n_{p,1}, n_{p,2}, \ldots, n_{p,p}]$ represents the index vector of the p-th order Volterra kernel coefficients corresponding to a p-dimensional Cartesian coordinate system. Note that the memory lengths N_p of the Volterra kernels can in general be different for each order p. Following [4], the corresponding diagonal coordinate representation of (2) is obtained by a coordinate transformation of the index vector $\mathbf{n}_p$. For the DCR we define the following two index vectors

$$\begin{aligned} \underline{\mathbf{r}}_p &= [r_{p,1}, r_{p,2}, \ldots, r_{p,p-1}], && (3)\\ \mathbf{r}_p(l) &= [l, r_{p,1}+l, \ldots, r_{p,p-1}+l], && (4) \end{aligned}$$

where $\underline{\mathbf{r}}_p$ references a certain diagonal that is parallel to the main diagonal ($r_{p,i} = 0\ \forall i$) of the corresponding p-dimensional Cartesian coordinate system described by the original index vector $\mathbf{n}_p$. Accordingly, $\mathbf{r}_p(l)$ represents a certain position on the diagonal $\underline{\mathbf{r}}_p$, i.e., it references a certain coefficient of the Volterra filter. Note that for the linear kernel $\underline{\mathbf{r}}_1 = [\ ]$, and $\mathbf{r}_1(l) = l$. Introducing the input signal of the diagonal $\underline{\mathbf{r}}_p$ according to

$$x_{\underline{\mathbf{r}}_p}(k) = x(k) \prod_{i=1}^{p-1} x(k - r_{p,i}), \tag{5}$$

where $x_{\underline{\mathbf{r}}_1}(k) = x(k)$, we rewrite (2) as

$$y_p(k) = \sum_{r_{p,1}=0}^{N_p-1} \cdots \sum_{r_{p,p-1}=r_{p,p-2}}^{N_p-1} y_{\underline{\mathbf{r}}_p}(k), \tag{6}$$

$$y_{\underline{\mathbf{r}}_p}(k) = \sum_{l=0}^{L_p(\underline{\mathbf{r}}_p)-1} h_{\mathbf{r}_p(l)}\, x_{\underline{\mathbf{r}}_p}(k-l). \tag{7}$$

where the length of $h_{\mathbf{r}_p(l)}$

$$L_p(\underline{\mathbf{r}}_p) = N_p - r_{p,p-1} \tag{8}$$

depends on both, the kernel order p, and the actual value of $r_{p,p-1}$. Note that for the linear kernel $y_1(k) = y_{\underline{\mathbf{r}}_1}(k)$ and $L_1(\underline{\mathbf{r}}_1) = N_1$. Obviously, $y_{\underline{\mathbf{r}}_p}(k)$ can be considered as the output of the linear FIR filter $h_{\mathbf{r}_p(k)}$ with the input signal $x_{\underline{\mathbf{r}}_p}(k)$:

$$y_{\underline{\mathbf{r}}_p}(k) = h_{\mathbf{r}_p(k)} * x_{\underline{\mathbf{r}}_p}(k). \tag{9}$$

Considering (6), $y_p(k)$ can therefore be interpreted as the output of a linear multiple input/single output (MISO) system, where each diagonal with index vector $\underline{\mathbf{r}}_p$ corresponds to one linear channel with input $x_{\underline{\mathbf{r}}_p}(k)$. Extending this interpretation to the computation of the output of the Volterra filter according to (1), $y(k)$ can be considered as the output of a special MISO system featuring a combination of P multichannel structures, where each channel corresponds to one particular diagonal of the DCR.

2.1. Application to cascaded structures

In the following, we show that the DCR is especially suited to represent certain nonlinear cascaded systems. Fig. 2 illustrates the configuration of a P-th order Volterra filter $h_{\mathbf{r}_p(l)}$ followed by a linear FIR filter with coefficients c_l. As the convolution is a linear

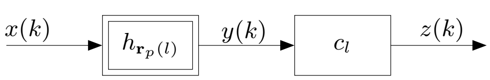

Fig. 2. Cascaded structure consisting of a P-th order Volterra filter followed by a linear FIR filter.

operation, the computation of the output of the cascaded structure directly follows from (1), (6), and (9):

$$z(k) = \sum_{p=1}^{P} \sum_{r_{p,1}=0}^{N_p-1} \cdots \sum_{r_{p,p-1}=r_{p,p-2}}^{N_p-1} z_{\underline{\mathbf{r}}_p}(k), \tag{10}$$

where the output $z_{\underline{\mathbf{r}}_p}(k)$ of each DCR-channel yields

$$z_{\underline{\mathbf{r}}_p}(k) = c_k * h_{\mathbf{r}_p(k)} * x_{\underline{\mathbf{r}}_p}(k) \tag{11}$$

$$= v_{\mathbf{r}_p(k)} * x_{\underline{\mathbf{r}}_p}(k). \tag{12}$$

We note from (10)-(12) that $z(k)$ can be considered as the output of a specific P-th order Volterra filter $v_{\mathbf{r}_p(k)}$ with input $x(k)$, where the number and the position of the diagonals is not changed compared to the Volterra filter $h_{\mathbf{r}_p(k)}$. However, the length of the filter in each DCR-channel with index vector $\underline{\mathbf{r}}_p$ is increased according to

$$\widetilde{L}_p(\underline{\mathbf{r}}_p) = L_p(\underline{\mathbf{r}}_p) + N_c - 1, \tag{13}$$

where N_c denotes the length of the linear filter.

In the AEC context (Fig. 1), the echo path to be modeled by the AEC is composed of a loudspeaker, the transfer function between loudspeaker and microphone, and the microphone itself. The transfer function between loudspeaker and microphone (i.e., the room impulse response) and the characteristics of the microphone can usually be considered as linear and, thus, can be modeled by an FIR filter. As the nonlinear behavior of the loudspeaker can be taken into account by using a second-order Volterra filter, the overall structure corresponds to Figure 2 for $P = 2$.

3. MULTIDELAY VOLTERRA FILTERS IN DCR

In this section we introduce an efficient computation of the output of a Volterra filter in DCR applying partitioned block methods for fast convolution in the DFT domain. Following [3], a block partitioned version of (9) is obtained by partitioning $h_{\mathbf{r}_p(l)}$ into $B(\underline{\mathbf{r}}_p)$ blocks of length N,

$$h_{\mathbf{r}_p(i),b} = h_{\mathbf{r}_p(l)}\big|_{l=i+bN}, \tag{14}$$

and defining the input signal of each partition according to

$$x_{\underline{\mathbf{r}}_p,b}(k) = x_{\underline{\mathbf{r}}_p}(k - bN). \tag{15}$$

Then, (9) can be rewritten as

$$y_{\underline{\mathbf{r}}_p}(k) = \sum_{b=0}^{B(\underline{\mathbf{r}}_p)-1} h_{\mathbf{r}_p(k),b} * x_{\underline{\mathbf{r}}_p,b}(k). \tag{16}$$

The number of partitions $B(\underline{\mathbf{r}}_p)$ of each diagonal at position $\underline{\mathbf{r}}_p$ has to be chosen such that

$$\left[B(\underline{\mathbf{r}}_p) - 1\right] N < L_p(\underline{\mathbf{r}}_p) \leq B(\underline{\mathbf{r}}_p) N. \tag{17}$$

Accounting for (13), we allow $L_p(\underline{\mathbf{r}}_p) > N_p$ in the following. The values for N_p, i.e., the maximum distance of a diagonal to the main diagonal of the p-order kernel, and the length along the main diagonal, i.e., $L_p(\underline{\mathbf{r}}_p = \mathbf{0})$, can then be used to specify the region of support of a Volterra filter. In order to exploit fast convolution techniques in the DFT domain via block processing for the computation of the convolutions appearing in (16) we define the following signal vectors composed of overlapping signal blocks:

$$\mathbf{y}_{\underline{\mathbf{r}}_p}(m) = \left[y_{\underline{\mathbf{r}}_p}(mR), \ldots, y_{\underline{\mathbf{r}}_p}(mR + N - 1)\right]^T, \tag{18}$$

$$\mathbf{x}_{\underline{\mathbf{r}}_p,b}(m) = \left[x_{\underline{\mathbf{r}}_p,b}(mR - N), \ldots, x_{\underline{\mathbf{r}}_p,b}(mR + N - 1\right]^T \tag{19}$$

where m represents a block time index with $k = mR$. The positive integer $R = N/\alpha$ equals the number of new samples of successive signal blocks. The parameter α is usually referred to as overlapping factor [3]. Note that $\mathbf{y}_{\underline{\mathbf{r}}_p}(m)$ has length N, while the input signal vector $\mathbf{x}_{\underline{\mathbf{r}}_p,b}(m)$ has the length $2N$. Furthermore, we combine the coefficients $h_{\mathbf{r}_p(l),b}$ of each partition to obtain the vectors

$$\mathbf{h}_{\underline{\mathbf{r}}_p,b} = \left[h_{\mathbf{r}_p(bN)}, h_{\mathbf{r}_p(bN+1)}, \ldots, h_{\mathbf{r}_p(bN+N-1)}\right]^T. \tag{20}$$

Note that

$$h_{\mathbf{r}_p(l),b} \equiv 0 \quad \text{for} \quad l + bN \geq L_p(\underline{\mathbf{r}}_p), \tag{21}$$

i.e., the definition of $\mathbf{h}_{\underline{\mathbf{r}}_p,b}$ according to (20) implies zero-padding for $b = B(\underline{\mathbf{r}}_p) - 1$ if $L_p(\underline{\mathbf{r}}_p) < B(\underline{\mathbf{r}}_p)N$. As we want to apply the overlap/save method [5], we define the DFT-domain vectors of length $M = 2N$ according to

$$\begin{aligned} \mathbf{Y}_{\underline{\mathbf{r}}_p}(m) &= \mathbf{F}_M \left[\mathbf{0}_{N\times 1}^T \; \mathbf{y}_{\underline{\mathbf{r}}_p}^T(m)\right]^T, & (22)\\ \mathbf{X}_{\underline{\mathbf{r}}_p,b}(m) &= \mathbf{F}_M \, \mathbf{x}_{\underline{\mathbf{r}}_p,b}(m), & (23)\\ \mathbf{H}_{\underline{\mathbf{r}}_p,b} &= \mathbf{F}_M \left[\mathbf{h}_{\underline{\mathbf{r}}_p,b}^T \; \mathbf{0}_{N\times 1}^T\right]^T, & (24) \end{aligned}$$

where $\mathbf{F}_M$ denotes the $M \times M$ DFT matrix and $\mathbf{0}_{N\times 1}$ represents the $N \times 1$ zero vector. Thus, we obtain

$$\mathbf{Y}_{\underline{\mathbf{r}}_p}(m) = \sum_{b=0}^{B(\underline{\mathbf{r}}_p)-1} \mathbf{G} \operatorname{diag}\left\{\mathbf{X}_{\underline{\mathbf{r}}_p,b}(m)\right\} \mathbf{H}_{\underline{\mathbf{r}}_p,b}, \tag{25}$$

with the abbreviation

$$\mathbf{G} = \mathbf{F}_M \begin{bmatrix} \mathbf{0}_{N\times N} & \mathbf{0}_{N\times N} \\ \mathbf{0}_{N\times N} & \mathbf{I}_{N\times N} \end{bmatrix} \mathbf{F}_M^{-1}, \tag{26}$$

where $\mathbf{0}_{N\times N}$ represents the $N \times N$ zero matrix and $\mathbf{I}_{N\times N}$ denotes the $N \times N$ identity matrix. The DFT-domain output of the Volterra filter is given by

$$\mathbf{Y}(m) = \sum_{p=1}^{P} \sum_{r_{p,1}=0}^{N_{p,p-1}} \cdots \sum_{r_{p,p-1}=r_{p,p-2}}^{N_p-1} \mathbf{Y}_{\underline{\mathbf{r}}_p}(m), \tag{27}$$

Taking the relation between circular and linear convolution into account [5], we notice that the first N elements of the time-domain correspondence $\mathbf{Y}(m)$, i.e.

$$\check{\mathbf{y}}(m) = \mathbf{F}_M^{-1}\, \mathbf{Y}(m), \tag{28}$$

are corrupted by time-domain aliasing, while the last N elements of $\check{\mathbf{y}}(m)$ result from a linear convolution and represent the desired output values $y(k)$. Thus, the time-domain output signal block of the Volterra filter, i.e.,

$$\mathbf{y}(m) = [y(mR), y(mR+1), \ldots, y(mR+N-1)]^T, \tag{29}$$

is finally obtained as

$$\mathbf{y}(m) = [\mathbf{0}_{N\times N} \; \mathbf{I}_{N\times N}]\; \mathbf{F}_M^{-1}\, \mathbf{Y}(m). \tag{30}$$

Note that for an overlapping factor $\alpha > 1$, only the first R elements represent new values of $y(k)$, whereas the remaining $N - R$ elements have already been computed in previous block time steps. However, choosing $\alpha > 1$ is beneficial for the adaptive implementation of the Volterra filter that is presented in the next section, as then, the adaptation of the kernel coefficients is performed α times more frequently, resulting in an increased convergence speed of the adaptive algorithm.

It should be mentioned that for the special case that all kernels have the same memory length (i.e., $L_i(\mathbf{0}) = L_j(\mathbf{0})$) and no partitioning is applied (i.e., $B(\underline{\mathbf{r}}_p) = 1$), the above algorithm reduces to the approach presented in [6]. Clearly, for the considered application to nonlinear acoustic echo cancellation this restrictions would lead to very inefficient system configurations, as $N_2 \ll N_1$ is sufficient for modeling the nonlinear behavior of loudspeakers in the AEC context [2].

4. ADAPTATION OF MULTIDELAY VOLTERRA FILTERS IN DCR

Aiming at an adaptive DFT-domain implementation of the multidelay Volterra filter, we define the time-domain signal vectors

$$\begin{aligned} \mathbf{d}(m) &= [d(mR), d(mR+1), \ldots, d(mR+N-1)]^T, & (31)\\ \mathbf{e}(m) &= [e(mR), e(mR+1), \ldots, e(mR+N-1)]^T, & (32) \end{aligned}$$

where according to Fig. 1, $e(k) = d(k) - y(k)$ represents the time-domain error signal. The corresponding DFT-domain signal vector are given by

$$\begin{aligned} \mathbf{D}(m) &= \mathbf{F}_M \left[\mathbf{0}_{N\times 1}^T \; \mathbf{d}^T(m)\right]^T, & (33)\\ \mathbf{E}(m) &= \mathbf{F}_M \left[\mathbf{0}_{N\times 1}^T \; \mathbf{e}^T(m)\right]^T. & (34) \end{aligned}$$

Next, we define the DFT-domain cost function to be minimized by the adaptive algorithm as

$$J(m) = \mathcal{E}\left\{\mathbf{E}^H(m)\mathbf{E}(m)\right\}, \tag{35}$$

where the superscript H denotes the Hermitian operator and $\mathcal{E}\{\}$ represents expectation. We notice from (25), (27) that $\mathbf{Y}(m)$ is linear with respect to the Volterra filter coefficients $\mathbf{H}_{\underline{\mathbf{r}}_p,b}$. Furthermore, (25) can be considered as the DFT-domain representation of the computation of the linear convolution between the linear filter $\mathbf{h}_{\underline{\mathbf{r}}_p}$ and the input $\mathbf{x}_{\underline{\mathbf{r}}_p}(m)$ applying the overlap/save method in its partitioned block version. Therefore, the results presented in [3] can directly applied to obtain an NLMS-type update equation:

$$\mathbf{H}_{\underline{\mathbf{r}}_p,b}(m+1) = \mathbf{H}_{\underline{\mathbf{r}}_p,b}(m) + \mu_p\, \widetilde{\mathbf{G}}_{\underline{\mathbf{r}}_p,b}\, \Delta_{\underline{\mathbf{r}}_p,b}(m), \tag{36}$$

where μ_p represents a positive step-size parameter. The update term $\Delta_{\underline{\mathbf{r}}_p,b}(m)$ is given by

$$\Delta_{\underline{\mathbf{r}}_p,b}(m) = \mathbf{S}^{-1}(m)\, \mathbf{E}(m) \operatorname{diag}\left\{\overline{\mathbf{X}}_{\underline{\mathbf{r}}_p,b}(m)\right\}, \tag{37}$$

where $\overline{\mathbf{A}}$ denotes the conjugate complex of $\mathbf{A}$. $\mathbf{S}^{-1}(m)$ is the inverse of a diagonal normalization matrix

$$\mathbf{S}(m) = \operatorname{diag}\left\{\left[S^{(0)}(m), S^{(1)}(m), \ldots, S^{(M-1)}\right]\right\}. \tag{38}$$

The computation of $\mathbf{S}(m)$ is discussed later in this section. The constraint matrix $\widetilde{\mathbf{G}}_{\underline{\mathbf{r}}_p,b}$ is given by

$$\widetilde{\mathbf{G}}_{\underline{\mathbf{r}}_p,b} = \mathbf{F}_M \operatorname{diag}\left\{\left[\mathbf{1}_{\underline{\mathbf{r}}_p,b} \; \mathbf{0}_{\underline{\mathbf{r}}_p,b}\right]\right\} \mathbf{F}_M^{-1}. \tag{39}$$

The row vector $\mathbf{1}_{\underline{\mathbf{r}}_p,b}$ contains only ones and its length depends on the actual diagonal $\underline{\mathbf{r}}_p$ and the partition number b:

$$\text{length}\left\{\mathbf{1}_{\underline{\mathbf{r}}_p,b}\right\} = \begin{cases} N, & \text{for } b < B(\underline{\mathbf{r}}_p) - 1 \\ L_p(\underline{\mathbf{r}}_p) - (b-1)N, & \text{for } b = B(\underline{\mathbf{r}}_p) - 1. \end{cases}$$

$\mathbf{0}_{\underline{\mathbf{r}}_p,b}$ represents a zero vector of appropriate length, such that $\widetilde{\mathbf{G}}_{\underline{\mathbf{r}}_p,b}$ is an $M \times M$ matrix:

$$\text{length}\left\{\mathbf{0}_{\underline{\mathbf{r}}_p,b}\right\} + \text{length}\left\{\mathbf{1}_{\underline{\mathbf{r}}_p,b}\right\} = M. \tag{40}$$

Note that the definition of $\widetilde{\mathbf{G}}_{\underline{\mathbf{r}}_p,b}$ in (39) implies that both, the time-domain constraint on $\mathbf{H}_{\underline{\mathbf{r}}_p,b}$ according to (24), i.e., the zero padding of $\mathbf{h}_{\underline{\mathbf{r}}_p,b}$, and the constraint on the region of support of the Volterra filter according to (21) are fulfilled by the update equation (36). Next, we consider the computation of $\mathbf{S}(m)$. In the following, $A^{(i)}$ always denotes the i-th element of a vector $\mathbf{A}$. For the discussion of the computation of $\mathbf{S}(m)$ it will be useful to define the vectors $\mathbf{X}_\nu(m)$ and $\mathbf{H}_\nu$, respectively, which are composed of the ν-th element of all vectors $\mathbf{X}_{\underline{\mathbf{r}}_p,b}(m)$ and $\mathbf{H}_{\underline{\mathbf{r}}_p,b}$, respectively. The elements of $\mathbf{X}_\nu(m)$ and $\mathbf{H}_\nu$ are arranged such that $Y^{(\nu)}(m)$, i.e., the ν-th element of $\mathbf{Y}(m)$ can be expressed by

$$Y^{(\nu)}(m) = \mathbf{H}_\nu^T\, \mathbf{X}_\nu(m). \tag{41}$$

The update term $\Delta_\nu(m)$ with respect to $\mathbf{H}_\nu(m)$ which corresponds to $\Delta_{\underline{\mathbf{r}}_p,b}(m)$ according to (37) yields

$$\Delta_\nu(m) = \frac{E^{(\nu)}(m)\,\overline{\mathbf{X}}_\nu(m)}{S^{(\nu)}(m)}. \tag{42}$$

Following the normalization technique commonly used for adaptive DFT-domain approaches, the normalization factor $S^{(\nu)}(m)$ is chosen independently for each DFT bin [3]. As the input signals of different diagonals with index vectors $\underline{\mathbf{u}}_i, \underline{\mathbf{v}}_i$ are in general not orthogonal for arbitrary excitation $x(k)$, i.e.,

$$\mathcal{E}\left\{x_{\underline{\mathbf{u}}_i}(k)x_{\underline{\mathbf{v}}_j}(k)\right\} \neq 0, \tag{43}$$

the elements of $\mathbf{X}_\nu(m)$ are also not orthogonal to each other, i.e.,

$$\mathcal{E}\left\{X_\nu^{(\rho)}(m)\overline{X}_\nu^{(\kappa)}(m)\right\} \neq 0, \quad \forall\, \rho, \kappa. \tag{44}$$

Discarding the constraint matrix $\widetilde{\mathbf{G}}_{\underline{\mathbf{r}}_p,b}$ in (36) and regarding (42), the reasoning for a complex-valued NLMS algorithm as given in [7] can directly be applied to obtain $S^{(\nu)}(m)$ according to

$$S^{(\nu)}(m) = \mathcal{E}\left\{\mathbf{X}_\nu^H(m)\mathbf{X}_\nu(m)\right\}, \tag{45}$$

i.e., the powers of the ν-th element of all vectors $\mathbf{X}_{\underline{\mathbf{r}}_p,b}(m)$ have to be added up for the computation of $S^{(\nu)}(m)$. If the normalization is performed using (45), the step-size parameter μ_p should be chosen from the interval $0 < \mu_p < 2$.

5. SIMULATION RESULTS

Our evaluation of the proposed algorithm is based on recorded speech data from a low-cost loudspeaker placed in an enclosure with low reverberation. As mentioned earlier, the nonlinear behavior of loudspeakers can be modeled be second-order Volterra filters and, thus, $P = 2$ has been chosen for the simulations. The performance is measured using the *Echo Return Loss Enhancement* ($ERLE$) which is defined by

$$ERLE = 10\log_{10}\frac{\mathrm{E}\left\{d^2(k)\right\}}{\mathrm{E}\left\{e^2(k)\right\}}\ [\mathrm{dB}]. \tag{46}$$

The parameters of the multidelay DFT-domain Volterra filter have been chosen to $N = 64$, $B(\underline{\mathbf{r}}_1) = 6$ and $L_1(\underline{\mathbf{r}}_1) = B(\underline{\mathbf{r}}_1)N$ for the linear kernel. For the quadratic kernel, the maximum distance of any diagonal to the main diagonal has been set to $N_2 = 20$, and no partitioning has been applied, i.e., $B(\underline{\mathbf{r}}_2) = 1$ implying $L_2(\underline{\mathbf{r}}_2 = \mathbf{0}) = N$. An overlapping factor $\alpha = 4$ has been used.

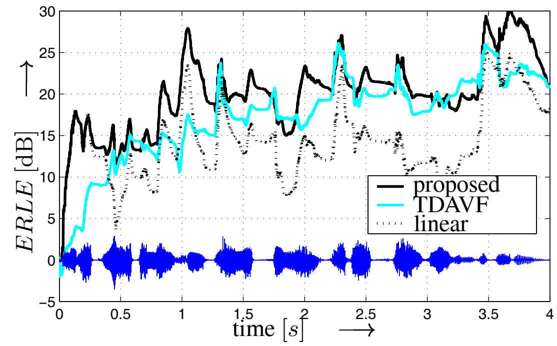

Fig. 3. Comparison of different nonlinear approaches and a linear DFT-domain algorithm for a realistic AEC scenario together with the input signal.

The resulting $ERLE$ curve is compared to a time-domain adaptive Volterra filter (TDAVF) with the same region of support as the DFT-domain approach applying an NLMS algorithm [1], and to a linear DFT-domain echo canceler in Fig. 3. We notice that extending the linear AEC to a second-order Volterra filter leads to an improved performance if the nonlinear distortion in the echo path is caused by loudspeakers, and that the proposed approach provides a faster convergence speed compared to the TDAVF.

6. CONCLUSION

We presented an efficient and fast-converging DFT-domain algorithm for the adaptation of P-th Volterra filters in DCR, where the region of support, i.e., the memory length and the number of diagonals, can be chosen different for each Volterra kernel. It has been shown that the DCR is especially suitable for cascaded structures, where a Volterra filter is followed by a linear filter. Simulation results with respect to nonlinear acoustic echo cancellation confirm the increase in convergence speed compared to a corresponding time-domain approach.

7. REFERENCES

[1] V. J. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*, John Wiley and Sons, New York, 2000.

[2] 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.

[3] E. Moulines et al., "The generalized multidelay adaptive filter: Structure and convergence analysis," *IEEE Trans. on Signal Processing*, vol. 43, no. 1, pp. 14–28, Jan. 1995.

[4] G. M. Raz and B. D. Van Veen, "Baseband Volterra filters for implementing carrier based nonlinearities," *IEEE Trans. on Signal Processing*, vol. 46, no. 1, pp. 103–114, Jan. 1998.

[5] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms and Applications*, Prentice Hall, New Jersey, 3rd edition, 1996.

[6] M. Morháč, "A fast algorithm of nonlinear Volterra filtering," *IEEE Trans. on Signal Processing*, vol. 39, no. 10, pp. 2353–2356, Oct. 1991.

[7] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 4th edition, 2001.