2016 IEEE INTERNATIONAL WORKSHOP ON MACHINE LEARNING FOR SIGNAL PROCESSING, SEPT. 13–16, 2016, SALERNO, ITALY

A BLOCK-BASED COMBINED SCHEME EXPLOITING SPARSITY IN NONLINEAR ACOUSTIC ECHO CANCELLATION

Danilo Comminiello[*†], *Michele Scarpiniti*[†], *Luis A. Azpicueta-Ruiz*[‡], *Jerónimo Arenas-García*[‡], and *Aurelio Uncini*[†]

[†] Dept. Inform. Eng., Electron. and Telecom.
"Sapienza" University of Rome
00184 Rome, Italy

[‡] Dept. Signal Theory and Communic.
Universidad Carlos III de Madrid
28911 Leganés, Spain

ABSTRACT

Nonlinear acoustic echo cancellation (NAEC) aims at estimating both the acoustic impulse response and the nonlinearities affecting the desired signal. Both the modeling processes show behaviors of sparse nature from an energy point of view. In this paper, we propose an adaptive NAEC algorithm that takes advantage of such sparsity behaviors to improve echo cancellation performance. The proposed scheme is characterized by two block-based adaptive combinations of proportionate adaptive filters, having different strategies, devoted respectively to the estimation of the linear and nonlinear responses. The proposed model is assessed in NAEC problems, where its advantages and effectiveness are shown.

Index Terms— Sparse Adaptive Filters, Adaptive Combination of Filters, Linear-in-the-Parameters Nonlinear Filters, Nonlinear Adaptive Filtering, Functional Links

1. INTRODUCTION

Acoustic echo occurs when a microphone acquires a delayed, and possibly distorted, replica of a signal reproduced by a loudspeaker, together with desired information. Such echo signal can be estimated and removed from the microphone signal by means of echo cancellers. Nonlinear acoustic echo cancellation (NAEC) represents one of the most complete and challenging problems in adaptive filtering, since it involves two main issues to be addressed: the estimation of the acoustic impulse response (AIR) and the estimation of the nonlinearities rebounding in the echo channel. The former issue has a mostly linear nature and can be tackled by designing suited adaptive filters. The latter problem requires the introduction of nonlinear models able to estimate those distortions that affect the signal captured by a microphone. Recently, several NAEC models have been proposed based on different approaches, including Bayesian learning [1], evolutionary strategies [2], hardware solutions [3], enhanced adaptive Volterra filters [4,5], kernel adaptive filters [6,7], spline adaptive filters [8], cascaded Kalman filters [9], and parallel Hammerstein architectures [10], among others.

Here, we propose an NAEC method based on the class of functional link adaptive filters (FLAFs) [11–14], in which the modeling of AIR and nonlinearities is split in two parallel filtering branches. We focus on the fact that both the modeling processes may show sparsity behaviors. Regarding the AIR estimate, it usually shows some degree of sparsity, since most of the energy of the AIR is contained in the first part of it [15]. This fact led proportionate adaptive algorithms to be often used in echo cancellation problems [15–20]. On the other hand, sparsity can be observed also in the estimate of nonlinearities, mainly due to nonlinear expansions of the input signal [13], leading to the employment of sparsity-aware algorithms also for the nonlinear filtering [13, 21–23].

In order to take always advantage of any sparseness, regardless of its degree, we use adaptive combinations of filters [24, 25], which present robust performance suited for different, and even unknown, time-varying sparseness degrees. Analyzing the sparsity behaviors in both the estimation of AIR and nonlinearities, it can be gather that coefficients with higher energy are adjacent to each other, so it is possible to select blocks with different sparseness degrees. For this reason, the proposed NAEC model is based on block-based combinations [26,27] with different combination strategies for the two estimation tasks. The proposed method is assessed in NAEC problems, where its effectiveness has been illustrated.

The rest of the paper is organized as follows: FLAFs and its split scheme for NAEC are introduced in Section 2. An analysis of the sparsity behaviors in functional link expansions is presented in Section 3, while in Section 4 the proposed method involving the two block-based combination approaches is introduced. In Section 5, experimental results are shown and, finally, in Section 6 our conclusions are drawn.

* Corresponding author e-mail: danilo.comminiello@uniroma1.it. The work of Luis A. Azpicueta-Ruiz is partially supported by Comunidad de Madrid under grant 'CASI-CAM-CM' (id. S2013/ICE-2845), by the Spanish Ministry of Economy and Competitiveness (under grant DAMA (TIN2015-70308-REDT) and grant TEC2014-52289-R), and by the European Union. The work of J. Arenas-García has been partly funded by MINECO project TEC2014-52289-R.

978-1-5090-0746-2/16/$31.00

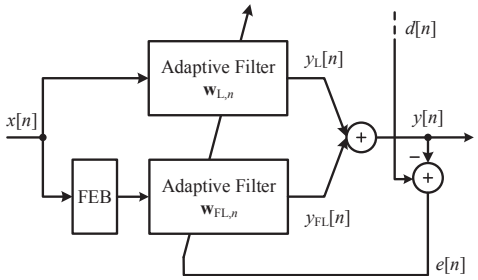

Fig. 1. The split functional link adaptive filter.

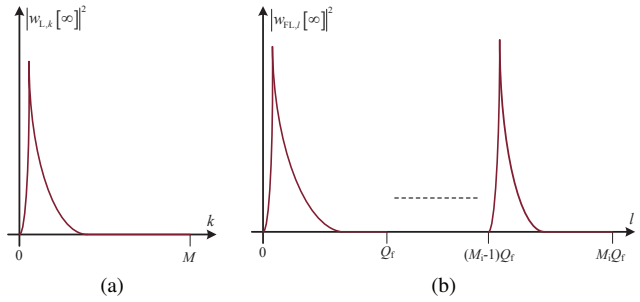

Fig. 2. Sparse behavior of $\mathbf{w}_{\mathrm{L},\infty}$ (a) and $\mathbf{w}_{\mathrm{FL},\infty}$ (b) from an energy point of view.

2. FUNCTIONAL LINK ADAPTIVE FILTERS FOR NAEC

In NAEC applications, the response of a system to be identified is produced by any combination of a linear and nonlinear components. For this reason, a split functional link adaptive filter (SFLAF) scheme was recently proposed [11] as a parallel architecture including a linear and a nonlinear branch, as depicted in Fig. 1. The former branch is simply composed of a linear adaptive filter, which is devoted to modeling the linear components of the echo path (including the AIR). This allows the nonlinear branch to focus on the modeling of the nonlinear components of that echo path. Such branch is a linear-in-the-parameters nonlinear filter, composed of a *functional expansion block* (FEB) and a subsequent adaptive filter.

At n-th time instant the SFLAF receives the input vector $\mathbf{x}_n \in \mathbb{R}^M = \left[\begin{array}{ccc} x[n] & \ldots & x[n-M+1] \end{array}\right]^{\mathrm{T}}$, where M is the input vector length that matches with the length of the filter $\mathbf{w}_{\mathrm{L},n} \in \mathbb{R}^M = \left[\begin{array}{ccc} w_{\mathrm{L},0}[n] & \ldots & w_{\mathrm{L},M-1}[n] \end{array}\right]^{\mathrm{T}}$. The adaptive filtering yields the output $y_{\mathrm{L}}[n] = \mathbf{x}_{\mathrm{L},n}^{\mathrm{T}}\mathbf{w}_{\mathrm{L},n-1}$. A selection of $M_{\mathrm{i}} \leq M$ samples of $\mathbf{x}_n$ is also processed on the nonlinear branch by an FEB, which consists of a series of functions satisfying universal approximation constraints. Such functions are called *functional links* and they are contained in a chosen set $\Phi = \left\{\varphi_0(\cdot), \ldots, \varphi_{Q_{\mathrm{f}}-1}(\cdot)\right\}$, where Q_{f} is the number of functional links. Each sample of the input vector is expanded by Φ, thus yielding the nonlinear expanded buffer $\mathbf{g}_n \in \mathbb{R}^{M_{\mathrm{e}}} = \left[\begin{array}{ccc} g_0[n] & \ldots & g_{M_{\mathrm{e}}-1}[n] \end{array}\right]^{\mathrm{T}}$, where $M_{\mathrm{e}} \geq M_{\mathrm{i}}$ represents the length of the expanded buffer. We adopt a nonlinear trigonometric series expansion to build the functional link set, such that:

$$\varphi_j(x[n-i]) = \begin{cases} \sin(p\pi x[n-i]), & j = 2p-2 \\ \cos(p\pi x[n-i]), & j = 2p-1 \end{cases} \quad (1)$$

where $p = 1, \ldots, P$ is the expansion index, being P the expansion order, and $j = 0, \ldots, Q_{\mathrm{f}}-1$ the functional link index. It is easy to verify that the set Φ is composed of $Q_{\mathrm{f}} = 2P$ functional links. It is worth noting that the expanded vector $\mathbf{g}_n$ is composed of nonlinear elements only, since the linear part of a system to be identified can be demanded to a separate filter, from which derives the name "split". Such expanded buffer is then fed into an adaptive filter $\mathbf{w}_{\mathrm{FL},n} \in \mathbb{R}^{M_{\mathrm{e}}} = \left[\begin{array}{ccc} w_{\mathrm{FL},0}[n] & w_{\mathrm{FL},1}[n] & \ldots \quad w_{\mathrm{FL},M_{\mathrm{e}}-1}[n] \end{array}\right]^{\mathrm{T}}$, thus providing the nonlinear output $y_{\mathrm{FL}}[n] = \mathbf{g}_n^{\mathrm{T}}\mathbf{w}_{\mathrm{FL},n-1}$. The SFLAF output results from the sum of the two branch outputs, i.e., $y[n] = y_{\mathrm{L}}[n] + y_{\mathrm{FL}}[n]$, and, thereby, the error signal is:

$$e[n] = d[n] - (y_{\mathrm{L}}[n] + y_{\mathrm{FL}}[n]) \quad (2)$$

which is used for the adaptation of both adaptive filters. Such filters can be adapted by using any adaptive algorithm, including proportionate algorithms that may take advantage of any sparsity of the system response.

3. EXPLOITING SPARSITY IN FUNCTIONAL LINK ADAPTIVE FILTERS

3.1. Sparsity in Linear and Nonlinear Modeling

NAEC application requires the estimate of a linear system, which is mainly represented by the AIR, and of a nonlinear system that introduces distortions in the echo signal. An impulse response may be considered sparse if a large fraction of its energy is concentrated in a small fraction of its duration [15]. That is the case of the AIR, since the most significant part of the response is restricted to its head, which contains the direct path and the early reflections of a reverberant environment. If we analyze the energy of the coefficient vector on the linear branch at steady state, $\mathbf{w}_{\mathrm{L},\infty}$, we find a sparse behavior, like that represented in Fig. 2(a), where most of the response has "inactive" coefficients having negligible energy values. The energy evolution shows a certain sparseness degree that may change according to the reverberation time, thus affecting the energy decay.

We also analyze the energy behavior of the coefficient vector $\mathbf{w}_{\mathrm{FL},\infty}$ to study sparsity in functional link expansions. Let us consider the expansion of one input sample $x[n]$ ($M_{\mathrm{i}} = 1$) with a certain expansion order P, such that our expanded vector $\mathbf{g}_n$ is composed of $M_{\mathrm{e}} = 2P = Q_{\mathrm{f}}$ elements. This means that the early coefficients of $\mathbf{w}_{\mathrm{FL},\infty}$ are

related to the functional links with small values of the expansion order ($p \to 1$), while the last ones are related to the functional links with $p \to P$, being p the expansion index. By analyzing the energy of $\mathbf{w}_{\text{FL},\infty}$, we can usually observe a decreasing exponential behavior. This means that early functional links generate the most significant samples of the expanded vector, while the last elements of $\mathbf{g}_n$ only produces small variations in the nonlinear modeling process. It is worth noting that $\mathbf{w}_{\text{FL},\infty}$ usually shows an exponentially decreasing energy as for the $\mathbf{w}_{\text{L},\infty}$, with a sparseness degree that in this case may vary according to the nonlinearity degree and to the chosen expansion order. If we generalize the above analysis for $M_{\text{i}} > 1$, the energy of $\mathbf{w}_{\text{FL},\infty}$ will be characterized by a periodic exponential decay, where the number of repetitions coincides with the number of input samples to the FEB, i.e., M_{i}, as represented in Fig. 2(b).

3.2. Full Proportionate SFLAF

In order to take advantage of any sparsity degree of the SFLAF's filters, we introduce a "full" regularizing mask that gives more prominence to those filter coefficients that have an active role in the linear and nonlinear filtering. We can adopt the same optimization procedure presented in [13], but applied to the filters on both the branches. We achieve the following adaptation rules for the full proportionate SFLAF:

$$\mathbf{w}_{\text{L},n} = \mathbf{w}_{\text{L},n-1} + \mu_{\text{L}} \frac{\mathbf{Q}_{\text{L},n}\mathbf{x}_n}{\mathbf{x}_n^{\text{T}}\mathbf{Q}_{\text{L},n}\mathbf{x}_n + \delta_{\text{PL}}} e\,[n] \qquad (3)$$

$$\mathbf{w}_{\text{FL},n} = \mathbf{w}_{\text{FL},n-1} + \mu_{\text{FL}} \frac{\mathbf{Q}_{\text{FL},n}\mathbf{g}_n}{\mathbf{g}_n^{\text{T}}\mathbf{Q}_{\text{FL},n}\mathbf{g}_n + \delta_{\text{PFL}}} e\,[n]\,, \qquad (4)$$

where μ_{L} and μ_{FL} are step-size parameters, δ_{PL} and δ_{PFL} are regularization factors, and $\mathbf{Q}_{\text{L},n}$ and $\mathbf{Q}_{\text{FL},n}$ are proportionate matrices, which aim at weighting the coefficient vectors proportionally to the contribution they provide to the modeling.

The proportionate matrices are diagonal matrices, whose diagonal elements are computed by using the filter coefficients estimated at the previous time instant: the larger a coefficient value, the higher the corresponding weighting. This means that significant coefficients are adapted faster. We choose the diagonal elements according to the improved proportionate normalized least mean-square (IPNLMS) algorithm [16], however nothing prevents to adopt other proportionate rules. The resulting diagonal elements of $\mathbf{Q}_{\text{L},n}$ and $\mathbf{Q}_{\text{FL},n}$ can be respectively written as:

$$q_{\text{L},k}\,[n] = \frac{1-\rho_{\text{L}}}{2M} + (1+\rho_{\text{L}})\,\frac{|w_{\text{L},k}\,[n-1]|}{2\,\|\mathbf{w}_{\text{L},n-1}\|_1 + \xi} \qquad (5)$$

$$q_{\text{FL},l}\,[n] = \frac{1-\rho_{\text{FL}}}{2M_{\text{e}}} + (1+\rho_{\text{FL}})\,\frac{|w_{\text{FL},l}\,[n-1]|}{2\,\|\mathbf{w}_{\text{FL},n-1}\|_1 + \xi} \qquad (6)$$

with $k = 0, \ldots, M-1$, $l = 0, \ldots, M_{\text{e}} - 1$, and ξ is a small positive constant that avoids divisions by zero. The proportionality factors $-1 \leq \{\rho_{\text{L}}, \rho_{\text{FL}}\} \leq 1$ balance the proportionality, since when their value is close to 1 a high sparseness degree is assumed, while, on the other hand, a low degree is supposed when these factors approach -1, reducing the adaptation to the normalized least mean-square (NLMS) scheme. According to the proportionality factors, the regularization parameters in (3) and (4) are, respectively, $\delta_{\text{PL}} = (1-\rho_{\text{L}})\,\delta/2M$ and $\delta_{\text{PFL}} = (1-\rho_{\text{FL}})\,\delta/2M_{\text{e}}$, where δ denotes a common regularization factor used for an NLMS adaptation.

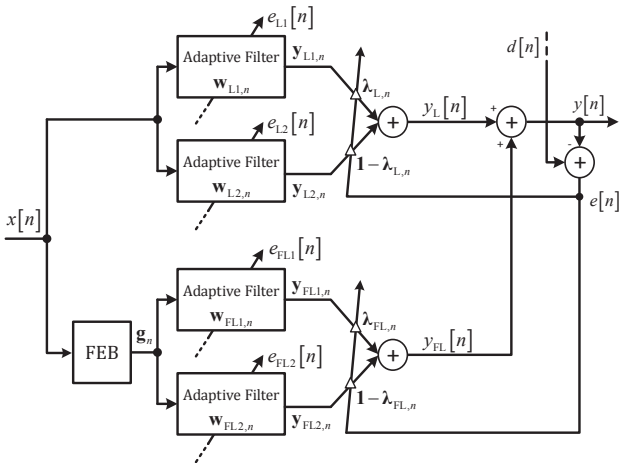

Fig. 3. The proposed filtering scheme where the filter block outputs are combined by the mixing parameter vectors.

4. BLOCK-BASED COMBINED PSFLAF ARCHITECTURE

The full PSFLAF takes advantage of the split architecture, depicted in Fig. 1, and of the proportionate algorithm. Here, we extend both the architecture and the learning algorithm to better exploit any sparsity. The new architecture is characterized by the combination of proportionate filters on both the linear and nonlinear branches, as represented in Fig. 3.

First, we focus on the linear branch, which is mainly devoted to the estimation of the AIR. In order to exploit any sparsity of the AIR, we perform an adaptive combination of filters, thus generalizing the learning regardless of the sparseness degree. In particular, as suggested in [27], we take into account a block-based combination approach that adopts only one adaptive mixing parameter for each group of adjacent coefficients, based on the typical behavior of an AIR. Therefore, we consider two adaptive filters on the linear branch, $\mathbf{w}_{\text{L}1,n}$ and $\mathbf{w}_{\text{L}2,n}$, and we divide them in L_{L} blocks, each one consisting of $N_{\text{L}} = M/L_{\text{L}}$ coefficients. The output signals for each filter on the linear branch are obtained as $y_{\text{L}j}\,[n] = \mathbf{x}_n^{\text{T}}\mathbf{w}_{\text{L}j,n-1}$, for $j = 1, 2$, where $\mathbf{w}_{\text{L}j,n}$ are updated according to (3), but using different proportionality factors $\rho_{\text{L}j}$ for the derivation of the proportionate matrices. In particular, we can choose a high proportionality factor for the first filter, i.e., $\rho_{\text{L}1}$

close to 1, and a lower one for the second filter, i.e., $\rho_{\text{L}2}$ close to -1, thus generalizing the learning regardless of how sparse is an AIR. The combined output can be expressed as

$$y_{\text{L}}[n] = \sum_{l=1}^{L_{\text{L}}} \lambda_{\text{L},l}[n]\, \mathbf{x}_n^{(l)\text{T}} \mathbf{w}_{\text{L}1,n-1}^{(l)} + (1-\lambda_{\text{L},l}[n])\, \mathbf{x}_n^{(l)\text{T}} \mathbf{w}_{\text{L}2,n-1}^{(l)} \tag{7}$$

where the superscript (l) represents the block index and $\lambda_{\text{L},l}[n]$ is the mixing parameter associated to the l-th block.

The l-th adaptive mixing parameter $\lambda_{\text{L},l}[n]$, with $l = 1, \ldots, L$, in (7) balances the combination between the l-th blocks of the two filters $\mathbf{w}_{\text{L}j}[n]$ ($j = 1, 2$), giving more importance to the best performing filter block [24]. Such awareness is obtained according to a mean-square error minimization. In particular, the adaptation of $\lambda_{\text{L},l}[n]$ is performed by using an auxiliary adaptive parameter for each block $a_{\text{L},l}[n]$, which is related to $\lambda_{\text{L},l}[n]$ by means of a sigmoidal function that keeps the mixing parameter in the range $[0, 1]$, and defined according to [28] as:

$$\lambda_{\text{L},l}[n] = \beta \left(\frac{1}{1 + e^{-a_{\text{L},l}[n]}} - \alpha \right), \tag{8}$$

where $\alpha = 1/\left(1 + e^4\right)$ and $\beta = 1/(1 - 2\alpha)$. The update rule for the auxiliary parameter for the l-th block is [27]:

$$\begin{aligned} a_{\text{L},l}[n] =& a_{\text{L},l}[n-1] + \frac{\mu_{\text{c}}}{\beta r_{\text{L},l}[n-1]} e[n]\, \Delta y_{\text{L},l}[n] \\ & \cdot (\lambda_{\text{L},l}[n] + \alpha\beta)(\beta - \alpha\beta - \lambda_{\text{L},l}[n]) \end{aligned} \tag{9}$$

for $l = 1, \ldots, L_{\text{L}}$, where

$$\Delta y_{\text{L},l}[n] = \mathbf{x}_n^{(l)\text{T}} \left(\mathbf{w}_{\text{L}1,n-1}^{(l)} - \mathbf{w}_{\text{L}2,n-1}^{(l)} \right). \tag{10}$$

In (9), μ_{c} is the step-size parameter of the adaptive combination, $r_{\text{L},l}[n] = \gamma r_{\text{L},l}[n-1] + (1-\gamma)\, \Delta y_{\text{L},l}^2[n]$ is the estimated power of $\Delta y_{\text{L},l}[n]$ that permits a normalized adaptation of $a_{\text{L},l}[n]$, and $0 \ll \gamma < 1$ is a smoothing factor.

On the other hand, in order to exploit sparsity in the nonlinear modeling, we propose a new block-based approach that takes into account the periodic nature of the energy decay in functional links, as depicted in Fig. 2(b). As said in Subsection 3.1, each group of Q_{f} samples shows a similar energy decay to that of the adjacent groups. According to this property, we may think to consider the mixing parameters associated to the L_{FL} blocks of one group of Q_{f} samples, and use them for all the M_{i} groups of Q_{f} samples. In this case, each blocks consists of $N_{\text{FL}} = Q_{\text{f}}/L_{\text{FL}}$ coefficients and the expanded buffer length can be also expressed as $M_{\text{e}} = L_{\text{FL}} N_{\text{FL}} M_{\text{i}}$. Considering the output signals for each filter on the nonlinear branch as $y_{\text{FL}j}[n] = \mathbf{g}_n^{\text{T}} \mathbf{w}_{\text{FL}j,n-1}$, for $j = 1, 2$, the output of the block-based combination on the nonlinear branch can be written as

$$\begin{aligned} y_{\text{FL}}[n] = \sum_{i=0}^{M_{\text{i}}-1} \Bigg(& \sum_{l=1}^{L_{\text{FL}}} \lambda_{\text{FL},l}[n]\, \mathbf{g}_n^{(iL_{\text{FL}}+l)\text{T}} \mathbf{w}_{\text{FL}1,n-1}^{(iL_{\text{FL}}+l)} \\ & + (1 - \lambda_{\text{FL},l}[n])\, \mathbf{g}_n^{(iL_{\text{FL}}+l)\text{T}} \mathbf{w}_{\text{FL}2,n-1}^{(iL_{\text{FL}}+l)} \Bigg) \end{aligned} \tag{11}$$

where it is possible to notice that the same L_{FL} mixing parameters are used for each group of Q_{f} samples. Similarly to the filters on the linear branch, the coefficient vectors $\mathbf{w}_{\text{FL}j,n}$, with $j = 1, 2$, are updated according to (4), but using different values for $\rho_{\text{FL}j}$ ($j = 1, 2$), thus generalizing the learning regardless of how sparse is the functional link expansion.

The adaptive mixing parameters $\lambda_{\text{FL},l}[n]$, for $l = 1, \ldots, L_{\text{FL}}$, in (11) are derived similarly to the linear case, i.e. by applying (10) and (9) to the filter outputs on the nonlinear branch. However, in the new block-based approach the block-difference output related to the l-th block is expressed as:

$$\Delta y_{\text{FL},l}[n] = \sum_{i=1}^{M_{\text{i}}} \mathbf{g}_n^{(i,l)\text{T}} \left(\mathbf{w}_{\text{FL}1,n-1}^{(i,l)} - \mathbf{w}_{\text{FL}2,n-1}^{(i,l)} \right), \tag{12}$$

where the superscript (i, l) denotes the l-th block related to the i-th entry. It is worth noting that eq. (12) takes into account the periodic nature of the functional link expansions.

Once achieved both the outputs from (7) and (11), we can derive the error signals used for the adaptation of pairs of filters, respectively, on the linear and nonlinear branches [4]:

$$e_{\text{L}j}[n] = d[n] - (y_{\text{L}j}[n] + y_{\text{FL}}[n]) \tag{13}$$

$$e_{\text{FL}j}[n] = d[n] - (y_{\text{L}}[n] + y_{\text{FL}j}[n]), \tag{14}$$

and also the overall error signal $e[n]$ by applying (2), as for the SFLAF. It is worth noting that such overall error signal is used for the adaptation of both the auxiliary parameters $a_{\text{L},l}[n]$, for $l = 1, \ldots, L_{\text{L}}$, and $a_{\text{FL},l}[n]$, for $l = 1, \ldots, L_{\text{FL}}$.

5. EXPERIMENTAL RESULTS

The performance of the proposed block-based combined scheme is assessed in an NAEC scenario characterized by a sparse AIR between the loudspeaker and the microphone, sampled at 8 kHz and truncated after $M = 512$ samples. The system to be identified is a Hammerstein scheme composed of a nonlinear block followed by a linear one. In order to simulate an asymmetric loudspeaker distortion, the nonlinear block applies a memoryless sigmoidal nonlinearity [11] to the input signal $x[n]$:

$$\overline{y}[n] = \zeta \left(\frac{1}{1 + e^{(-\eta z[n])}} - \frac{1}{2} \right) \tag{15}$$

where:

$$z[n] = \frac{3}{2} x[n] - \frac{3}{10} x^2[n]. \tag{16}$$

In (15), the parameter ζ is the sigmoid gain and it is set equal to $\zeta = 2$, while η represents the sigmoid slope, chosen as:

$$\eta = \begin{cases} 4, & z[n] > 0 \\ \frac{1}{2}, & z[n] \leq 0 \end{cases}. \tag{17}$$

The resulting signal $\overline{y}[n]$ is then convolved by the AIR. Additive Gaussian noise is added to provide 30 dB of *signal-to-noise ratio* (SNR), thus yielding the desired signal $d[n]$.

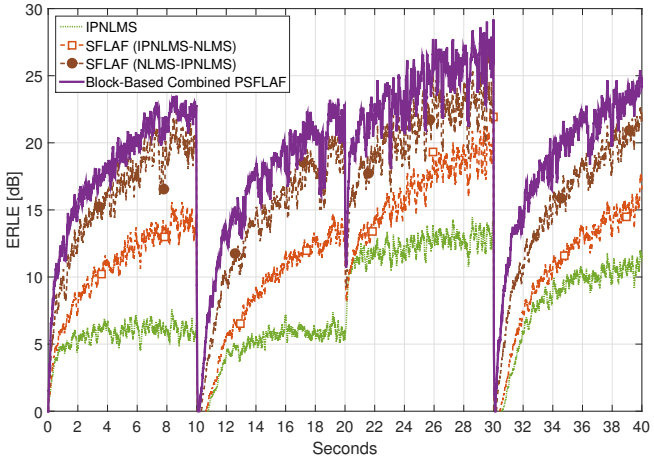

Fig. 4. Performance behavior in terms of the ERLE in case of colored noise as input.

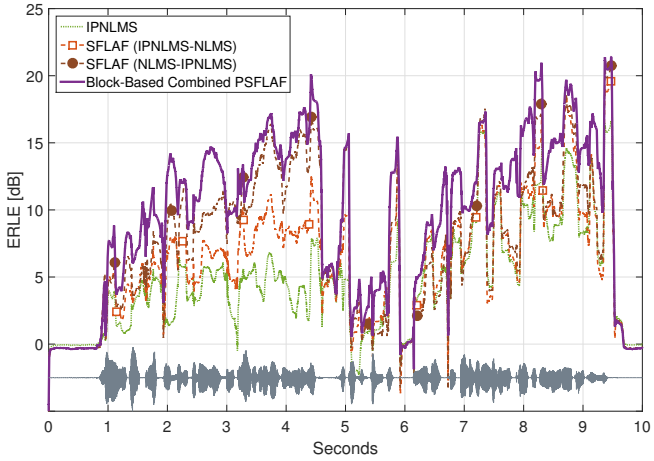

Fig. 5. Performance behavior in terms of the ERLE in case of female speech as input.

We first consider a colored noise as input signal $x\,[n]$, generated by means of a first-order autoregressive model, whose transfer function is $\sqrt{1-\theta^2}/\left(1-\theta z^{-1}\right)$, with $\theta = 0.8$, fed with an i.i.d. Gaussian random process. The length of this experiment is 40 seconds. In order to evaluate the effectiveness of the proposed scheme in changing environments, we introduce abrupt changes for both the linear and nonlinear modeling. In particular, after 10 seconds from the start, an echo path change is simulated by shifting the AIR circularly to the right by 20 samples. Then, at the 20-th second, we also introduce a change in the nonlinearity by setting $\eta = 5/2$ for $z\,[n] \leq 0$ in eq. (17). Another change of the AIR is then introduced, similarly to the first path change, at the 30-th second.

We compare the performance of the proposed algorithm with an IPNLMS algorithm, which is completely linear, an SFLAF architecture with an IPNLMS on the linear branch, denoted as SFLAF (IPNLMS-NLMS), and an SFLAF architecture with an IPNLMS on the nonlinear branch (that is the one proposed in [13]), denoted as SFLAF (NLMS-IPNLMS). We use the following parameter setting: input buffer length $M_{\text{I}} = M$, step-size parameters $\mu_{\text{L}} = 0.1$ and $\mu_{\text{FL}} = 0.1$, regularization parameter $\delta = 10^{-3}$, and expansion order $P = 10$. The proportionality factors are set equal to 0 for the non-combined architectures, while in the combined scheme the proportionality factors represent the peculiar parameter of each combination. As a matter of fact, in order to exploit the sparsity of both AIR and functional links, we set $\rho_{\text{L}1} = \rho_{\text{FL}1} = 0.9$ and $\rho_{\text{L}2} = \rho_{\text{FL}2} = -1$ in order to specialize the architecture regardless of any sparseness degree. It should be noted that we select the value of $\rho_{\text{L}2}$ and $\rho_{\text{FL}2}$ to be exactly equal to -1 instead of -0.9, since this leads to an NLMS adaptation instead of an IPNLMS scheme, thus yielding similar performance with reduced computational cost. The block-based combinations are adapted by using the same step-size value $\mu_{\text{c}} = 0.5$, smoothing factor $\gamma = 0.9$ and initial values $a_l\,[0] = 0$ and $r_l\,[0] = 1$. For the block-based combination on the linear branch we use $L_{\text{L}} = 8$ blocks yielding a block size of $N_{\text{L}} = M/L_{\text{L}} = 64$ coefficients, while for the block-based combination on the nonlinear branch we use a smaller number of blocks $L_{\text{FL}} = 4$, each one consisting of $N_{\text{FL}} = 2P/L_{\text{FL}} = 5$ coefficients.

Performance is evaluated in terms of *echo return loss enhancement* (ERLE), which is defined in dB as $\text{ERLE}\,[n] = 10\log_{10}\left(\text{E}\left\{d^2\,[n]\right\}/\text{E}\left\{e^2\,[n]\right\}\right)$. Results in Fig. 4 show that the proposed scheme provides superior performance than the other filters, in particular in the convergence rate after each change on the environment.

We also assess the effectiveness of the proposed block-based combined scheme when the input signal is a female speech of 10 seconds. At half of the length, we introduce a change of both the linear and nonlinear paths. We use the same parameter setting of the previous experiment, thus achieving the results depicted in Fig. 5 in terms of the ERLE. Even in this case it is evident the improvement produced by the proposed method with respect to the other filters, reaching gains close to 5 dBs.

6. CONCLUSIONS

In this paper, we have proposed a new nonlinear filtering architecture that takes advantage of any sparseness degree of both the acoustic impulse response and the functional link expansions, specially suited for NAEC. The proposed model is based on adaptive combinations of proportionate filters on both the linear and the nonlinear branches, having different block-based approaches, thus yielding an improvement of the overall modeling performance. Experimental results have shown the effectiveness and robustness of our proposal with both colored noise and speech as input signal. Future research will include the combination of filters using hierarchical schemes that would further generalize this model.

7. REFERENCES

[1] 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.

[2] C. Huemmer, C. Hofmann, R. Maas, A. Schwarz, and W. Kellermann, "The elitist particle filter based on evolutionary strategies as novel approach for nonlinear acoustic echo cancellation," in *IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*, Florence, Italy, May 2014, pp. 1329–1333.

[3] P. Shah, I. Lewis, S. Grant, and S. Angrignon, "Nonlinear acoustic echo cancellation using voltage and current feedback," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 23, no. 10, pp. 1589–1599, Oct. 2015.

[4] L. A. Azpicueta-Ruiz, M. Zeller, A. R. Figueiras-Vidal, J. Arenas-García, and W. Kellermann, "Adaptive combination of Volterra kernels and its application to nonlinear acoustic echo cancellation," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, no. 1, pp. 97–110, Jan. 2011.

[5] L. A. Azpicueta-Ruiz, M. Zeller, A. R. Figueiras-Vidal, W. Kellermann, and J. Arenas-García, "Enhanced adaptive Volterra filtering by automatic attenuation of memory regions and its application to acoustic echo cancellation," *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2745–2750, June 2013.

[6] J. M. Gil-Cacho, M. Signoretto, T. Van Waterschoot, M. Moonen, and S. H. Jensen, "Nonlinear acoustic echo cancellation based on a sliding-window leaky kernel affine projection algorithm," *IEEE Trans. Audio, Speech, Language Process.*, vol. 21, no. 9, pp. 1867–1878, Sept. 2013.

[7] S. Van Vaerenbergh, L. A. Azpicueta-Ruiz, and D. Comminiello, "A split kernel adaptive filtering architecture for nonlinear acoustic echo cancellation," in *24th Eur. Signal Process. Conf. (EUSIPCO)*, Budapest, Hungary, Aug. 2016.

[8] M. Scarpiniti, D. Comminiello, R. Parisi, and A. Uncini, "Comparison of Hammerstein and Wiener systems for nonlinear acoustic echo cancelers in reverberant environments," in *17th Int. Conf. Digital Signal Process. (DSP)*, Corfù, Greece, July 6-8 2011, pp. 1–6.

[9] M. Z. Ikram, "Non-linear acoustic echo cancellation using cascaded Kalman filtering," in *IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*, Florence, Italy, May 4-9 2014, pp. 1334–1338.

[10] C. Hofmann, C. Huemmer, and W. Kellermann, "Significance-aware Hammerstein group models for nonlinear acoustic echo cancellation," in *IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*, Florence, Italy, May 2014, pp. 5975–5979.

[11] D. Comminiello, M. Scarpiniti, L. A. Azpicueta-Ruiz, J. Arenas-García, and A. Uncini, "Functional link adaptive filters for nonlinear acoustic echo cancellation," *IEEE Trans. Audio, Speech, Language Process.*, vol. 21, no. 7, pp. 1502–1512, July 2013.

[12] D. Comminiello, M. Scarpiniti, R. Parisi, and A. Uncini, "Convergence properties of nonlinear functional link adaptive filters," *Electronics Lett.*, vol. 49, no. 14, pp. 873–875, July 2013.

[13] D. Comminiello, M. Scarpiniti, L. A. Azpicueta-Ruiz, J. Arenas-García, and A. Uncini, "Nonlinear acoustic echo cancellation based on sparse functional link representations," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 7, no. 22, pp. 1172–1183, July 2014.

[14] D. Comminiello, M. Scarpiniti, S. Scardapane, R. Parisi, and A. Uncini, "Improving nonlinear modeling capabilities of functional link adaptive filters," *Neural Networks*, vol. 69, pp. 51–59, Sept. 2015.

[15] Y. Huang, J. Benesty, and J. Chen, *Acoustic MIMO Signal Processing*, Signals and Communication Technology. Springer-Verlag Berlin Heidelberg, 2006.

[16] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," in *IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*, Orlando, FL, May 2002, vol. 2, pp. 1881–1884.

[17] P. A. Naylor, J. Cui, and M. Brookes, "Adaptive algorithms for sparse echo cancellation," *Signal Process.*, vol. 86, no. 6, pp. 1182–1192, June 2004.

[18] C. Paleologu, S. Ciochină, and J. Benesty, "An efficient proportionate affine projection algorithm for echo cancellation," *IEEE Signal Process. Lett.*, vol. 17, no. 2, pp. 165–168, Feb. 2010.

[19] H. Zhao, Y. Yu, S. Gao, X. Zeng, and Z. He, "Memory proportionate APA with individual activation factors for acoustic echo cancellation," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 22, no. 6, pp. 1047–1055, June 2014.

[20] Y. Kopsinis, S. Chouvardas, and S. Theodoridis, "Sparsity-aware learning in the context of echo cancellation: A set theoretic estimation approach," in *22nd Eur. Signal Process. Conf. (EUSIPCO)*, Lisbon, Portugal, Sept. 1-5 2014, pp. 1846–1850.

[21] D. Comminiello, M. Scarpiniti, L. A. Azpicueta-Ruiz, J. Arenas-García, and A. Uncini, "A nonlinear architecture involving a combination of proportionate functional link adaptive filters," in *23rd Eur. Signal Process. Conf. (EUSIPCO)*, Nice, France, Sept. 2015, pp. 2919–2923.

[22] T. G. Burton and R. A. Goubran, "A generalized proportionate subband adaptive second-order Volterra filter for acoustic echo cancellation in changing environments," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, no. 8, pp. 2364–2373, Nov. 2011.

[23] F. Albu and K. Nishikawa, "The kernel proportionate NLMS algorithm," in *21st Eur. Signal Process. Conf. (EUSIPCO)*, Marrakech, Morocco, Sept. 9-13 2013, pp. 1–5.

[24] J. Arenas-García, A. R. Figueiras-Vidal, and A. H. Sayed, "Mean-square performance of a convex combination of two adaptive filters," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1078–1090, Mar. 2006.

[25] J. Arenas-García, L. A. Azpicueta-Ruiz, M. T. M. Silva, V. H. Nascimento, and A. H. Sayed, "Combinations of adaptive filters: Performance and convergence properties," *IEEE Signal Process. Mag.*, vol. 33, no. 1, pp. 120–140, Jan. 2016.

[26] L. A. Azpicueta-Ruiz, M. Lázaro-Gredilla, A. R. Figueiras-Vidal, and J. Arenas-García, "A block-based approach to adaptively bias the weights of adaptive filters," in *IEEE Int. Workshop Machine Learning for Signal Process. (MLSP)*, Beijing, China, Sept. 18-21 2011, pp. 1–6.

[27] J. Arenas-García and A. R. Figueiras-Vidal, "Adaptive combination of proportionate filters for sparse echo cancellation," *IEEE Trans. Audio, Speech, Language Process.*, vol. 17, no. 6, pp. 1087–1098, Aug. 2009.

[28] D. Comminiello, M. Scarpiniti, R. Parisi, and A. Uncini, "Combined adaptive beamforming schemes for nonstationary interfering noise reduction," *Signal Process.*, vol. 93, no. 12, pp. 3306–3318, Dec. 2013.