A study about Chebyshev nonlinear filters ☆

Alberto Carini [a,*,1], Giovanni L. Sicuranza [b]

[a] DiSBeF, University of Urbino, 61029 Urbino, Italy
[b] DIA, University of Trieste, 34127 Trieste, Italy

ARTICLE INFO

Accepted 7 November 2015

Keywords:
Nonlinear filters
Chebyshev nonlinear filters
Perfect periodic sequences
Cross-correlation method

ABSTRACT

The paper studies a novel family of nonlinear filters based on Chebyshev polynomials of the first kind, the Chebyshev nonlinear filters. This family shares many of the characteristics of the recently introduced Legendre and even mirror Fourier nonlinear filters, but has also peculiar properties. Chebyshev nonlinear filters belong to the class of linear-in-the-parameters nonlinear filters. Their basis functions are polynomials, specifically, products of Chebyshev polynomial expansions of the input signal samples. According to the Stone–Weierstrass theorem, they are universal approximators for causal, time-invariant, finite-memory, continuous, nonlinear systems. Their basis functions are mutually orthogonal for white input signals with a particular nonuniform distribution. They admit perfect periodic sequences, i.e., periodic input sequences that guarantee the mutual orthogonality of the basis functions on a finite period. Using perfect periodic input signals, an unknown nonlinear system and its most relevant basis functions can be identified with the cross-correlation method. It is shown in the paper that the perfect periodic sequences of Chebyshev nonlinear filters are simply related to those of even mirror Fourier nonlinear systems. Experimental results involving a real nonlinear system illustrate the potentialities of these filters.

1. Introduction

Linear-in-the-parameters (LIP) nonlinear filters include the most popular finite-memory and infinite-memory nonlinear filters. The filters belonging to this class are characterized by the property that their output depends linearly on the coefficients. The filters find application in speech [1], audio [2,3], telecommunication [4], image processing [1], biological system modeling [5], and many other fields. The LIP class includes truncated Volterra filters [1], still actively studied and used in applications [6–10], but also Wiener nonlinear filters [1], Hammerstein filters [1,11–14], memory and generalized memory polynomial filters [15,16], filters based on functional expansions of the input samples, as functional link artificial neural networks (FLANN) [17–20] and radial basis function networks [21]. A review of finite-memory LIP nonlinear filters can be found in [22]. LIP nonlinear filters with infinite-memory have also been studied [23–27] and used in applications.

The Wiener nonlinear filters [1] were introduced to overcome the limitations of Volterra filters, whose basis functions are never orthogonal, not even for a white input signal. In Wiener filters the basis functions are orthogonal for white Gaussian input signals, which means that for these inputs a fast convergence speed of gradient descent algorithms can be expected and the cross-correlation

☆ This work has been supported in part by DiSBeF Research Grant (Grant No. 118/2015-DiSBeF).
* Corresponding author. Tel.: +39 0722 304412; fax: +39 0722 4475.
E-mail addresses: alberto.carini@uniurb.it (A. Carini), sicuranza@univ.trieste.it (G.L. Sicuranza).
1 Member of EURASIP.

method can be applied to efficiently identify an unknown nonlinear system. Nevertheless, it should be noted that the Wiener basis functions depend on the standard deviation of the Gaussian input and that the cross-correlation method applied to stochastic inputs often requires millions of samples for an accurate estimate of the filter coefficients [5].

Recently, the finite-memory LIP class has been enriched with novel sub-classes of nonlinear filters that guarantee the orthogonality of the basis functions for white uniform input signals in the range $[-1, +1]$: the Fourier nonlinear (FN) filters [28,29], the even mirror Fourier nonlinear (EMFN) filters [29,30], and the Legendre nonlinear (LN) filters [31,32]. FN and EMFN filters are based on trigonometric function expansions of the input signal samples, and do not include a linear term among the basis functions. In contrast, LN filters are based on Legendre polynomial expansions of the input signal samples and have a linear term formed by the first order basis functions. The basis functions of FN, EMFN, and LN filters form algebras that satisfy all the requirements of the Stone–Weierstrass approximation theorem [33]. Thus, they can arbitrarily well approximate any causal, time-invariant, finite-memory, continuous, nonlinear system. While [29] suggested that EMFN filters should be preferred to FN filters, because they often provide a much more compact representation of nonlinear systems, it was found in [34] that also FN filters can be very useful for modeling nonlinear systems. It has been shown that EMFN filters can also be better models than LN and Volterra filters in the presence of strong nonlinearities, while Volterra and LN filters provide better results for weak or medium nonlinearities [29,32].

In [32,35–37], it has been proved that perfect periodic sequences (PPSs) can be developed for the identification of EMFN and LN filters. A periodic sequence is called perfect for a modeling filter if all cross-correlations between two of its basis functions, estimated over a period, are zero. PPSs guarantee the orthogonality of the basis functions on a finite period. An unknown system can be effectively modeled, using a PPS as input signal, by means of the cross-correlation method, i.e., simply computing the cross-correlations between the basis functions and the system output.

The most relevant basis functions, i.e., those that guarantee the most compact representation of the nonlinear system according to some information criterion, can also be easily estimated.

In this paper we study a novel sub-class of finite-memory LIP nonlinear filters based on Chebyshev polynomials of the first kind, the Chebyshev nonlinear (CN) filters. CN filters share many of the characteristics of EMFN and LN filters, but have also distinctive peculiarities that make them interesting alternatives to EMFN and LN filters. CN filters are products of the first kind Chebyshev polynomial expansions of the input samples that satisfy all the requirements of the Stone–Weierstrass approximation theorem. Therefore, they can arbitrarily well approximate any causal, time-invariant, finite-memory, continuous, nonlinear system, as well as Volterra, FN, EMFN, and LN filters. As the Volterra and the LN filters, CN filters include a linear term among the basis functions, allowing them to effectively model weak and medium nonlinearities. The basis functions of CN filters are not orthogonal for white uniform input signals in $[-1, +1]$, as those of EMFN and LN filters. On the contrary, they are orthogonal for white signals with a particular distribution that can be easily obtained with a sine transformation from a white uniform distribution. Thus, for these white input signals they provide fast convergence speed of gradient descent adaptation algorithms and efficient identification algorithms. Finally, CN filters also admit PPSs, which can be used for their efficient identification using the cross-correlation approach. Interestingly, the PPSs of CN filters are related to those of EMFN filters with a simple sine transformation.

Chebyshev polynomials have been already used in various fields of nonlinear signal processing. In [38,39] a nonlinear system is converted into a time-variant linear system with respect to a transformation function composed of Chebyshev polynomials. In [40], baseband modeling of nonlinear devices such as RF amplifiers is studied using a frequency-domain Volterra kernel approximation based on Chebyshev polynomials. Artificial neural networks exploiting functional expansions of the input signals using Chebyshev polynomials have been proposed for nonlinear dynamic system identification [41,42], nonlinear channel equalization [43–45], nonlinear adaptive filtering [46], and modeling of loudspeakers [47] and nonlinear audio effects [48]. Nevertheless, in contrast with the nonlinear filters discussed in this paper, the approaches of the literature do not consider cross-terms, i.e., products of Chebyshev basis functions involving samples with different time delay, which can be very important for modeling nonlinear systems [22]. Their basis functions are not complete under product and do not form an algebra. In contrast to the filters here proposed, the previous filters based on Chebyshev polynomials do not satisfy the conditions of the Stone–Weierstrass approximation theorem and thus are not universal approximators.

The paper is organized as follows. Section 2 reviews Chebyshev polynomials of the first kind and their properties. Section 3 derives the CN filters and discusses their properties, with particular attention to orthogonality and PPSs. Section 4 provides experimental results involving a real nonlinear device and test sets available in the literature. Concluding remarks follow in Section 5.

The following notation is used throughout the paper. $\mathbb{R}$ is the set of real numbers, $\mathbb{R}_1$ is the unit interval $[-1, +1]$, $\mathbb{N}$ is the set of natural numbers, δ_{ij} is the Kronecker delta, and $\langle x(n)\rangle_L$ indicates time average over L successive samples of $x(n)$.

2. Chebyshev polynomials of the first kind

The Chebyshev polynomials of the first kind [49,50] are a family of orthogonal polynomials generated by the following recursive relation:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \tag{1}$$

with $T_0(x) = 1$, $T_1(x) = x$ and $T_n(x)$ the Chebyshev polynomial of order n. The polynomials are orthogonal in $\mathbb{R}_1$

with respect to the weighting function $\frac{1}{\pi\sqrt{1-x^2}}$, since

$$\int_{-1}^{+1} T_n(x)T_m(x)\frac{1}{\pi\sqrt{1-x^2}}\,\mathrm{d}x = \begin{cases} 0, & n \neq m \\ 1, & n = m = 0 \\ 1/2, & n = m \neq 0. \end{cases} \qquad (2)$$

For any $x \in \mathbb{R}_1$ also $T_n(x) \in \mathbb{R}_1$ and it has maxima and minima equal to $+1$ and -1, respectively. Thus, the $T_n(x)$ are equiripple functions in $\mathbb{R}_1$.

The Chebyshev polynomials of the first kind can also be defined as the only polynomials that satisfy the following property:

$$T_n[\cos(\theta)] = \cos(n\theta). \qquad (3)$$

Furthermore, it is

$$T_n[\sin(\theta)] = \begin{cases} (-1)^m \sin[(2m+1)\theta], & n = 2m+1 \\ (-1)^m \cos[2m\theta], & n = 2m. \end{cases} \qquad (4)$$

Any degree N polynomial, $p(x)$, can be expressed as a linear combination of Chebyshev polynomials:

$$p(x) = \sum_{n=0}^{N} c_n T_n(x). \qquad (5)$$

Most importantly, any real continuous function $f(x)$ can be arbitrarily well approximated with a linear combination of Chebyshev polynomials. This property can be proved by resorting to the Stone–Weierstrass theorem [33]:

> "Let $\mathcal{A}$ be an algebra of real continuous functions on a compact set K. If $\mathcal{A}$ separates points on K and if $\mathcal{A}$ vanishes at no point of K, then the uniform closure $\mathcal{B}$ of $\mathcal{A}$ consists of all real continuous functions on K".

A family $\mathcal{A}$ of real functions is said to be an algebra if $\mathcal{A}$ is closed under addition, multiplication, and scalar multiplication, i.e., if (i) $f+g \in \mathcal{A}$, (ii) $f \cdot g \in \mathcal{A}$, and (iii) $cf \in \mathcal{A}$, for all $f \in \mathcal{A}$, $g \in \mathcal{A}$ and for all real constants c. The Chebyshev polynomials and their linear combinations form an algebra on the compact $\mathbb{R}_1$ that satisfies all the requirements of the Stone–Weierstrass theorem. Indeed, the family of functions is closed under addition, scalar multiplication, and multiplication, since

$$2T_m(x)T_n(x) = T_{m+n}(x) + T_{|m-n|}(x), \qquad (6)$$

it separates points (e.g., with $T_1(x)$) and vanishes at no point (e.g., with $T_0(x)$).

The approximation of a continuous function $f(x)$ with a linear combination of Chebyshev polynomials, $p(x)$, up to a degree N is very close to a min–max approximation [50]. Indeed, the approximation error can be expressed as:

$$\epsilon(x) = f(x) - p(x) = \sum_{N+1}^{+\infty} c_n T_n(x). \qquad (7)$$

If the function is continuous and differentiable, the sequence of coefficients c_n converges rapidly to 0, such that $\epsilon(x) \simeq c_{N+1}T_{N+1}(x)$, which is an equiripple function for the properties of Chebyshev polynomials.

Table 1 summarizes the first Chebyshev polynomials of the first kind.

Table 1
Chebyshev polynomials of the first kind.

$T_0(x) = 1$
$T_1(x) = x$
$T_2(x) = 2x^2 - 1$
$T_3(x) = 4x^3 - 3x$
$T_4(x) = 8x^4 - 8x^2 + 1$
$T_5(x) = 16x^5 - 20x^3 + 5x$

3. Chebyshev nonlinear filters

In this section, we first introduce the family of CN filters. Then, we discuss an orthogonality property and we show how PPSs can be developed and used for system identification.

3.1. The family of nonlinear filters

We are interested in developing a family of nonlinear filters capable to arbitrarily well approximate any causal, time-invariant, finite-memory, continuous, nonlinear system, whose input–output relationship can be expressed by a nonlinear function f of the N most recent input samples,

$$y(n) = f[x(n), x(n-1), \ldots, x(n-N+1)], \qquad (8)$$

where the input signal $x(n)$ is assumed to take values in $\mathbb{R}_1$, $y(n) \in \mathbb{R}$ is the output signal, and N is the system memory.

We proceed with the same approach used for introducing EMFN and LN filters [29,32]. Eq. (8) can be interpreted as a multidimensional function in the $\mathbb{R}_1^N$ space, where each dimension corresponds to a delayed input sample. We want to expand the nonlinear function $f[x(n), x(n-1), \ldots, x(n-N+1)]$ with a series of basis functions f_i,

$$f[x(n), x(n-1), \ldots, x(n-N+1)] = \sum_{i=1}^{+\infty} c_i f_i[x(n), x(n-1), \ldots, x(n-N-1)], \qquad (9)$$

where $c_i \in \mathbb{R}$, and f_i is a continuous function from $\mathbb{R}_1^N$ to $\mathbb{R}$, for all i. Moreover, we want the basis functions to satisfy the requirements of the Stone–Weierstrass theorem. To this purpose, we first write the Chebyshev polynomials, i.e., the 1-dimensional basis functions, for $x = x(n), x(n-1), \ldots, x(n-N+1)$:

$$\begin{array}{l} 1, T_1[x(n)], T_2[x(n)], T_3[x(n)], \ldots \\ 1, T_1[x(n-1)], T_2[x(n-1)], T_3[x(n-1)], \ldots \\ \vdots \\ 1, T_1[x(n-N+1)], T_2[x(n-N+1)], T_3[x(n-N+1)], \ldots \end{array}$$

Then, to guarantee completeness of the algebra under multiplication, we multiply the terms having different variables in any possible manner, taking care of avoiding repetitions. It is easy to verify that this family of real functions and their linear combinations constitute an algebra on the compact $\mathbb{R}_1^N$ that satisfies all the requirements of the Stone–Weierstrass theorem. Indeed, the set of functions is closed under addition, multiplication (because of (6)) and scalar multiplication. The algebra vanishes at no point due to the presence of $T_0[x(n)] = 1$.

Moreover, it separates points, since two separate points must have at least one different coordinate $x(n-k)$ and the linear term $T_1[x(n-k)] = x(n-k)$ separates these points. As a consequence, the nonlinear filters formed by a linear combination of these basis functions are able to arbitrarily well approximate any causal, time-invariant, finite-memory, continuous, nonlinear system.

We define the order of an N-dimensional basis function as the sum of the orders of the constituent 1-dimensional basis functions. Avoiding repetitions, the following basis functions are formed:

The basis function of order 0 is the constant 1.

The basis functions of order 1 coincide with the N 1-dimensional basis functions of order 1, i.e., the linear terms:

$$x(n), x(n-1), \ldots, x(n-N+1).$$

The basis functions of order 2 are the N 1-dimensional basis functions of order 2 and the basis functions originated by the product of two 1-dimensional basis functions of order 1:

$$\begin{aligned} &T_2[x(n)], T_2[x(n-1)], \ldots, T_2[x(n-N+1)],\\ &x(n)x(n-1), \ldots, x(n-N+2)x(n-N+1)\\ &x(n)x(n-2), \ldots, x(n-N+3)x(n-N+1)\\ &\vdots\\ &x(n)x(n-N+1). \end{aligned}$$

Thus, there are $N \cdot (N+1)/2$ basis functions of order 2.

Similarly, the basis functions of order 3 are the N 1-dimensional basis functions of order 3, the basis functions originated by the product between an 1-dimensional basis function of order 2 and an 1-dimensional basis function of order 1, and the basis functions originated by the product of three 1-dimensional basis functions of order 1. This constructive rule can be iterated for any order P.

The constructive rule for generating the basis functions follows the same multiplicative rule used for generating the basis functions of Volterra filters, with the only difference that the product of k equal samples, $x^k(n-i)$, is replaced by $T_k[x(n-i)]$. Thus, the two classes of filters have the same number of basis functions of order P, memory N. The linear combination of all the Chebyshev basis functions with order ranging from 0 to P and memory length of N samples defines a CN filter of order P, whose number of terms is

$$\binom{N+P}{N}. \tag{10}$$

3.2. An orthogonality property

By exploiting the orthogonality property of the Chebyshev polynomials in (2), it can be verified that the basis functions are orthogonal in $\mathbb{R}_1^N$ with weighting function $\frac{1}{\pi\sqrt{1-x^2(n)}} \cdots \frac{1}{\pi\sqrt{1-x^2(n-N+1)}}$. Taking two different basis functions f_i and f_j, the orthogonality condition is written as

$$\begin{aligned} &\int_{-1}^{+1} \cdots \int_{-1}^{+1} f_i[x(n), \ldots, x(n-N+1)]\\ &\cdot f_j[x(n), \ldots, x(n-N+1)]\\ &\cdot \frac{dx(n) \cdots dx(n-N+1)}{\pi\sqrt{1-x^2(n)} \cdots \pi\sqrt{1-x^2(n-N+1)}} = 0, \end{aligned} \tag{11}$$

which immediately follows since the basis functions are products of Chebyshev polynomials which satisfy (2). Since the weighting function is not a constant, the basis functions are not orthogonal for a white uniform distribution of the input signal in $\mathbb{R}_1$, as for EMFN and LN filters. Nevertheless, the basis functions are orthogonal for a white distribution of the input signal in $\mathbb{R}_1$ having probability density function

$$p_x(x) = \frac{1}{\pi\sqrt{1-x^2}}. \tag{12}$$

Indeed, in (11) the factor

$$\frac{1}{\pi\sqrt{1-x^2(n)} \cdots \pi\sqrt{1-x^2(n-N+1)}}$$

can be interpreted as the joint probability density function of the N-tuple $[x(n), \ldots, x(n-N+1)]$. As a consequence, a fast convergence of the gradient descent adaptation algorithms, used for nonlinear systems identification, is expected using a white input signal with probability density function (12). This signal can be obtained by transforming a signal $u(n)$, white uniform in $\mathbb{R}_1$, with the following mapping:

$$x(n) = \sin\left[\frac{\pi}{2}u(n)\right]. \tag{13}$$

Indeed, the probability density of $x(n)$, $p_x(x)$, is related to the probability density of $u(n)$, $p_u(u)$, as follows:

$$p_x(x) = \left|\frac{du}{dx}\right| p_u(u). \tag{14}$$

By inserting in (14) the inverse of (13),

$$u(n) = \frac{2}{\pi}\arcsin[x(n)], \tag{15}$$

and $p_u(u) = \frac{1}{2}$, it is immediate to obtain (12).

3.3. Perfect periodic sequences

The CN filters admit also perfect periodic sequences (PPSs), i.e., periodic sequences that guarantee the orthogonality of the basis functions on a finite period L, such that

$$\langle f_i[x(n), \ldots, x(n-N+1)] f_j[x(n), \ldots, x(n-N+1)] \rangle_L = 0, \tag{16}$$

for all $i \neq j$. Indeed, PPSs for CN filters can be obtained by transforming the PPSs of EMFN filters with the mapping in (13). This can be easily proved. Indeed, the 1-dimensional basis functions of EMFN filters are equal to $\sin\left[\frac{k\pi}{2}x\right]$, for order $k = 2m+1$ and $m \in \mathbb{N}$, and are equal to $\cos\left[\frac{k\pi}{2}x\right]$, for order $k = 2m$ and $m \in \mathbb{N}$. According to (4), a Chebyshev basis function excited by the sequence $\sin[\pi/2u(n)]$ produces the same output (apart from a sign change) of the corresponding EMFN basis function excited by $u(n)$. Thus, the orthogonality of the EMFN basis functions for a PPS input $u(n)$ implies the orthogonality of the Chebyshev basis functions for $x(n) = \sin\left[\frac{\pi}{2}u(n)\right]$.

PPSs are very useful for identifying causal, time-invariant, finite-memory, continuous, nonlinear systems. Let us assume that the input–output relationship of the nonlinear system is expressed as a linear combination of Chebyshev basis functions up to order K and memory of N

samples,

$$y(n) = \sum_{l} g_l f_l(n). \tag{17}$$

Using a PPS input for a CN filter of order K and memory N, the coefficients g_l can be efficiently estimated by computing the cross-correlation between the system output and each basis function over a multiple m of the sequence period L,

$$\hat{g}_l = \frac{\langle f_l(n) y(n) \rangle_{mL}}{\langle f_l^2(n) \rangle_{mL}}. \tag{18}$$

Let us assume that the nonlinear system is identified with a PPS of order K and memory N. As discussed in [32,36], when the nonlinear system in (8) is a linear combination of basis functions with memory N and maximum order greater than K, the identification is affected by an error that influences mainly the coefficients of the higher-order basis functions, while, in general, has only a marginal effect on the coefficients of the lower-order basis functions. When the system to be identified is a linear combination of basis functions with order K but memory greater than N, the identification is also affected by an error, which influences mainly the coefficients of basis functions associated with the most recent samples $x(n)$, $x(n-1), \ldots$, while, in general, the coefficients of basis functions associated with less recent samples $x(n-N+1)$, $x(n-N+2), \ldots$ are only marginally affected.

Exploiting the orthogonality of the basis functions on a PPS period, we can easily rank them according to the mean square error (MSE) they produce. For the l-th basis function, the MSE reduction is

$$\delta\text{MSE}_l = \frac{\langle f_l(n) y(n) \rangle_{mL}^2}{\langle f_l^2(n) \rangle_{mL}}. \tag{19}$$

Eq. (19) can be combined with any information criterion to obtain a compact representations for the nonlinear system. The Bayesian information criterion [51] will be used in the experimental results of this paper.

In general, the error in the approximation of a nonlinear system with the CN filter is not approximately equal to a single basis function, as in (7). Thus, we cannot claim that the CN filters provide a min–max approximation of the nonlinear system. Nevertheless, in many simulations identifying synthetic systems we have observed that CN filters provide approximately an equiripple error.

4. Experimental results

In this section we highlight the properties of the novel family of nonlinear filters by means of three experiments involving real devices. Specifically, in the first two experiments we consider the identification of an audiophile vacuum tube preamplifier, Behringer Tube Ultragain Mic 100. The preamplifier has a gain setting that can be used for introducing different levels of nonlinear distortion. Different input signals have been fed to the preamplifier and the output signals have been recorded at 8 kHz using a notebook. In the third experiment, two data sets proposed for benchmarking in nonlinear system identification [52] and composed of data recorded on real nonlinear systems are used to assess the performance of the proposed filter. The characteristics, the computational complexity and the performance of the CN filters are carefully analyzed in comparison to those of the well-known Volterra filters, the EMFN filters described in [29,30], and the LN filters introduced in [31,32]. The perfect sequences for EMFN and LN filters, used in the second experiment, have been derived in [35,36] and [32,37], respectively, and are available in [53].

4.1. First experiment

In the first experiment we want to highlight the usefulness of the orthogonality condition in (11). Thus, we identify the preamplifier with a gradient descent algorithm using a white input sequence in $\mathbb{R}_1$ having probability density function (12). With the selected settings, the preamplifier introduces, on a sinusoidal input at 200 Hz, a second and third order harmonic distortion of 5.6% and 20.2%, respectively. The harmonic distortion is defined as the ratio, in percent, between the magnitude of each harmonic and that of the fundamental frequency. The signal to noise ratio is greater than 65 dB. At 8 kHz sampling frequency, the system has memory length of around 15 samples. Thus, the system has been identified with a linear filter with memory of 15 samples, a Volterra, an EMFN, a LN, and a CN filter all with memory of 15 samples, order 3, and 816 coefficients. In order to do a fair comparison between the filters, the basis functions of all filters have been normalized to have a unit power and the preamplifier has been identified with the standard LMS algorithm,

$$e(n) = d(n) - \mathbf{h}^T(n)\mathbf{x}(n), \tag{20}$$

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \mu e(n)\mathbf{x}(n), \tag{21}$$

where $d(n)$ is the unknown system output, $e(n)$ is the a priori estimation error, $\mathbf{x}(n)$ is the vector collecting the LIP filter basis functions, $\mathbf{h}(n)$ is the vector collecting the corresponding coefficients of the LIP filter, and μ is the step-size. This algorithm can be easily applied to any linear-in-the parameters nonlinear filter, since its output is linear with respect to the coefficients themselves. Moreover, the step-size, which has been assumed equal for all the coefficients, has been carefully tuned for each of the adaptive filters. Specifically, using the acquired signals, for each filter the nonlinear system has been identified with different step-sizes $\mu = 0.01a^{-i}$, with $a = 10^{1/4}$ and $i \in \mathbb{N}$, and the corresponding learning curves have been plot on the same diagram. For each adaptive filter the step-size that guarantees the minimum steady-state MSE (apart from a dB fraction) with the fastest convergence speed has been annotated. Indeed, in gradient descent algorithms the steady-state MSE is the sum of three contributes: (i) the additive noise, (ii) the modeling error of the specific adaptive filter, and (iii) the excess MSE generated by the gradient noise. While the first two contributes do not depend on the step-size, the latter is negligible compared to the first two for a sufficiently small step-size. As a matter of fact, Fig. 1 shows the learning curves of MSE for

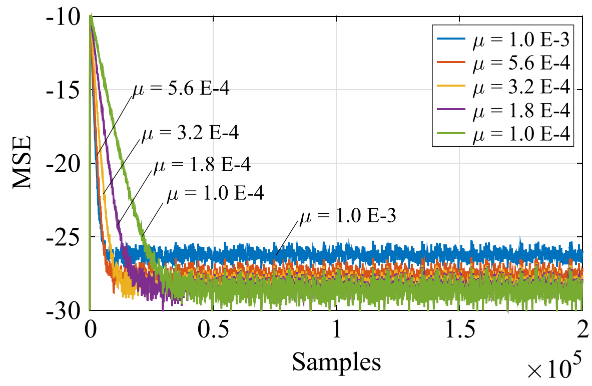

Fig. 1. Learning curves for the CN filter.

the CN filter with different step-sizes. Each learning curve is the ensemble average of 50 runs of the LMS algorithm applied to different data segments. Moreover, the learning curves have been smoothed using a box filter of 100 samples length. For a step-size $\mu \geq 5.6 \cdot 10^{-4}$ the steady-state MSE is larger than the minimum one, while for $\mu \leq 3.2 \cdot 10^{-4}$ the steady-state MSE almost coincides for all curves.

The learning curves of the different filters have then been compared with the annotated step-sizes. Fig. 2 shows the result of this comparison. Again the learning curves presented are the ensemble average of 50 runs of the LMS algorithm and have been smoothed using a box filter of 100 samples length. The step-size used for each learning curve is reported in the legend. Given the strong nonlinearity considered in the experiment, the linear filter is inadequate to model the preamplifier. In contrast, all nonlinear filter structures provide almost the same steady-state error in these experimental conditions. For the selected input signal, the linear and the CN filter have orthogonal basis functions and provide a fast convergence speed of the LMS algorithm. On the contrary, the Volterra, the EMFN and the LN filters do not share this orthogonality property for the selected input and, indeed, their convergence speed is slower than that of the linear and CN filters. Nevertheless, the EMFN and the LN filters have orthogonal basis functions for white uniform input signals. Thus, in the current conditions they provide faster convergence speed than the Volterra filter, since their basis functions are closer to the orthogonality condition. Indeed, the input signal autocorrelation matrix estimated over 1 000 000 samples has condition number 18.8 for the EMFN filter, 25.1 for the LN filter, and 509 for the Volterra filter.

4.2. Second experiment

In the second experiment, we consider the identification of the preamplifier under different distortion conditions using perfect sequences. Twenty different settings of the gain control have been considered, ranging from the lowest till the highest possible setting. The second, third and total harmonic distortions on a 200 Hz signal at the maximum used amplitude at the different settings are represented in Fig. 3. Clearly, the last settings that have total harmonic distortion close or greater than 100% provide extreme distortion conditions. The preamplifier has been identified with EMFN, LN, and CN filters using three PPSs, suitable to the three filters with order 3, memory 20, and with a period of 655 408 samples. The perfect sequence for the CN filter has been obtained by transforming with (13) the PPS for EMFN filters. The coefficients of the filters were estimated with the cross-correlation method in (18) and the most relevant basis functions were selected according to the Bayesian information criterion, minimizing

$$B(\nu) = L \log_e[\sigma_\epsilon^2(\nu)] + \nu \log_e[L] \quad (22)$$

where $\sigma_\epsilon^2(\nu)$ is the variance of the residual error associated to the first ν most relevant terms of the model and L is the number of data used for the model estimation. The preamplifier has also been identified with a Volterra filter on the same data used for the CN filter identification. The cross-correlation method cannot be applied to the Volterra filter estimation since its basis functions are not orthogonal. Thus, the Volterra filter has been identified with the method of [54], which is one of the most computationally efficient identification methods for LIP nonlinear systems available in the literature. In all conditions, the signal to noise ratio was greater than 65 dB.

Fig. 4 shows the number of selected terms and the percentage of unexplained power (i.e., the ratio in percent between the residual MSE and the power of the output signal). The percentage of unexplained power is very low, close to 1%, till settings 16, then it tends to increase indicating that a third order model in inadequate to represent the nonlinear system at those high nonlinear distortions. CN, LN, and Volterra filters are all polynomial filters and each filter can be converted into one of the other representations. Thus, for the same input signal the filters should provide very similar results, with just a possible little change in the number of selected basis functions. This is confirmed in Fig. 4 for Volterra and CN filters, which are estimated for the same input signal. Also LN filters, which are estimated for a different PPS input, provide very similar results to Volterra and CN filters for small and medium nonlinear distortions. On the contrary, because of the different input signals, they produce larger errors for higher distortions. When the Volterra, LN, and CN filters are estimated on the same signal, they provide very similar results. EMFN filters for low and medium distortions originate slightly worse results than the other filters, because they lack a linear term, but for higher distortions they are able to provide better results than the other filters, also when estimated on the same signal.

There is here a significant difference between the effort necessary to estimate the CN, LN, and EMFN filters, and that for estimating the Volterra filters. Obtaining the CN, LN, and EMFN filters using PPSs and the cross-correlation method required only a few hours of computer time. In contrast, computing the results for Volterra filters with the method in [54] requested days of simulations on the same computer. As a matter of fact, if T indicates the number of samples used for the identification, B the number of candidate basis functions, and S the number of selected basis functions, the computational cost of the method of [54]

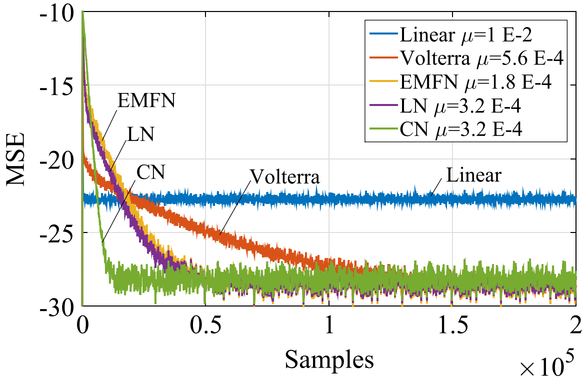

Fig. 2. Comparison of the learning curves for the linear, Volterra, EMFN, LN, and CN filters.

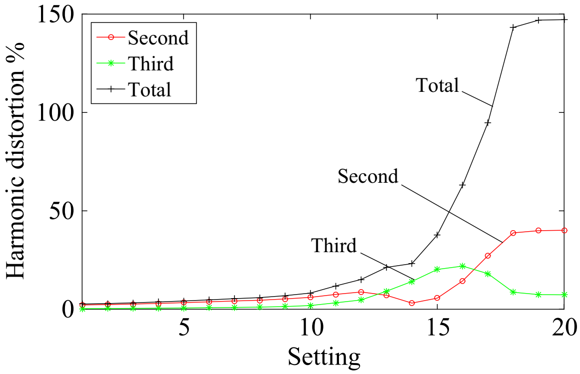

Fig. 3. Second, third and total harmonic distortion.

has order of TBS^2 operations, while the cross-correlation method requires only *TB* operations.

4.3. Third experiment

In the third experiments, we compare the performance of the CN, Volterra, LN, and EMFN filters on two data sets available in the literature [52] and currently used for benchmarking in nonlinear system identification [55]. The first data set is recorded from coupled electric drives, i.e., two electric motors driving a pulley with flexible belt. The second data set is generated by a fluid level control system, i.e., two cascaded tanks with free outlets fed by a pump. Two measured data are considered for each data set. More details on the data sets can be found in [52]. For the input signals of these data sets, the basis functions of CN, Volterra, LN, and EMFN filters are never orthogonal. Thus, all filters have been identified with the method of [54] considering a maximum order 3 and memory length 25 for the filters. Table 2 provides for the four filters the number of selected basis functions with the Bayesian information criterion and the percentage of unexplained power. While CN, Volterra, Legendre, and EMFN filters of order 3, memory length 25 have 3276 basis functions, a subset of 30–60 basis functions is sufficient here to accurately model the nonlinear system with percentage of unexplained power around or below 1%. In Table 2, all polynomial filters, i.e., CN, Volterra, and EMFN filters, provide

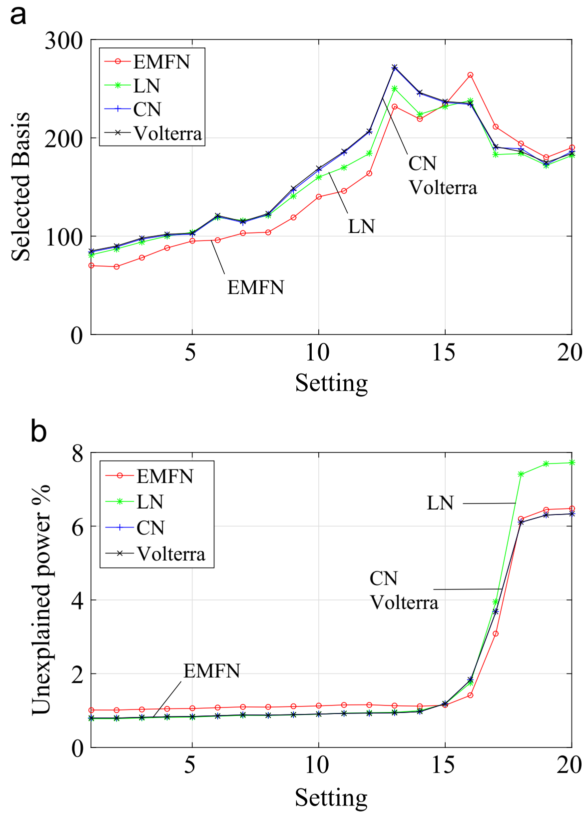

Fig. 4. Number of selected bases (a) and unexplained power (b) for EMFN, LN, CN, and Volterra filters.

Table 2
Identification results on benchmarking data sets.

Data set	Chebyshev	Volterra	Legendre	EMFN
Number of selected bases				
Drive 1	46	62	51	52
Drive 2	52	49	46	47
Tank 1	28	28	24	39
Tank 2	54	44	46	50
Unexplained power %				
Drive 1	0.91	0.73	0.84	0.81
Drive 2	0.32	0.35	0.35	0.60
Tank 1	0.82	0.81	0.82	0.95
Tank 2	1.03	1.04	1.04	1.07

very similar modeling performance, with just a difference in the number of selected basis functions, which depends on the specific system. For “drive 1”, the Volterra filter appears to provide better performance than the other filters in terms of unexplained power, but the number of selected basis functions is much larger than CN and LN filters. Selecting the most significant 100 basis functions all filters provide the same percentage of unexplained power (around 0.5%). Apart from the data signal “drive 2”, also the EMFN filter provides in these conditions similar performance as the other filters. Indeed, even if the EMFN filter lacks a linear term, its basis functions are particularly suited for modeling strong nonlinearities, as those of the nonlinear systems considered in the data sets of this experiment. The main advantage of CN, LN and EMFN filters with respect to Volterra filters comes from the possibility of selecting as input signal a PPS. Exploiting the orthogonality of basis functions, the PPS input facilitates system identification and basis functions selection.

5. Conclusion

A novel family of nonlinear filters based on Chebyshev polynomials of the first kind has been described. According to the Stone–Weierstrass theorem, the novel filters are universal approximators for causal, time-invariant, finite-memory, continuous, nonlinear systems. Their basis functions are mutually orthogonal for white input signals having a particular non uniform distribution. Moreover they admit perfect periodic sequences which are related to those of even mirror Fourier nonlinear systems. Experimental results involving a real nonlinear system and two data sets available in the literature, and used as benchmarking of nonlinear systems, have been provided to highlight the potentialities of these filters.

References

[1] V.J. Mathews, G.L. Sicuranza, Polynomial Signal Processing, Wiley, New York, 2000.
[2] U. Zölzer, DAFX: Digital Audio Effects, John Wiley & Sons, New York, USA, 2011.
[3] N.V. George, G. Panda, Advances in active noise control: a survey, with emphasis on recent nonlinear techniques, Signal Process. 93 (2) (2013) 363–377.
[4] F.-L. Luo (Ed.), Digital Front-End in Wireless Communications and Broadcasting, Cambridge University Press, Cambridge, UK, 2011.
[5] V. Marmarelis, Nonlinear Dynamic Modeling of Physiological Systems, John Wiley & Sons, Hoboken, NJ, USA, 2004.
[6] M. Zeller, L.A. Azpicueta-Ruiz, W. Kellermann, Online estimation of the optimum quadratic kernel size of second-order Volterra filters using a convex combination scheme, in: Proceedings of ICASSP 2009, International Conference on Acoustic, Speech, Signal Processing, Taipei, Taiwan, 2009, pp. 2965–2968.
[7] E.L.O. Batista, O.J. Tobias, R. Seara, A sparse-interpolated scheme for implementing adaptive Volterra filters, IEEE Trans. Signal Process. 58 (4) (2010) 2022–2035.
[8] L.A. Azpicueta-Ruiz, M. Zeller, A.R. Figueiras-Vidal, J. Arenas-Garcia, W. Kellermann, Adaptive combination of Volterra kernels and its application to nonlinear acoustic echo cancellation, IEEE Trans. Audio Speech Lang. Process. 19 (1) (2011) 97–110.
[9] A. Carini, V.J. Mathews, G.L. Sicuranza, Efficient NLMS and RLS algorithms for a class of nonlinear filters using periodic input sequences, in: Proceedings of ICASSP 2011, International Conference on Acoustic, Speech, Signal Processing, Prague, Czech Republic, 2011, pp. 4280–4283.
[10] H. Zhao, J. Zhang, A novel adaptive nonlinear filter-based pipelined feedforward second-order Volterra architecture, IEEE Trans. Signal Process. 57 (1) (2009) 237–246, http://dx.doi.org/10.1109/TSP.2008.2007105.
[11] J. Jeraj, V.J. Mathews, A stable adaptive Hammerstein filter employing partial orthogonalization of the input signals, IEEE Trans. Signal Process. 54 (4) (2006) 1412–1420.
[12] G.L. Sicuranza, A. Carini, On the accuracy of generalized Hammerstein models for nonlinear active noise control, in: Proceedings of 2006 IEEE Instrumentation and Measurement Technology Conference, Sorrento, Italy, 2006, pp. 1411–1416.
[13] E.L.O. Batista, R. Seara, A new perspective on the convergence and stability of NLMS Hammerstein filters, in: Proceedings of ISPA 2013, The Eighth International Symposium on Image and Signal Processing and Analysis, Trieste, Italy, 2013, pp. 336–341.
[14] M. Gasparini, L. Romoli, S. Cecchi, F. Piazza, Identification of Hammerstein model using cubic splines and FIR filtering, in: Proceedings of ISPA 2013, The Eighth International Symposium on Image and Signal Processing and Analysis, Trieste, Italy, 2013, pp. 347–352.
[15] J. Kim, K. Konstantinou, Digital predistortion of wideband signals based on power amplifier model with memory, Electron. Lett. 37 (2001) 1417–1418.
[16] D.R. Morgan, Z. Ma, J. Kim, M.G. Zierdt, J. Pastalan, A generalized memory polynomial model for digital predistortion of RF power amplifiers, IEEE Trans. Signal Process. 54 (2006) 3852–3860.
[17] Y.H. Pao, Adaptive Pattern Recognition and Neural Networks, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1989.
[18] H. Zhao, J. Zhang, Adaptively combined FIR and functional link artificial neural network equalizer for nonlinear communication channel, IEEE Trans. Neural Netw. 20 (4) (2009) 665–674, http://dx.doi.org/10.1109/TNN.2008.2011481.
[19] J. Patra, R. Pal, R. Baliarsingh, G. Panda, Nonlinear channel equalization for QAM signal constellation using artificial neural networks, IEEE Trans. Syst. Man Cybern. Part B: Cybern. 29 (2) (1999) 262–271, http://dx.doi.org/10.1109/3477.752798.
[20] N.V. George, G. Panda, On the development of adaptive hybrid active noise control system for effective mitigation of nonlinear noise, Signal Process. 92 (2) (2012) 509–516.
[21] M.D. Buhmann, Radial Basis Functions: Theory and Implementations, University Press, Cambridge, 1996.
[22] G.L. Sicuranza, A. Carini, On a class of nonlinear filters, in: M.G.I. Tabus, K. Egiazarian (Ed.), Festschrift in Honor of Jaakko Astola on the Occasion of his 60th Birthday, TICSP Report 47, 2009, pp. 115–144.
[23] E. Roy, R.W. Stewart, T.S. Durrani, Theory and applications of adaptive second order IIR Volterra filters, in: Proceedings of ICASSP 1996, International Conference on Acoustic, Speech, Signal Processing, Atlanta, GA, May 7–10, 1996, pp. 1597–1601.
[24] H. Zhao, X. Zeng, J. Zhang, Adaptive reduced feedback FLNN nonlinear filter for active control of nonlinear noise processes, Signal Process. 90 (2010) 834–847.
[25] G.L. Sicuranza, A. Carini, On the BIBO stability condition of adaptive recursive FLANN filter with application to nonlinear active noise control, IEEE Trans. Audio Speech Lang. Process. 20 (1) (2012) 234–245.
[26] G.L. Sicuranza, A. Carini, Recursive controllers for nonlinear active noise control, in: Proceedings of ISPA 2011, The Seventh International Symposium on Image and Signal Processing and Analysis, Dubrovnik, Croatia, Sept. 4–6, 2011, pp. 9–14.
[27] G.L. Sicuranza, A. Carini, A novel class of BIBO stable recursive nonlinear filters, in: Proceedings of ICASSP 2014, International Conference on Acoustic, Speech, Signal Processing, Florence, Italy, May 4–9, 2014, pp. 7984–4988.
[28] A. Carini, G.L. Sicuranza, A new class of FLANN filters with application to nonlinear active noise control, in: Proceedings of EUSIPCO 2012, European Signal Processing Conference, Bucharest, Romania, August 27–31, 2012, pp. 1950–1954.
[29] A. Carini, G.L. Sicuranza, Fourier nonlinear filters, Signal Process. 94 (2014) 183–194.
[30] A. Carini, G.L. Sicuranza, Even mirror Fourier nonlinear filters, in: Proceedings of ICASSP 2013, International Conference on Acoustic, Speech, Signal Processing, Vancouver, Canada, 2013, pp. 5608–5612.
[31] A. Carini, S. Cecchi, M. Gasparini, G.L. Sicuranza, Introducing Legendre nonlinear filters, in: Proceedings of ICASSP 2014, International Conference on Acoustic, Speech, Signal Processing, Florence, Italy, 2014, pp. 7989–7993.
[32] A. Carini, S. Cecchi, L. Romoli, G.L. Sicuranza, Legendre nonlinear filters, Signal Process. 109 (2015) 84–94.
[33] W. Rudin, Principles of Mathematical Analysis, McGraw-Hill, New York, 1976.
[34] N. Vanli, S. Kozat, A comprehensive approach to universal piecewise nonlinear regression based on trees, IEEE Trans. Signal Process. 62 (20) (2014) 5471–5486, http://dx.doi.org/10.1109/TSP.2014.2349882.
[35] A. Carini, G.L. Sicuranza, Perfect periodic sequences for identification of even mirror Fourier nonlinear filters, in: Proceedings of ICASSP 2014, International Conference on Acoustic, Speech, Signal Processing, Florence, Italy, 2014, pp. 8009–8013.
[36] A. Carini, G.L. Sicuranza, Perfect periodic sequences for even mirror Fourier nonlinear filters, Signal Process. 104 (2014) 80–93.
[37] A. Carini, S. Cecchi, L. Romoli, G.L. Sicuranza, Perfect periodic sequences for Legendre nonlinear filters, in: Proceedings of EUSIPCO 2014, European Signal Processing Conference, Lisbon, Portugal, 2014, pp. 2400–2404.
[38] K. Komatsu, S. Takata, A computation method of formal linearization for time-variant nonlinear systems via Chebyshev interpolation, in:

Proceedings of the 41st IEEE Conference on Decision and Control, vol. 4, 2002, pp. 4173–4178. http://dx.doi.org/10.1109/CDC.2002.1185024.
[39] K. Komatsu, S. Takata, A formal linearization for a general class of time-varying nonlinear systems and its applications, IEICE Trans. Fundam. Electron. Commun. Comput. Sci. E87-A (9) (2004) 2203–2209.
[40] P. Singerl, G. Kubin, Chebyshev approximation of baseband Volterra series for wideband RF power amplifiers, in: Proceedings of ISCAS 2005, IEEE International Symposium on Circuits and Systems, vol. 3, 2005, pp. 2655–2658. http://dx.doi.org/10.1109/ISCAS.2005.1465172.
[41] J. Patra, A. Kot, Y.Q. Chen, Chebyschev functional link artificial neural networks for nonlinear dynamic system identification, in: Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, vol. 4, 2000, pp. 2655–2660. http://dx.doi.org/10.1109/ICSMC.2000.884395.
[42] M. Li, Y. He, Nonlinear system identification using adaptive Chebyshev neural networks, in: Proceedings of ICIS 2010, IEEE International Conference on Intelligent Computing and Intelligent Systems, vol. 1, 2010, pp. 243–247. http://dx.doi.org/10.1109/ICICISYS.2010.5658578.
[43] J. Patra, W.B. Poh, N. Chaudhari, A. Das, Nonlinear channel equalization with QAM signal using Chebyshev artificial neural network, in: Proceedings of IJCNN '05, IEEE International Joint Conference on Neural Networks, vol. 5, 2005, pp. 3214–3219. http://dx.doi.org/10.1109/IJCNN.2005.1556442.
[44] H. Zhao, M. Xie, X. Zeng, A novel structure of conjunction decision feedback equalizer for nonlinear channels, in: Proceedings of APCCAS 2006, IEEE Asia Pacific Conference on Circuits and Systems, 2006, pp. 1760–1763. http://dx.doi.org/10.1109/APCCAS.2006.342159.
[45] H. Zhao, J. Zhang, Functional link neural network cascaded with Chebyshev orthogonal polynomial for nonlinear channel equalization, Signal Process. 88 (2008) 1946–1957.
[46] H. Zhao, J. Zhang, Pipelined Chebyshev functional link artificial recurrent neural network for nonlinear adaptive filter, IEEE Trans. Syst. Man Cybern. Part B: Cybern. 40 (1) (2010) 162–172.
[47] B. Bank, Computationally efficient nonlinear Chebyshev models using common-pole parallel filters with the application to loudspeaker modeling, in: Proceedings of AES 130th Convention, 2011, pp. 956–961.
[48] A. Novak, L. Simon, P. Lotton, J. Gilbert, Chebyshev model and synchronized swept sine method in nonlinear audio effect modeling, in: Proceedings of the 13th International Conference on Digital Audio Effects, 2010.
[49] T.J. Rivlin, Chebyshev Polynomials: from Approximation Theory to Algebra and Number Theory, John Wiley & Sons, New York, USA, 1990.
[50] A. Gil, J. Segura, N. Temme, Numerical Methods for Special Functions, Society for Industrial and Applied Mathematics, Philadelphia, 2007.
[51] G. Schwartz, Estimating the dimension of a model, Ann. Stat. 6 (1978) 416–464.
[52] T. Wigren, J. Schoukens, Three free data sets for development and benchmarking in nonlinear system identification, in: 2013 European Control Conference (ECC), 2013, pp. 2933–2938.
[53] A. Carini, Perfect Periodic Sequences. URL ⟨http://www2.units.it/ipl/res_PSeqs.htm⟩, 2015.
[54] K. Li, J.-X. Peng, G.W. Irwin, A fast nonlinear model identification method, IEEE Trans. Autom. Control 50 (8) (2005) 1211–1216.
[55] M. Scarpiniti, D. Comminiello, R. Parisi, A. Uncini, Novel cascade spline architectures for the identification of nonlinear systems, IEEE Trans. Syst. I: Reg. Pap. 62 (7) (2015) 1825–1835.