# Efficient Nonlinear Acoustic Echo Cancellation by Partitioned-Block Significance-Aware Hammerstein Group Models

Christian Hofmann, Michael Guenther, Christian Huemmer, and Walter Kellermann
University of Erlangen-Nuremberg, Multimedia Communications and Signal Processing,
Cauerstr. 7, D-91058 Erlangen, Germany
hofmann@LNT.de,michael.guenther@FAU.de,huemmer@LNT.de,wk@LNT.de

***Abstract*—A powerful and efficient model for nonlinear echo paths of hands-free communication systems is given by the recently proposed Significance-Aware Hammerstein Group Model (SA-HGM). Such a model learns memoryless loudspeaker nonlinearities on a small temporal support of the echo path (preferably the direct-sound region) and extrapolates the nonlinearities for the entire echo path afterwards. In this contribution, an efficient frequency-domain realization of the significance-aware concept for nonlinear acoustic echo cancellation is proposed. The proposed method exploits the benefits of partitioned-block frequency-domain adaptive filtering and will therefore be referred to as Partitioned-Block Significance-Aware Hammerstein Group Model (PBSA-HGM). This allows to efficiently model a long nonlinear echo path by a linear partitioned-block frequency-domain adaptive filter after a parametric memoryless nonlinear preprocessor, the parameters of which are estimated via a nonlinear Hammerstein Group Model (HGM) with the short temporal support of a single block only.**

## I. Introduction

Since the first adaptive linear echo canceler for network echoes in telephone lines [1], linear echo cancellation has evolved to a key ingredient of almost any full-duplex speech communication system. This has resulted in a multitude of approaches to efficiently model, parametrize and estimate even complex linear systems, such as the acoustic echo paths in hands-free wideband telecommunication scenarios [2]. With increasing nonlinear distortions produced by miniaturized amplifiers and loudspeakers in modern portable devices, dedicated nonlinear echo-path models have emerged as an important topic of research and motivated sophisticated approaches for nonlinear Acoustic Echo Cancellation (AEC) based on Volterra filters [3]–[5], artificial neural networks [6], [7], functional link adaptive filters [8] or kernel methods [9], [10]. A very simple, yet effective model for nonlinear acoustic echo paths is the cascade of a memoryless preprocessor (modeling loudspeaker signal distortions) and a subsequent linear system (modeling sound propagation through air) [11]. Due to its simplicity, this so-called Hammerstein Model (HM) has been frequently employed [12]–[18] and will also be used in this contribution. So will be a group of $B$ parallel HMs, referred to as Hammerstein Group Model (HGM). The recently proposed efficient Significance-Aware Hammerstein Group Model (SA-HGM) [17] combines the advantages of HMs and HGMs.

In contrast to previous work, this contribution generalizes the concept of Significance-Aware (SA) filtering [17], [18] to efficient frequency-domain realizations by combining SA filtering with Partitioned Block Frequency-Domain Filtering (PBFDF) to provide a highly efficient nonlinear AEC. The remainder of this paper is structured as follows: after introducing the notation in Section II, the concept of SA filtering and its application to HGMs is reviewed in Section III. Afterwards, we propose an efficient frequency-domain implementation of an SA-HGM in Section IV and verify its AEC performance by a comparison to a linear filter and the computationally more complex HGM in Section V. Finally, the manuscript is summarized in Section VI.

## II. Notation

Vectors are written as boldface lower case letters, matrices are typeset with capital boldface letters, and the corresponding frequency-domain quantities are marked by an additional underline. Furthermore, $a(k) * b(k)$ and $a(k) \circledast b(k)$ denote linear and cyclic convolution between time sequences $a(k)$ and $b(k)$, respectively, where $k$ is the discrete-time sample index. Besides, $\mathbf{A} \odot \mathbf{B}$ denotes element-wise multiplication (Hadamard product), $\langle \mathbf{a}, \mathbf{b} \rangle$ stands for the scalar product $\mathbf{a}^{\mathrm{T}}\mathbf{b}$, where $(\cdot)^{\mathrm{T}}$ is transposition, and $(\cdot)^*$ denotes complex conjugation. Special matrices with an individual symbol are the $S \times S$ identity matrix $\mathbf{I}_S$, the $S \times S$ all-zero matrix $\mathbf{0}_S$, and the windowing matrices

$$\mathbf{W}_{01} = \begin{bmatrix} \mathbf{0}_S & \mathbf{0}_S \\ \mathbf{0}_S & \mathbf{I}_S \end{bmatrix} \quad \text{and} \quad \mathbf{W}_{10} = \begin{bmatrix} \mathbf{I}_S & \mathbf{0}_S \\ \mathbf{0}_S & \mathbf{0}_S \end{bmatrix}, \qquad (1)$$

setting the first or second half of a length-$2S$ vector to zero, respectively. Furthermore, $\mathrm{DFT}_M\{\cdot\}$ and $\mathrm{IDFT}_M\{\cdot\}$ denote the $M$th-order Discrete Fourier Transform (DFT) and its inverse, respectively, which are implemented efficiently as fast Fourier transforms in practice.

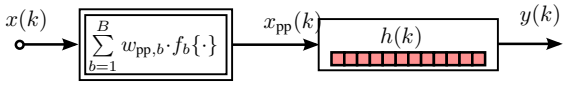
## III. Echo-path Models and Significance-Aware Filtering

In this section, commonly used echo path models are introduced as a prerequisite before briefly reviewing the authors' previous work on SA filtering [17].

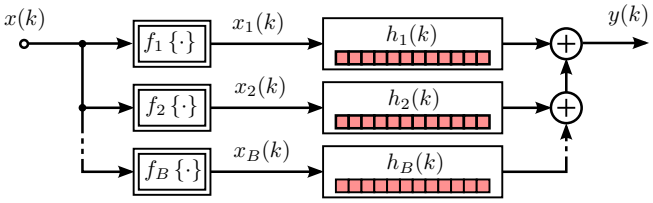
### A. Linear Echo Path Models

As the propagation of acoustic waves between loudspeaker and microphone can be modeled very accurately by a linear system, acoustic echo paths for high-quality playback and

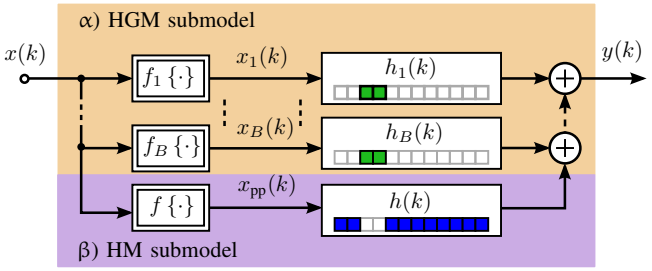
978-0-9928-6265-7/16/$31.00 ©2016 IEEE




(a) HM with parametric preprocessor, performing weighted superposition of $B$ static nonlinearities.




(b) Hammerstein group model (HGM) with $B$ branches of HMs.




(c) An SA-HGM, consisting of the **HGM submodel (block α)** modeling only the direct-path lags of the RIR and a **Hammerstein submodel (block β)** which covers the remaining temporal support of the RIR.

Figure 1. Block diagram of (a) an HM with parametric preprocessor, (b) an HGM, and (c) an SA-HGM. The temporal support captured by a particular linear system is indicated by non-white squares of a schematic impulse-response vector below the system's symbol.

recording equipment can be described in the discrete time domain by a Room Impulse Response (RIR) $h(k)$. In this case, the input signal $x(k)$ and the output signal $y(k)$ are related by the convolutive product

$$y(k) = x(k) * h(k). \tag{2}$$

However, miniaturization and the need for energy efficiency of portable devices lead to nonlinearly distorting playback equipment (e.g., loudspeakers and amplifiers) and thereby render linear echo path models insufficient.

### B. Hammerstein Model (HM)

Nonlinearities caused by the playback equipment as part of the acoustic echo path can be modeled by a memoryless nonlinearity $f\{\cdot\}$ (a memoryless preprocessor) prior to the linear RIR $h(k)$ [11]. This structure with the input/output relation

$$y(k) = f\{x(k)\} * h(k) = x_{\text{pp}}(k) * h(k) \tag{3}$$

is commonly referred to as a Hammerstein Model (HM). Actual implementations of Hammerstein echo path models may employ parametric preprocessors of the form [17], [19]

$$x_{\text{pp}}(k) = \sum_{b=1}^{B} w_{\text{pp},b} \cdot f_b\{x(k)\}, \tag{4}$$

where $B$ fixed nonlinear basis functions $f_b\{\cdot\}$, $b = 1, \ldots, B$, are weighted by expansion coefficients $w_{\text{pp},b}$ (denoted as preprocessor coefficients) and superimposed. This structure is visualized by the block diagram in Fig. 1(a).

Note that the system output of such an HM does not exhibit a simple linear dependency on each of the model parameters (filter coefficients $h(k)$ and preprocessor coefficients $w_{\text{pp},b}$), but depends on the products of filter coefficients and preprocessor coefficients. In general, employing the Minimum Mean Squared Error (MMSE) criterion for deriving a joint estimator for $h(k)$ and $w_{\text{pp},b}$ is not analytically tractable [16]. Thus, we combine an HM estimating only the linear system $h(k)$ with an HGM from which the HM's preprocessor coefficients $w_{\text{pp},b}$ are estimated (more details will be given in Sections III-C and III-D).

### C. Hammerstein Group Model (HGM)

Inserting (4) into (3) and exploiting the linearity of the convolution allows to rewrite the input/output relation of the HM as

$$y(k) = \sum_{b=1}^{B} \underbrace{f_b\{x(k)\}}_{x_b(k)} * \underbrace{(w_{\text{pp},b} \cdot h(k))}_{h_b(k)}. \tag{5}$$

Obviously, the HM is realized in (5) as the sum of $B$ parallel branches with HMs with fixed nonlinearities $f_b\{\cdot\}$ and linearly dependent impulse responses $h_b(k)$, referred to as kernels in the following. Such a group of parallel HMs, as visualized in Fig. 1(b), can be categorized as HGM[1]. As main advantage of the HGM, the fixed nonlinearities lead to an input/output relation which is linear in the parameters (filter coefficients). This allows the application of classical Least-Mean-Square (LMS)-type algorithms [20] for estimating the HGM kernels $h_b(k)$. Prominent examples for HGMs are the so-called power filters [21], forming a subset of Volterra filters [22].

### D. Significance-Aware Hammerstein Group Model (SA-HGM)

The SA-HGM [17] exploits the aforementioned equivalence of an HM and an HGM for modeling one particular nonlinear system. Furthermore, the SA-HGM exploits that the nonlinearities between loudspeaker and microphone signals can be observed best for time lags of significant energy transmission (e.g. the direct-path region of the impulse response).

The input/output relation of an entire SA-HGM system (so-far non-adaptive) is visualized in Fig. 1(c). Therein, an **HGM submodel (block α)** only models a small temporal region of the RIR (green filter taps), whereas the remaining temporal support of the RIR is then modeled by a single **HM submodel (block β)**. To adaptively identify such a structure, estimates $\hat{h}_1(k), \ldots, \hat{h}_B(k)$ and $\hat{h}(k)$ for the impulse responses $h_1(k), \ldots, h_B(k)$ and $h(k)$ are adapted, respectively, and the HGM submodel's kernel estimates $\hat{h}_b(k)$ are employed to estimate the HM submodel's memoryless nonlinear preprocessor. In particular, as shown in [17], a least squares estimate for the preprocessor coefficients can be obtained by

$$\hat{w}_{\text{pp},b} = \frac{\langle \hat{\mathbf{h}}_1, \hat{\mathbf{h}}_b \rangle}{\langle \hat{\mathbf{h}}_1, \hat{\mathbf{h}}_1 \rangle}, \tag{6}$$

where $\hat{\mathbf{h}}_b = [h_b(0), h_b(1), \ldots, h_b(M-1)]^{\text{T}}$ is the HGM submodel's estimate for the kernel of the $b$th branch[2]. An in-depth description and visualization of the adaptation method

[1] Note that HGMs allow for independent kernels $h_b(k)$ in general.

[2] A time-dependence of $\hat{\mathbf{h}}_b$ due to their adaptive identification has been omitted for the conciseness of notation.

for a time-domain SA-HGM is given in [17]. This method allows an efficient estimation of the preprocessor coefficients and leads to significantly less computational complexity than a time-domain HGM with the full temporal support (see Fig. 1(b)).

## IV. Partitioned-Block SA-HGM (PBSA-HGM)

Different to time-domain adaptive filters, the complexity for filter adaptation in the frequency domain is not determined by the length of the time domain support of the filter, but by the DFT order. This disqualifies classical frequency-domain adaptive filters, as the HGM submodel in Fig. 1(c) would have the same complexity as the HGM with full temporal support in Fig. 1(b). This reveals partitioned block frequency domain methods [23]–[26] as reasonable realization for the SA-HGM structure from [17].

### A. Partitioned Block Frequency-Domain Filtering (PBFDF)

A linear convolution, as in (2) and (3), can be realized despite large filter lengths $L$ with a low input/output delay efficiently by block-based processing methods like PBFDF [23]–[26], also known as multidelay convolution. In the following, only a uniform partitioning with frame shift $S$ and frame size $M = 2S < L$ will be considered.

In this case, the input signal $x(k)$, the impulse response $h(k)$, and the output signal $y(k)$ are partitioned into length-$M$ vectors

$$\mathbf{x}(\nu) = [x(\nu S - M + 1), \ldots, x(\nu S)]^{\mathrm{T}} \tag{7}$$

$$\mathbf{h}^{(n)} = [h(nS), \ldots, h\,(nS + S - 1)\,, 0, \ldots, 0]^{\mathrm{T}} \tag{8}$$

$$\mathbf{y}(\nu) = [\,0, \ldots, 0,\, y(\nu S - S + 1), \ldots, y\,(\nu S)]^{\mathrm{T}}\,, \tag{9}$$

respectively, where $\nu$ is the frame index for block processing and $n$ is the index of the impulse response partition. After a DFT, transformed versions of the signal vectors and impulse response partitions are referred to as

$$\underline{\mathbf{x}}^{(n)}(\nu) = \mathrm{DFT}_M\{\mathbf{x}(\nu - n)\} \tag{10}$$

$$\underline{\mathbf{h}}^{(n)} = \mathrm{DFT}_M\{\mathbf{h}^{(n)}\}. \tag{11}$$

As outlined in Fig. 2 for $N = 2$ partitions, performing fast DFT-domain convolution between each pair of $\mathbf{h}^{(n)}$ and $\mathbf{x}(\nu - n)$ and summing up the respective partial results yields

$$\mathbf{y}(\nu) = \mathbf{W}_{01}\,\mathrm{IDFT}_M \left\{ \sum_{n=0}^{N-1} \underline{\mathbf{x}}^{(n)}(\nu) \odot \underline{\mathbf{h}}^{(n)} \right\}, \tag{12}$$

where $N = \left\lceil \frac{L}{S} \right\rceil$ is the number of non-zero impulse response partitions and where the windowing matrix $\mathbf{W}_{01}$ according to (1) suppresses additionally computed samples, most of which contain cyclic convolution artifacts. Note that such a PBFDF scheme is computationally efficient, because each input signal frame's DFT has to be computed only once.

### B. Realization of the PBSA-HGM

In this section, a novel, efficiency-increased frequency-domain realization of the SA-HGM will be developed based on PBFDF and the Frequency-Domain Normalized Least-Mean-Square (FNLMS) algorithm. One partition can be thought of as covering two adjacent squares of the impulse responses in Fig. 1 (e.g., green squares in HGM branches), in the following.

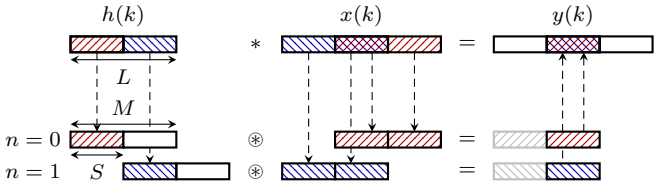



Figure 2. Example for partitioned block convolution using $N = 2$ blocks for partitioning the IR $h(k)$: the summation during convolution is split into $N = 2$ partial sums of length $S$. Each partition of $h(k)$ produces a sequence of $S$ samples of its contribution to the convolutive product in an overlap-save manner in the DFT domain, where the block length is $M = 2S$. Summing up the second half of these partial results (free from cyclic convolution artifacts) yields a block of $S$ output samples.

*1) Estimation of the RIR of the HM Submodel:* Partitioning the impulse response estimate according to the PBFDF scheme of Section IV-A leads to time-domain partition vectors $\hat{\mathbf{h}}^{(n)}(\nu)$ for the $n$th partition of room impulse response estimate at frame $\nu$, which become $\underline{\hat{\mathbf{h}}}^{(n)}(\nu)$ after a DFT. With this, an echo signal estimate vector can be obtained according to (12) by

$$\hat{\mathbf{y}}_{\mathrm{HM}}(\nu) = \mathbf{W}_{01}\,\mathrm{IDFT}_M \left\{ \sum_{n=0}^{N-1} \underline{\mathbf{x}}_{\mathrm{pp}}^{(n)}(\nu) \odot \underline{\hat{\mathbf{h}}}^{(n)}(\nu) \right\}. \tag{13}$$

Its deviation from the vector of actual microphone signal samples, $\mathbf{y}(\nu)$, is the error signal vector, the DFT of which computes as

$$\underline{\hat{\mathbf{e}}}_{\mathrm{HM}}(\nu) = \mathrm{DFT}_M\left\{(\mathbf{y}(\nu) - \hat{\mathbf{y}}_{\mathrm{HM}}(\nu))\right\}. \tag{14}$$

Iteratively minimizing the two-norm of $\underline{\hat{\mathbf{e}}}_{\mathrm{HM}}(\nu)$ leads to the multidelay block frequency domain adaptive filter [27] update rule of (15) on the next page, where $\mathbf{S}^{(n)}(\nu)$ is the inverse of the diagonal matrix of the input signal energies in each DFT bin at frame $\nu - n$ and where $\mu$ is the adaptation stepsize.

*2) Estimation of the HGM Submodel:* Furthermore, the partitioning directly suggests to use the partition describing the strongest energy transmission for the SA filtering: the high-energy partition of the HM is replaced by an HGM with $B$ branches. Each branch $b$ contains a DFT-domain estimate $\underline{\hat{\mathbf{h}}}_b^{(d)}(\nu)$ of the $b$th kernel of the HGM, modeling the system at the time lags covered by partition $d$ with maximum energy. With these notations, the PBSA-HGM yields the echo estimate

$$\hat{\mathbf{y}}_{\mathrm{SA}}(\nu) = \mathbf{W}_{01}\,\mathrm{IDFT}_M \left\{ \sum_{b=1}^{B} \underline{\mathbf{x}}_b^{(d)}(\nu) \odot \underline{\hat{\mathbf{h}}}_b^{(d)}(\nu) + \sum_{n \in \{0,\ldots,N-1\} \setminus d} \underline{\mathbf{x}}_{\mathrm{pp}}^{(n)}(\nu) \odot \underline{\hat{\mathbf{h}}}^{(n)}(\nu) \right\}, \tag{16}$$

where the first sum in (16) corresponds to the output of the HGM submodel (**block $\alpha$** in Fig. 1(c)) and the second sum corresponds to the remaining temporal support of the RIR, captured by the HM submodel (**block $\beta$** in Fig. 1(c)). Minimizing the resulting error signal

$$\underline{\hat{\mathbf{e}}}_{\mathrm{SA}}(\nu) = \mathrm{DFT}_M\left\{(\mathbf{y}(\nu) - \hat{\mathbf{y}}_{\mathrm{SA}}(\nu))\right\} \tag{17}$$

yields the update rule of (18) on the next page, where $\mathbf{S}_b^{(d)}(\nu)$ is the inverse of the diagonal matrix of the branch signal

energies in each DFT bin in frame $\nu - n$ and where $\mu$ is the adaptation stepsize. The reason for the unconventional application of the windowing in (18) to the actual impulse response partitions, unlike to the update term in (15), will become obvious in the following paragraph.

*3) Estimating the HM's Preprocessor:* Finally, the HM's preprocessor coefficients can be recomputed employing the HGM. Although the unitarity of the DFT allows an evaluation of (6) with DFT-domain vectors as well, we propose to compute

$$w_{\text{pp},b} = \left\langle \hat{\mathbf{h}}_1^{(d)}(\nu), \hat{\mathbf{h}}_b^{(d)}(\nu) \right\rangle \Big/ \left\langle \hat{\mathbf{h}}_1^{(d)}(\nu), \hat{\mathbf{h}}_1^{(d)}(\nu) \right\rangle \tag{19}$$

from the time-domain kernels $\hat{\mathbf{h}}_b^{(d)}(\nu)$. Note that this does not imply additional Fourier transforms to obtain $\hat{\mathbf{h}}_b^{(d)}(\nu)$ as long as the filter constraint (enforcing zeros in $\hat{\mathbf{h}}_b^{(d)}(\nu)$) is not applied to the update, but to the filter coefficients themselves (as done in (18)). Thus, the scalar products have to be performed only between real-valued sub-vectors of length $S$.

Thereby, the estimation of the nonlinear system has been split up into two beneficially interacting subproblems (HM and HGM adaptation), each of which now only linearly depends on the coefficients to be estimated and can therefore be solved elegantly by state-of-the-art adaptive filtering algorithms. The beneficial interaction is achieved by the preprocessor coefficient refinement based on the HGM and by the extension of the temporal support of the HGM employing partitions of the HM - analogously to the time-domain solution in [17]. Note that the computational savings of a PBSA-HGM over a Partitioned-Block HGM (PB-HGM) with full temporal support include the multiplications of the filtering operations and filter adaptation for $N-1$ blocks and $B-1$ branches, while all branch signals $x_b(k)$ still have to be transformed to the DFT domain. For a large number of blocks, the overhead due to the DFTs for the HGM's branch signals becomes vanishingly small, such that the ratio of the computational complexity of a PBSA-HGM to a PB-HGM approaches a factor $C \approx (N+B)/(N \cdot B)$. For $N = 10$ partitions and $B = 5$ branches, this leads to a complexity ratio of $C \approx \frac{3}{10}$.

## V. Experimental Verification

### A. Experimental Setup

For the evaluation, about 130 seconds of double-talk free speech are played back and recorded by a smartphone in hands-free mode in a living-room like environment (Setup A), as well as 80 seconds in an anechoic environment (Setup B). Setups C and D are synthesized by convolving the anechoic nonlinear recordings obtained from Setup B with measured RIRs. For these setups, a partitioned-block linear filter (PB-HGM with a single linear branch), a PB-HGM (full temporal support), and a PBSA-HGM will be compared. In the following experiments, processing is done at a sampling rate of $f_s = 16\,\text{kHz}$. The acoustic echo paths are estimated with a memory length corresponding to $L = 1024$ taps and the input signals and impulse responses are partitioned into $N = 4$ blocks with a relative frame shift of $S = 256$ taps (implying $M = 512$). As in [17], the HGMs consist of $B = 5$ branches, with odd-order Legendre polynomials up to order 9 as nonlinear basis functions $f_b\{\cdot\}$ in (4). The kernels of both the full PB-HGM and the PBSA-HGM are adapted with a stepsize of $\mu = 0.1$ (see (15) and (18)) while a stepsize of $\mu = 0.5$ was chosen for the reference algorithm, consisting of the linear filter, to account for the lack of multiple adaptation branches. Furthermore, the adaptation process is refined using the robust statistics described in [24], [28], which hardly affects the computational complexity, but robustifies the system identification against outliers in the error signal. The AEC performance is quantified by computing the Echo-Return-Loss Enhancement (ERLE)

$$\text{ERLE} = 10\log_{10}\left\{\frac{\mathcal{E}\{y^2(k)\}}{\mathcal{E}\{e^2(k)\}}\right\}\text{dB}, \tag{20}$$

where the numerator is the power of the microphone signal $y(k)$ and the denominator is the power of the error signal $e(k)$ produced by the model to be evaluated (linear filter, PB-HGM, PBSA-HGM).

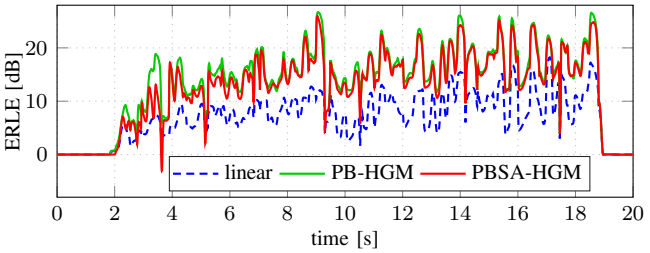



Figure 3. Initial adaptation behavior of different PBFDF AEC methods for Setup B during the first 20 s of the experiment.

### B. Experimental Results

Figure 3 shows the ERLE during the initial adaptation phase for different AEC structures, where the instantaneous estimates of the signal powers in (20) are calculated from short intervals of about 200 ms.

As expected, the full 5-branch PB-HGM (green curve) outperforms the linear filter without nonlinear preprocessor (dashed blue curve). Furthermore, the system identification achieved by the PBSA-HGM (red curve) is only slightly reduced compared to the full PB-HGM. Additionally, average

$$\underline{\hat{\mathbf{h}}}^{(n)}(\nu+1) = \underline{\hat{\mathbf{h}}}^{(n)}(\nu) + \text{DFT}_M\left\{\mathbf{W}_{10}\,\text{IDFT}_M\left\{\frac{\mu}{N}\mathbf{S}^{(n)}(\nu)\left(\hat{\underline{\mathbf{e}}}_{\text{HM}}(\nu)\odot\left(\underline{\mathbf{x}}_{\text{pp}}^{(n)}(\nu)\right)^*\right)\right\}\right\} \quad \forall \text{ partitions } n \tag{15}$$

$$\hat{\underline{\mathbf{h}}}_b^{(d)}(\nu+1) = \text{DFT}_M\Big\{\underbrace{\mathbf{W}_{10}\,\text{IDFT}_M\left\{\hat{\underline{\mathbf{h}}}_b^{(d)}(\nu) + \mu\,\mathbf{S}_b^{(d)}(\nu)\left(\hat{\underline{\mathbf{e}}}_{\text{SA}}(\nu)\odot\left(\underline{\mathbf{x}}_b^{(d)}(\nu)\right)^*\right)\right\}}_{\hat{\mathbf{h}}_b^{(d)}(\nu+1)}\Big\} \quad \forall \text{ branches } b \tag{18}$$

Table I
PERFORMANCE COMPARISON OF PBFDF AEC METHODS AND RELATIVE COMPUTATIONAL COMPLEXITY OF THESE FREQUENCY-DOMAIN (FD) AND TIME-DOMAIN (TD) IMPLEMENTATIONS.

| | average ERLE in dB for Setup | | | | relative complexity | |
|---|---|---|---|---|---|---|
| | A | B | C | D | FD | TD |
| linear | 8.0 | 8.2 | 8.7 | 8.6 | 1.0 | 15.0 |
| full HGM | 13.1 | 14.8 | 14.4 | 14.0 | 4.1 | 52.5 |
| SA-HGM | 12.7 | 14.1 | 13.8 | 13.4 | 2.6 | 17.2 |

ERLE values, computed over the entire sequence, are given in Table I for all setups defined in Section V. Obviously, the full PB-HGM performs best for all investigated setups and the PBSA-HGM consistently follows the PB-HGM performance closely (only $0.4$ dB to $0.7$ dB less ERLE). This efficacy confirms the successful extension of the SA-HGM to frequency-domain adaptive filters.

Regarding computation time (second last column in Table I) on an Intel Core i7 CPU, employing the PBSA-HGM instead of the full PB-HGM reduces the elapsed wall-clock time of the MATLAB simulation (MATLAB R2015b) by about 37% for the chosen parameter set. Note that the computational savings become even larger for more partitions ($M \ll L$), where the adaptation of a single-partition HGM contributes even less to the overall complexity (recall Section IV-B3). In general, the frequency-domain implementations lead to significantly reduced wall-clock time in comparison to their time-domain equivalents [17] with according filter length, for which the relative numbers are given in the last column of Table I.

## VI. CONCLUSIONS

The SA-HGM algorithm has been extended to an even more efficient partitioned-block frequency-domain version. Therein, the SA concept seamlessly integrates into the partitioning concept by facilitating the latter to implement the temporal splitting of the acoustic echo path. The larger the number of partitions, the more efficient the proposed PBSA-HGM becomes. While the echo reduction performance of the proposed PBSA-HGM is very close to the performance of a PB-HGM with full temporal support, the computational complexity is reduced for the PBSA-HGM by about 37% in the experiments.

## REFERENCES


[1] M. M. Sondhi, "An adaptive echo canceller," *Bell System Technical Journal*, vol. 46, no. 3, pp. 497–511, 1967.

[2] P. Dreiseitel, E. Hänsler, and H. Puder, "Acoustic echo and noise control - a long lasting challenge," in *Proc. European Signal Process. Conf. (EUSIPCO)*, Rhodes, Greece, Sept. 1998, pp. 945–952.

[3] E. Thomas, "Some considerations on the application of the Volterra representation of nonlinear networks to adaptive echo cancellers," *Bell System Technical Journal*, vol. 50, no. 8, pp. 2797–2805, Oct. 1971.

[4] A. Stenger, L. Trautmann, and R. Rabenstein, "Nonlinear acoustic echo cancellation with 2nd order adaptive Volterra filters," in *Proc. IEEE Intl. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, vol. 2, Phoenix, AZ, Mar 1999, pp. 877–880.

[5] M. Zeller, L. A. Azpicueta-Ruiz, J. Arenas-Garcia, and W. Kellermann, "Adaptive Volterra filters with evolutionary quadratic kernels using a combination scheme for memory control," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1449 – 1464, Apr. 2011.

[6] A. N. Birkett and R. A. Goubran, "Nonlinear echo cancellation using a partial adaptive time delay neural network," in *Proc. IEEE Workshop Neural Networks Signal Process. (NNSP)*, Cambridge, MA, Sept. 1995, pp. 449–458.

[7] L. Ngja and J. Sjobert, "Nonlinear acoustic echo cancellation using a Hammerstein model," in *Proc. IEEE Intl. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, Seattle, WA, May 1998, pp. 1229–1232.

[8] D. Comminiello *et al.*, "Functional link adaptive filters for nonlinear acoustic echo cancellation," *IEEE Trans. Audio, Speech, Language Process.*, vol. 21, no. 7, pp. 1502–1512, July 2013.

[9] G. Li, C. Wen, W. X. Zheng, and Y. Chen, "Identification of a class of nonlinear autoregressive models with exogenous inputs based on kernel machines," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2146–2159, May 2011.

[10] J. Kivinen, A. J. Smola, and R. C. Williamson, "Online learning with kernels," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 165 – 176, Aug. 2004.

[11] A. Birkett and R. Goubran, "Limitations of handsfree acoustic echo cancellers due to nonlinear loudspeaker distortion and enclosure vibration effects," in *Proc. IEEE Workshop Applications Signal Process. Audio Acoustics (WASPAA)*, New Paltz, NY, Oct. 1995, pp. 103–106.

[12] S. Malik and G. Enzner, "State-Space Frequency-Domain Adaptive Filtering for Nonlinear Acoustic Echo Cancellation," *IEEE Trans. Audio, Speech, Language Process.*, vol. 20, no. 7, pp. 2065–2079, Sept. 2012.

[13] S. Shimauchi and Y. Haneda, "Nonlinear acoustic echo cancellation based on piecewise linear approximation with amplitude threshold decomposition," in *Proc. Int. Workshop Acoustic Echo, Noise Control (IWAENC)*, Aachen, Germany, Sept. 2012, pp. 1–4.

[14] S. Malik and G. Enzner, "Variational Bayesian inference for nonlinear acoustic echo cancellation using adaptive cascade modeling," in *Proc. IEEE Intl. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, Kyoto, Japan, Mar 2012, pp. 37–40.

[15] S. Malik and G. Enzner, "A variational Bayesian learning approach for nonlinear acoustic echo control," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 5853–5867, Dec. 2013.

[16] C. Huemmer *et al.*, "The elitist particle filter based on evolutionary strategies as novel approach for nonlinear acoustic echo cancellation," in *Proc. IEEE Intl. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, Florence, Italy, May 2014, pp. 1315–1319.

[17] C. Hofmann, C. Huemmer, and W. Kellermann, "Significance-aware Hammerstein group models for nonlinear acoustic echo cancellation," in *Proc. IEEE Intl. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, Florence, Italy, May 2014, pp. 5934–5938.

[18] C. Huemmer, C. Hofmann, R. Maas, and W. Kellermann, "The significance-aware EPFES to estimate a memoryless preprocessor for nonlinear acoustic echo cancellation," in *Proc. IEEE Global Conf. Signal Information Process. (GlobalSIP)*, Atlanta, GA, Dec. 2014, pp. 557–561.

[19] A. Stenger and W. Kellermann, "Adaptation of a memoryless preprocessor for nonlinear acoustic echo cancelling," *Signal Process.*, vol. 80, no. 9, pp. 1747–1760, Sept. 2000.

[20] S. Haykin, *Adaptive Filter Theory*, 4th ed. Upper Saddle River (NJ), USA: Prentice Hall, 2002.

[21] F. Küch, A. Mitnacht, and W. Kellermann, "Nonlinear acoustic echo cancellation using adaptive orthogonalized power filters," in *Proc. IEEE Intl. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, vol. 3, Philadelphia, PA, Mar 2005, pp. 105–108.

[22] F. Küch, "Adaptive polynomial filters and their application to nonlinear acoustic echo cancellation," Ph.D. dissertation, University Erlangen-Nuremberg, Germany, Nov. 2005.

[23] W. G. Gardner, "Efficient Convolution without Input-Output Delay," *J. Audio Eng. Soc.*, vol. 43, no. 2, pp. 127–136, Mar 1995.

[24] H. Buchner, J. Benesty, T. Gänsler, and W. Kellermann, "An outlier-robust extended multidelay filter with application to acoustic echo cancellation," in *Proc. Int. Workshop Acoustic Echo, Noise Control (IWAENC)*, Kyoto, Japan, 2003, pp. 19–22.

[25] M. Zeller and W. Kellermann, "Self-configuring system identification via evolutionary frequency-domain adaptive filters," in *Proc. Int. Workshop Acoustic Echo, Noise Control (IWAENC)*, Aug. 2010.

[26] F. Küch, E. Mabande, and G. Enzner, "State-space architecture of the partitioned-block-based acoustic echo controller," in *Proc. IEEE Intl. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, Florence, Italy, May 2014, pp. 1295–1299.

[27] J.-S. Soo and K. Pang, "Multidelay block frequency domain adaptive filter," *IEEE Trans. Acoustics, Speech, Signal Process.*, vol. 38, no. 2, pp. 373–376, Feb. 1990.

[28] T. Gänsler, "A double-talk resistant subband echo canceller," *Signal Process.*, vol. 65, no. 1, pp. 89–101, 1998.