Functional Link Adaptive Filters for Nonlinear Acoustic Echo Cancellation

Danilo Comminiello, *Member, IEEE*, Michele Scarpiniti, *Member, IEEE*, Luis A. Azpicueta-Ruiz, Jerónimo Arenas-García, *Senior Member, IEEE*, and Aurelio Uncini, *Member, IEEE*

***Abstract*—This paper introduces a new class of nonlinear adaptive filters, whose structure is based on Hammerstein model. Such filters derive from the *functional link adaptive filter* (FLAF) model, defined by a nonlinear input expansion, which enhances the representation of the input signal through a projection in a higher dimensional space, and a subsequent adaptive filtering. In particular, two robust FLAF-based architectures are proposed and designed *ad hoc* to tackle nonlinearities in acoustic echo cancellation (AEC). The simplest architecture is the *split FLAF*, which separates the adaptation of linear and nonlinear elements using two different adaptive filters in parallel. In this way, the architecture can accomplish distinctly at best the linear and the nonlinear modeling. Moreover, in order to give robustness against different degrees of nonlinearity, a *collaborative FLAF* is proposed based on the adaptive combination of filters. Such architecture allows to achieve the best performance regardless of the nonlinearity degree in the echo path. Experimental results show the effectiveness of the proposed FLAF-based architectures in nonlinear AEC scenarios, thus resulting an important solution to the modeling of nonlinear acoustic channels.**

***Index Terms*—Functional links, nonlinear channel modeling, nonlinear acoustic echo cancellation, collaborative adaptive filters.**

I. Introduction

THE presence of nonlinearities in acoustic echo paths affects the performance of a conventional acoustic echo canceler (AEC) compromising the quality requirements of speech communications. In recent years, this topic has become even more sensible matter of interest due to the growing spread of low-cost commercial hands-free systems, which are often composed of poor quality elements, most of all electronic components, such as amplifiers and loudspeakers, and covering materials, such as plastic shells. These devices may cause significant nonlinearities in acoustic impulse responses (AIRs), thus leading to perceptual quality degradation of speech [1], [2]. In order to tackle this problem, nonlinear acoustic echo cancelers (NAECs) are employed, thus resulting in nonlinear path modeling and speech enhancement.

In recent years, different structures have been investigated in order to model the nonlinearities rebounding on acoustic echo paths. A prevalent technique is based on the use of *nonlinear transformations*, in an attempt to model different kinds of distortions [3]–[6]. A raised-cosine function is used in [3] to model both soft-clipping and hard-clipping nonlinearities. In [4], a two-parameter sigmoid function is proposed, whose slope and amplitude can be updated during the learning process. Another adaptive sigmoid function is used in [5] to evaluate NAEC performance as reverberation time changes. A more flexible solution is proposed in [6] by using spline functions, that are smooth parametric curves defined by interpolation of properly control points collected in a look-up table [7]. Block-based Wiener-Hammerstein models using nonlinear functions are also investigated [8], [9].

Despite NAECs using nonlinear functions provide good performance, the most popular nonlinear model for echo canceling applications is based on *adaptive Volterra filters* (VFs) [10], [11]. The generic structure of VFs derives from the well-known Taylor series, and it can be considered as a straightforward generalization of linear adaptive filters [10]. VFs can model a large range of nonlinearities, both with memory and memoryless [11]. However, acoustic echo cancellation, as well as other hands-free applications, requires large adaptive filters in order to model the AIR [12]. As the computational cost is proportional to the number of filter taps [13], the adaptation of VFs can become prohibitively expensive, compromising real-time implementation. Thus, in recent years, Volterra models with reduced computational complexity have been investigated to make real-time implementation possible [12], [14].

Another promising nonlinear modeling technique is based on the so-called *functional links*. The functional link is a functional operator which allows to represent an input pattern in a feature space where its processing turns out to be enhanced. The functional links have been initially proposed by Pao [15] with the aim of developing a class of single-layer feedforward neural networks, known as *functional link artificial neural networks* (FLANNs). Pao has shown that FLANN may be conveniently used for function approximation and pattern recognition with faster convergence rate and lesser computational load than a multi-layer perceptron ANN [15].

Manuscript received June 05, 2012; revised October 07, 2012, March 01, 2013; accepted March 01, 2013. Date of publication March 27, 2013; date of current version April 15, 2013. The work of D. Comminiello, M. Scarpiniti, and A. Uncini was supported in part by the Italian National Project: "Computational Analysis of Acoustic Scene for Machine Listening Systems and Augmented Reality Communications," under Grant C26A11BC43. The work of J. Arenas-García and L. A. Azpicueta-Ruiz was supported in part by MICINN projects TEC-2011-22480 and PRI-PIBIN-2011-1266. This was performed in part while D. Comminiello was a visiting Ph.D. student at the Department of Signal Theory and Communications. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Woon-Seng Gan.

D. Comminiello, M. Scarpiniti, and A. Uncini are with the Department of Information Engineering, Electronics and Telecommunications, "Sapienza" University of Rome, 00184 Rome, Italy (e-mail: danilo.comminiello@uniroma1.it; michele.scarpiniti@uniroma1.it; aurel@ieee.org).

L. A. Azpicueta-Ruiz and J. Arenas-García are with the Department of Signal Theory and Communications, Universidad Carlos III de Madrid, 28911 Leganés, Spain (e-mail: lazpicueta@tsc.uc3m.es; jarenas@tsc.uc3m.es).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASL.2013.2255276

In this paper, we adopt a nonlinear adaptive filtering model based on functional links, called *functional link adaptive filter* (FLAF), which exploits the nonlinear modeling capabilities of functional links and the filtering properties of linear adaptive algorithms. This model is different from the original formulation of FLANN but uses the same approach of recently proposed FLANNs (see for example [16], [17]). In order to avoid any ambiguity we have renamed the model as FLAF. FLAFs are computationally less expensive than former FLANNs and VFs, thus resulting an effective tool to model nonlinearities affecting speech signals in acoustic applications, such as NAEC.

We propose two novel FLAF-based architectures to specifically address some of the inherent properties of AEC, expanding the proposals preliminarily presented in [18]. The first architecture is represented by the *split FLAF* (SFLAF), in which the linear and the nonlinear elements are adapted separately in order to improve convergence performance. The second FLAF-based scheme is proposed to deal with the lack of *a priori* knowledge regarding the presence of nonlinear distortions or the ratio between the power of linear and nonlinear distortions. Such information is relevant in the choice of an AEC, since when an echo path is roughly linear, or contains negligible nonlinearities, an NAEC could perform worse than a conventional AEC due to the gradient noise introduced by the nonlinear filter [1], [19]. The proposed solution to this problem is represented by *collaborative FLAF* (CFLAF), which draws inspiration from recent works for combining adaptive filters [20]–[22]. Combined filters have been used in AEC applications [1], [23], as well as in other adaptive signal processing areas [24]. Differently from combined systems, the proposed collaborative architecture allows an exchange of information between the involved filters.

The paper is organized as follows: Section II introduces a comprehensive description of FLAF, including its approach to nonlinear modeling, how to choose a proper functional link expansion, and its main advantages. The proposed FLAF-based architectures are described in Section III. Section IV contains the experimental results proving the effectiveness and the robustness of the proposed architectures in several NAEC scenarios. Finally, in Section V our conclusions are drawn.

A. Notation

In this paper, matrices are represented by boldface capital letters and vectors are denoted by boldface lowercase letters. Time-varying vectors and matrices show discrete-time index as a subscript index, while in time-varying scalar elements the time index is denoted in square brackets. A regression vector is represented as $\mathbf{x}_n \in \mathbb{R}^M = [x[n]\; x[n-1] \ldots x[n-M+1]]^T$, where M is the overall vector length and $x[n-i]$ is individual entry at the generic time instant $n-i$. A generic coefficient vector, in which all elements depend on the same time instant, is denoted as $\mathbf{w}_n \in \mathbb{R}^M = [w_0[n]\; w_1[n] \ldots w_{M-1}[n]]^T$, where $w_i[n]$ is the generic i-th individual entry at n-th time instant. When the coefficient vector is a realization of a time-invariant process the time index is omitted. All vectors are represented as column vectors.

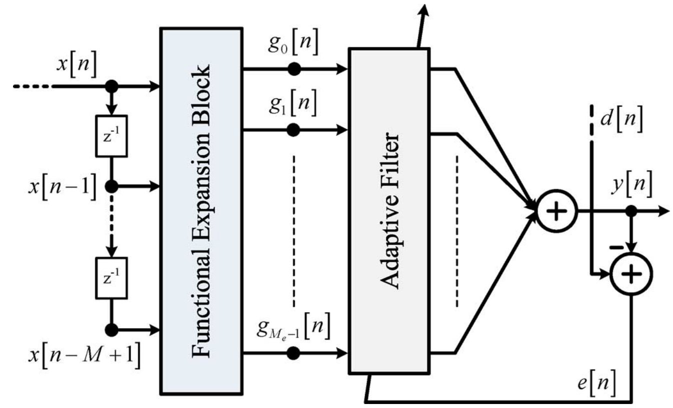

Fig. 1. The functional link adaptive filter.

II. Functional Link Adaptive Filters

The main idea which underpins the FLAF approach is that of enhancing the original input signal right from the start by representing it in a space of higher dimension [15]. This process derives directly from the machine learning theory, and more exactly from Cover's Theorem on the separability of patterns [25]. The functional link adaptive filtering is carried out in two stages: a nonlinear functional expansion of the input and a subsequent linear filtering, as depicted in Fig. 1.

At n-th time instant FLAF receives an input buffer $\mathbf{x}_n \in \mathbb{R}^M = [x[n]\; x[n-1] \ldots x[n-M+1]]^T$, where M is the input buffer length. Then, FLAF processes the input buffer by means of a *functional expansion block* (FEB). The FEB consists of a series of functions, which might be a subset of a complete set of orthonormal basis functions, satisfying universal approximation constraints [26]. The term "functional links" actually refers to the functions contained in the chosen set $\Phi = \{\varphi_0(\cdot), \varphi_1(\cdot), \ldots, \varphi_{Q-1}(\cdot)\}$, where Q is the number of functional links. The FEB processes the input buffer by passing each element of the buffer $\mathbf{x}_n$ as argument for the chosen functions, each yielding a subvector $\bar{\mathbf{g}}_{i,n} \in \mathbb{R}^Q$:

$$\bar{\mathbf{g}}_{i,n} = \begin{bmatrix} \varphi_0(x[n-i]) \\ \varphi_1(x[n-i]) \\ \vdots \\ \varphi_{Q-1}(x[n-i]) \end{bmatrix} \quad (1)$$

The described process, for $i = 0, \ldots, M-1$, results in an *expanded buffer* $\mathbf{g}_n$:

$$\begin{aligned} \mathbf{g}_n &= \begin{bmatrix} \bar{\mathbf{g}}_{0,n}^T & \bar{\mathbf{g}}_{1,n}^T & \ldots & \bar{\mathbf{g}}_{M-1,n}^T \end{bmatrix}^T \\ &= [g_0[n] \quad g_1[n] \quad \ldots \quad g_{M_\mathrm{e}-1}[n]]^T \end{aligned} \quad (2)$$

which is a concatenation of subvectors and whose individual entries are denoted as $g_m[n]$, with $m = 0, \ldots, M_\mathrm{e} - 1$ where $M_\mathrm{e} \geq M$ is the expanded buffer length.

Although no new *ad hoc* information has been inserted, the expansion process enhances the nonlinear modeling. Once achieved the expanded buffer, the functional link adaptive filtering process is completed by simply linearly filtering the expanded buffer. The use of an adaptive filter after the FEB, already implemented in previous works [16], [17], represents

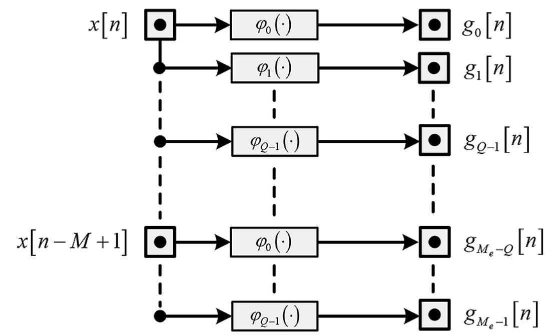

Fig. 2. Functional expansion in memoryless FLAF.

a significant difference from FLANNs in their original formulation [15] and some of their recent use [27], and it provides several advantages to FLAF, as described in Subsection II.C.

A. FLAF With Trigonometric Expansion

An important choice in the FEB design concerns the expansion type, i.e., the basis functions, or a subset of it, to assign for each functional link. This choice mostly depends on the application and in particular on the signals involved in the processing.

Basis functions, satisfying universal approximation constraints [26], may be a subset of orthogonal polynomials, like Chebyshev, Legendre and trigonometric polynomials, or just approximating functions, such as sigmoid and Gaussian functions. Among these, one of the most popular functional expansion consists of trigonometric polynomial functions [17], [27], [28], which provide the best compact representation of any nonlinear function in the mean square sense [28]. Trigonometric functions are also computationally cheaper than power series-based polynomials. Due to these properties we adopt the trigonometric expansion to model nonlinear speech signals.

Taking into account the i-th sample of the input buffer, a generic set of functional links for an FLAF using trigonometric basis expansion can be written as:

$$\varphi_j(x[n-i]) = \begin{cases} x[n-i], & j=0 \\ \sin(p\pi x[n-i]), & j=2p-1 \\ \cos(p\pi x[n-i]), & j=2p \end{cases} \tag{3}$$

where $p = 1,\ldots,P$ is the expansion index, being P the *expansion order*, and $j = 0,\ldots,Q-1$ is the functional link index. In particular, the functional link index denotes the sequence of output samples resulting from the nonlinear expansion of each input sample, aside from the kind of transformation implemented by any functional link. The functional link set Φ in this case is composed of $Q = 2P+1$ functional links, where 1 refers to the linear element $\varphi_0(x[n-i])$, which is the replica of the current i-th input sample. In this way, the expanded buffer $\mathbf{g}_n$ contains both linear and nonlinear elements.

B. Memory and Memoryless FLAF

Another important choice in the FLAF design concerns the memory of the input buffer, which bears on the correspondence between samples of the input buffer and those of the expanded buffer. The choice of taking into account some memory is strictly related to the nature of the system to identify. In fact, it depends on the type of nonlinearity which deteriorates the input signal, and more precisely on whether the nonlinearity is *instantaneous*, i.e., it is independent from the time instant, or *dynamic*, i.e., the nonlinearity depends even on the time instant.

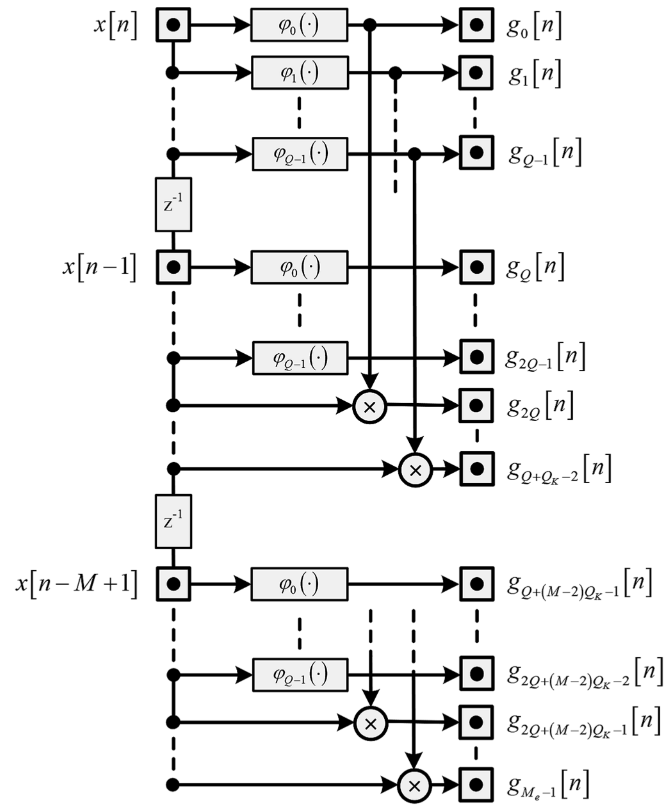

Fig. 3. Functional expansion in FLAF with memory.

The kind of nonlinearity that is more frequently encountered in acoustic systems is the *memoryless* (or instantaneous) one [12]. Modeling of memoryless nonlinearities is very important since many complex nonlinear systems can be broken down into a linear system and an instantaneous nonlinear operator. Memoryless nonlinearities require a *memoryless FLAF* which generates a straightforward relation between the input buffer and the expanded buffer, as depicted in Fig. 2. The set of functional links described by (3) is a typical example of a memoryless expansion, since each functional link depends only on the current i-th input sample.

A set of memoryless functional links provides a satisfying approximation of a continuous multivariate function, whether the nonlinearity is instantaneous or dynamic. However, in case of nonlinear dynamic systems, better results may be achieved exploiting the flexibility of the FEB. In particular, it is possible to add further functional links to a memoryless structure in order to take into account the memory of a certain dynamic nonlinearity. We refer to such set $\Phi_K = \{\varphi_0(\cdot),\ldots,\varphi_{Q-1}(\cdot),\varphi_Q(\cdot),\ldots,\varphi_{Q_K-1}(\cdot)\}$ as a set of *functional links with memory*, where K denotes the memory order and $Q_K > Q$ is the overall number of functional links with memory. A way of considering the memory of a nonlinearity is that of taking into account the outer products of the i-th input sample with the functional links of the K previous input samples, as depicted in Fig. 3.

In a FLAF with memory, the memory order K determines the length of the additional functional links, i.e., the depth of the outer products between the i-th input sample and the functional links related to the previous input samples. Fig. 3, similarly to [17], shows an expansion with memory order $K = 1$.

C. FLAF Coefficient Adaptation

Once chosen the set of basis functions, the problem focuses on finding out the coefficients of the FLAF weight vector $\mathbf{w}_n \in \mathbb{R}^{M_e}$, defined as:

$$\mathbf{w}_n = [w_0[n] \quad w_1[n] \quad \ldots \quad w_{M_e-1}[n]]^T, \tag{4}$$

in order to yield the best possible approximation of the nonlinear model within a small error value. The explicit representation of the FLAF error signal $e[n]$ is:

$$\begin{aligned} e[n] &= d[n] - y[n] \\ &= d[n] - \mathbf{g}_n^T \mathbf{w}_{n-1} \end{aligned} \tag{5}$$

whose minimization depends on a proper estimate of the weights of the filter. In order to adjust the coefficient vector $\mathbf{w}_n$ it is possible to use any adaptive algorithm based on gradient descent rule [13]. In this work we use adaptive algorithms based on stochastic gradient rule to adapt the filter coefficients since the goal of this paper is to introduce new FLAF-based architectures. However, a wide range of adaptive algorithms can be easily used to adapt FLAF coefficients.

The joint use of functional links and stochastic gradient adaptation entails several attractive advantages. Firstly, FLAF has a hugely flexible architecture due to its scalable nonlinear expansion and to its scalable structural complexity. The former property allows to choose *a priori* a suitable series of functional links according to the application of interest. Also, the introduction of high-order functions entails a robust generalizing ability [29], [30]. On the other hand, the scalable complexity allows to deal with high dimension input signals, modeling the FEB structure in order to find the right trade-off between performance and computational complexity, according to application requirements and available computational resources. The flexibility of FLAF architecture allows an easy integration of any *a priori* knowledge of a certain nonlinear system.

FLAF might also show some drawbacks, often due by the extreme flexibility of the architecture that may also cause a biased convergence [31]. Such problems can be addressed introducing more robust FLAF architectures.

III. Robust FLAFs for NAEC

FLAF performance may incur in a biased solution due to the fact that the FEB expands the whole input buffer. A control over the expanded buffer seems to be problematic, since the choice of the input buffer length is bound up with an accurate estimate of the acoustic impulse response. Additionally, the setting of the optimal parameters of the adaptive filter, such as the step size, may result critical. This may occur when the presence of nonlinearities in the echo path varies in time, as it is often the rule in AEC. In that sense, improvements can be achieved modifying the FLAF structure up to yield robust filtering architectures expressly suited for NAEC application. According to this, we propose two FLAF-based architectures:

- the *split FLAF* decouples the adaptation speed for the linear and nonlinear elements of the echo path, thus solving convergence drawbacks;
- the *collaborative FLAF* adaptively biases the acoustic channel modeling whether the path is affected by nonlinearities or not.

A. The Split Functional Link Adaptive Filter

In the acoustic channel modeling, a first significant improvement of the FLAF performance can be achieved separating the adaptation of linear and nonlinear elements of the expanded buffer. In particular, it is possible to consider two different adaptive filters in parallel, one completely linear and the other purely nonlinear. In literature other parallel structures involving functional links are proposed (e.g., in [19], [27]), but none of them decoupled linear and nonlinear elements. In the SFLAF the linear filter receives the whole input buffer and aims exclusively at estimating the echo path. On the other hand, the nonlinear filter receives an expanded buffer devoid of linear elements, generated by a purely nonlinear set of functional links:

$$\varphi_j(x[n-i]) = \begin{cases} \sin(p\pi x[n-i]), & j = 2p-2 \\ \cos(p\pi x[n-i]), & j = 2p-1 \end{cases} \tag{6}$$

where again $j = 0, \ldots, Q-1$ and $p = 1, \ldots, P$. In this case, the functional link set is composed of $Q = 2P$ functional links. Equation (6) describes a memoryless expansion. An expansion with memory can be easily achieved adding the cross-product terms as depicted in Fig. 3. Therefore, the nonlinear FLAF focuses only on the modeling of nonlinearities affecting the echo signal. In this way it is possible to distinguish two different filterings with two different settings of the parameters, such that each filter can accomplish its task at best.

Using this structure, the FEB can receive the whole input buffer or just a portion of it. This is the reason why the nonlinear input buffer length is denoted with $M_i \leq M$. The possibility of choosing the expanded buffer length for the SFLAF may also be beneficial from a computational point of view compared with the FLAF described in Section II and with FLANN [17], [27], [28], [32]. This architecture, to which we will refer as *Split Functional Link Adaptive Filter* (SFLAF), is depicted in Fig. 4, where it is possible to notice that the SFLAF output signal results from the sum of the output of the linear filter and the output of the nonlinear FLAF:

$$y[n] = y_L[n] + y_{FL}[n] \tag{7}$$

where

$$\begin{aligned} y_L[n] &= \mathbf{x}_n^T \mathbf{w}_{L,n-1} \\ y_{FL}[n] &= \mathbf{g}_n^T \mathbf{w}_{FL,n-1} \end{aligned} \tag{8}$$

being $\mathbf{w}_{L,n} \in \mathbb{R}^M = [w_0[n] \quad w_1[n] \quad \ldots \quad w_{M-1}[n]]^T$ the coefficient vector of the linear filter, and $\mathbf{w}_{FL,n} \in \mathbb{R}^{M_e} = [w_0[n] \quad w_1[n] \quad \ldots \quad w_{M_e-1}[n]]^T$ the coefficient vector of

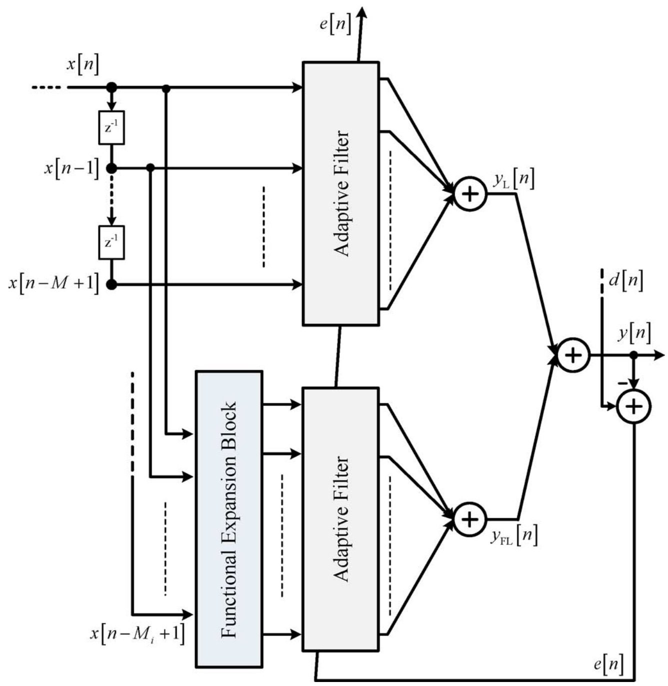

Fig. 4. The Split Functional Link Adaptive Filter.

the nonlinear FLAF. From Fig. 4, it is possible to gather that both linear and nonlinear filters are adapted using the overall error signal $e[n] = d[n] - y[n]$. However, each filter can be adapted using a different adaptation rule and different parameter settings. This architecture provides a significant improvement in terms of convergence performance and great flexibility which allows to achieve a good compromise between computational load and signal enhancement.

The computational complexity of the SFLAF is subjected to several parameters, apart from the adaptive algorithm adopted. Considering an NLMS for the adaptation of both the linear and the nonlinear part, the SFLAF requires $3M + 2$ multiplications and $2M$ additions for the linear path, while for the memoryless nonlinear path the cost is $3M_{\mathrm{e}} + PM_{\mathrm{i}} + P + 1$ multiplications, $3M_{\mathrm{e}} + 1$ additions and $2P$ function evaluations (i.e., sines and cosines). Such function evaluations can be efficiently implemented with lookup tables. A nonlinear path with memory needs an additional cost of $(M_{\mathrm{i}} - K)KP + \sum_{k=1}^{K-1} kP$ multiplications. As it is possible to notice, the overall complexity can vary significantly according to the buffer lengths M_{i} and M_{e}, the expansion order P and the memory order K.

B. The Collaborative Functional Link Adaptive Filter

Changes proposed in the SFLAF gives robustness to an NAEC, due to the possibility to make the right choice for the critical parameters of the filter. Some drawbacks may linger on when the nonlinearity degree varies in time. In particular, a non-optimal filtering may occur when the linear-to-nonlinear power ratio (LNLR) changes over time. It is well known [1], [19] that NAEC performance may result inferior than that of a conventional linear AEC when the desired signal is not affected by any nonlinearity, or when the nonlinearity level is negligible. In that case, the nonlinear filter only brings some gradient noise in the filtering process, thus degrading NAEC

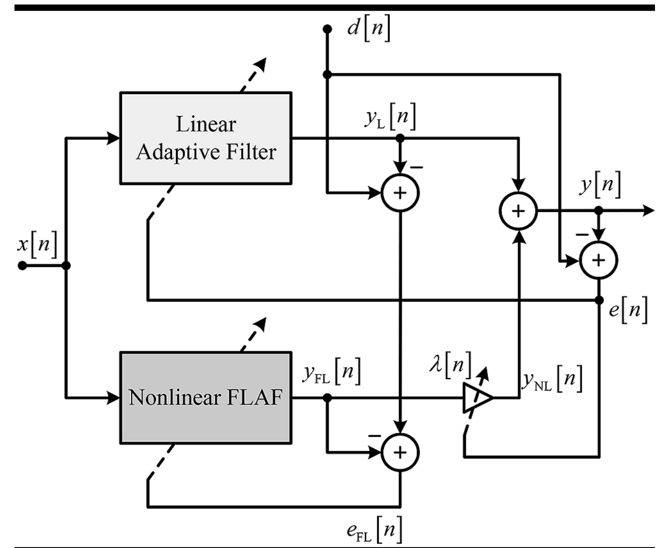

Fig. 5. The Collaborative Functional Link Adaptive Filter.

performance. In addition, the LNLR is unknown *a priori* and it is time-varying for nonstationary signals, like speech. Thereby, it is not possible *a priori* to know whether an NAEC will improve or deteriorate the cancellation. This problem, along with the expensive computational cost of an NAEC, affects the strategies of many companies that provide teleconferencing services, which often choose to drop the use of nonlinear echo cancelers even at the expense of communication quality.

With the aim of designing an NAEC robust to the changes of LNLR, we propose a collaborative architecture, based on the convex combination of adaptive filters [21]. Such *collaborative FLAF* (CFLAF) is capable of activating or deactivating automatically the nonlinear filtering path, thus adaptively biasing the acoustic channel modeling according to the nonlinearity residual in the echo path.

The CFLAF is depicted in Fig. 5, in which it is possible to notice that the overall output signal $y[n]$ is different from (7):

$$y[n] = y_{\mathrm{L}}[n] + \lambda[n] y_{\mathrm{FL}}[n] \qquad (9)$$

where $\lambda[n]$ is a *shrinkage parameter* which allows to either keep or remove the output of the nonlinear FLAF as required by the filtering scenario. The nonlinear path of the CFLAF, including the adaptive switch $\lambda[n]$, can be also interpreted as a convex combination between the nonlinear FLAF with a null *virtual* filter, which we refer to as *all-zero kernel* (AZK) since its coefficients are static and set to zero [1]. In fact, in (9) the term related to the AZK, and thus weighted with $(1 - \lambda[n])$, is omitted as its contribution is null.

Due to the fact that linear and nonlinear filterings have different tasks, each filter in CFLAF is updated using different error signals in order to completely exploit the collaborative structure. In particular, the linear filter $\mathbf{w}_{\mathrm{L},n}$ pursues the minimization of the overall error signal $e[n] = d[n] - y[n]$, since the output contribution of the linear filter is always present. In addition, the estimate of the AIR results more accurate when we consider the overall error signal instead of the local linear error signal. Differently, the nonlinear FLAF $\mathbf{w}_{\mathrm{FL},n}$ is updated using the local error $e_{\mathrm{FL}}[n]$ from which the linear output $y_{\mathrm{L}}[n]$ is subtracted, as it is always taken into account by the linear filtering:

$$e_{\mathrm{FL}}[n] = d[n] - (y_{\mathrm{L}}[n] + y_{\mathrm{FL}}[n]). \tag{10}$$

The shrinkage (or *mixing*) parameter $\lambda[n]$ can be adapted, considering the convex constraints, i.e., $0 \leq \lambda[n] \leq 1$, through the adaptation of an auxiliary parameter, $a[n]$, related to $\lambda[n]$ by means of a sigmoidal function [21]:

$$\lambda[n] = \mathrm{sgm}(a[n]) = \frac{1}{(1+e^{-a[n]})}. \tag{11}$$

$\lambda[n]$ is computed adapting $a[n]$ according to a gradient descent rule as $a[n+1] = a[n] + \Delta a[n]$, where $\Delta a[n]$ results from a *normalized least mean squares* (NLMS) adaptation:

$$\begin{aligned}\Delta a[n] &= -\frac{1}{2}\mu_a \frac{\partial e^2[n]}{\partial a[n]} \\ &= -\frac{\mu_a}{r[n]} e[n] \frac{\partial(d[n] - y_{\mathrm{L}}[n] - \lambda[n] y_{\mathrm{FL}}[n])}{\partial \lambda[n]} \frac{\partial \lambda[n]}{\partial a[n]} \\ &= \frac{\mu_a}{r[n]} e[n] y_{\mathrm{FL}}[n] \lambda[n](1-\lambda[n]) \end{aligned} \tag{12}$$

where:

$$r[n] = \beta r[n-1] + (1-\beta) y_{\mathrm{FL}}^2[n] \tag{13}$$

is a rough low-pass filtered estimate of the power of the signal of interest [33]. The parameter β is a smoothing factor which ensures that $r[n]$ is adapted faster than any filter component. The value of $a[n]$ is kept within $[4, -4]$ for practical reasons [21].

This FLAF-based architecture is named "collaborative" because, differently from combined architectures [20]–[22], it allows an exchange of information between the involved filters. As it is possible to notice from Fig. 5, the error signal with which the linear filter is adapted may contain the output contribution of the nonlinear filter, due to (9). On the other hand, the error with which the nonlinear FLAF is adapted always contains the linear output contribution, as represented by (10).

The proposed CFLAF is robust to any nonlinearity level, since when the echo path is merely linear $\lambda[n]$ converges towards 0 and the whole scheme behaves like a purely linear filter, thus avoiding any gradient noise from the nonlinear FLAF. On the other hand, when the echo path conveys nonlinearities the shrinkage parameter approaches 1 according to the nonlinearity level in the echo path. Note that when $\lambda[n] = 1$ the CFLAF architecture works like the SFLAF one. As regards the computational complexity the CFLAF implies an overcharge, with respect to the SFLAF, of only 10 multiplications, 5 additions and 1 function evaluation.

IV. Experimental Results

In this section we evaluate the performance of proposed FLAF-based architectures in NAEC scenarios. The first three sets of experiments are conducted in a simulated teleconferencing environment with a reverberation time of $T_{60} \approx 120$ ms. The AIR between the loudspeaker and the microphone is simulated by means of Matlab *Roomsim* toolbox [34], using a 8 kHz sampling rate; the AIR is truncated after $M = 300$ samples, which is also the length of the linear filters.

In order to simulate an asymmetric loudspeaker distortion, we apply the following memoryless sigmoidal nonlinearity to the far-end signal before convolving it with the AIR:

$$\bar{y}[n] = \gamma \left(\frac{1}{1+e^{(-\rho q[n])}} - \frac{1}{2} \right) \tag{14}$$

where:

$$q[n] = \frac{3}{2}x[n] - \frac{3}{10}x^2[n]. \tag{15}$$

In (14), the parameter γ is the sigmoid *gain* and it is set equal to $\gamma = 2$, while ρ represents the sigmoid *slope* and it is chosen as:

$$\rho = \begin{cases} 4, & q[n] > 0 \\ \frac{1}{2}, & q[n] \leq 0 \end{cases}. \tag{16}$$

With the aim of simulating dynamic systems, it is possible to replace $q[n]$ in (14) with a sparse fifth order nonlinearity:

$$\begin{aligned} q_{\mathrm{d}}[n] = &\frac{3}{2}x[n] - \frac{3}{10}x^2[n] + \frac{9}{5}x[n]x[n-1] \\ &+ \frac{1}{2}x[n]x[n-2] - \frac{2}{5}x[n]x[n-3] \\ &- \frac{3}{2}x[n-1]x[n-2] + \frac{9}{10}x[n]x[n-1]x[n-3] \\ &+ \frac{1}{2}x[n-1]x[n-2]x[n-3] - \frac{1}{10}x^2[n-1] \\ &+ \frac{1}{5}x^2[n-2] - \frac{1}{10}x^2[n-3] + \frac{3}{10}x[n-3]x[n-4] \\ &- \frac{6}{5}x^2[n-4] + \frac{1}{5}x[n-1]x[n-5] \\ &+ \frac{3}{10}x[n-3]x[n-5] + \frac{6}{5}x^2[n-5]. \end{aligned} \tag{17}$$

Two kinds of far-end input signal are used: a colored noise and female speech input. The colored signal is generated by means of a first-order autoregressive model, whose transfer function is $\sqrt{1-\theta^2}/(1-\theta z^{-1})$, with $\theta = 0.8$, fed with an independent and identically distributed (i.i.d.) Gaussian random process. In both the cases, additive Gaussian noise is added at the output of the echo path in order to provide 20 dB of *signal to noise ratio* (SNR). The length of the experiments is 10 seconds.

Performance of all architectures is evaluated in terms of *echo return loss enhancement* (ERLE), which is defined as:

$$\mathrm{ERLE}[n] = 10\log_{10}\left(\frac{\mathrm{E}\{d^2[n]\}}{\mathrm{E}\{e^2[n]\}}\right) \tag{18}$$

where the operator $\mathrm{E}\{\cdot\}$ denotes the mathematical expectation, which is estimated by averaging over 100 experiments with respect to input and noise (in the case of colored input) or just noise when the input signal is speech.

A. Evaluation of the Split FLAF

First of all it is important to show the performance improvement brought by the SFLAF compared to the FLAF. We also take into account a 2nd order VF and linear filter in order to have respectively a nonlinear and a linear reference. Since a full 2nd order VF could introduce an excessive number of nonlinear elements for the memoryless case, we also add the performance

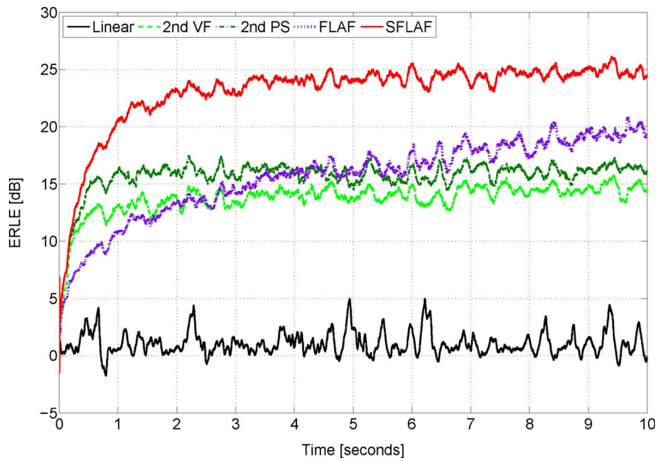

Fig. 6. Performance comparison in terms of ERLE between a SFLAF, a FLAF, a 2nd order VF, a 2nd order PS and a linear filter, in case of colored input.

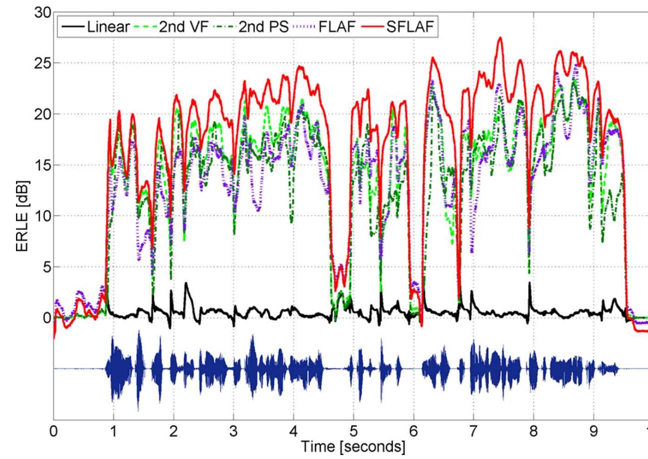

Fig. 7. Performance comparison in terms of ERLE between a SFLAF, a FLAF, a 2nd order VF, a 2nd order PS and a linear filter, in case of female speech input.

of a 2nd order VF with only the main diagonal, i.e., a 2nd order power series (PS) filter, in order to provide a complete comparison. We use the same parameter setting for both the FLAF and the SFLAF. All the adaptive filters are updated using an NLMS algorithm. Due to the use of trigonometric functional expansion, we normalized the input signal in order to expand a signal whose amplitude is comprised in the range $[-1,1]$. The input buffer length is set to $M_{\rm i}=M$, i.e., we use a sufficient number of linear elements to estimate the AIR both for the FLAF and the SFLAF. After several experiments, we choose an expansion order of $P=5$, which, in this case, results a good trade-off between performance and complexity. The step size parameter $\mu_{\rm L}=0.2$ is set for the adaptation of the FLAF and of the linear filter of the SFLAF. In order to exploit the SFLAF advantage of having two separate filters, we adopt a different step size parameter for the nonlinear filter of the SFLAF equal to $\mu_{\rm FL}=0.5$, thus reacting faster to nonlinearities, whose nature is highly nonstationary. We use $\mu_{\rm L}$ to adapt the linear kernel of the 2nd order VF and $\mu_{\rm FL}$ to adapt the corresponding quadratic kernel, while the linear filter is adapter using $\mu_{\rm L}$. Both the VF kernels and the linear filter have length equal to M. All the filters are adapted using the same regularization parameter $\delta=20\sigma_x^2$, where σ_x^2 represents the variance of the input signal.

In order to have an unbiased comparison, we choose a memoryless scenario, which involves (14) and (15), and we adopt memoryless architectures for both the FLAF and the SFLAF. We compare the performance of the above mentioned architectures in terms of ERLE, both for the colored input and for the female speech input.

Results are respectively depicted in Figs. 6 and 7 in which the performance improvement brought by the SFLAF compared to the other filtering architectures is evident. In the case of speech input the improvement is less conspicuous since nonlinearities do not affect the signal in a uniform way due to the time-varying nature of the speech signal. It has also to be considered that, apart from the step size, it is also possible to change the input buffer length of the nonlinear path of the SFLAF. This might improve the performance when nonlinearities are mild and require a small number of nonlinear elements for an optimal modeling.

B. Comparison Between Memory and Memoryless SFLAF

Flexibility is one of the most important feature of FLAF-based architectures. As a matter of fact, it is possible to design the FEB in order to model at best nonlinearities in the echo path, as explained in Section II. Actually, when nonlinearities in the echo path are highly dynamic, it is possible to choose a functional expansion with memory. In order to have a clear evaluation of advantages and drawbacks of using a functional link expansion with memory rather than a memoryless one, we consider a dynamic scenario described by (14) and (17). We take into account the same SFLAF, VF, and linear filter of the previous set of experiments, using also the same settings. In addition, we evaluate the performance of an SFLAF with memory order of $K=2$ in order to have a fair comparison with the VF.

Results are shown in Figs. 8 and 9 respectively for colored and speech input. It is possible to notice that the difference between the SFLAF with memory and the memoryless one is quite small. In the speech case (Fig. 9), slight improvements occur only in proximity of signals peaks. It is also possible to notice that the difference between FLAF-based architectures and VF is reduced compared with the previous experiments, and this is due to the fact that a full VF works better in dynamic scenarios. Despite this, even in this case FLAF-based architectures show the best performance. Therefore, using an SFLAF with memory may be important in enhancing the speech, although it involves a larger computational load compared with the memoryless SFLAF. This is the reason why there is not a universal best choice, nevertheless a proper SFLAF design can be based on available resources and quality requirements.

C. Evaluation of the Collaborative FLAF

In the following set of experiments we demonstrate the effectiveness of the CFLAF architecture. We consider a changing environment in which for the first 5 seconds the echo path is merely linear and then a memoryless nonlinearity occurs (the same used in Subsection IV.A). We evaluate the performance in terms of ERLE of a CFLAF, an SFLAF and a linear filter. Both FLAF-based architectures adopt a memoryless functional expansion. We use the same parameter settings of the previous experimental sets, and, in addition, we choose $a[0]=0, r[0]=$

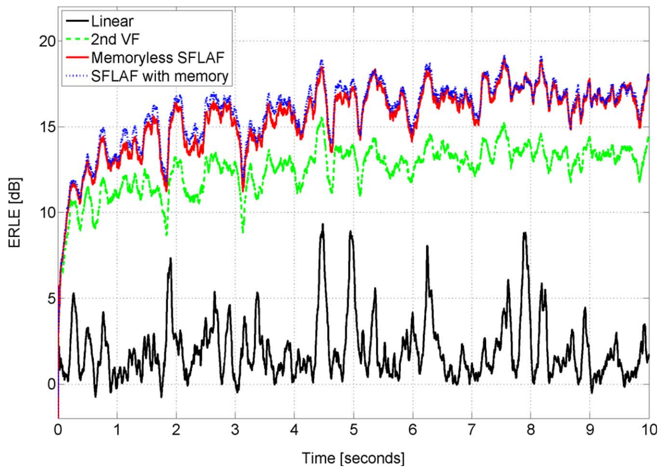

Fig. 8. Performance comparison in terms of ERLE between an SFLAF with memory, a memoryless SFLAF, a 2nd order VF and a linear filter, in case of colored input corrupted by a dynamic nonlinearity.

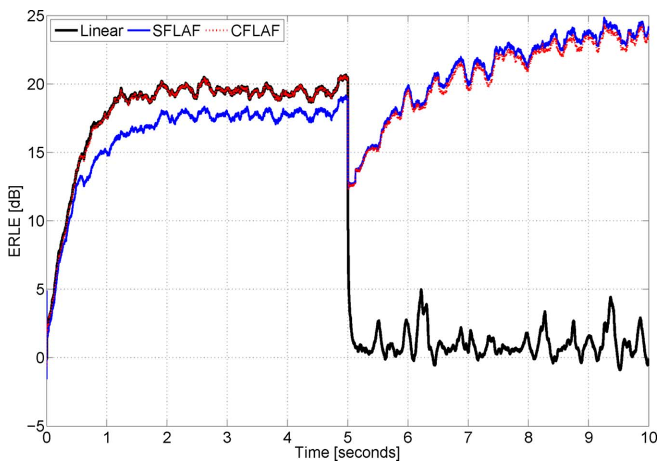

Fig. 9. Performance comparison in terms of ERLE between an SFLAF with memory, a memoryless SFLAF, a 2nd order VF and a linear filter, in case of female speech input corrupted by a dynamic nonlinearity.

$1, \mu_a = 0.5, \beta = 0.9$ related to the adaptive combination of the CFLAF.

Fig. 10 evaluates the CFLAF behavior in terms of ERLE when the far-end input is a colored signal. In the first 5 seconds the echo path is not affected by nonlinearities and the CFLAF follows the linear filter, which results the best performing model. In this stage it is evident the performance gap between SFLAF and CFLAF which is due to the gradient noise introduced by the nonlinear elements of the SFLAF. On the other hand, in the second half of the experiment the CFLAF behaves like the SFLAF as expected. In Fig. 11 the evolution of the shrinkage parameter $\lambda[n]$ (defined in (11)) of the CFLAF is depicted. It is possible to notice how the parameter switches from 0 to 1 when the channel becomes nonlinear. Also, in the first part of Fig. 11 the convergence rate is not very fast due to the non-optimal initial convergence of the linear filter. On the other hand, when the nonlinearity is introduced the shrinkage parameter rapidly converges to 1 since the AIR is already well-estimated by the linear filter. The same experiment is repeated in case of speech input, and a similar result is shown in Fig. 12.

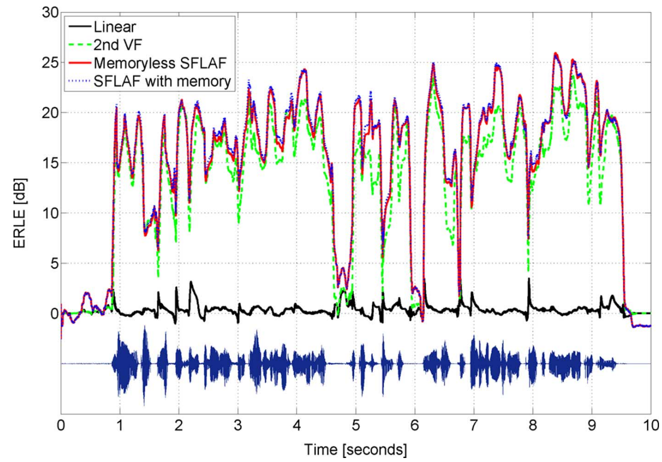

Fig. 10. Evaluation of the Collaborative FLAF in terms of ERLE in case of colored input. In the first 5 seconds the echo path is purely linear, while in the second half of the experiment a nonlinearity occurs.

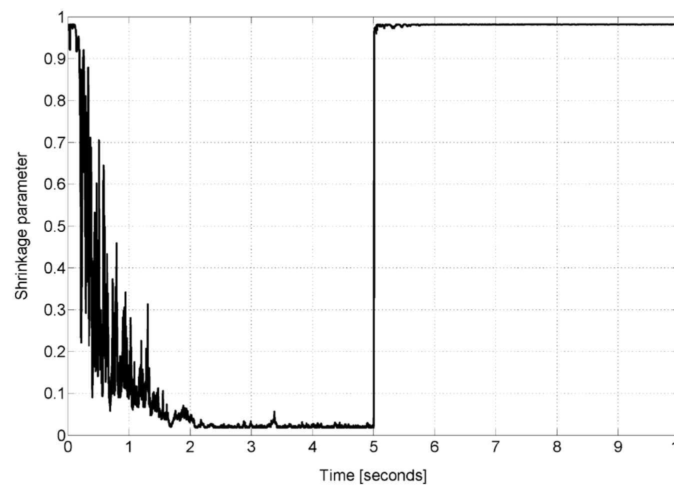

Fig. 11. Shrinkage parameter behavior in case of colored input. In the first 5 seconds the echo path is purely linear, while in the second half of the experiment a nonlinearity occurs.

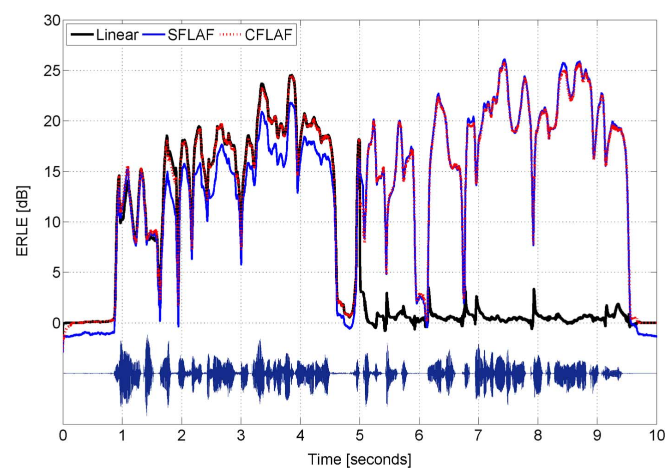

Fig. 12. Evaluation of the Collaborative FLAF in terms of ERLE in case of female speech input. In the first 5 seconds the echo path is purely linear, while in the second half of the experiment a nonlinearity occurs.

D. Evaluation of FLAF-Based Architectures in a Real Scenario

In order to provide a complete evaluation of the proposed FLAF-based architectures we take into account a real AEC scenario. The experiment takes place in a typical office room with

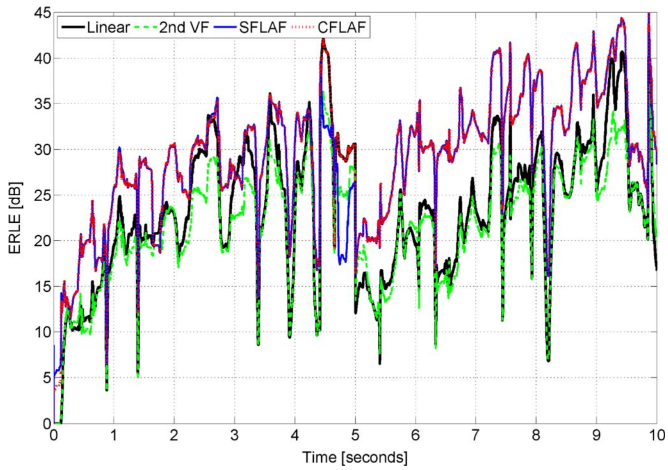

Fig. 13. Evaluation of the FLAF-based architectures in case of male speech input in a real scenario. In the first 5 seconds the echo path is mainly linear, while in the second half the LNLR is lower due to an increase of the loudspeaker volume.

a quite low level of background noise. We use a commercial loudspeaker but a professional microphone to avoid any further introduction of nonlinearities. The microphone is located 40 cm far from the loudspeaker, which is a typical distance in a desktop teleconferencing scenario. The input signal is a male speech recorded at 16 kHz sampling rate. The length of the experiment is 10 seconds. In the first 5 seconds the signal is reproduced at a low/medium volume level to provide a mainly linear echo path. At second 5 we turn up the volume of the loudspeaker with an increase of 6 dB and we keep this level until the end of the experiment. The volume increase may cause a larger loudspeaker distortion.

We estimate the AIR by using a linear filter length of $M = 1100$ samples (which is higher with respect to the previous experiments also due to the higher sampling rate). For the FLAF-based architectures (SFLAF and CFLAF) we use the same parameter setting of the previous experiment: step sizes $\mu_{\mathrm{L}} = 0.2$ and $\mu_{\mathrm{FL}} = 0.5$, regularization parameter $\delta = 20\sigma_x^2$, expansion order $P = 5$. Both the SFLAF and the CFLAF are memoryless (i.e., the memory order $K = 0$). Due to the large filter length and to the low nonlinear conditions we limit the input buffer length for both the FLAF-based architectures to the value $M_{\mathrm{i}} = M/20$. We also show the performance behavior of the linear reference filter, an NLMS, and the nonlinear reference filter, a second order VF, whose parameter settings are the same used in Subsection IV.A.

Again we evaluate performance in terms of ERLE. However, in this case there is no possibility to average over several experiments but we use a moving average filter to smooth curves and make the graphic visualization similar to the previous experiments.

Results are shown in Fig. 13. In the first 5 seconds of the experiments, where the LNLR is quite high, the ERLE evolution of the FLAF-based nonlinear architectures is very close to those of linear and nonlinear references and sometimes it is superior, especially in correspondence of the signal peaks. The presence of such nonlinear component in the echo path may be caused by the male speech which usually conveys a larger amount of nonlinearities at low frequencies. In the range from 4 to 5 seconds it is possible to notice that the linear filter performs better than the SFLAF. In this case the CFLAF stops following the SFLAF behavior and performs like the linear filter. This prove the effectiveness of the collaborative filtering of the CFLAF even for real scenarios. As regards the second part of the experiment (from second 5 on), where the LNLR is lower with respect to the first part, it is worth noting that, since the experiment takes place in non-critical conditions, nonlinearities can be still considered mild. Therefore, a second order VF with the complete quadratic kernel may overfit the solution introducing some gradient noise that does not facilitate any significant performance improvement. On the other hand, even in the presence of mild nonlinearities FLAF-based architectures show advantages compared with the reference filters.

V. Conclusion

In this paper a new class of nonlinear adaptive algorithms based on the FLAF model has been introduced for nonlinear modeling of acoustic channels. Due to its flexible architecture, FLAF-based schemes represent an effective solution to model nonlinearities that affect speech signals. This is the reason why we have proposed two robust FLAF-based architectures for NAEC application. The first architecture is the *split FLAF*, whose strength is based on its separation between linear and nonlinear elements, thus performing two different adaptive filtering: a linear one aiming at the estimate of the AIR and a nonlinear one whose only task is to model nonlinearities. This allows to achieve a significant improvement in terms of convergence performance. The second proposed architecture is based on the adaptive combination of filters and it is robust against different degrees of nonlinearity. In particular, such *collaborative FLAF* allows to avoid any gradient noise caused by the estimation of the nonlinear part when the system is nearly linear. In this way, it is possible to guarantee good modeling performance, whether nonlinearities are present or not. Therefore, FLAF-based architectures can be considered as effective solutions for acoustic channel modeling, even when the nature of the echo path is not known *a priori*. This result paves the way for the development of more sophisticated filtering architectures able to model a wide range of nonlinear channels.

References

[1] 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, Lang. Process.*, vol. 19, no. 1, pp. 97–110, Jan. 2011.

[2] N. Birkett and R. A. Goubran, "Limitations of handsfree acoustic echo cancellers due to nonlinear loudspeaker distortion and enclosure vibration effects," in *Proc. IEEE ASSP Workshop Appl. Signal Process. Audio Acoust. (WASPAA'95)*, New Paltz, NY, USA, 1995, pp. 103–106.

[3] D. Hongyun and W.-P. Zhu, "Compensation of loudspeaker nonlinearity in acoustic echo cancellation using raised-cosine function," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 53, no. 11, pp. 1190–1194, Nov. 2006.

[4] J. Fu and W.-P. Zhu, "A nonlinear acoustic echo canceller using sigmoid transform in conjunction with RLS algorithm," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 55, no. 10, pp. 1056–1060, Oct. 2008.

[5] R. Parisi, R. Russo, M. Scarpiniti, and A. Uncini, "Performance of acoustic nonlinear echo cancellation in the presence of reverberation," in *Proc. Int. Symp. Front. Res. Speech Music (FRSM'09)*, Gwalior, India, 2009, pp. 106–111.

[6] A. Uncini, A. Nalin, and R. Parisi, "Acoustic echo cancellation in the presence of distorting loudspeakers," in *Proc. Eur. Signal Process. Conf. (EUSIPCO'02)*, Tolouse, France, 2002, vol. 1, pp. 535–538.

[7] M. Scarpiniti, D. Comminiello, R. Parisi, and A. Uncini, "Nonlinear spline adaptive filtering," *Signal Process.*, vol. 93, no. 4, pp. 772–783, Apr. 2013.

[8] J.-P. Costa, A. Lagrange, and A. Arliaud, "Acoustic echo cancellation using nonlinear cascade filters," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP'03)*, Hong Kong, 2003, vol. 5, pp. 389–392.

[9] M. Scarpiniti, D. Comminiello, R. Parisi, and A. Uncini, "Comparison of Hammerstein and Wiener systems for nonlinear acoustic echo cancelers in reverberant environments," in *Proc. IEEE Int. Conf. Digital Signal Process. (DSP'11)*, Corfù, Greece, 2011, pp. 1–6.

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

[11] A. Guerin, G. Faucon, and R. Le Bouquin-Jeannes, "Nonlinear acoustic echo cancellation based on Volterra filters," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 6, pp. 672–683, Nov. 2003.

[12] A. Stenger, L. Trautmann, and R. Rabenstein, "Nonlinear acoustic echo cancellation with 2nd order adaptive Volterra filters," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP'99)*, Phoenix, AZ, USA, 1999, vol. 2, pp. 877–880.

[13] A. H. Sayed*, Adaptive Filters*. Hoboken, NJ, USA: Wiley, Apr. 2008.

[14] A. Fermo, A. Carini, and G. L. Sicuranza, "Simplified Volterra filters for acoustic echo cancellation in GSM receivers," in *Proc. Eur. Signal Process. Conf. (EUSIPCO'00)*, Tampere, Finland, 2000.

[15] Y.-H. Pao*, Adaptive Pattern Recognition and Neural Networks*. Reading, MA, USA: Addison-Wesley, 1989.

[16] D. P. Das and G. Panda, "Active mitigation of nonlinear noise processes using a novel filtered-s LMS algorithm," *IEEE Trans. Speech Audio Process.*, vol. 12, no. 3, pp. 313–322, May 2004.

[17] G. L. Sicuranza and A. Carini, "A generalized FLANN filter for nonlinear active noise control," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 8, pp. 2412–2417, Nov. 2011.

[18] D. Comminiello, L. A. Azpicueta-Ruiz, M. Scarpiniti, A. Uncini, and J. Arenas-Garcia, "Functional link based architectures for nonlinear acoustic echo cancellation," in *Proc. IEEE J. Workshop Hands-Free Speech Commun. Microph. Arrays (HSCMA'11)*, Edinburgh, U.K., May 2011, pp. 180–184.

[19] D. Comminiello, M. Scarpiniti, R. Parisi, and A. Uncini, "A functional link based nonlinear echo canceller exploiting sparsity," in *Proc. Int. Workshop Acoust. Echo, Noise Control (IWAENC'10)*, Tel Aviv, Israel, 2010.

[20] J. Arenas-García, M. Martínez-Ramón, A. Navia-Vázquez, and A. R. Figueiras-Vidal, "Plant identification via adaptive combination of transversal filters," *Signal Process.*, vol. 86, no. 9, pp. 2430–2438, Sep. 2006.

[21] 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.

[22] S. Kozat and A. Singer, "Multi-stage adaptive signal processing algorithms," in *Proc. IEEE Sens. Array, Multichannel Signal Process. Workshop (SAM'00)*, Cambridge, MA, USA, 2000, pp. 380–384.

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

[24] D. P. Mandic, P. Vayanos, C. Boukis, B. Jelfs, S. Goh, T. Gautama, and T. Rutkowski, "Collaborative adaptive learning using hybrid filters," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP'07)*, Honolulu, HI, USA, 2007, vol. 3, pp. 921–924.

[25] S. Haykin*, Neural Networks and Learning Machines*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, Nov. 2008.

[26] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Math. Control, Signals, Syst. (MCSS)*, vol. 2, no. 4, pp. 303–314, Dec. 1989.

[27] H. Zhao and J. Zhang, "Adaptively combined FIR and functional link artificial neural network equalizer for nonlinear communication channel," *IEEE Trans. Neural Netw.*, vol. 20, no. 4, pp. 665–674, Apr. 2009.

[28] J. C. Patra, R. N. Pal, B. N. Chatterji, and G. Panda, "Identification of nonlinear dynamic systems using functional link artificial neural networks," *IEEE Trans. Syst., Man, Cybern., B: Cybern.*, vol. 29, no. 2, pp. 254–262, Apr. 1999.

[29] M. S. Klassen and Y. H. Pao, "Characteristics of the functional link net: A higher order delta rule net," in *Proc. IEEE Ann. Int. Conf. Neural Net. (ICNN'88)*, San Diego, CA, USA, 1988, vol. 1, pp. 507–513.

[30] A. Namatame and Y. Kimata, "Improving the generalizing capabilities of a back-propagation network," *Int. J. Neural Netw.*, vol. 1, no. 2, pp. 86–94, 1989.

[31] L. Stankovć, "Performance analysis of the adaptive algorithm for bi-asto-variance tradeoff," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1228–1234, May 2004.

[32] J. C. Patra, W. C. Chin, P. K. Meher, and G. Chakraborty, "Legendre-FLANN-based nonlinear channel equalization in wireless communication system," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC'08)*, Singapore, 2008, pp. 1826–1831.

[33] L. A. Azpicueta-Ruiz, A. R. Figueiras-Vidal, and J. Arenas-García, "A normalized adaptation scheme for the convex combination of two adaptive filters," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP'08)*, Las Vegas, NV, USA, 2008, pp. 3301–3304.

[34] D. R. Campbell, K. J. Palomaki, and G. J. Brown, "Roomsim, a MATLAB simulation of "shoebox" room acoustics for use in teaching and research," *Comput. Inf. Syst.*, vol. 9, no. 3, pp. 48–51, 2005.

Danilo Comminiello (M'12) was born in Potenza, Italy, in 1981. He received the M.Sc. degree in telecommunication engineer in 2008 and the Ph.D. degree in information and communication engineering in 2012, both from "Sapienza" University of Rome, Italy. He is currently a Postdoctoral Research Fellow with the Department of Information Engineering, Electronics and Telecommunications (DIET), "Sapienza" University of Rome, Italy. From September 2007 to May 2008 he was an Intern Service Engineer at Ericsson Telecomunicazioni S.P.A., Rome, Italy. From November 2008 to October 2011, during his Ph.D. course, he collaborated with Fondazione Ugo Bordoni (FUB), Rome, Italy. From October 2010 to January 2011 he was a Visiting Researcher with the Department of Signal Theory and Communications, Universidad Carlos III de Madrid, Leganés, Spain.

His current research interests include signal processing and machine learning topics, particularly focused on audio and speech intelligent systems. He is an affiliate member of the IEEE Machine Learning for Signal Processing Technical Committee and of the IEEE Audio and Acoustic Signal Processing Technical Committee.

Michele Scarpiniti (M'09) was born in Leonberg, Germany, in 1978. He received the "Laurea" degree in electrical engineering with honors from the University of Rome "La Sapienza," Italy, in 2005 and the Ph.D. in information and communication engineering in 2009. From March 2008 he is an Assistant Professor of Circuit Theory with the Department of Information Engineering, Electronics and Telecommunications, University of Rome "La Sapienza," Italy.

His present research interests include ICA and blind signal processing, adaptive filters, audio processing and neural networks for signal processing. He is a member of IEEE, a member of the "Audio Engineering Society" (AES) and a member of the "Societ Italiana Reti Neuroniche" (SIREN).

Luis A. Azpicueta-Ruiz was born in Guadalajara, Spain, in 1978. He received the Telecommunication Engineer degree in 2004 from Universidad Politécnica de Madrid, Madrid, Spain and the Ph.D. degree in telecommunication technologies (honors) from Universidad Carlos III de Madrid, Leganés, Spain, in 2011. He is an Lecturer of Electroacoustics and Acoustic Engineering with the Department of Signal Theory and Communications at Universidad Carlos III, Madrid.

His present research interests are focused in the fields of adaptive signal processing and their applications, mainly in audio and acoustic processing.

Jerónimo Arenas-García (SM'12) received a Ph.D. in telecommunication technologies (Hons.) from Universidad Carlos III de Madrid, Spain (2004). After a postdoctoral stay at the Technical University of Denmark, he returned to Universidad Carlos III, where he is now a Lecturer of Digital Signal and Information Processing.

His research interests are focused on statistical learning theory, particularly in adaptive algorithms and advanced machine learning techniques. He has coauthored more than 75 papers on these topics. Dr. Arenas-García is member of the Machine Learning for Signal Processing TC of the IEEE SPS, and serves as Associate Editor for the *IEEE Signal Processing Letters*.

Aurelio Uncini (M'88) received the Laurea degree in electronic engineering from the University of Ancona, Italy, on 1983 and the Ph.D. degree in electrical engineering in 1994 from the University of Bologna, Italy. From 1984 to 1986 he was with the "Ugo Bordoni" Foundation, Rome, Italy, engaged in research on digital processing of speech signals and automatic speech recognition. From 1986 to 1987 he was at Italian Ministry of Communication in Rome. From 1987 to 1993 he has been a free researcher affiliated at the Department of Electronics and Automatics—University of Ancona and where from 1994 to 1998 he was assistant professor. From 1999 to 2004, he was Associate Professor at the Department of Information Engineering, Electronics and Telecommunications of the University of Rome "La Sapienza," Italy. At present time he is Full Professor at the same department where he is teaching Circuits Theory, Adaptive Algorithm for Signal Processing and Digital Audio Processing and where he is the founder and director of the "Intelligent Signal Processing and Circuits" (ISPAC) group.

He his author of several papers in the field of circuits theory, optimization algorithms for circuits design, neural networks and signal processing. His present research interests include also adaptive filters, adaptive audio and array processing, machine learning for signal processing, blind signal processing and multi-sensors data fusion. Prof. Uncini is a member of the Institute of Electrical and Electronics Engineers (IEEE), of the Associazione Elettrotecnica ed Elettronica Italiana (AEI), of the International Neural Networks Society (INNS) and of the Societ Italiana Reti Neuroniche (SIREN).