Legendre nonlinear filters ☆

Alberto Carini [a,*], Stefania Cecchi [b], Laura Romoli [b], Giovanni L. Sicuranza [c]

[a] DiSBeF, University of Urbino Carlo Bo, Urbino, Italy
[b] DII, Università Politecnica delle Marche, Ancona, Italy
[c] DIA, University of Trieste, Trieste, Italy

ARTICLE INFO

Accepted 31 October 2014

Keywords:
Nonlinear filters
Nonlinear system identification
Linear-in-the-parameters nonlinear filters
Legendre nonlinear filters
Perfect periodic sequences
Cross-correlation method

ABSTRACT

The paper discusses a novel sub-class of linear-in-the-parameters nonlinear filters, the Legendre nonlinear filters. The novel sub-class combines the best characteristics of truncated Volterra filters and of the recently introduced even mirror Fourier nonlinear filters, in particular: (i) Legendre nonlinear filters can arbitrarily well approximate any causal, time-invariant, finite-memory, continuous, nonlinear system; (ii) their basis functions are polynomials, specifically, products of Legendre polynomial expansions of the input signal samples; (iii) the basis functions are also mutually orthogonal for white uniform input signals and thus, in adaptive applications, gradient descent algorithms with fast convergence speed can be devised; (iv) perfect periodic sequences can be developed for the identification of Legendre nonlinear filters. A periodic sequence is perfect for a certain nonlinear filter if all cross-correlations between two different basis functions, estimated over a period, are zero. Using perfect periodic sequences as input signals permits the identification of the most relevant basis functions of an unknown nonlinear system by means of the cross-correlation method. Experimental results involving identification of real nonlinear systems illustrate the effectiveness and efficiency of this approach and the potentialities of Legendre nonlinear filters.

1. Introduction

The paper discusses a novel sub-class of finite memory linear-in-the-parameters (LIP) nonlinear filters. LIP nonlinear filters constitute a very broad filter class, which includes most of the commonly used finite-memory and infinite-memory nonlinear filters. The class is characterized by the property that the filter output depends linearly on the filter coefficients. It includes the well known truncated Volterra filters [1], which are still actively studied and used in applications [2–9], but also other popular polynomials filters, as the Hammerstein filters [1,10–13], the memory and generalized memory polynomial filters [14,15], and non-polynomial filters based on functional expansions of the input samples, as the functional link artificial neural networks (FLANN) [16] and the radial basis function networks [17]. The interested reader can refer to [18] for a review under a unified framework of finite-memory LIP nonlinear filters. Infinite-memory LIP nonlinear filters have also been studied [19–24] and used in applications.

Recently, the finite memory LIP class has been enriched with two novel sub-classes: the Fourier Nonlinear (FN) filters [25,26] and the Even Mirror Fourier Nonlinear (EMFN) filters [26,27]. FN and EMFN filters can be originated from the truncation of a multidimensional Fourier

☆ This work has been supported in part by DiSBeF Research Grant (no. 84/2013-DiSBeF-24_5). Parts of this paper have been published in the Proceedings of ICASSP 2014 (A. Carini, S. Cecchi, M. Gasparini, G.L. Sicuranza, Introducing Legendre nonlinear filters, Florence, Italy, 2014), and in Proceedings of EUSIPCO 2014 (A. Carini, S. Cecchi, L. Romoli, G.L. Sicuranza, Perfect periodic sequences for Legendre nonlinear filters, Lisbon, Portugal, 2014).
* Corresponding author. Tel.: +39 0722 304412; fax: +39 0722 4475.
E-mail addresses: alberto.carini@uniurb.it (A. Carini), s.cecchi@univpm.it (S. Cecchi), l.romoli@univpm.it (L. Romoli), sicuranza@univ.trieste.it (G.L. Sicuranza).

series expansion of a periodic repetition or an even mirror periodic repetition, respectively, of the nonlinear function they approximate. FN and EMFN filters are based on trigonometric function expansions, as the FLANN filters, but in contrast to the latter, their basis functions form an algebra that satisfies all the requirements of the Stone–Weierstrass approximation theorem [28]. Consequently, they can arbitrarily well approximate any causal, time-invariant, finite-memory, continuous, nonlinear system. EMFN filters provide a much more compact representation of nonlinear systems than FN filters [26], and thus should be the preferred choice. It has been shown that EMFN filters can also be better models than Volterra filters in the presence of strong nonlinearities, while Volterra filters provide better results for weak or medium nonlinearities [26]. An interesting property of EMFN (and FN) filters, which is not shared by Volterra filters, derives from orthogonality of the basis functions for white uniform input signals in the range $[-1, +1]$. This property is particularly appealing since it allows the derivation of gradient descent algorithms with fast convergence speed and of efficient identification algorithms. In [29,30], it was shown that perfect periodic sequences (PPSs) can be developed for the identification of EMFN filters. PPSs have been extensively studied and proposed as inputs for linear system identification [31] and in this context they have found application in signal processing [32], information theory [33], communications [34,35], and acoustics [36]. 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. By applying as input signal a PPS, it is possible to model an unknown system exploiting the cross-correlation method, i.e., computing the cross-correlation 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.

The novel sub-class of finite memory LIP nonlinear filters discussed in this paper is that of Legendre nonlinear (LN) filters, first introduced in [37]. LN filters combine the best characteristics of truncated Volterra and EMFN filters, as detailed in the following. First of all, the basis functions of LN filters are polynomials, as for Volterra filters. More specifically, they are products of Legendre polynomial expansions of the input samples that satisfy all the requirements of the Stone–Weierstrass approximation theorem. Therefore, LN filters are universal approximators, as well as Volterra, FN, and EMFN filters. With the term "universal approximators" we mean that these filters can arbitrarily well approximate any causal, time-invariant, finite-memory, continuous, nonlinear system. Secondly, the basis functions of LN filters are orthogonal for white uniform input signals in $[-1, +1]$, which means that they share all the benefits offered by FN and EMFN filters in terms of convergence speed of gradient descent adaptation algorithms and efficient identification algorithms. As a matter of fact, it is shown in Section 5 that the 2-norm condition number of the autocorrelation matrix of the input data vector for the Volterra filter is always larger than that of the EMFN and LN filters. As a consequence, EMFN and LN filters always provide a better convergence speed than a Volterra filter for white uniform input signals. Finally, as it was first shown in [38], PPSs can also be developed and used for the identification of LN filters. Indeed, they easily allow an efficient estimation of the most compact representation of the unknown nonlinear system, by using the cross-correlation approach and some information criterion. All these advantages come at the expense of a very small increase of the implementation complexity with respect to Volterra filters. All these aspects are considered in detail in the paper.

It is worth noting that LN filters are based on polynomial basis functions including the linear function, and thus their modeling capabilities are similar to those of Volterra filters. Therefore, LN filters can provide more compact models than EMFN filters for weak or medium nonlinearities. Moreover, identifying LN filters using PPSs is one of the most efficient methods for the identification of Volterra filters. Indeed, once the LN filter has been identified, it can be easily transformed into a Volterra filter representation exploiting the properties of Legendre polynomials.

The approach used in this paper to introduce the LN filter class can be applied to any family of orthogonal polynomials defined on a finite interval. Legendre polynomials are specifically considered since they have been already used for nonlinear filtering. Indeed, they have found application in Hammerstein models [39,40], FLANN filters [41–43], and neural networks [44]. Nevertheless, it should be noticed that the approaches of the literature do not make use of cross-terms, i.e., products among basis functions involving samples with different time delay, which can be very important for modeling nonlinear systems [18]. The corresponding basis functions do not form an algebra, because they are not complete under product. Thus, in contrast to the filters proposed in this paper, those previously considered are not universal approximators for causal, time-invariant, finite-memory, continuous, nonlinear systems.

Compared with the early conference contributions [37,38], in this paper we present an organic and detailed introduction of LN filters and their properties, discussing with particular attention PPSs for LN filters. Differently from [37], LN filters are introduced in this paper starting from a normalized set of Legendre polynomials. In contrast to [38], full proofs of properties of the PPSs for LN filters are here presented. Moreover, a discussion about the advantages and disadvantages of using LN filters and PPSs for system identification is also included in this paper.

The rest of the paper is organized as follows. Section 2 reviews basic concepts about LIP nonlinear filters, the Stone–Weierstrass theorem, and Legendre polynomials. Section 3 derives the LN filters and discusses their properties. Section 4 discusses PPSs for LN filters and their use for system identification. Section 5 presents experimental results that illustrate the advantages of LN filters and PPSs. Concluding remarks follow in Section 6.

Throughout the paper the mathematical notation of Table 1 is used. Moreover, sets are represented with curly brackets, intervals with square brackets, while the following convention for brackets $\{[(\cdots\{[()]\}\cdots)]\}$ is used elsewhere.

Table 1
Mathematical notation.

$\mathrm{leg}_i(x)$	Legendre polynomial of order i
$\mathrm{len}_i(x)$	Normalized Legendre polynomial of order i
δ_{ij}	The Kronecker delta
$\mathbb{R}$	The set of real numbers
$\mathbb{R}_1$	The interval $[-1,+1]$
$\mathbb{N}^+$	The set of positive integers
$\langle f(n)\rangle_L$	The average of L consecutive samples of $f(n)$

2. Basic concepts

In this section, some basic concepts about LIP nonlinear filters, the Stone–Weierstrass theorem, and Legendre polynomials are reviewed.

2.1. LIP nonlinear filters

The input–output relationship of a time-invariant, finite-memory, causal, continuous, nonlinear system can be expressed by a nonlinear function f of the N most recent input samples,

$$y(n)=f[x(n),x(n-1),\ldots,x(n-N+1)], \tag{1}$$

where the input signal $x(n)$ is assumed to take values in the range $\mathbb{R}_1=\{x\in\mathbb{R},\ \text{with}\ |x|\le 1\}$, $y(n)\in\mathbb{R}$ is the output signal, and N is the system memory. $f[x(n),\ldots,x(n-N+1)]$ is a multidimensional function in the $\mathbb{R}_1^N$ space: each dimension corresponds to a delayed input sample. In what follows, $f[x(n),\ldots,x(n-N+1)]$ is expanded with a series of basis functions f_i,

$$f[x(n),x(n-1),\ldots,x(n-N+1)] = \sum_{i=1}^{+\infty} c_i f_i[x(n),x(n-1),\ldots,x(n-N+1)], \tag{2}$$

where $c_i\in\mathbb{R}$, and f_i is a continuous function from $\mathbb{R}_1^N$ to $\mathbb{R}$, for all i. Every choice of the set of basis functions f_i defines a different LIP nonlinear filter that can be used to approximate the nonlinear systems in (1). Among all possible candidates, our interest falls on filters able to arbitrarily well approximate every time-invariant, finite-memory, continuous, nonlinear system. To this purpose, the Stone–Weierstrass theorem is exploited [28]:

> 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.

According to the theorem any algebra of real continuous functions on the compact $\mathbb{R}_1^N$ which separates points and vanishes at no point is able to arbitrarily well approximate the continuous function f in (1). 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.

2.2. Legendre polynomials

Nonlinear filters based on Legendre polynomials are developed in this paper. Legendre polynomials are orthogonal polynomials in $\mathbb{R}_1$ with

$$\int_{-1}^{1} \mathrm{leg}_i(x)\mathrm{leg}_j(x)\,dx=\frac{2}{2i+1}\delta_{ij}. \tag{3}$$

Differently from [37], in this paper LN filters are developed starting from a normalized set of Legendre polynomials, $\mathrm{len}_i(x)$, such that

$$\int_{-1}^{1} \mathrm{len}_i(x)\mathrm{len}_j(x)\,dx=2\delta_{ij}. \tag{4}$$

This choice simplifies the application of the cross-correlation method and the derivation of PPSs. The normalized set of Legendre polynomials is obtained by applying the following recursive relation:

$$\mathrm{len}_{i+1}(x)=\frac{2i+1}{i+1}\sqrt{\frac{2i+3}{2i+1}}x\,\mathrm{len}_i(x)-\frac{i}{i+1}\sqrt{\frac{2i+3}{2i-1}}\mathrm{len}_{i-1}(x) \tag{5}$$

with $\mathrm{len}_0(x)=1$ and $\mathrm{len}_1(x)=\sqrt{3}x$. With this choice, $\mathrm{len}_i(x)=\sqrt{2i+1}\,\mathrm{leg}_i(x)$. The first six normalized Legendre polynomials are listed in Table 2 with the number of multiplications and additions necessary to compute them.

Note that according to (4),

$$\int_{-1}^{1} \mathrm{len}_i(x)\,dx=\int_{-1}^{1} \mathrm{len}_i(x)\mathrm{len}_0(x)\,dx=0 \tag{6}$$

for all $i>0$.

The product of two Legendre polynomials of order i and j, respectively, can be expressed as a linear combination of Legendre polynomials up to the order $i+j$ [45]:

$$\mathrm{len}_i(x)\mathrm{len}_j(x)=\sqrt{2i+1}\sqrt{2j+1}\sum_{m=|i-j|}^{i+j} \times\frac{A(s-i)A(s-j)A(s-m)}{A(s)}\frac{\sqrt{2m+1}}{2s+1}\mathrm{len}_m(x) \tag{7}$$

where $s=(i+j+m)/2$ and

$$A(t)=\frac{1\cdot 3\cdot\ldots\cdot(2t+1)}{1\cdot 2\cdot\ldots\cdot t}=\frac{(2t)!}{2^t(t!)^2}. \tag{8}$$

The set of Legendre polynomials satisfies all the requirements of Stone–Weierstrass theorem on the compact $\mathbb{R}_1$. Thus, any continuous function from $\mathbb{R}_1$ to $\mathbb{R}$ can be

Table 2
Normalized Legendre polynomials and the number of multiplications and additions necessary to compute them.

Polynomial	×	+
$\mathrm{len}_0(x)=1$	0	0
$\mathrm{len}_1(x)=\sqrt{3}x$	1	0
$\mathrm{len}_2(x)=\frac{\sqrt{5}}{2}(3x^2-1)$	2	1
$\mathrm{len}_3(x)=\frac{\sqrt{7}}{2}x(5x^2-3)$	3	1
$\mathrm{len}_4(x)=\frac{\sqrt{9}}{8}(35x^4-30x^2+3)$	4	2
$\mathrm{len}_5(x)=\frac{\sqrt{11}}{8}x(63x^4-70x^2+15)$	5	2

arbitrarily well approximated with a linear combination of Legendre polynomials.

3. LN filters

A set of Legendre basis functions that allow us to arbitrarily well approximate any nonlinear system (1) is now developed by interpreting $f[x(n), x(n-1), \ldots, x(n-N+1)]$ as a multidimensional function in the $\mathbb{R}_1^N$ space. In order to extend the Legendre expansion to the N-dimensional case, the normalized Legendre polynomials are considered for $x = x(n)$, $x(n-1), \ldots, x(n-N+1)$:

$$
\begin{aligned}
&1, \mathrm{len}_1[x(n)], \mathrm{len}_2[x(n)], \mathrm{len}_3[x(n)], \ldots \\
&1, \mathrm{len}_1[x(n-1)], \mathrm{len}_2[x(n-1)], \mathrm{len}_3[x(n-1)], \ldots \\
&\vdots \\
&1, \mathrm{len}_1[x(n-N+1)], \mathrm{len}_2[x(n-N+1)], \mathrm{len}_3[x(n-N+1)], \ldots
\end{aligned}
$$

Then, the terms having different variables are multiplied in any possible manner, avoiding repetitions, to guarantee completeness of the algebra under multiplication. The family of so derived real functions and their linear combinations constitute an algebra on the compact $[-1,1]$ that satisfies all the requirements of the Stone–Weierstrass theorem. Indeed, the set is closed under addition, multiplication (because of (7)) and scalar multiplication. The algebra vanishes at no point (the set includes the constant 1) and separates points (because two separate points must have at least one different coordinate $x(n-k)$ and the linear term $\mathrm{len}_1[x(n-k)]$ separates them). As a consequence, the LN filters originated by these basis functions are able to arbitrarily well approximate any time-invariant, finite-memory, continuous, nonlinear system.

The order of an N-dimensional basis function is defined as the sum of the orders of the constituent 1-dimensional basis functions. The basis function of order 0 is the constant 1.

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

$$\mathrm{len}_1[x(n)], \mathrm{len}_1[x(n-1)], \ldots, \mathrm{len}_1[x(n-N+1)].$$

The basis functions of order 2 are the N 1-dimensional basis functions of the same order and the basis functions originated by the product of two 1-dimensional basis functions of order 1. Avoiding repetitions, there are $N \cdot (N+1)/2$ basis functions of order 2:

$$
\begin{aligned}
&\mathrm{len}_2[x(n)], \mathrm{len}_2[x(n-1)], \ldots, \mathrm{len}_2[x(n-N+1)], \\
&\mathrm{len}_1[x(n)]\mathrm{len}_1[x(n-1)], \ldots, \mathrm{len}_1[x(n-N+2)]\mathrm{len}_1[x(n-N+1)] \\
&\mathrm{len}_1[x(n)]\mathrm{len}_1[x(n-2)], \ldots, \mathrm{len}_1[x(n-N+3)]\mathrm{len}_1[x(n-N+1)] \\
&\vdots \\
&\mathrm{len}_1[x(n)]\mathrm{len}_1[x(n-N+1)].
\end{aligned}
$$

Similarly, the basis functions of order 3 are the N 1-dimensional basis functions of the same order, 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:

$$
\begin{aligned}
&\mathrm{len}_3[x(n)], \mathrm{len}_3[x(n-1)], \ldots, \mathrm{len}_3[x(n-N+1)], \\
&\mathrm{len}_2[x(n)]\mathrm{len}_1[x(n-1)], \ldots, \mathrm{len}_2[x(n-N+2)]\mathrm{len}_1[x(n-N+1)] \\
&\vdots \\
&\mathrm{len}_2[x(n)]\mathrm{len}_1[x(n-N+1)], \\
&\mathrm{len}_1[x(n)]\mathrm{len}_2[x(n-1)], \ldots, \mathrm{len}_1[x(n-N+2)]\mathrm{len}_2[x(n-N+1)], \\
&\vdots \\
&\mathrm{len}_1[x(n)]\mathrm{len}_2[x(n-N+1)], \\
&\mathrm{len}_1[x(n)]\mathrm{len}_1[x(n-1)]\mathrm{len}_1[x(n-2)], \ldots, \mathrm{len}_1[x(n-N+3)] \\
&\mathrm{len}_1[x(n-N+2)]\mathrm{len}_1[x(n-N+1)], \\
&\vdots \\
&\mathrm{len}_1[x(n)]\mathrm{len}_1[x(n-N+2)]\mathrm{len}_1[x(n-N+1)].
\end{aligned}
$$

This constructive rule can be iterated for any order P.

The basis functions of order P can also be obtained with the following procedure:

(i) Multiply in every possible way the basis functions of order $P-1$ by those of order 1.
(ii) Delete repetitions.
(iii) Apply the following substitution rule for products between factors having the same time index:

$$\mathrm{len}_i(x)\mathrm{len}_1(x) = \mathrm{len}_i(x)\sqrt{3}x \longrightarrow \mathrm{len}_{i+1}(x).$$

In the last passage, the property in (7) has been exploited neglecting all polynomials of order less than $i+1$.

This procedure for generating the basis functions of order P from those of order $P-1$ is the same applied for Volterra filters, thus the two classes of filters have the same number of basis functions of order P, memory N. In our case, the linear combination of all the Legendre basis functions of the same order P defines an LN filter of uniform order P, having $\binom{N+P-1}{P}$ terms, where N is the memory length. The linear combination of all the basis functions with order ranging from 0 to P and memory length of N samples defines an LN filter of nonuniform order P, whose number of terms is

$$\mathcal{N}_T(P,N) = \binom{N+P}{N}. \tag{9}$$

In what follows, $S_f(P,N)$ indicates the set of basis functions of order less than or equal to P and memory N, with cardinality $\mathcal{N}_T(P,N)$. $S_{f,n}(P,N)$ indicates the subset of $S_f(P,N)$ formed by the basis functions that are function of $x(n)$, which can be proved to have cardinality $\mathcal{N}_T(P-1,N)$. $f_l(n)$ indicates the l-th LN basis function estimated at time n, with l ranging between 1 and the cardinality of the set $f_l(n)$ belongs to.

Given the orthogonality property of the Legendre polynomials, the LN basis functions are orthogonal in $\mathbb{R}_1^N$, i.e., for $i \neq j$ it is

$$\int_{-1}^{+1} \cdots \int_{-1}^{+1} f_i[x(n), \ldots, x(n-N+1)] \cdot f_j[x(n), \ldots, x(n-N+1)] \cdot dx(n) \cdots dx(n-N+1) = 0. \tag{10}$$

Indeed, the basis functions are product of Legendre polynomials which satisfy (4) and (6). Because of this orthogonality property, the expansion of $f[x(n), \ldots, x(n-N+1)]$ with the proposed basis functions is a generalized Fourier series expansion [46]. Moreover, from (4) the basis functions are orthonormal for a white uniform distribution of the input signal in $\mathbb{R}_1$,

$$\int_{-1}^{+1} \cdots \int_{-1}^{+1} f_i[x(n),\ldots,x(n-N+1)] \cdot f_j[x(n),\ldots,x(n-N+1)] \cdot p[x(n),\ldots,x(n-N+1)] \cdot dx(n) \cdots dx(n-N+1) = \delta_{ij}, \quad (11)$$

where $p[x(n),\ldots,x(n-N+1)]$ is the probability density of the N-tuple $[x(n),\ldots, x(n-N+1)]$, equal to the constant $1/2^N$. As a consequence, as for FN and EMFN filters, a fast convergence of the gradient descent adaptation algorithms, used for nonlinear systems identification, is expected for white uniform input signals in $[-1,+1]$. For the same orthogonality property, an unbiased estimate for the coefficients of the LN filter approximating (1) can be easily found using the cross-correlation method [47]. In fact, the coefficient g_l of the basis function $f_l(n)$ is given by

$$g_l = E[f_l(n)y(n)], \quad (12)$$

where the expectation can be estimated using time averages.

To obtain a reasonable estimate for the coefficients using (12), a huge number of samples (on the order of millions) is needed [47, p. 77]. To overcome this problem, in the next section we introduce PPSs for LN filters, i.e., periodic sequences that guarantee the orthogonality of the basis functions on a finite time interval.

4. PPS for LN filters and system identification

In this section we first derive PPSs for LN filters and then discuss their use in system identification.

4.1. PPS for LN filters

The development of PPSs for LN filters follows the approach of [29,30].

Let us consider a sequence $x_0, x_1, \ldots, x_{L-1}$ of period L. Such a sequence is perfect for an LN filter of order K and memory N if all cross-correlations between two different basis functions, estimated over a period, are zero:

$$\langle f_l(n) \cdot f_m(n) \rangle_L = 0, \quad (13)$$

for all $f_l(n) \in S_{f,n}(K,N)$, $f_m(n) \in S_f(K,N)$ with $f_l(n) \neq f_m(n)$. Together with (13) it is also convenient to impose

$$\langle f_l(n) \cdot f_l(n) \rangle_L = 1 \quad (14)$$

for all $f_l(n) \in S_{f,n}(K,N)$ to guarantee orthonormality of the set of basis functions over a period. It is proved in Appendix that the system of nonlinear equations defined in (13) and (14) is equivalent to the following simpler system:

$$\langle f_l(n) \rangle_L = 0, \quad (15)$$

for all $f_l(n) \in S_{f,n}(2K,N)$.

In (15) we have $Q = \mathcal{N}_T(2K-1,N)$ equations in L variables $x_0, x_1, \ldots, x_{L-1}$. For sufficiently large L, the system is underdetermined and may have infinite solutions. In our experiments, a solution for (15) was always found. Any algorithm for solving nonlinear equation systems can be used. One of the most effective algorithms that were tested is the Newton–Raphson method, implemented as in [48, Chapter 9.7], starting from a random distribution of $x_0, x_1, \ldots, x_{L-1}$ in $\mathbb{R}_1$, with the only modification of reflecting the variables $x_0, x_1, \ldots, x_{L-1}$ in $\mathbb{R}_1$ when they exceeded the range. Employing numerical methods, only an approximate solution is obtained. Nevertheless, the cross-correlations between basis functions can be made as small as desired, selecting an appropriate precision in the stop-condition of the Newton–Raphson method. The number of iterations necessary for this method to converge depends on the selected precision and on the ratio L/Q.

Since the number of equations Q increases exponentially with the order K and geometrically with the memory N, even for low orders and memory lengths, Q can be unacceptably large. The number of equations and variables can be reduced imposing a specific structure to the periodic sequence. For example, the following conditions have been found to almost halve the number of equations and variables [30]:

(1) Symmetry: when the PPS is formed with the terms $a_1, a_2, \ldots, a_M$ and the reversed ones $a_M, a_{M-1}, \ldots, a_1$, for any couple of symmetric basis functions, only one of them has to be considered.
(2) Oddness: when the PPS is formed with the terms $a_1, a_2, \ldots, a_M$ and the negated ones $-a_1, -a_2, \ldots, -a_M$, all odd basis functions have a priori zero average.
(3) Oddness-1: when the PPS is formed with the terms $a_1, a_2, \ldots, a_M$ and those obtained by alternatively negating one every two terms, $a_1, -a_2, a_3, -a_4, \ldots, -a_M$, all Odd-1 functions have a priori zero average.

By definition, Odd-1 are all those basis functions that change their sign by alternatively negating one every two sample, e.g., $\mathrm{len}_1[x(n)]\mathrm{len}_1[x(n-1)]$. Similarly Odd-2, Odd-4, …, functions can be considered. Two or more conditions can also be considered together. The reduction in the number of equations takes to a longer period of the resulting PPS, but is often determinant to solve the system in (15) since the Newton–Raphson algorithm memory and the processing time requirements grow with Q^3. Another strategy to reduce the computational complexity of the system in (15) for large orders and memory lengths resorts to the use of simplified models, as done for Volterra filters in [49].

It is also possible to develop sequences that are perfect at the same time for LN filters and EMFN filters. To this purpose, together with the equations of the nonlinear system in (15), we have to impose also the similar equations for EMFN filters [26]. The number of equations Q in the nonlinear system increases, and so do the number of independent variables and the period of the sequence. The number of iterations necessary for Newton–Raphson method to converge increases, but convergence was always found for a sufficiently large period of sequence. A PPS for LN and EMFN filters will be used in the experimental results of Section 5.

4.2. Identification using PPS

This subsection describes how PPSs can be used to identify a time-invariant, finite-memory, causal, continuous, nonlinear system. When the input–output relationship of the nonlinear system is expressed as a linear combination of LN basis functions up to order K and memory N

$$y(n) = \sum_l g_l f_l(n), \tag{16}$$

the coefficients g_l can be estimated with a PPS input by computing the cross-correlation between the output of the system and each basis function over a period mL,

$$\hat{g}_l = \langle f_l(n) y(n) \rangle_{mL}, \tag{17}$$

where $m \in \mathbb{N}^+$ and L is the PPS period.

4.2.1. Influence of order K

Assume that a PPS for LN filters of order K, memory N, is used to identify a system formed by a linear combination of LN basis functions with memory N but maximum order greater than K, i.e.,

$$y(n) = \sum_l g_l f_l(n) + O_{K+1}(n), \tag{18}$$

with $f_l(n)$ having maximum order K and where $O_{K+1}(n)$ is a linear combination of basis functions of order greater than K. In these conditions, the coefficient $\hat{g}_l$ is affected by an error generated by $O_{K+1}(n)$. The error affects mainly the coefficients of the higher-order basis functions and, in general, only mildly the coefficients of the lower-order basis functions. This can be easily justified by considering $O_{K+1}(n)$ a linear combination of basis functions of order $K+1$. All the coefficients $\hat{g}_l$ of the basis functions of order less than K are not affected by the error, because for the PPS construction rules the cross-correlation of their basis functions with $O_{K+1}(n)$ is zero. Only the coefficients of the basis functions of order K are affected, because their cross-correlation with $O_{K+1}(n)$ is generally different from zero.

4.2.2. Influence of memory N

Assume next that a PPS for LN filters of order K, memory N, is used to identify a system that is a linear combination of LN basis functions with order K but memory greater than N, i.e.,

$$y(n) = \sum_l g_l f_l(n) + M_{N+1}(n), \tag{19}$$

with $f_l(n)$ having order K, memory N, and where $M_{N+1}(n)$ is a linear combination of basis functions of memory greater than N. In this case, the coefficients $\hat{g}_l$ are again affected by an error generated by $M_{N+1}(n)$. This error affects mainly the coefficients of the basis functions associated with the most recent samples, $x(n), x(n-1), \ldots$, and, in general, only mildly the coefficients of the basis functions associated with the less recent samples, $x(n-N+1)$, $x(n-N+2)$, … . To explain this property, let us identify the following system:

$$y(n) = \sum_{i=0}^{N} g_l \mathrm{len}[x(n-i)] \tag{20}$$

using a PPS for an LN filter of order 1, memory N. In this case, the estimate of g_0 in (17) is affected by an error, since $\langle \mathrm{len}[x(n)] \cdot \mathrm{len}[x(n-N)] \rangle_{mL}$ is in general different from zero. In contrast, $g_1, \ldots, g_{N-1}$ are not affected by this error since $\langle \mathrm{len}[x(n-i)] \cdot \mathrm{len}[x(n-N)] \rangle_{mL} = 0$ for all $1 \le i \le N-1$ according to the construction rules of the PPS.

4.3. Most relevant basis functions and information criteria

The orthonormality of the basis functions on a PPS period simplifies the identification of the most relevant basis functions, which maximize the mean square error (MSE) reduction. For the l-th basis function, the MSE reduction is

$$\delta \mathrm{MSE}_l = \langle f_l(n) y(n) \rangle_{mL}^2. \tag{21}$$

A compact representation for the nonlinear system can be obtained by combining (17) and (21) with some information criterion. Common criteria, exploited in the experiments of Section 5, are Akaike's information criterion (AIC) [50], the Final Prediction Error (FPE) [50], Khundrin's law of iterated logarithm criterion (LILC) [51], and the Bayesian information criterion (BIC) [52].

5. Experimental results

To show the potentialities of LN filters, the identification of a real-world nonlinear device, i.e., a Behringer Ultra Feedback Distortion FD300 guitar pedal is considered. The pedal provides a drive potentiometer that controls the amount of introduced distortion. In order to reduce as much as possible the strength of the nonlinearity, which is particularly sharp in these kind of devices, the potentiometer was set to the allowed minimum level. At the maximum used volume, the pedal introduced on a 1 kHz sinusoidal input a second order and a third order harmonic distortion[1] of 4.3% and 12.5%, respectively. The signal to noise ratio on the output signal was of 48 dB.

In what follows, experimental results about the identification of the pedal (i) with the LMS algorithm and white uniform noise input and (ii) with the cross-correlation method and a PPS input are presented. The input signals have been fed to the pedal at 8 kHz sampling frequency and the corresponding output has been recorded with a notebook.

5.1. Identification with LMS algorithm

The pedal is first identified with an LMS algorithm using a white uniform input signal. The nonlinear system memory length is lower than 20 samples, and thus the system has been identified with the LMS algorithm using (i) a linear filter of 20 sample memory, and (ii) an LN, (iii) an EMFN, and (iv) a Volterra filter, all without the constant term and with memory of 20 samples, order 3, and 1770 coefficients. The same step-size has been used in the identification of all coefficients.

Different nonlinear filters have different modeling abilities, and thus different steady state Mean-Square-Errors (MSE) and different convergence properties. A difficult choice is that of the step-size of the adaptation algorithm that guarantees a fair comparison between the different filters. To cope with this problem, the learning curves of the different filters have been compared by

[1] The harmonic distortion is defined as the percentage ratio between the magnitude of each harmonic and that of the fundamental frequency.

choosing for each filter the step-size that obtains the minimum steady-state MSE with the fastest convergence speed. Indeed, the steady-state MSE is the sum of three contributes: (i) the additive noise, (ii) the modeling error, and (iii) the excess MSE generated by the gradient noise. The last contribute depends on the choice of the step-size and, for sufficiently small step-sizes, is negligible compared to the other two contributes. Thus, using the recorded signals, for each filter structure the nonlinear system has been identified with a set of step-sizes uniformly distributed on a geometric scale. The corresponding learning curves have been plot on the same diagram and the largest step-size that reaches the minimum visible steady-state MSE (with tolerance a fraction of dB error) has been annotated. Fig. 1 illustrates this approach and shows the learning curves of MSE for the LN filter using the LMS algorithm with different step-sizes. Each learning curve is the ensemble averaged over 50 simulations applied to non-overlapping data segments. Moreover, the learning curves have been smoothed using a box filter of 100 sample memory length. From Fig. 1, it can be noticed that for a step-size $\mu \geq 5.6 \times 10^{-4}$ the steady-state MSE is larger than the minimum one, while for $\mu \leq 4.3 \times 10^{-4}$ almost the same steady-state MSE is obtained for all curves.

Using the annotated step-sizes, the learning curves of the four filters have been compared in Fig. 2. The step-sizes used for each filter are reported in the legend. The linear filter and the EMFN and LN filters have orthogonal basis functions for white uniform input signals, and thus provide a fast convergence speed of the LMS algorithm. Since the Volterra filter does not share this orthogonality property, its convergence speed is much slower than that of the other filters. In the time interval of Fig. 2, the Volterra filter does not reach the steady state conditions. For this reason, at time $n = 50,000$ it provides a larger MSE than LN filters, even though better than EMFN filters. EMFN and LN filters always provide a better convergence speed than the Volterra filter for white uniform input signals, whatever the identified system is. The better convergence speed can be rigorously proved by estimating the 2-norm condition number (the ratio of the largest singular value to the smallest) of the input data vector autocorrelation matrix. At time n, the input data vector is the vector collecting the values of basis functions $f_l(n)$. For EMFN and LN filters, the condition number is 1 (or close to 1, if estimated on a finite data) because of the orthonormality of the basis functions. On the contrary, for Volterra filters the condition number is always greater than 1 and it grows with the order and the memory length of the Volterra filter. As a matter of fact, Table 3 provides the estimated value over 2,000,000 samples of the condition number for different orders and memory lengths of the Volterra filter. The larger the condition number, the slower is the convergence speed of the LMS adaptation. Thus, Table 3 shows that, already for low orders and low memory lengths, the Volterra filter is expected to provide a worst convergence speed than EMFN and LN filters.

Returning to Fig. 2, the linear filter is unable to cope with the strong nonlinearity of the pedal and provides here the worst steady-state performance, but it has also the fastest convergence speed due to the orthogonality of basis functions and the reduced number of adapted coefficients. In this experiment, the LN filter and the EMFN filters show similar convergence speed but, thanks to the presence of the linear terms, the steady-state MSE of LN filter is lower than that of the EMFN filter. However, it has to be pointed out that, with the same pedal but higher distortion levels, the EMFN filter has been found to provide lower MSEs than the LN filter. There is no perfect filter for every condition. For causal, finite memory, continuous nonlinear systems, both EMFN and LN filters are eligible candidates for the identification of the unknown system and both maximize the convergence speed of the adaptation in the presence of a white uniform input signal. LN filters are better fitted to model mild or medium nonlinearities, due to the presence of the linear terms, while EMFN filters are interesting candidates for stronger nonlinearities [26].

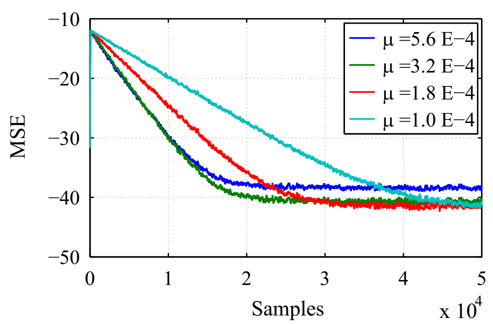

Fig. 1. Learning curves of LN filters for different values of the step-size.

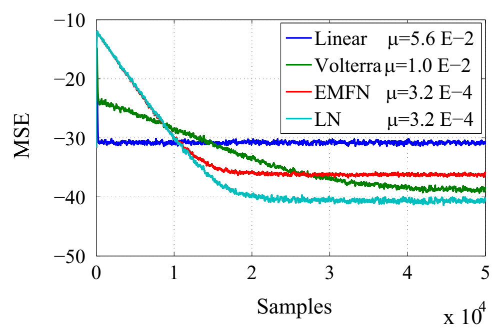

Fig. 2. Learning curves of linear, Volterra, EMFN, and LN filters.

Table 3
Number of basis functions and condition number of the input data vector autocorrelation matrix for a Volterra filter and a white uniform input noise.

Order	Memory length	Number of bases	Condition number
2	5	20	7.3
2	10	65	13.5
2	15	135	19.8
2	20	230	26.1
3	5	55	46.4
3	10	285	106.8
3	15	815	187.4
3	20	1770	293.8

5.2. Identification with the cross-correlation method

In the second experiment, we consider the identification of the pedal with a PPS. Since the system has a memory length lower than 20 samples, a PPS suitable for the identification of LN and EMFN filters of order 3, memory 20, exploiting oddness, oddness-1, oddness-2, oddness-4, and symmetry, and with period of $L=1{,}089{,}328$ samples has been used to identify the pedal.

The pedal is identified with LN and EMFN filters on a PPS period using the cross-correlation method. On the same data, a Volterra filter has been identified with the method of [53]. Different information criteria have been used to select the most relevant basis functions. Specifically the AIC (with parameter 4), the FPE, the LILC, and the BIC information criteria have been considered. Table 4 summarizes the number of terms selected by these information criteria, and the corresponding MSE for the three nonlinear filters. For a linear filter of memory 20 the MSE is 1.05×10^{-2}. LN and Volterra filters provide almost identical results. Indeed, both are polynomial models and the basis functions of the LN filter are a linear combination of those of Volterra filter (and vice versa). Thus, both filters provide the same MSE, and there is only a little difference in the number of selected basis functions. In contrast, the EMFN filter gives slightly worse results because it lacks a linear term. Nevertheless, it has to be pointed out that, at higher distortion levels of the pedal, the EMFN filter provides better results than the LN filter. Among the information criteria, the BIC criteria is the most conservative (maybe also the most appropriate in these conditions, since the same MSE is obtained with all criteria) and it halves the number of filter coefficients used to model the pedal.

Table 4
Results of identification of Behringer FD300.

Filter	Information criterion	Selected bases	MSE
LN	AIC(4)	1115	9.35E−4
	FPE	1271	9.34E−4
	LILC	1060	9.35E−4
	BIC	880	9.36E−4
EMFN	AIC(4)	926	2.62E−3
	FPE	1067	2.62E−3
	LILC	871	2.62E−3
	BIC	734	2.62E−3
Volterra	AIC(4)	1116	9.35E−4
	FPE	1272	9.34E−4
	LILC	1061	9.35E−4
	BIC	881	9.36E−4

Fig. 3 shows the order and the diagonal number of the first 400 selected basis functions for LN and EMFN filters (those for the Volterra filter have not been included since they are almost identical to those of LN filters). The "diagonal number" of a basis function is defined as the maximum time difference between the samples involved in its expression (for example, len[$x(n)$]len[$x(n-5)$] has diagonal number 5, len[$x(n)$] has diagonal number 0). Low diagonal numbers are selected in the first terms. Thus, if a very compact representation is desired, in this case the system could be modeled with a simplified LN, EMFN or Volterra filter with maximum diagonal number around 10.

5.3. Discussion on LN filters and identification methods

The main advantage of cross-correlation method is the remarkable computational efficiency. Indeed, it has a computational cost of TB operations, i.e., multiplications

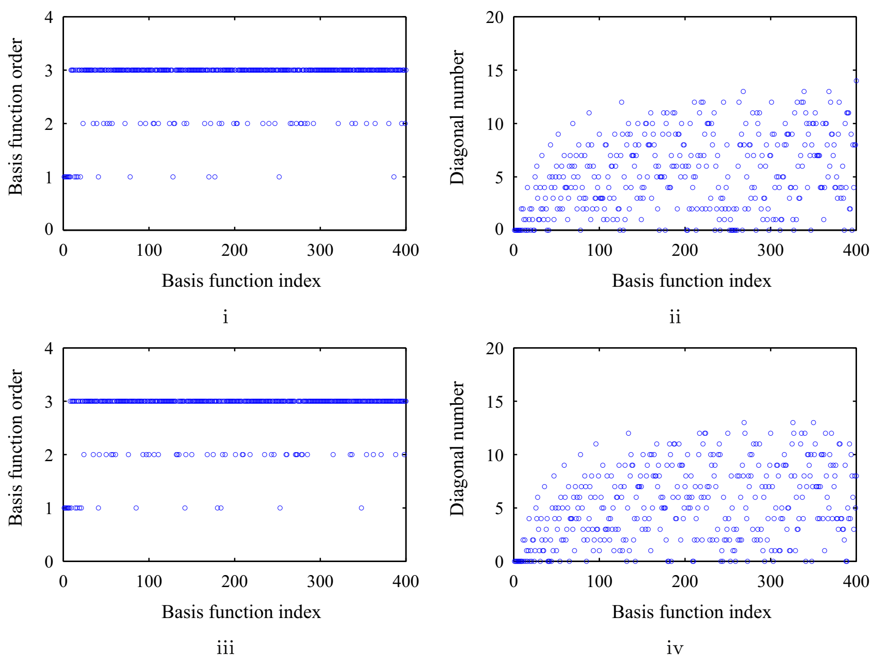

Fig. 3. Order and diagonal number of the first 400 selected basis functions for (i)–(ii) the LN filter, (iii)–(iv) the EMFN filter.

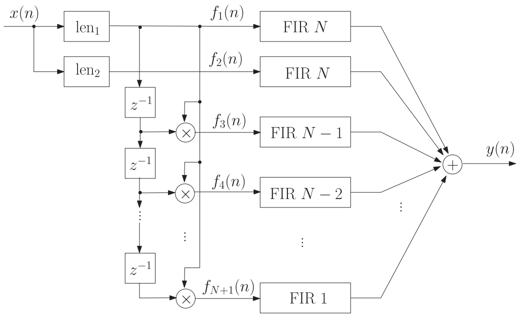

Fig. 4. Block diagram of a second order LN filter.

and additions, with T being the number of samples used for the identification, B being the number of candidate basis functions. In the second experiment, the method of [53] was chosen for comparison purposes because it is one of the most computationally efficient identification methods for LIP nonlinear systems available in the literature. The method of [53] can be applied to any input signal but has a computational cost of order TBS^2 operations, with S being the number of selected basis functions. Thus, the cross-correlation approach, taking advantage of the properties of the PPS, reduces the computational cost by a factor S^2. In the experiments, the execution of the cross-correlation method required a processing time of few minutes. In contrast, the method of [53] requested hours of simulation.

Given the very long period $L = 1,089,328$ of the PPS used in the experiments, it could be argued that a white uniform noise could be used for the identification of the LN system with the cross-correlation method obtaining similar performance. In reality, the mean-square difference between the coefficients identified with the cross-correlation method on 1,089,328 white uniform noise samples and the coefficients obtained with the least square approach on the same data is 1.5×10^{-4}. In contrast, the mean-square difference between the coefficients identified with the cross-correlation method on a PPS period of $1,089,328$ samples and the coefficients obtained with the least square approach on the same data is 1.1×10^{-18}. Moreover, periodic sequences with shorter periods can be developed, considering sequences perfect for LN (EMFN) filters only, or perfect for some simplified structures. For example, a PPS for linear and EMFN filters of order 3, memory 20 and period of 201,412 samples was developed in [30] and a PPS for LN filters of order 3, memory 20, and period of 357,956 samples was developed in [37].

Another advantage of the cross-correlation approach is the simplicity in the ranking and in the selection of the most relevant basis functions, which allows an easy integration with any information criteria. This property is not shared by the LMS identification.

A disadvantage of the LN filters is related to the complexity of the basis functions in comparison with those of the Volterra filters. In reality, the computational complexity increment is very limited, since it is related only to the calculation of the polynomials $\mathrm{len}_1[x(n)], \mathrm{len}_2[x(n)], \ldots, \mathrm{len}_P[x(n)]$, being P the filter order (and, in most cases, $P \le 3$). According to Table 2 the computation of each of these polynomials requires only few multiplications and additions. Indeed, all the LN basis functions can be expressed as products of the values and delayed values of the polynomials $\mathrm{len}_1[x(n)], \ldots, \mathrm{len}_P[x(n)]$. For example, Fig. 4 shows the block diagram for the implementation of a second order LN filter of memory N. The block diagram is formally identical to the implementation of a second order Volterra filter, apart from the presence of the blocks len_1 and len_2 which compute the corresponding Legendre polynomials. Thus, for any memory length N, the computational complexity of a second order LN filter is the same of a second order Volterra filter apart from three multiplications and one addition necessary to implement len_1 and len_2.

Once an LN filter has been identified, it can be easily converted into a Volterra filter, replacing each of the Legendre polynomials with its expression in Table 2, computing the products between the polynomials involved in each basis function, and reducing the resulting expression to a sum of products. For example,

$$
\begin{aligned}
&A\,\mathrm{len}_1[x(n)] + B\,\mathrm{len}_3[x(n)] + C\,\mathrm{len}_2[x(n)]\,\mathrm{len}_1[x(n-1)] \\
&\quad = A\sqrt{3}x(n) + B\frac{\sqrt{7}}{2}x(n)\left(5x^2(n)-3\right) \\
&\qquad + C\frac{\sqrt{5}}{2}\left(3x^2(n)-1\right)\sqrt{3}x(n-1) \\
&\quad = \left(\sqrt{3}A - \frac{3\sqrt{7}}{2}B\right)x(n) - \frac{\sqrt{15}}{2}Cx(n-1) \\
&\qquad + \frac{5\sqrt{7}}{2}Bx^3(n) + \frac{3\sqrt{15}}{2}Cx^2(n)x(n-1).
\end{aligned}
$$

As a matter of fact, for the properties of Legendre polynomials there is a bijective correspondence between LN and Volterra filters.

6. Conclusions

A sub-class of polynomial, finite-memory, LIP nonlinear filters, the LN filters, has been discussed. LN filters are universal approximators, according to the Stone–Weierstrass theorem, for causal, time-invariant, finite-memory, continuous, nonlinear systems, as well as the Volterra filters and the EMFN filters. The basis functions of LN filters are mutually orthogonal for white uniform input signals, as those of EMFN filters. Thanks to this orthogonality property, gradient descent algorithms with fast convergence speed can be devised. The orthogonality property can also be guaranteed on a finite period using PPSs, which allow an efficient identification of LN filters with the cross-correlation approach. In contrast to the EMFN filters, the basis functions of LN filters include the linear terms. Consequently, these filters are better fitted than EMFN filters for modeling weak or medium nonlinearities. In summary, the proposed filters combine the best characteristics of Volterra and EMFN filters.

Examples of PPSs can be downloaded from http://www.units.it/ipl/res_PSeqs.htm.

Appendix A. Equivalence between (13), (14) and (15)

In this Appendix, the equivalence between the system of nonlinear equations in (15) with $f_l(n) \in S_{f,n}(2K,N)$ and that in (13) and (14) with $f_l(n) \in S_{f,n}(K,N)$, $f_m(n) \in S_f(K,N)$, and $f_l(n) \neq f_n(n)$ is proved.

First, it is shown that (15) implies (13) and (14). Note that if (15) is met for $f_l(n) \in S_{f,n}(2K,N)$, it is also true for $f_l(n) \in S_f(2K,N)$.

For (7), the product between two different basis function $f_l(n) \in S_{f,n}(K,N)$ and $f_m(n) \in S_f(K,N)$ is a linear combination of basis functions belonging to $S_f(2K,N)$ and different from $f_0(n) = 1$ (because the summation in (7) starts for $m > 0$). Thus, (15) implies (13). Moreover, for any $f_l(n)$, $f_l^2(n)$ is equal to 1 plus a linear combination of other basis functions belonging to $S_f(2K,N)$. Indeed, the summation in (7) starts for $m = 0$ and the coefficient of $\text{len}_0(x)$ is 1. Thus, (15) implies also (14).

Next, (13) and (14) are assumed true and (15) is proved. In most cases, any $f_l(n) \in S_{f,n}(2K,N)$ can be directly written as the product of two basis functions belonging to $S_f(K,N)$ and, from (13), it is $\langle f_l(n) \rangle_L = 0$. The only cases where this is not possible are when $f_l(n)$ has a factor $\text{len}_t(x)$ with $t > K$. There can be just one of these factors in $f_l(n) \in S_{f,n}(2K,N)$. These cases are treated in two steps:

(a) First consider the case $t < 2K$ and take the product of two Legendre functions of order i and j, respectively, such that $i \neq j$ and $i + j = t$. According to (7), $\text{len}_t(x)$ can be expressed as $\text{len}_i(x) \cdot \text{len}_j(x)$ minus a linear combination of Legendre polynomials from order $|i-j| > 0$ till $t-1$. By applying recursively this reduction rule we can express $\text{len}_t(x)$ as a linear combination of product of couples of Legendre polynomials of order less than or equal to K. Thus, from (13), it is $\langle f_l(n) \rangle_L = 0$.

(b) In case $t = 2K$, consider the product $\text{len}_K(x) \cdot \text{len}_K(x)$. According to (7), $\text{len}_{2K}(x)$ can be expressed as $\text{len}_K(x) \cdot \text{len}_K(x)$ minus a linear combination of Legendre polynomials from order 0 till $2K-1$. In this linear combination the coefficient of $\text{len}_0(x) = 1$ is 1. Thus, from (14) and the proof in (a), we find that $\langle f_l(n) \rangle_L = 0$.

References

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

[2] G.A. Glentis, P. Koukoulas, N. Kalouptsidis, Efficient algorithms for Volterra system identification, IEEE Trans. Signal Process. 47 (11) (1999) 3042–3057.

[3] L. Tan, J. Jiang, Adaptive Volterra filters for active noise control of nonlinear processes, IEEE Trans. Signal Process. 49 (8) (2001) 1667–1676.

[4] F. Küch, W. Kellermann, Partitioned block frequency-domain adaptive second-order Volterra filter, IEEE Trans. Signal Process. 53 (2) (2005) 564–575.

[5] E.L.O. Batista, O.J. Tobias, R. Seara, A fully LMS adaptive interpolated Volterra structure, in: Proceedings of ICASSP 2008, International Conference on Acoustic, Speech, Signal Processing, Las Vegas, NV, 2008, pp. 3613–3616.

[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 (11) (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] 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.

[11] 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.

[12] E.L.O. Batista, R. Seara, A new perspective on the convergence and stability of NLMS Hammerstein filters, in: Proceedings of ISPA 2013, 8th International Symposium on Image and Signal Processing and Analysis, Trieste, Italy, 2013, pp. 336–341.

[13] M. Gasparini, L. Romoli, S. Cecchi, F. Piazza, Identification of Hammerstein model using cubic splines and FIR filtering, in: Proceedings of ISPA 2013, 8th International Symposium on Image and Signal Processing and Analysis, Trieste, Italy, 2013, pp. 347–352.

[14] J. Kim, K. Konstantinou, Digital predistortion of wideband signals based on power amplifier model with memory, Electron. Lett. 37 (2001) 1417–1418.

[15] 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.

[16] Y.H. Pao, Adaptive Pattern Recognition and Neural Networks, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1989.

[17] M.D. Buhmann, Radial Basis Functions: Theory and Implementations, University Press, Cambridge, 1996.

[18] G.L. Sicuranza, A. Carini, On a class of nonlinear filters, in: M.G.I. Tabus, K. Egiazarian (Eds.), Festschrift in Honor of Jaakko Astola on the Occasion of his 60th Birthday, vol. TICSP Report #47, 2009, pp. 115–144.

[19] 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.

[20] X.Z.H. Zhao, J. Zhang, Adaptive reduced feedback FLNN nonlinear filter for active control of nonlinear noise processes, Signal Process. 90 (2010) 834–847.

[21] G.L. Sicuranza, A. Carini, A generalized FLANN filter for nonlinear active noise control, IEEE Trans. Audio Speech Lang. Process. 19 (8) (2011) 2412–2417.

[22] 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.

[23] G.L. Sicuranza, A. Carini, Recursive controllers for nonlinear active noise control, in: Proceedings of ISPA 2011, 7th International Symposium on Image and Signal Processing and Analysis, Dubrovnik, Croatia, September 4–6, 2011, pp. 9–14.

[24] 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.

[25] 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.

[26] A. Carini, G.L. Sicuranza, Fourier nonlinear filters, Signal Process. 94 (2014) 183–194.

[27] 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.

[28] W. Rudin, Principles of Mathematical Analysis, McGraw-Hill, New York, 1976.

[29] 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.

[30] A. Carini, G.L. Sicuranza, Perfect periodic sequences for even mirror Fourier nonlinear filters, Signal Process. 104 (2014) 80–93.
[31] A. Milewski, Periodic sequences with optimal properties for channel estimation and fast start-up equalization, IBM J. Res. Dev. 27 (5) (1983) 426–431.
[32] C. Antweiler, Multi-channel system identification with perfect sequences – theory and applications, in: R. Martin, U. Heute, C. Antweiler (Eds.), Advances in Digital Speech Transmission, John Wiley & Sons, Ltd., Chichester, UK, 2008, pp. 171–198.
[33] I. Mercer, Merit factor of Chu sequences and best merit factor of polyphase sequences, IEEE Trans. Inf. Theory 59 (9) (2013) 6083–6086.
[34] M. Soltanalian, P. Stoica, Design of perfect phase-quantized sequences with low peak-to-average-power ratio, in: Proceedings of EUSIPCO 2012, European Signal Processing Conference, 2012, pp. 2576–2580.
[35] J.S. Pereira, H. da Silva, Orthogonal perfect discrete Fourier transform sequences, IET Signal Process. 6 (2) (2012) 107–113.
[36] C. Antweiler, G. Enzner, Perfect sequence lms for rapid acquisition of continuous-azimuth head related impulse responses, in: Proceedings of WASPAA '09, IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2009, pp. 281–284.
[37] 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.
[38] 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.
[39] J.M. Gil-Cacho, T. van Waterschoot, M. Moonen, S.H. Jensen, Linear-in-the-parameters nonlinear adaptive filters for loudspeaker modeling in acoustic echo cancellation, Internal Report 13-110, ESAT-SISTA, K.U.Leuven, Leuven, Belgium.
[40] L. Romoli, M. Gasparini, S. Cecchi, A. Primavera, F. Piazza, Adaptive identification of nonlinear models using orthogonal nonlinear functions, in: Proceedings of AES 48th International Conference, Munich, Germany, September 21–23, 2012, pp. 1597–1601.
[41] J. Patra, W. Chin, P. Meher, G. Chakraborty, Legendre-FLANN-based nonlinear channel equalization in wireless communication system, in: Proceedings of SMC 2008, IEEE International Conference on Systems, Man and Cybernetics, 2008, pp. 1826–1831.
[42] K. Das, J. Satapathy, Legendre neural network for nonlinear active noise cancellation with nonlinear secondary path, in: Proceedings of IMPACT 2011, International Conference on Multimedia, Signal Processing and Communication Technologies, 2011, pp. 40–43.
[43] N. George, G. Panda, A reduced complexity adaptive Legendre neural network for nonlinear active noise control, in: Proceedings of IWSSIP 2012, 2012, pp. 560–563.
[44] H. Zhao, J. Zhang, M. Xie, X. Zeng, Neural Legendre orthogonal polynomial adaptive equalizer, in: Proceedings of ICICIC'06, First International Conference on Innovative Computing, Information and Control, vol. 2, 2006, pp. 710–713.
[45] J.C. Adams, On the expression of the product of any two Legendre's coefficients by means of a series of Legendre's coefficients, Proc. R. Soc. Lond. 27 (1878) 63–71.
[46] R.L. Herman, An Introduction to Fourier and Complex Analysis with Applications to the Spectral Analysis of Signals, 5th ed., 2012. URL: ⟨http://people.uncw.edu/hermanr/mat367/FCABook⟩.
[47] V. Marmarelis, Nonlinear Dynamic Modeling of Physiological Systems, John Wiley & Sons, Hoboken, NJ, USA, 2004.
[48] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in C: The Art of Scientific Computing, Cambridge University Press, New York, NY, USA, 1995.
[49] A. Fermo, A. Carini, G.L. Sicuranza, Low-complexity nonlinear adaptive filters for acoustic echo cancellation in GSM handset receivers, Eur. Trans. Telecommun. 14 (2) (2003) 161–169.
[50] H. Akaike, A new look at the statistical model identification, IEEE Trans. Signal Process. 19 (6) (1974) 716–723.
[51] E.J. Hannan, B.G. Quinn, The determination of the order of an autoregression, J. R. Stat. Soc. Ser. B 41 (2) (1979) 190–195.
[52] G. Schwartz, Estimating the dimension of a model, Ann. Stat. 6 (1978) 416–464.
[53] K. Li, J.-X. Peng, G.W. Irwin, A fast nonlinear model identification method, IEEE Trans. Autom. Control 50 (8) (2005) 1211–1216.