Orthogonal Periodic Sequences for the Identification of Functional Link Polynomial Filters

Alberto Carini, *Senior Member, IEEE,* Simone Orcioni, *Senior Member, IEEE,* Alessandro Terenzi, and Stefania Cecchi, *Member, IEEE*

***Abstract*—The paper introduces a novel family of deterministic periodic signals, the orthogonal periodic sequences (OPSs), that allow the perfect identification on a finite period of any functional link polynomials (FLiP) filter with the cross-correlation method. The class of FLiP filters is very broad and includes many popular nonlinear filters, as the well-known Volterra and the Wiener nonlinear filters. The novel sequences share many properties of the perfect periodic sequences (PPSs). As the PPSs, they allow the perfect identification of FLiP filters with the cross-correlation method. But, while PPSs exist only for orthogonal FLiP filters, the OPSs allow also the identification of non-orthogonal FLiP filters, as the Volterra filters. In OPSs, the modeled system input can be any persistently exciting sequence and can also be a quantized sequence. Moreover, OPSs can often identify FLiP filters with a sequence period and a computational complexity much smaller that PPSs. The provided experimental results, involving the identification of real devices and of a benchmark model, highlight the potentialities of the proposed OPSs in modeling unknown nonlinear systems.**

***Index Terms*—Ortogonal Periodic Sequences, Functional Link Polynomial filters, nonlinear filters.**

I. Introduction

THE paper discusses a novel family of deterministic signals that can be used to identify any Functional Link Polynomial (FLiP) filter with the cross-correlation method. FLiP filters [1], [2] are a class of nonlinear filters that includes many of the most popular linear-in-the-parameters (LIP) nonlinear filters used in theory and practice [3]–[17]. The class of FLiP filters is very broad. It includes the well known Volterra filters [18]–[20] and Wiener nonlinear (WN) filters [18], [21], that derive from the truncation of the Volterra and Wiener series, respectively. It includes also the even mirror Fourier nonlinear (EMFN) filters [22], the Legendre nonlinear (LN) filters [23], the Chebyshev nonlinear (CN) filters [24] and many others [2]. These models have recently produced interest not only in signal processing but also in computational intelligence and machine learning [25]. FLIP filters are formed by a linear combination of basis functions. These basis functions are products of nonlinear expansions of delayed input samples that follow the constructive rule of Volterra filters. The FLiP basis functions form algebras that satisfy all requirements of the Stone-Weierstass theorem [2]. Therefore, the FLiP filters are universal approximators, i.e., they can arbitrarily well approximate any discrete-time, time invariant, finite memory, continuous nonlinear system [2]. Some families of FLiP filters provide an orthogonal representation for some stochastic inputs. For example, WN filters have orthogonal basis functions for white Gaussian inputs, EMFN and LN filters for white uniform inputs, and CN filters for a particular nonuniform distribution [24]. In those cases, the coefficients can be calculated by projecting the output of the filter on each of the basis functions. The projection can be performed by estimating the expected value of the product of the filter output and the basis function. Computing these expected values with time averages gives origin to the *cross-correlation* method, which is based on the calculation of the cross-correlation between the basis functions and the system output. Since the system output is a polynomial function of the input, the estimate of these expected values implies the computation of high order input moments with time averages. Unfortunately, this presents the drawback of requiring millions of samples to obtain an accurate estimation [26], [27]. For many years, it has been possible to identify orthogonal FLiP filters with the cross-correlation method only with the stochastic inputs.

As an alternative to stochastic inputs, appropriate deterministic input signals have also been applied to system identification with the cross-correlation method. Among these, the perfect periodic sequences (PPSs) [28], [29] are deterministic sequences with an ideal autocorrelation function that is periodic impulsive. They have been first proposed as inputs for linear system identification [30]–[33] but in recent years they have been extended also to nonlinear filters. In the latter case, by definition a periodic sequence is perfect when the cross-correlation between any two different basis functions, estimated over a period, is zero. This definition is the natural extension of the linear case, where the input samples themselves can be considered as basis functions. PPSs suitable for the identification of many families of orthogonal FLiP filters, e.g, the EMFN [34], [35], the LN [23], [36], the CN [37] and the WN filters [38], have been developed, and a general methodology for their derivation has been discussed in [2]. These PPSs have been obtained by imposing the cross-correlation between any two different basis functions to be zero, and by solving the resulting system of nonlinear equa-

A. Carini is with the Department of Engineering and Architecture, University of Trieste, Trieste, Italy e-mail: {acarini@units.it}.

S. Orcioni, A. Terenzi, and S. Cecchi are with the Department of Information Engineering, Università Politecnica delle Marche, Ancona, Italy. E-mail: {s.orcioni, s.cecchi}@staff.univpm.it, a.terenzi@pm.univpm.it.

Parts of this work have been published in A. Carini, S. Orcioni, and S. Cecchi, “Introducing the Orthogonal Periodic Sequences for the Identification of Functional Link Polynomial Filters,” in ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019, pp. 5486–5490.

This work was supported in part by the Department of Information Engineering Research Grant.

tions with an iterative approach. This directly constraints the PPS samples, which are in general real numbers in $[-1, +1]$.

Extending the results of [39], this paper contributes to the subject by proposing a novel family of deterministic periodic sequences, the orthogonal periodic sequences (OPSs), which can be used to identify FLiP filters on a finite time interval with the cross-correlation method. The OPSs share many similarities with the PPSs. As the PPSs, (i) they have been developed for the identification of FLiP filters and (ii) they allow the perfect estimation of a FLiP filter on a finite time interval with the cross-correlation method. But OPSs have a more general use than PPSs. In fact, they can identify any orthogonal or non-orthogonal FLiP filter, including the Volterra filters[1]. Moreover, they could also be developed for any family of LIP nonlinear filters that admits a filterbank implementation (e.g., Hammerstein filters or functional link artificial neural networks (FLANN)), since these filters are often special cases of FLiP filters or combinations of FLiP filters. In contrast to PPSs, in OPSs the input sequence does not need to be perfect periodic: it can have any arbitrary persistently exciting distribution and can also be a quantized sequence. Once chosen the input sequence and the FLiP filter used for the identification, a set of OPSs can be developed. Each OPS is designed to estimate one of the so-called "diagonals" of the FLiP filter using the cross-correlation method. It will be shown that OPSs can often identify FLiP filters with a sequence period and, thus, a computational complexity much smaller than PPSs.

It has to be pointed out that the identification procedure using OPSs differs from the classical methods of Lee-Schetzen [21], [40], which identify WN filters using white noise, and of Korenberg [41], which determine the coefficients of a data dependent orthogonal representation. The methods in [21], [40] and [41] are based on a Gram-Schmidt orthogonalization of the Volterra series and they require filter conversions. The identification using OPSs does not apply any Gram-Schmidt orthogonalization nor require filter conversions to identify the desired filter.

In the experimental results, the identification using OPS will be compared with other approached based on the cross-correlation method and with the least-square (LS) identification. While there is a huge amount of alternative identification approaches that can be applied to nonlinear filters, in this paper we will consider only those most related to the proposed approach, i.e., all those approaches suitable for FLiP filters that are based on the cross-correlation method, which has an unmatched low computational-complexity, and the LS method that is used as baseline. The experimental results are indeed intended to show that OPSs can identify FLiP filters as well as PPSs and the LS method. However, in the last experiment we will also compare the results obtained with the proposed approach with those of a parametric method based on multitones.

Given the low computational complexity of the cross-correlation method and the relatively small period of OPSs, the use of OPSs provides one of the most efficient identification methodologies for FliP filters. Nevertheless, we must point out the following aspects: i) Nonlinear system identification using OPSs is possible only if the unknown nonlinear system can be represented with a FLiP filter. ii) The order and memory length of the FLiP filter should be a priori known or estimated with a trial and error procedure. iii) Identification with OPSs requires the application of a deterministic input signal. Alternative recent methodologies that do not require the latter can be found in [42]–[49].

The main original contributions of this paper are the following: i) We introduce the concept of OPSs and explain how they can be developed for any FLiP filter. ii) We analyze the nonlinear system identification using OPSs. The cases where the identified model underestimates the memory or the order of the nonlinear system are specifically addressed. The analysis considers also the effect of different kinds of additive output noises, i.e., white Gaussian, colored Gaussian, and non Gaussian noise. iii) We show that nonlinear system identification using OPSs is one of the most efficient identification methods for FLiP filters.

The rest of the paper is organized as follows. In Section II FLiP filters are reviewed. The OPSs are developed in Section III. FLiP filter identification using OPSs is discussed in Section IV. Section V provides experimental results about the identification of real nonlinear devices and of a benchmark model and compares OPSs with other identification methods based on cross-correlation and with LS identification.

The following notation is used throughout the paper: $\mathbb{R}$ is the set of real numbers, $\mathbb{R}_1$ is the interval $[-1, +1]$, $<a(n)>_L$ is the sum of $a(n)$ over a period of L consecutive samples, $E[\cdot]$ indicates expectation, $\delta(n)$ is the unit pulse sequence.

II. Functional Link Polynomial Filters

The FLiP filters are a broad class of LIP nonlinear filters and are universal approximators: they can arbitrarily well approximate any discrete-time, time-invariant, finite memory, continuous nonlinear system,

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

where f is a continuous N-dimensional function from $\mathbb{R}_1^N$ to $\mathbb{R}$, and $x(n) \in \mathbb{R}_1$.

FLiP filters can be proved to be universal approximators according to the Stone-Weierstrass theorem [50]:

> "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 an algebra when $\mathcal{A}$ is closed under addition, multiplication, and scalar multiplication.

The basis functions of FLiP filters are formed following the constructive rule of Volterra filters starting from an ordered set of univariate functions

$$\{g_0[\xi], g_1[\xi], g_2[\xi], ...\} \tag{2}$$

satisfying the requirements of Stone-Weierstass theorem. In (2) $g_0[\xi]$ is a function of order 0, usually the constant 1, $g_{2i+1}[\xi]$

[1]To the authors knowledge, they are the only sequences that allow the exact direct estimation of Volterra filters in a finite time interval with the cross-correlation method.

TABLE I
FLiP BASIS FUNCTIONS

Order 0:	$g_0[x(n)] = 1.$
Order 1:	$g_1[x(n)], \dots, g_1[x(n-N+1)].$
Order 2:	$g_2[x(n)], \dots, g_2[x(n-N+1)],$ $g_1[x(n)]g_1[x(n-1)], \dots, g_1[x(n-N+2)]g_1[x(n-N+1)],$ $g_1[x(n)]g_1[x(n-2)], \dots, g_1[x(n-N+3)]g_1[x(n-N+1)],$ $\vdots$ $g_1[x(n)]g_1[x(n-D)], \dots, g_1[x(n-N+D+1)]g_1[x(n-N+1)],$
Order 3:	$g_3[x(n)], \dots, g_3[x(n-N+1)],$ $g_2[x(n)]g_1[x(n-1)], \dots, g_2[x(n-N+2)]g_1[x(n-N+1)],$ $\vdots$ $g_2[x(n)]g_1[x(n-D)], \dots, g_2[x(n-N+D+1)]g_1[x(n-N+1)],$ $g_1[x(n)]g_2[x(n-1)], \dots, g_1[x(n-N+2)]g_2[x(n-N+1)],$ $\vdots$ $g_1[x(n)]g_2[x(n-D)], \dots, g_1[x(n-N+D+1)]g_2[x(n-N+1)],$ $g_1[x(n)]g_1[x(n-1)]g_1[x(n-2)], \dots$ $g_1[x(n-N+3)]g_1[x(n-N+2)]g_1[x(n-N+1)],$ $\vdots$ $g_1[x(n)]g_1[x(n-D+1)]g_1[x(n-D)], \dots$ $g_1[x(n-N+D+1)]g_1[x(n-N+2)]g_1[x(n-N+1)],$

for any $i \in \mathbb{N}$ is an odd function of order $2i+1$, $g_{2i}[\xi]$ for any $i \in \mathbb{N}$ is an even function of order $2i$.

For $N = 1$, the set of basis functions obtained setting $\xi = x(n)$ in (2) can arbitrarily well approximate the nonlinear system in (1).

For $N > 1$, a set of FLiP basis functions capable of arbitrarily well approximating (1) can be developed by

1) writing the functions in (2) for $\xi = x(n)$, $x(n-1)$, …, $x(n-N+1)$, and then
2) multiplying the terms of different variable in all possible manners, as in the constructive rule of Volterra filters, taking care of avoiding repetitions.

It can be verified that this set of basis functions and their linear combinations form an algebra that separates points on $\mathbb{R}_1^N$ and vanishes in no point (for the presence of g_0) and thus satisfies all requirements of Stone-Weierstrass theorem.

The *order* of a FLiP basis function is defined as the sum of the orders of the constituent factors $g_i(\xi)$, and the *diagonal number* of a basis function is the maximum time difference between the involved input samples. For sake of clarity, the FLiP basis functions up to order 3, diagonal number D, and memory N are given in Table I.

A FLiP filter of order K, memory N, diagonal number D is the linear combination of all FLiP basis functions, with order, memory, and diagonal number up to K, N, and D, respectively. For example, a FLiP filter of order 3, memory N and diagonal number D, is the linear combination of all basis function in Table I.

FLiP filters can be implemented in the form of a filter bank,

$$y(n) = \sum_{p=0}^{R-1} \sum_{m=0}^{N_p-1} h_p(m) f_p(n-m), \tag{3}$$

where $f_p(n)$ are the zero lag basis functions (e.g., the first elements of the rows of Table I), i.e., $f_0(n) = g_0[x(n)] = 1$, $f_1(n) = g_1[x(n)]$, $f_2(n) = g_2[x(n)]$, $f_3(n) = g_1[x(n)]g_1[x(n-1)]$, …, $f_{2+D}(n) = g_1[x(n)]g_1[x(n-D)]$, $f_{3+D}(n) = g_3[x(n)]$, …; N_p is the memory length for the basis function $f_p(n)$ and is equal to N minus the diagonal number of $f_p(n)$; R is the total number of zero lag basis functions, i.e.,

$$R = \binom{D+K}{D+1} + 1. \tag{4}$$

The FLiP filter has N_D coefficients with [1]

$$N_D = \binom{D+K+1}{D+1} + \binom{D+K}{D+1}(N-1-D). \tag{5}$$

Following the naming conventions of Volterra filters, each sequence $h_p(m)$ with $0 \le p \le R-1$ is called a *diagonal* of the FLiP filter.

Any choice of the univariate functions $g_i(\xi)$ takes to a different family of nonlinear filters. FLiP filters comprise many well known families of nonlinear filters, specifically

- the Volterra filters, where $g_i(\xi) = \xi^i$;
- the WN filters, which also derive for the truncation of the Wiener series, where $g_i(\xi)$ are the Hermite polynomials of variance σ_x^2 (according to the definition in [51]),
$$\{1, \xi, \xi^2 - \sigma_x^2, \xi^3 - 3\sigma_x^2\xi, \xi^4 - 6\sigma_x^2\xi^2 + 3\sigma_x^4, \dots\}; \tag{6}$$
- the LN filters, where $g_i(\xi)$ are Legendre polynomials,
$$\{1, \xi, (3\xi^2-1)/2, \xi(5\xi^2-3)/2, (35\xi^4-30\xi^2+3)/8, \dots\}; \tag{7}$$

and others, as discussed in [2]. Some of these filters are *orthogonal FLiP filters*, i.e., have basis functions that are orthogonal for a specific distribution of the input signal samples. For example, the basis functions of WN filters are orthogonal for a zero mean white Gaussian input with variance σ_x^2, those of LN filters are orthogonal for a white uniform input in $\mathbb{R}_1$. If the input distribution guarantees the orthogonality of the basis functions, the coefficients of the filter can in theory be estimated with the cross-correlation method, computing the cross-correlation between the basis functions and the system output. Unfortunately, applying stochastic inputs the cross-correlation method often requires millions of samples to accurately estimate the FLiP filter coefficients. Nevertheless, it has been shown that the orthogonal FLiP filters also admit PPSs, i.e., deterministic periodic sequences that guarantee the orthogonality of the basis functions over a period. Using a PPS input, an orthogonal FLiP filter can still be identified with the cross-correlation method, with the coefficients $h_i(j)$ given by

$$h_i(j) = \frac{< y(n) f_i(n-j) >_L}{< f_i^2(n) >_L}, \tag{8}$$

where L is a period of the PPS.

PPSs for FLiP filters of order K, memory N, diagonal number D have been obtained by considering a periodic sequence with L unknown samples and imposing the orthogonality of all basis functions, i.e., imposing

$$< f_{i_1}(n-j_1) f_{i_2}(n-j_2) >_L = 0, \tag{9}$$

for all i_1, i_2, j_1, j_2. For sufficiently large L, this is an underdetermined system of nonlinear equations that may have infinite solutions, and a solution has always been computed solving (9) with the Newton-Raphson method [52]. It must be pointed out that the procedure for solving the system is very slow, since the Newton-Raphson method is an iterative method and each iteration has a computational cost proportional to $\overline{Q}^3$, with $\overline{Q}$ the number of nonlinear equations, and $\overline{Q} \simeq R(N_D+1)$. It should be noted that the value of $\overline{Q}$ influences also the period of the PPS, since L should be sufficiently larger than $\overline{Q}$ and L is often in the range $[2\overline{Q}, 6\overline{Q}]$. In contrast, the OPSs introduced in the next section can be derived solving a simpler linear system and can often identify the FLiP filter with a much smaller period L.

III. Orthogonal Periodic Sequences

By definition, an OPS is a periodic sequence that cross-correlated with the filter output provides one of the diagonals of the FLiP filter. Let us consider a FLiP filter and a periodic input sequence $x(n)$ of period L. The only condition imposed to the input sequence is to persistently excite the FLiP filter. This condition guarantees the invertibility of all input data matrices that will be introduced in the following. The condition is satisfied when the input sequence samples are aleatory taken from a Gaussian distribution, a white uniform distribution, or other random distribution. The sequence could also be a quantized sequence, provided that it is quantized with a sufficiently large number of levels. For example, it was shown in [53] that an independent, identically distributed sequence must take at least $K+1$ distinct values to persistently excite an order K Volterra filter. For obtaining the experimental results of Section V, 10 bits quantized sequences have been considered.

Let us consider the i-th diagonal of the FLiP filter, $h_i(j)$ with $0 \le j \le N_i - 1$. In what follows we want to develop the OPS $z_i(n)$ of period L such that

$$h_i(j) = < y(n) z_i(n-j) >_L, \tag{10}$$

for $0 \le j \le N_i - 1$.

Consider first the case of $i = 0$ and $f_0(n) = 1$. Inserting (3) in (10), for $j = 0$

$$h_0(0) = \sum_{p=0}^{R-1} \sum_{m=0}^{N_p-1} h_p(m) < f_p(n-m) z_0(n) >_L . \tag{11}$$

Thus, to be (10) true it must be

$$< f_0(n) z_0(n) >_L = < z_0(n) >_L = 1, \tag{12}$$

and

$$< f_p(n-m) z_0(n) >_L = 0, \tag{13}$$

for all $0 \le m \le N_p - 1$, and $0 < p \le R-1$. Any sequence $z_0(n)$ that satisfies the $Q_0 = N_D$ linear equations in (12) and (13) is an OPS that can compute $h_0(n)$, i.e., the constant term of the FLiP filter.

Consider now $i > 0$. Inserting (3) in (10), we have

$$h_i(j) = \sum_{p=0}^{R-1} \sum_{m=0}^{N_p-1} h_p(m) < f_p(n-m) z_i(n-j) >_L . \tag{14}$$

To be (10) true for $j = 0$, it must be

$$\begin{aligned} < f_0(n) z_i(n) >_L &= < z_i(n) >_L = 0, && (15)\\ < f_i(n) z_i(n) >_L &= 1, && (16)\\ < f_i(n-m_i) z_i(n) >_L &= 0, && (17)\\ < f_p(n-m_p) z_i(n) >_L &= 0, && (18) \end{aligned}$$

for all $1 < m_i \le N_i - 1$, $0 \le m_p \le N_p - 1$ and $0 < p \le R-1$ with $p \ne i$. To be (10) true also for $j > 0$, together with (15)–(18) it must also be

$$< f_p(n) z_i(n-j) >_L = < f_p(n+j) z_i(n) >_L = 0. \tag{19}$$

for all $0 < j \le N_i - 1$. Thus, $z_i(n)$ must satisfy the linear equation system

$$\begin{aligned} < z_i(n) >_L &= 0, && (20)\\ < f_i(n) z_i(n) >_L &= 1, && (21)\\ < f_i(n-m_i) z_i(n) >_L &= 0, && (22)\\ < f_p(n-m_p) z_i(n) >_L &= 0, && (23) \end{aligned}$$

for all $-(N_i-1) < m_i \le N_i - 1$ and $m_i \ne 0$, $-(N_i-1) \le m_p \le N_p - 1$ and $0 < p \le R-1$ with $p \ne i$.

The system in (20)–(23) has Q_i equations and L variables (the samples of $z_i(n)$), with

$$Q_i = N_D + (R-1)(N_i-1). \tag{24}$$

For $L \ge Q_i$, the system is critically determined or underdetermined and for persistently exciting inputs it always admits a solution. It can be written in matrix form as follows,

$$\mathbf{S}\mathbf{z} = \mathbf{d} \tag{25}$$

where $\mathbf{z}$ is a vector collecting the samples of $z_i(n)$, $\mathbf{d}$ is a vector of all zeros apart from the element 1 corresponding to (21), and $\mathbf{S}$ is a fat or square matrix. Each row of $\mathbf{S}$ is formed by the samples of a basis function $f_p(n-m_p)$, with n ranging along the row from 0 to L, and p and m_p changing along the columns with $0 \le p \le R-1$ and $-(N_i-1) \le m_p \le N_p - 1$. The minimum norm solution of the system is

$$\mathbf{z} = \mathbf{S}(\mathbf{S}\mathbf{S}^T)^{-1}\mathbf{d}. \tag{26}$$

The elements of the matrix $\mathbf{S}\mathbf{S}^T$ are formed by cross-correlations between basis functions with different time delays, i.e., are

$$< f_{p_1}(n-m_{p_1}) f_{p_2}(n-m_{p_2}) >_L \tag{27}$$

where $0 \le p_1, p_2 \le R-1$, $-(N_i-1) \le m_{p_1} \le N_{p_1} - 1$, and $-(N_i-1) \le m_{p_2} \le N_{p_2} - 1$. $\mathbf{S}\mathbf{S}^T$ is a block matrix whose entries are Toeplitz matrices with varying dimensions. For its particular structure it admits efficient algorithms for its inversions. For example, the product $(\mathbf{S}\mathbf{S}^T)^{-1}\mathbf{d}$ can be efficiently computed with the algorithm presented in [54], which was successfully adopted in our research. Working with nonlinear basis functions, $\mathbf{SS}^T$ could have a bad conditioning, but for sufficiently large L we have always been able to find a solution with sufficient accuracy working with double precision arithmetic.

When $L \geq Q = \max_i Q_i = N_D + (R-1)(N-1)$, it becomes possible to develop a set of OPSs $z_i(n)$ for $0 \leq i \leq R-1$, which allows to estimate with the same input sequence all diagonals of the FLiP filter. Furthermore, the same input sequence could be used for estimating different types of FLiPs filters by developing different sets of OPSs.

It should be noted that Q is in general much lower than the number of nonlinear equations $\overline{Q} \simeq R(N_D+1)$ that have to be solved for deriving a PPS for the same FLiP filter (N_D is generally much larger than N). Computing an entire set of OPSs $z_i(n)$ for $0 \leq i \leq R-1$ requires often less computations than a single iteration of the algorithm used to develop PPSs.

The approach for developing OPSs presented in this Section could be easily adapted to identify other families of LIP nonlinear filters. The main characteristic of FLiP filters that is exploited in developing the OPSs and deriving equations (11) and (14) is the filter bank form in (3). Thus, it is possible to develop OPSs for any LIP nonlinear filter that has a filter bank implementation and admits persistently exciting inputs. Most of these LIP filters are specific cases of FLiP filters or combinations of FLiP filters.

IV. Nonlinear system identification with OPSs

In absence of measurement noise, a set of OPSs suitable for the identification of a FLiP filter allows the exact identification with the cross-correlation method of any nonlinear system that can be modeled with the chosen FLiP filter. In case of an under-estimation of the nonlinear system or in presence of a measurement noise, the identification will be affected by an error. In what follows, we assume that a set of OPSs for a FLiP filter of order K, memory N, diagonal number D is used to estimate a nonlinear system whose output is corrupted by a measurement noise $\overline{\nu}(n)$,

$$y(n) = f[x(n), x(n-1), \dots, x(n-N_0+1)] + \overline{\nu}(n). \quad (28)$$

The nonlinear system in (28), is assumed to be a FLiP filter of memory N_{Sys}, order K_{Sys}, diagonal number D_{Sys}, with N_{Sys}, K_{Sys}, D_{Sys} possibly larger than N, K, and D, respectively. In what follows we separately study the effect on the identified model of (i) an underestimation of the memory of the nonlinear system, (ii) an underestimation of the order or diagonal number of the nonlinear system, and (iii) the measurement noise. Eventually, we discuss the computational complexity of OPS identification in comparison with PPS and least-square methods.

A. Memory underestimation

First, the case of an underestimation of the memory of the system is considered. We assume $N_{\text{Sys}} = N + \Delta_N$, with $\Delta_N > 0$, while $K_{\text{Sys}} \leq K$ and $D_{\text{Sys}} \leq D$, and we neglect the effect of the noise assuming $\nu(n) = 0$. In this case, the system in (28) can be written as

$$y(n) = \sum_{p=0}^{R-1} \sum_{m=0}^{N_p+\Delta_N-1} \tilde{h}_p(m) f_p(n-m). \quad (29)$$

where $\tilde{h}_p(n)$ are the coefficients to be estimated with (10). The estimation of the system performed with a set OPSs for a filter of memory N, order K, diagonal number D is affected by an aliasing error, which influences the first Δ_N terms of the diagonals. In fact, for $j \in [0, \Delta_N - 1]$ the estimation of $h_i(j)$ with $< y(n) z_i(n-j) >_L$ is affected by the cross-correlations $< f_p(n-k) z_i(n-j) >_L$ with $k \in [N_p + j, N_p + \Delta_N - 1]$. On the contrary, the cross-correlations $< f_p(n-k) z_i(n-j) >_L$ are zero when $k - j < N_p$ for the orthogonality conditions imposed by (20)-(23). Thus, in case of a memory underestimation the identification is biased.

B. Order or diagonal number underestimation

We next consider the case of an order or diagonal number underestimation. We assume $K_{\text{Sys}} > K$ or $D_{\text{Sys}} > D$, while $N_{\text{Sys}} \leq N$ and $\nu(n) = 0$. In this case, the system in (28) can be written as

$$y(n) = \sum_{p=0}^{R-1} \sum_{m=0}^{N_p-1} \tilde{h}_p(m) f_p(n-m) + \Delta_{K,D}(n) \quad (30)$$

where $\tilde{h}_p(n)$ are the coefficients to be estimated with (10) and $\Delta_{K,D}(n)$ is the linear combination of all basis functions of order greater than K or diagonal number greater than D. In this case, all coefficients estimated with (10) are affected by an error caused by the basis functions of order greater than K or diagonal number greater than D unaccounted in the development of the OPSs, and the identification is again biased. The error on the coefficient $h_i(j)$ is equal to $< \Delta_{K,D}(n) z_i(n-j) >_L$. Even though this error in reality is deterministic, working with large periods and a large number of neglected basis functions in $\Delta_{K,D}(n)$, according to the law of large numbers we can assume this error to be stochastic and Gaussian distributed. Thus, in these conditions the effect of this error can be deemed similar to a measurement noise.

C. Effect of the measurement noise

Eventually, the effect of a measurement noise $\overline{\nu}(n)$ on the coefficients identification is studied considering $N_{\text{Sys}} \leq N$, $K_{\text{Sys}} \leq K$, and $D_{\text{Sys}} \leq D$. Thus, the nonlinear system to be modeled is

$$y(n) = \sum_{p=0}^{R-1} \sum_{m=0}^{N_p-1} \tilde{h}_p(m) f_p(n-m) + \overline{\nu}(n), \quad (31)$$

where $\tilde{h}_p(n)$ are the coefficients to be estimated with the values $h_i(j)$ given in (10). We separately consider the case of a white or colored Gaussian noise, and of a non Gaussian noise.

1) Gaussian noise: We consider first the case where the output is affected by a colored Gaussian noise

$$\overline{\nu}(n) = h_\nu(n) * \nu(n), \tag{32}$$

where $\nu(n)$ is a zero-mean, variance σ_ν^2, white Gaussian noise, and $h_\nu(n)$ is the causal, finite memory N_ν, impulse response of the forming filter that generates the measurement noise, where

$$\sum_{n=0}^{N_\nu-1} h_\nu(n)^2 = 1. \tag{33}$$

As a particular case, when $h_\nu(n) = \delta(n)$, the output of the system is corrupted by a zero-mean, variance σ_ν^2, white Gaussian measurement noise.

Using (10) and exploiting the properties of OPSs,

$$h_i(j) = \tilde{h}_i(j) + < (h_\nu(n) * \nu(n)) z_i(n-j) >_L . \tag{34}$$

Since $\nu(n)$ is zero mean, $E[h_i(j)] = \tilde{h}_i(j)$ and the identification method is unbiased.

We next estimate the mean square deviation (MSD) of the coefficients. The MSD of the coefficients of $f_i(n-j)$ is defined as

$$\mathrm{MSD}_{i,j} = E[(h_i(j) - \tilde{h}_i(j))^2]. \tag{35}$$

For OPSs, from (10) we have

$$\mathrm{MSD}_{i,j} = E[(< (h_\nu(n) * \nu(n)) z_i(n-j) >_L)^2]. \tag{36}$$

It is proved in the Appendix that

$$\mathrm{MSD}_{i,j} = \sigma_\nu^2 \sum_{m=-N_\nu+1}^{L-1} < h_\nu(n-m) z_i(n-j) >_L^2 . \tag{37}$$

The sum in (37) can be interpreted as the energy of the sequence

$$[z_i(-j), z_i(-j+1), \ldots, z_i(L-j-1)], \tag{38}$$

filtered with the finite impulse response (FIR) filter with impulse response $h_\nu(-n)$. When the nonlinear system output is corrupted by a white Gaussian noise, i.e., when $h_\nu(n) = \delta(n)$, the mean square deviation reduces to

$$\mathrm{MSD}_{i,j} = \sigma_\nu^2 < z_i(n)^2 >_L, \tag{39}$$

which is independent of the delay j.

From (37) and (39) it is evident that $\mathrm{MSD}_{i,j}$ is proportional to the noise power σ_ν^2. It can also be observed that $\mathrm{MSD}_{i,j}$ is inversely proportional to $< f_i^2(n) >_L$, the energy of the basis functions over a period, because according to (21) $< z_i^2(n) >_L$ is inversely proportional to $< f_i^2(n) >_L$. Since for constant power, the energy of the basis functions over a period increases proportionally to L, $\mathrm{MSD}_{i,j}$ is also inversely proportional to the period L. Thus, we can improve the accuracy of the estimation by increasing the period L or by computing the cross-correlation over multiple periods.

To compare the different OPSs on equal terms, we introduce the noise gain G_ν, which is defined as the average value over all basis functions of

$$G_{\nu,i,j} = \frac{\mathrm{MSD}_{i,j}}{\sigma_\nu^2} < f_i^2(n-j) >_L, \tag{40}$$

and evaluate it in particular for a white measurement noise. From (39) we have that G_ν is the average value over all basis functions of

$$G_{\nu,i,j} =< z_i^2(n) >_L \cdot < f_i^2(n) >_L . \tag{41}$$

For PPSs, it can be proved that $G_{\nu,i,j}$ and G_ν are always 1, independently of the considered filter or the period L of the sequence. On the contrary, for OPSs we show in Section V that G_ν changes with the chosen FLiP filter, the distribution of the input samples, and the period L. For a specific filter and input sample distribution, G_ν can greatly vary with L, because the choice of L influences the power of the designed OPS $z_i(n)$. When $L = Q$, i.e., the minimum period of the OPS, we have found G_ν can assume very large values that make the identification with OPSs useless. On the contrary, when $L \gg Q$, G_ν assumes reasonable values. In orthogonal FLiP filters, when the input samples are taken from the distribution ideally guaranteeing the orthogonality of the basis functions (e.g., Gaussian for WN filters or uniform for LN filters), when L tends to infinity, G_ν tends to 1, the ideal value we have with PPSs. This property can be easily expected since the longer is L, the closer the input sequence is to a PPS.

2) Non Gaussian noise: At this point it is interesting to note how the OPS identification behaves when the noise is zero mean but non-Gaussian. Using (10) and exploiting the properties of OPSs, we now have

$$h_i(j) = \tilde{h}_i(j) + < \overline{\nu}(n) z_i(n-j) >_L . \tag{42}$$

If $\overline{\nu}(n)$ is zero mean, $E[h_i(j)] = \tilde{h}_i(j)$ and the identification method is also in this case unbiased.

According to equation (10) and (31), for the central limit theorem the error $h_i(j) - \tilde{h}_i(j)$ will still have a Gaussian distribution, with

$$\mathrm{MSD}_{i,j} = E[(< \overline{\nu}(n) z_i(n-j) >_L)^2]. \tag{43}$$

depending on the auto-correlation function of the noise and on the OPS sequence $z_i(n)$. We must also point out that the only protection with respect to outliers is the averaging performed in the cross-correlation with OPSs. The interested reader, is referred to [47]–[49], [55] for some recent methods that are robust against outliers.

D. Computational cost of identification

From (10), the computational cost of a filter identification with OPS is around LN_D operations, i.e., multiplications and additions, if the cross-correlations are computed in time domain. It is of $(2\log_2(L)+1)LR$ operations if the cross-correlations are computed in DFT domain, assuming a FFT cost $L\log_2(L)$ operations. These computational costs are much lower than that of a LS identification on the same data, which is order of LN_D^2 operations. Also for PPSs we have a computational cost of LN_D operations in time domain and $(2\log_2(L)+1)LR$ operations in DFT domain, provided in (8) we neglect the cost of computing the basis functions and of the normalization. Nevertheless, it will be shown in Section V that OPSs can provide identification performance similar to the PPSs with much shorter periods.

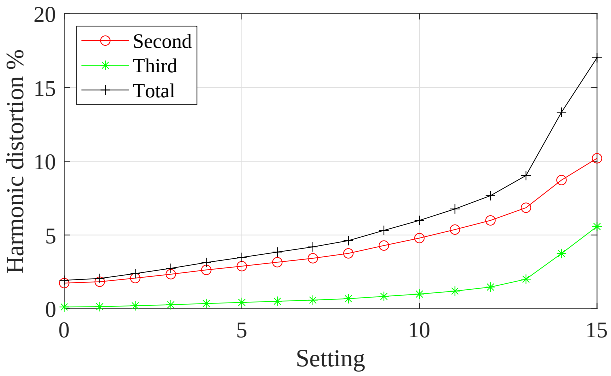


Fig. 1. First experiment: Seconds, third, and total harmonic distortion.

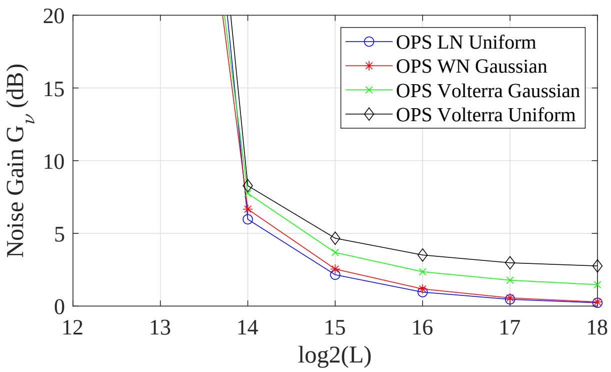


Fig. 2. First experiment: Noise Gain of OPSs for LN, WN and Volterra filters.

V. Experimental results

In this section, we provide some experimental results that illustrate the ability of OPSs to identify FLiP filters and we compare the cross-correlation method based on OPSs with that based on PPSs and stochastic signals, and moreover with the LS method. We consider three sets of experiments, with the first two involving the identification of real nonlinear devices. In the first set of experiments, we consider the identification of the nonlinear device with different FLiP filters having maximum diagonal number. In the second set of experiments, to reduce the complexity of the model the nonlinear device is identified with a simplified FLiP filter, whose memory and diagonal number decrease with the kernel order. In the third experiment, we consider the identification of a nonlinear system often used for benchmark purposes, the Duffing oscillator. In this experiment, the model is a simplified Volterra filter composed only of odd kernels. We use the measured Volterra kernels to estimate the Duffing model parameters and we compare the results obtained with proposed approach, which is non-parametric, with those of a parametric method based on multi-tones.

A. First Experiment

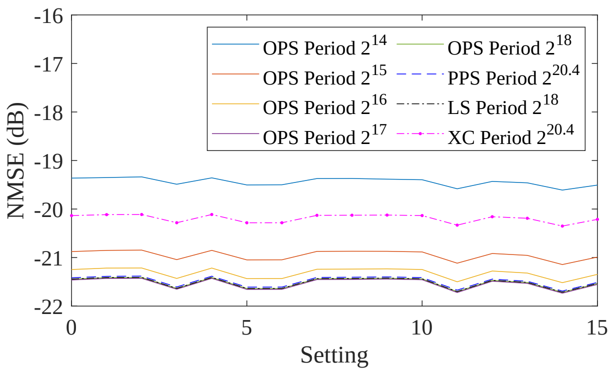
The first experiment concerns the identification of a Behringer MIC100 tube pre-amplifier. The acquisition was performed at 16 kHz sampling frequency and the identification was obtained with LN, WN and Volterra filters of order 3, memory 25, with full diagonals ($D = N - 1 = 24$). For the OPSs, we used periodic inputs with the following periods: $\{11700, 2^{14}, 2^{15}, 2^{16}, 2^{17}, 2^{18}\}$, with samples having uniform or Gaussian distribution. PPSs for Legendre and WN filters with period of $1,393,024 \simeq 2^{20.4}$ samples were also considered. Moreover, $1,000,000$ samples of zero mean white Gaussian and white uniform noise were also used as input. All inputs with samples having uniform distribution had signal power equal to $1/3$, while those with uniform distribution had power equal to $1/12$. The maximum peak amplitude was the same for both distributions and around 1. The input signals at 16 kHz were up-sampled in Matlab at 48 kHz to allow playback and recording on the Focusrite Scarlett 2i2 audio interface and the recorded signals were downsampled back at 16 kHz. Thus, in this experiment we truly identify a chain composed by the upsampler, the Scarlett digital to analog converter (DAC), the MIC100 preamplifier, the Scarlett analog to digital (ADC), and the downsampler. The output signal to noise ratio was more the 60 dB.

Sixteen different settings were considered with increasing level of nonlinear distortion. Fig. 1 provides an indication of the level of nonlinearity at the different settings. It plots the harmonic distortion for a tone having the same maximum amplitude of the input signals. The harmonic distortion is defined as the ratio in percent between the power of the harmonics and that of the fundamental.

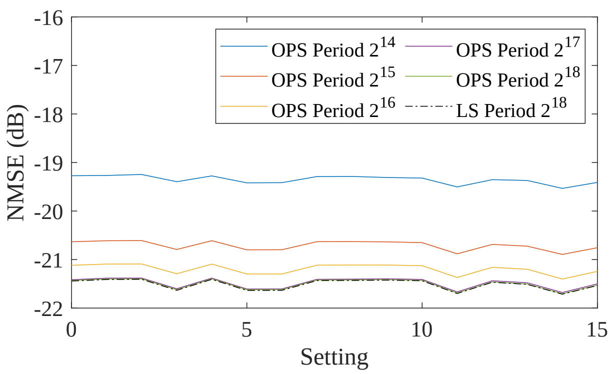
The Noise Gain for all OPSs is shown in Fig. 2. It can be appreciated that the Noise Gain tends to decrease with the period of the OPS. For the period of $11,700$, the minimum possible OPS period for the considered system, the noise gain is unacceptably high, the identification performance is consequently very bad and is not plotted in the following. For the orthogonal FLiP filters, i.e., LN and WN filters, with an input sample distribution in accordance with that ensuring the orthogonality of the basis functions, the Noise Gain tends to zero for large periods L. On the contrary, with the Volterra filter for large periods the Noise Gain tends to a constant value, which is anyway acceptably small (2–3 dB).

The MIC100 tube pre-amplifier was identified with LN, WN, and Volterra filters with OPSs considering both the Gaussian and uniform periodic inputs. All filters were also identified with the LS method over a segment of 2^{18} samples having Gaussian or uniform distribution. The LN and WN filters were also identified with the cross-correlation method using PPSs and stochastic inputs having a uniform and Gaussian distribution, respectively.

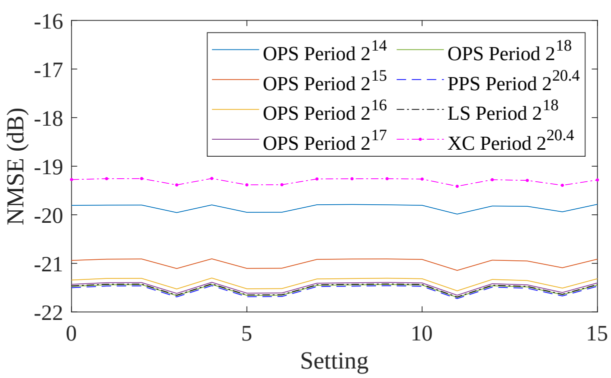
All identified filters were then tested on $200,000$ samples of a stochastic input different from that used for identification, but having an input sample distribution matching that used for the identification. Figure 3 shows the resulting normalized mean square error (NMSE) in the identification with the different methods. The NMSE is the mean square error normalized by the power of the output signal. It is evident that the OPSs are capable of obtaining very low levels of NMSE in all filters (i.e., LN, WN and Volterra with uniform and Gaussian distribution) and all settings. For periods equal to or larger than 2^{17} the OPSs provide results equivalent to those obtained with PPS or with the LS method. On the contrary,



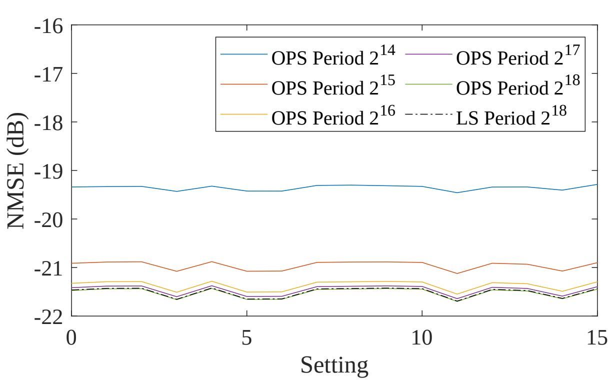
(a)



(b)



(c)



(d)

Fig. 3. First experiment: NMSEs for (a) LN filter and (b) Volterra filter on uniform distribution input, and for (c) WN filter and (d) Volterra filter on Gaussian distribution input.

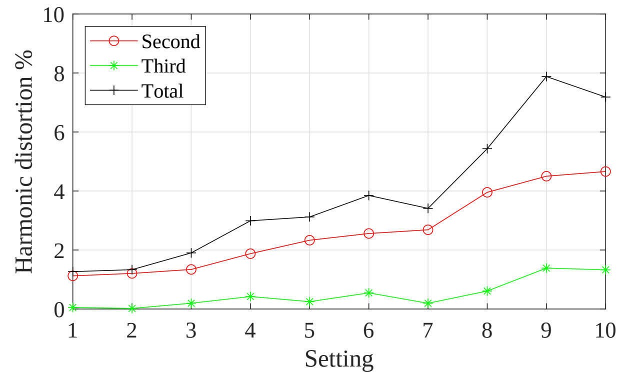


Fig. 4. Second Experiment: Seconds, third, and total harmonic distortion.

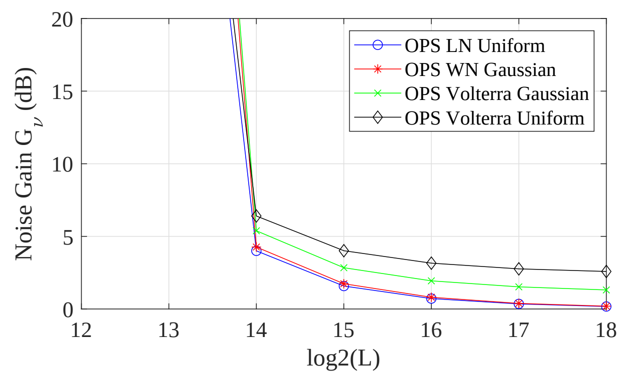
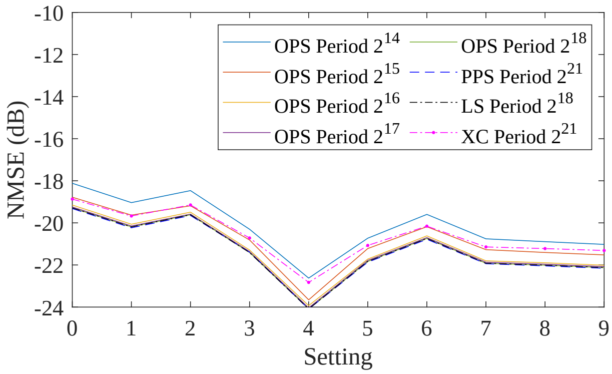


Fig. 5. Second Experiment: Noise Gain of OPSs for LN, WN and Volterra filters.

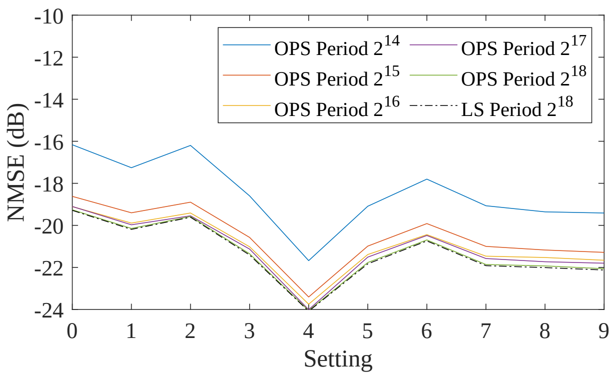
the results of the cross-correlation method based on stochastic inputs, indicated with XC in Fig. 3, provide poor results in all conditions. It is thus possible to appreciate the great improvement obtained in recent years in the identification with the cross-correlation method, passing from stochastic inputs to deterministic sequences as the PPS and OPS. Furthermore, it is important to underline that OPSs reach these good results with a lower computational complexity than PPSs and the LS method. The OPSs with period 2^{18} reduce the computational complexity of the identification by at least a factor 5 in comparison with the PPS that have period $2^{20.4}$.

B. Second Experiment

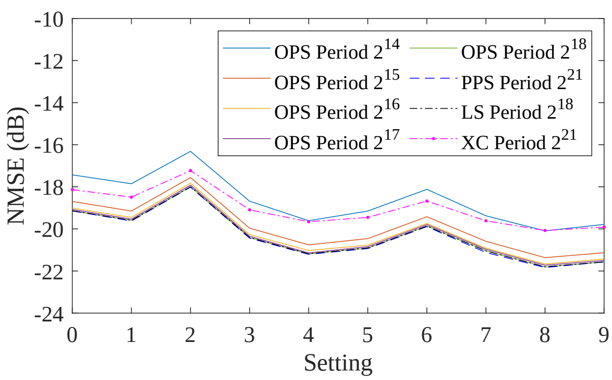
In the second experiment, the identification of a Fender Hot Rod Deluxe vacuum tube power amplifier working at 44.1 kHz sampling frequency was performed. The guitar amplifier was loaded with a Two Notes Torpado Load Captor 8 avoiding the speaker influence. The acquisitions were performed using a National Instruments Compact Rio chassis (cRIO-9024) equipped with a 2-channel voltage analog output NI-9260 and with a 3-channel voltage analog input NI-9232, connected to a desktop PC. Both generation and acquisition boards were configured to use the same sample clock with a sample rate of 44,100 Hz. The National Instruments software LabView was used to generate the test signals and to acquire the measurements.



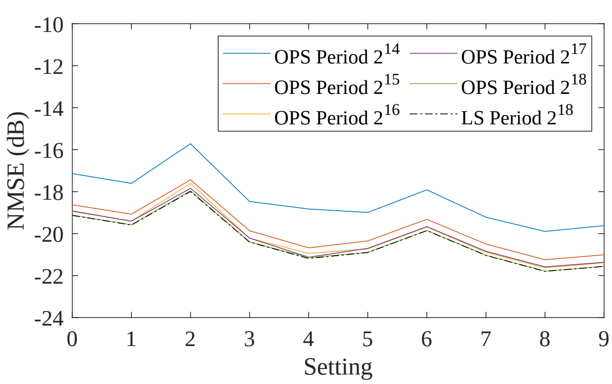
(a)



(b)



(c)



(d)

Fig. 6. Second Experiment: NMSEs for (a) LN filter and (b) Volterra filter on uniform distribution input, and for (c) WN filter and (d) Volterra filter on Gaussian distribution input.

The amplifier was identified again with LN, WN, and Volterra filters of order 3, but with kernels having different memory and diagonal number:

- kernel 1 has memory 64,
- kernel 2 has memory 40 and diagonal number 15,
- kernel 3 has memory 32 and diagonal number 12.

For the OPSs, we used periodic inputs with periods: $\{9573, 2^{14}, 2^{15}, 2^{16}, 2^{17}, 2^{18}\}$, and samples having uniform or Gaussian distribution. PPSs for Legendre and WN filters suitable to identify the same filter with period of $2,097,656 \simeq 2^{21}$ samples were also considered. Moreover, more than $1,000,000$ random samples with Gaussian and uniform distribution were also applied as input. The input signal powers and the peak amplitude for the samples having Gaussian and uniform distribution were the same of the first experiment.

Ten different settings were considered with different level of nonlinear distortion and different gain on the input signal. To give an indication of the level of nonlinear distortion, Fig. 4 plots the harmonic distortion at the different settings. The settings are sorted for increasing second order harmonic distortion. The output signal to noise ratio ranged from 16 dB till 50 dB according to the different settings.

The Noise Gain for the considered signals is shown in Fig. 5. For the Period 9573, the minimum OPS period that allows us the identification of the considered filters, the noise gain is unacceptably high, identification performance is consequently very bad and is not plotted in the following.

The LN, WN, and Volterra filters were identified on the periodic signals with the cross-correlation method using OPSs. The LN and WN filters were also identified with the cross-correlation method using PPSs and stochastic inputs. All filters were identified also with the LS method over 2^{18} input samples having Gaussian or uniform distribution. The resulting filters were then tested on a different segment of $200,000$ random samples having the same distribution used for the identification. Figure 6 shows the NMSE in the identification with the different methods. It is evident that also in this case the OPSs are capable of obtaining very low levels of NMSE in all filters (i.e., LN, WN and Volterra with uniform and Gaussian distribution) and all settings. The performance obtained with OPSs are similar to those of PPSs and LS method. The performance of the cross-correlation method with stochastic inputs, indicated with XC in the plots, is worse in all conditions. It is important to underline that the good results of the OPSs are reached with a lower computational complexity than PPSs, as reported in the Figures. Considering an OPS of period 2^{16}, the identification has a computational cost at least 32 time lower than the identification with PPSs.

C. Third Experiment

The third experiment considers the identification of an analog nonlinear system, the Duffing oscillator model, which has often been used as benchmark for nonlinear system

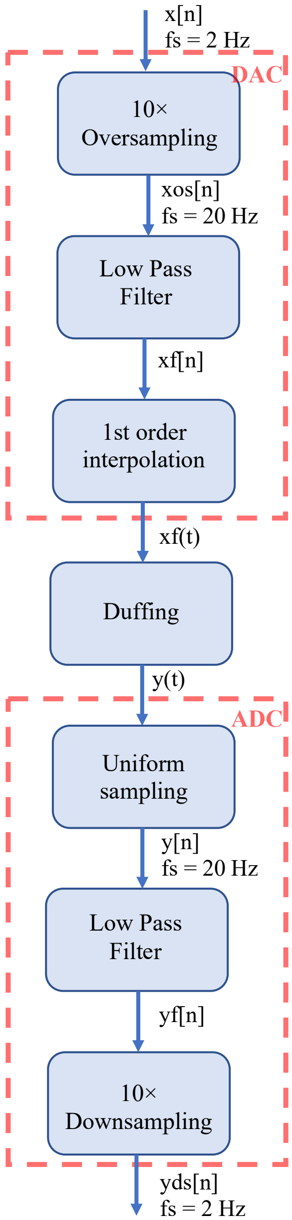


Fig. 7. Third Experiment: the identified system.

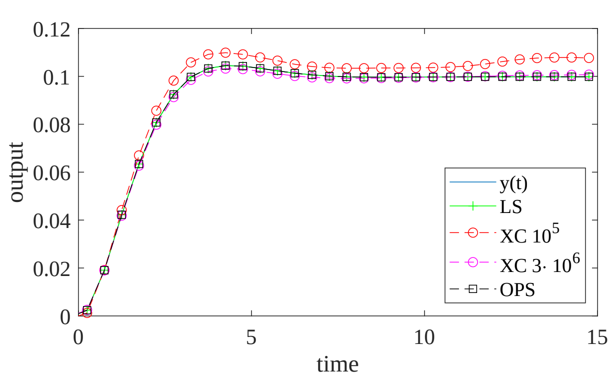


Fig. 8. Third Experiment: Step response of the nonlinear system identified with a third order Volterra filter.

identification, especially for continuous time system:

$$\begin{cases} m\dfrac{d^2y}{dt^2}(t) + c\dfrac{dy}{dt}(t) + k_1 y(t) + k_3 y^3(t) = u(t), \\ y(0) = 0, \\ \dfrac{dy(0)}{dt}(0) = 0. \end{cases} \tag{44}$$

In (44), $u(t)$ and $y(t)$ are the continuous-time input and output, respectively, and m, c, k_1, and k_3 are the oscillator parameters.

The Duffing oscillator admits a Volterra series expansion around the zero equilibrium point when the input is small [56]–[58]. In what follows, we consider $m = 1$, $c = 1.4$, $k_1 = 1$ and $k_3 = 0.1$, and we aim at identifying the system with a discrete-time Volterra filter, and to estimate the model parameters m, c, k_1, and k_3 from the knowledge of the Volterra kernels. The system in (44) is an analog system and we simulate the digital measurement chain as described in Figure 7. Specifically, considering t measured in s, we work with a sampling frequency of 2 Hz, which has been found adequate in our experiments. We simulate the effect of the digital-to-analog converter by i) upsampling the discrete-time input signal by a factor 10, ii) filtering it by discrete time low-pass filter of order 400 with 0.8 Hz passband, iii) linearly interpolating the resulting signal (to find the input values required by the ordinary differential equations solver). We then apply the interpolated signal to the Duffing oscillator (using the *ode15s* solver of Matlab). The analog-to-digital converter is simulated by i) estimating the Duffing output signal on a uniform time grid at 20 Hz sampling frequency (it requires an interpolation of the *ode15s* output samples, which are not uniformly distributed), ii) low-pass filtering the resulting signal with a filter having order 400 and 0.8 Hz passband, ii) downsampling the filter output with a 2 Hz frequency. The delay introduced by the digital filters has been compensated before the model identification.

Different signals have been applied and used to identify the Duffing oscillator with a Volterra filter of order 3, memory 40, diagonal number 19, with only odd kernels since the even kernels are zero [56], and thus with a total of 5780 coefficients. The system has been identified as follows: i) with the cross-correlation method considering an OPS for Volterra filters of period $131,072$ and noise gain 1.62; ii) with the cross-correlation method applied to stochastic inputs composed by $131,072$ and $3,000,000$ samples of a zero mean white Gaussian noise, estimating first a WN filter and then converting it to a Volterra filter; iii) with the LS method over $131,072$ samples of zero mean white Gaussian noise; iv) with multi-tone signals, described later, that allow a parametric identification of the model. All signals have the same variance equal to $1/12$. The Volterra models (case i, ii, iii) were then used to obtain the step-response of the Duffing oscillator for a step of amplitude 0.1. Fig. 8 compares the estimated step-responses with the actual step-response of the system in Figure 7. In the Figure, $y(t)$ is the actual step-response, LS, XC, OPS indicate the step-responses obtained with the LS method, with the cross-correlation method using a stochastic input and using an OPS, respectively. We can see that system identified with OPS provides equally good results as the LS method. The system identified with the cross-correlation method with a stochastic input over $131,072$ samples provides a much worse estimation. This estimation improves for $3,000,000$ samples but it is still worse than the other methods. The cross-correlations with stochastic inputs, indeed, reach the ideal values only for very large numbers of input samples [26], [59].

In this experiment, for the same identification performance, the OPS allows us to identify the model with a computational complexity which is 24 times lower than cross-correlation method with stochastic input, and is several magnitude orders lower than the LS method.

We next consider the estimation of the Duffing model parameters m, c, k_1, and k_3 under different noise conditions. These parameters can be estimated from the knowledge of the Volterra kernels. In fact, as shown in [60], the Fourier transform of the first order kernel of the Duffing model is

$$H_1(j\omega) = \frac{1}{k_1 + c(j\omega) - m(j\omega)^2}, \tag{45}$$

and the main diagonal of multidimensional Fourier transform of the third order kernel is

$$H_3(j\omega, j\omega, j\omega) = -k_3[H_1(j\omega)]^3 H_1(3j\omega). \tag{46}$$

Thus, m, c, k_1 have been estimated by curve fitting the Fourier transform of the first order Volterra kernel in the band $[0, 0.8]$ Hz. Then k_3 has been obtained from (46) and averaged over a few frequency samples around the zero frequency.

For comparison purposes, we have also estimated the same parameters m, c, k_1, and k_3 with a multi-tone approach. According to [61], the outputs for a dual-tone signal and a single-tone signal are sufficient to estimate the Duffing oscillator parameters. The method of [61] estimates the output spectrum on eight frequencies using the dual-tone and applies a single-tone for obtaining a further measure. With these nine values, it obtains a nonlinear equation system in nine variables, i.e., k_3 and the Fourier transform of the first order kernel at the eight frequencies. Parameters m, c, and k_1 are obtained by curve fitting the Fourier transform. Nevertheless, at the considered settings we have obtained poor results with the method of [61] due to a poor estimation of the tones harmonics. To solve this problem, we have estimated the Duffing parameters using four dual-tone signals with frequencies equal to those of the fundamentals and harmonics of [61], i.e., with angular frequencies, $[0.4\omega_n, 0.5\omega_n]$, $[1.2\omega_n, 1.5\omega_n]$, $[0.6\omega_n, 1.4\omega_n]$, $[0.3\omega_n, 1.3\omega_n]$, where $\omega_n = \sqrt{\frac{k_1}{m}}$. All signals have power $1/12$. A single tone sequence with angular frequency $0.5\omega_n$ and power $1/24$ was also applied as input. From the values of output spectra in correspondence to the tone frequencies, using the formulas (22a), (22b), and (22i) of [61] it is possible to obtain a nonlinear equation system of nine equations in nine variables that allows to accurately estimate k_3 and the Fourier transform of the first order Volterra kernel at the tone frequencies. From the latter, it is possible to estimate the parameters m, c, and k_1 by curve fitting as in [61].

Table II provides the ensemble average of the Duffing oscillator parameters estimated over 100 identifications with the OPS, the least-square method (LS) and the multi-tone approach (MT), when the output signal is corrupted with a zero mean white Gaussian noise at different signal-to-noise (SNR) ratios. For all methods, the parameters are estimated over the same number of input samples. Table III provides the same results for a non-Gaussian output noise, specifically a Gaussian mixture composed by a zero mean Gaussian noise with probability 0.95 and a second zero mean Gaussian noise with variance 100 times larger than the first and probability 0.05. We estimated also the same parameters for the cross-correlation method applied to stochastic inputs, but these results are not reported since the estimation of k_3 is much worse than with the other methods. From Table II and III we can notice that the estimate of m, c, and k_1 is very reliable, while the estimate of k_3 is generally affected by a larger error. The proposed estimate with OPSs, provides results as good as those of the LS method and of the multi-tone method, confirming the good characteristics of the proposed approach.

TABLE II
ENSEMBLE AVERAGES OF DUFFING PARAMETERS ESTIMATED UNDER DIFFERENT GAUSSIAN MEASUREMENT NOISES

		SNR			
		0 dB	20 dB	40 dB	60 dB
Method	OPS	$m = 1.011$ $c = 1.381$ $k_1 = 1.011$ $k_3 = 0.071$	$m = 0.997$ $c = 1.404$ $k_1 = 0.998$ $k_3 = 0.092$	$m = 0.999$ $c = 1.404$ $k_1 = 0.999$ $k_3 = 0.098$	$m = 0.999$ $c = 1.404$ $k_1 = 0.999$ $k_3 = 0.097$
	LS	$m = 1.020$ $c = 1.400$ $k_1 = 1.010$ $k_3 = 0.061$	$m = 1.003$ $c = 1.403$ $k_1 = 1.001$ $k_3 = 0.086$	$m = 0.999$ $c = 1.403$ $k_1 = 0.999$ $k_3 = 0.093$	$m = 0.999$ $c = 1.403$ $k_1 = 0.999$ $k_3 = 0.093$
	MT	$m = 1.003$ $c = 1.400$ $k_1 = 0.998$ $k_3 = 0.061$	$m = 1.000$ $c = 1.400$ $k_1 = 0.999$ $k_3 = 0.088$	$m = 1.001$ $c = 1.400$ $k_1 = 0.998$ $k_3 = 0.091$	$m = 1.000$ $c = 1.400$ $k_1 = 0.998$ $k_3 = 0.091$

TABLE III
ENSEMBLE AVERAGES OF DUFFING PARAMETERS ESTIMATED UNDER DIFFERENT NON-GAUSSIAN MEASUREMENT NOISES

		SNR			
		0 dB	20 dB	40 dB	60 dB
Method	OPS	$m = 0.964$ $c = 1.423$ $k_1 = 0.969$ $k_3 = 0.150$	$m = 0.994$ $c = 1.404$ $k_1 = 0.997$ $k_3 = 0.092$	$m = 0.999$ $c = 1.404$ $k_1 = 0.999$ $k_3 = 0.097$	$m = 0.999$ $c = 1.404$ $k_1 = 0.999$ $k_3 = 0.097$
	LS	$m = 0.978$ $c = 1.402$ $k_1 = 0.988$ $k_3 = 0.172$	$m = 0.997$ $c = 1.406$ $k_1 = 0.998$ $k_3 = 0.094$	$m = 0.999$ $c = 1.403$ $k_1 = 0.999$ $k_3 = 0.094$	$m = 0.999$ $c = 1.403$ $k_1 = 0.999$ $k_3 = 0.093$
	MT	$m = 0.999$ $c = 1.400$ $k_1 = 0.999$ $k_3 = 0.094$	$m = 1.000$ $c = 1.400$ $k_1 = 0.998$ $k_3 = 0.091$	$m = 1.000$ $c = 1.400$ $k_1 = 0.998$ $k_3 = 0.091$	$m = 1.000$ $c = 1.400$ $k_1 = 0.998$ $k_3 = 0.091$

VI. CONCLUSION

The paper presents a novel family of deterministic input signals, the OPSs, that allows the identification of FLiP filters on a finite time interval with the cross-correlation method. System identification with OPSs requires applying a periodic input to the unknown system, recording the corresponding output for at least one period[2], computing each of the FLiP filter diagonals from the cross-correlation between the output and the OPS corresponding to the specific diagonal. The OPSs share many similarities with the PPSs. In contrast to PPSs, OPSs allow also the identification of non-orthogonal FLiP filters, as the popular Volterra filter. They can also identify FLiP filters with a period and a computational complexity much lower than PPSs. The achievable performance of OPSs has been theoretically studied and has been verified with experimental results. With more than 1,300 filter identifications, it has been shown that OPSs can achieve performance similar to PPSs and to LS method with a much smaller computational complexity. It should also be noted that OPSs could be developed also for other families of LIP nonlinear filters, specifically all families of LIP filters that share the filterbank structure of FLiP filters, as for example the Hammerstein or the FLANN filters. In fact, these filters can often be interpreted as particular cases of FLiP filters, or parallel connections of FLiP filters. Future works will concern the use of OPSs in machine learning and artificial intelligence systems.

The OPSs used in the experiments of Section V and other example of OPSs can be found at [62].

APPENDIX
PROOF OF (37)

By replacing in (36) the expression of $h_\nu(n) * \nu(n)$,

$$\sum_{m=0}^{N_\nu-1} h_\nu(m)\nu(n-m) = \sum_{m=n-N_\nu+1}^{n} \nu(m)h_\nu(n-m),$$

we have

$$\begin{aligned}\mathrm{MSD}_{i,j} &= \\ &= E[< \sum_{m=n-N_\nu+1}^{n} \nu(m)h_\nu(n-m)z_i(n-j) >_L^2] \\ &= E[(\sum_{n=0}^{L-1}\sum_{m=n-N_\nu+1}^{n} \nu(m)h_\nu(n-m)z_i(n-j))^2] \\ &= E[(\sum_{n=0}^{L-1}\sum_{m=-N_\nu+1}^{L-1} \nu(m)h_\nu(n-m)z_i(n-j))^2],\end{aligned}$$

since $h_\nu(n) = 0$ for $n \notin [0, N_\nu - 1]$. By exchanging the order of the two summations we obtain

$$\begin{aligned}\mathrm{MSD}_{i,j} &= \\ &= E[(\sum_{m=-N_\nu+1}^{L-1} \nu(m)\sum_{n=0}^{L-1} h_\nu(n-m)z_i(n-j))^2] \\ &= E[(\sum_{m=-N_\nu+1}^{L-1} \nu(m) < h_\nu(n-m)z_i(n-j) >_L)^2].\end{aligned}$$

Computing the expectation of the squared summation in the last expression, all autocorrelation functions of the white noise $\nu(n)$ vanish except those equal to the variance of $\nu(n)$, letting us to obtain (37).

[2]In practice, when we apply the periodic input starting from zero, we must wait the transitory period to end before starting recording, with the transitory period which is equal to the memory N of the FLiP filter.

ACKNOWLEDGMENT

The authors would like to thank Giovanni L. Sicuranza for sharing his expertise and wisdom and for reviewing the first version of this paper.

REFERENCES

[1] G. L. Sicuranza and A. Carini, “Nonlinear system identification using quasi-perfect periodic sequences,” *Signal Processing*, vol. 120, pp. 174–184, Mar. 2016.
[2] A. Carini, S. Cecchi, and S. Orcioni, *Orthogonal LIP Nonlinear Filters*. Oxford, UK: Butterworth-Heinemann, 2018, ch. 2, pp. 15–46.
[3] L. A. Azpicueta-Ruiz, M. Zeller, A. R. Figueiras-Vidal, J. Arenas-Garcia, and W. Kellermann, “Adaptive combination of Volterra kernels and its application to nonlinear acoustic echo cancellation,” *IEEE Transactions on Audio, Speech and Language Processing*, vol. 19, no. 1, pp. 97–110, 2011.
[4] N. Vanli and S. Kozat, “A comprehensive approach to universal piecewise nonlinear regression based on trees,” *IEEE Transactions on Signal Processing*, vol. 62, no. 20, pp. 5471–5486, 2014.
[5] H. Vogt, G. Enzner, and A. Sezgin, “State-space adaptive nonlinear self-interference cancellation for full-duplex communication,” *IEEE Transactions on Signal Processing*, vol. 67, no. 11, pp. 2810–2825, 2019.
[6] F. C. Pinheiro and C. G. Lopes, “A low-complexity nonlinear least mean squares filter based on a decomposable Volterra model,” *IEEE Transactions on Signal Processing*, vol. 67, no. 21, pp. 5463–5478, 2019.
[7] C. Hofmann and W. Kellermann, *Recent Advances on LIP Nonlinear Filters and Their Applications: Efficient Solutions and Significance-Aware Filtering*. Oxford, UK: Butterworth-Heinemann, 2018, pp. 71 – 102.
[8] M. Scarpiniti, D. Comminiello, R. Parisi, and A. Uncini, *Spline Adaptive Filters: Theory and Applications*. Oxford, UK: Butterworth-Heinemann, 2018, pp. 47 – 69.
[9] V. Patel and N. V. George, “Design of dynamic linear-in-the-parameters nonlinear filters for active noise control,” in *24th European Signal Processing Conference (EUSIPCO)*, Aug. 2016, pp. 16–20.
[10] R. Claser, V. H. Nascimento, and Y. V. Zakharov, “A low-complexity RLS-DCD algorithm for Volterra system identification,” in *24th European Signal Processing Conference (EUSIPCO)*, Aug. 2016, pp. 6–10.
[11] O. Karakuş, E. Kuruoğlu, and M. A. Altınkaya, “Bayesian Volterra system identification using reversible jump MCMC algorithm,” *Signal Processing*, vol. 141, pp. 125–136, 2017.
[12] E. Hoerr and R. C. Maher, “Using Volterra series modeling techniques to classify black-box audio effects,” in *Audio Engineering Society Convention 147*, 2019.
[13] V. Bhateja, M. Misra, and S. Urooj, “Non-linear polynomial filters for edge enhancement of mammogram lesions,” *Computer methods and programs in biomedicine*, vol. 129, pp. 125–134, 2016.
[14] L. Tronchin, “The emulation of nonlinear time-invariant audio systems with memory by means of Volterra series,” *Journal of the Audio Engineering Society*, vol. 60, no. 12, pp. 984–996, 2013.
[15] L. Tronchin and V. L. Coli, “Further investigations in the emulation of nonlinear systems with Volterra series,” *Journal of the Audio Engineering Society*, vol. 63, no. 9, pp. 671–683, 2015.
[16] L. Tronchin, V. L. Coli, and F. F. Gionfalo, “Modelling nonlinearities on musical instruments by means of Volterra series,” in *Audio Engineering Society Convention 142*, 2017.
[17] F. Oliveira, E. Batista, and R. Seara, “Effective hardware implementation of Volterra filters based on reduced-rank approaches,” *Electronics Letters*, vol. 54, no. 16, pp. 1005–1007, 2018.
[18] V. J. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*. New York: John Wiley and Sons, 2000.
[19] V. Marmarelis, *Nonlinear Dynamic Modeling of Physiological Systems*. Hoboken, NJ, USA: John Wiley & Sons, 2004.
[20] S. Orcioni, “Improving the approximation ability of Volterra series identified with a cross-correlation method,” *Nonlinear Dynamics*, vol. 78, no. 4, pp. 2861–2869, dec 2014.
[21] M. Schetzen, “Nonlinear system modeling based on the Wiener theory,” *Proceedings of the IEEE*, vol. 69, no. 12, pp. 1557–1573, 1981.
[22] A. Carini and G. L. Sicuranza, “Fourier nonlinear filters,” *Signal Processing*, vol. 94, pp. 183–194, Jan. 2014.

[23] A. Carini, S. Cecchi, L. Romoli, and G. L. Sicuranza, "Legendre nonlinear filters," *Signal Processing*, vol. 109, pp. 84–94, Apr. 2015.

[24] A. Carini and G. L. Sicuranza, "A study about Chebyshev nonlinear filters," *Signal Processing*, vol. 122, pp. 24–32, May 2016.

[25] W. Zhao, T. H. Beach, and Y. Rezgui, "Efficient least angle regression for identification of linear-in-the-parameters models," *Proc. R. Soc. A Math. Phys. Eng. Sci.*, vol. 473, no. 2198, pp. 1–22, feb 2017.

[26] S. Orcioni, M. Pirani, and C. Turchetti, "Advances in Lee-Schetzen method for Volterra filter identification," *Multidimens. Syst. Signal Process.*, vol. 16, no. 3, pp. 265–284, 2005.

[27] S. Orcioni, A. Terenzi, S. Cecchi, F. Piazza, and A. Carini, "Identification of Volterra models of tube audio devices using multiple-variance method," *Journal of the Audio Engineering Society*, vol. 66, no. 10, pp. 823–838, Oct. 2018.

[28] V. Ipatov, "Ternary sequences with ideal periodic autocorrelation properties," *Radio Eng. Electronics and Physics*, vol. 24, pp. 75–79, 1979.

[29] H. Luke, "Sequences and arrays with perfect periodic correlation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 24, no. 3, pp. 287–294, May 1988.

[30] A. Milewski, "Periodic sequences with optimal properties for channel estimation and fast start-up equalization," *IBM J. Res. Development*, vol. 27, no. 5, pp. 426–431, Sep. 1983.

[31] C. Antweiler and M. Dörbecker, "Perfect sequence excitation of the NLMS algorithm and its application to acoustic echo control," in *Annales des Telecommunications*, vol. 49, no. 7-8, 1994, pp. 386–397.

[32] C. Antweiler, "Multi-channel system identification with perfect sequences – theory and applications –," in *Adv. Digit. Speech Transm.* Chichester, UK: John Wiley & Sons, Ltd, 2008, pp. 171–198.

[33] C. Antweiler and G. Enzner, "Perfect sequence lms for rapid acquisition of continuous-azimuth head related impulse responses," in *2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2009, pp. 281–284.

[34] A. Carini and G. L. Sicuranza, "Perfect periodic sequences for identification of even mirror Fourier nonlinear filters," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 7959–7963.

[35] ——, "Perfect periodic sequences for even mirror Fourier nonlinear filters," *Signal Processing*, vol. 104, pp. 80–93, Nov. 2014.

[36] A. Carini, S. Cecchi, L. Romoli, and G. L. Sicuranza, "Perfect periodic sequences for Legendre nonlinear filters," in *22nd European Signal Processing Conference (EUSIPCO)*, Sep. 2014, pp. 2400–2404.

[37] A. Carini and G. L. Sicuranza, "A study about Chebyshev nonlinear filters," *Signal Processing*, vol. 122, pp. 24–32, May 2016.

[38] A. Carini, L. Romoli, S. Cecchi, and S. Orcioni, "Perfect periodic sequences for nonlinear Wiener filters," in *24th European Signal Processing Conference (EUSIPCO)*, Aug. 2016, pp. 1788–1792.

[39] A. Carini, S. Orcioni, and S. Cecchi, "Introducing the Orthogonal Periodic Sequences for the Identification of Functional Link Polynomial Filters," in *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019, pp. 5486–5490.

[40] Y. W. Lee and M. Schetzen, "Measurement of the Wiener kernels of a non-linear system by cross-correlation," *International Journal of Control*, vol. 2, no. 3, pp. 237–254, 1965.

[41] M. J. Korenberg, S. B. Bruder, and P. J. McLlroy, "Exact orthogonal kernel estimation from finite data records: Extending Weiner's identification of nonlinear systems," *Annals of Biomedical Engineering*, vol. 16, no. 2, pp. 201–214, 1988.

[42] D. Comminiello and J. C. Príncipe, *Adaptive learning methods for nonlinear system modeling*. Butterworth-Heinemann, 2018.

[43] E. L. O. Batista, O. J. Tobias, and R. Seara, "A sparse-interpolated scheme for implementing adaptive Volterra filters," *IEEE Transactions on Signal Processing*, vol. 58, no. 4, pp. 2022–2035, 2009.

[44] K. Sun, S. Mou, J. Qiu, T. Wang, and H. Gao, "Adaptive fuzzy control for nontriangular structural stochastic switched nonlinear systems with full state constraints," *IEEE Transactions on Fuzzy systems*, vol. 27, no. 8, pp. 1587–1601, 2018.

[45] K. Sun, J. Qiu, H. R. Karimi, and H. Gao, "A novel finite-time control for nonstrict feedback saturated nonlinear systems with tracking error constraint," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–12, 2019.

[46] K. Sun, L. Liu, J. Qiu, and G. Feng, "Fuzzy adaptive finite-time fault-tolerant control for strict-feedback nonlinear systems," *IEEE Transactions on Fuzzy Systems*, 2020. [Online]. Available: https://doi.org/10.1109/TFUZZ.2020.2965890

[47] V. Stojanovic and V. Filipovic, "Adaptive input design for identification of output error model with constrained output," *Circuits, Systems, and Signal Processing*, vol. 33, no. 1, pp. 97–113, 2014.

[48] V. Stojanovic and N. Nedic, "Robust Kalman filtering for nonlinear multivariable stochastic systems in the presence of non-gaussian noise," *International Journal of Robust and Nonlinear Control*, vol. 26, no. 3, pp. 445–460, 2016.

[49] ——, "Robust identification of oe model with constrained output using optimal input design," *Journal of the Franklin Institute*, vol. 353, no. 2, pp. 576–593, 2016.

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

[51] S. Roman, *The Umbral Calculus*. San Diego: Academic Press, 1984.

[52] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in C : the art of scientific computing*. New York, NY, USA: Cambridge University Press, 1995.

[53] R. D. Nowak and B. D. Van Veen, "Random and pseudorandom inputs for Volterra filter identification," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 2124–2135, 1994.

[54] G.-O. Glentis and N. Kalouptsidis, "Efficient algorithms for the solution of block linear systems with Toeplitz entries," *Linear Algebra and its Applications*, vol. 179, pp. 85 – 104, 1993.

[55] V. Filipovic, N. Nedic, and V. Stojanovic, "Robust identification of pneumatic servo actuators in the real situations," *Forschung im Ingenieurwesen*, vol. 75, no. 4, pp. 183–196, 2011.

[56] Y. Guo, L. Guo, S. Billings, D. Coca, and Z. Lang, "Volterra series approximation of a class of nonlinear dynamical systems using the Adomian decomposition method," *Nonlinear Dynamics*, vol. 74, no. 1-2, pp. 359–371, 2013.

[57] Z. Peng and Z. Lang, "On the convergence of the Volterra-series representation of the Duffing's oscillators subjected to harmonic excitations," *Journal of Sound and Vibration*, vol. 305, no. 1-2, pp. 322–332, 2007.

[58] G. Tomlinson, G. Manson, and G. Lee, "A simple criterion for establishing an upper limit to the harmonic excitation level of the Duffing oscillator using the Volterra series," *Journal of Sound and Vibration*, vol. 190, no. 5, pp. 751–762, 1996.

[59] A. Carini, S. Orcioni, A. Terenzi, and S. Cecchi, "Nonlinear System Identification Using Wiener Basis Functions and Multiple-Variance Perfect Sequences," *Signal Processing*, vol. 160, pp. 137–149, feb 2019.

[60] S. Billings and J. Peyton Jones, "Mapping non-linear integro-differential equations into the frequency domain," *International Journal of Control*, vol. 52, no. 4, pp. 863–879, 1990.

[61] A. Chatterjee, "Parameter estimation of Duffing oscillator using Volterra series and multi-tone excitation," *International Journal of Mechanical Sciences*, vol. 52, no. 12, pp. 1716–1722, 2010.

[62] A. Carini, S. Orcioni, and S. Cecchi, "Orthogonal periodic sequences," http://www2.units.it/ipl/res_OPSeqs.htm, 2019, [Online; accessed 2019-11-15].

Alberto Carini (S'89–A'97–M'02–SM'06) received the Laurea degree (summa cum laude) in electronic engineering and the Dottorato di Ricerca (Ph.D.) degree in information engineering from the University of Trieste, Trieste, Italy, in 1994 and 1998, respectively. In 1996 and 1997, during his doctoral studies, he was a Visiting Scholar with the University of Utah, Salt Lake City. From 1997 to 2003, he was a DSP Engineer with Telit Mobile Terminals SpA, Trieste, where he led audio processing R&D activities. In 2003, he was with Neonseven srl, Trieste, as an audio and DSP expert. From 2001 to 2004, he collaborated with the University of Trieste as a Contract Professor of digital signal processing. From 2004 to 2018, he was an Associate Professor with the University of Urbino, Urbino, Italy. From 2018, he has been an Associate Professor with the University of Trieste, Trieste, Italy. His research interests include system identification, adaptive filtering, nonlinear filtering, nonlinear equalization, acoustic echo cancellation, active noise control, room response equalization, and signal processing for road surface monitoring. He is a Member of EURASIP. His professional activities are as follows: Member-at-Large, SPS Conference Board (2009–2010); Member, SPS Signal Processing Theory and Methods Technical Committee (2006–2011); and Program Co-Chair signal processing area, 2009 International Symposium on Image and Signal Processing and Analysis (ISPA) and 2013 International Symposium on ISPA. He is currently an Editorial Board Member of Elsevier Signal Processing and a Member of SPS Signal Processing Theory and Methods Technical Committee. He received the Zoldan Award for the best Laurea degree.

Simone Orcioni (S'89–'95 M'05–SM'18) received the Laurea degree in Electronics Engineering from Università Politecnica delle Marche, Italy, in 1992 and the Ph.D. degree in 1995. From 1997 to 1999 he was Post-Doctoral Fellow at the same university. Since 2000 he is an Assistant Professor, teaching courses in analog and digital electronics, and publishing a text book. In 2017 he was Guest Professor at Ubiquitous Computing Lab (UC-Lab), HTWG Konstanz, and lecturer of a course. He has published more than forty papers in indexed journals. He was Guest Editor for EURASIP Journal on Embedded Systems, and Frontiers in Energy Research, Program Committee Member of International Conferences, Program Chair of three International Conferences, Editor of four International Books, and Inventor in two patents.

He has been working in statistical device modeling and simulation, analog circuit design, cyber-physical system, and nonlinear system identification. His current research interests include circuits and systems for biomedical applications, nonlinear filtering and stochastic system modeling.

Alessandro Terenzi was born in Senigallia, Italy, in 1991. He received the Laurea degree (with honors) in electronic engineering in July 2016 at the Polytechnic University of Marche (Italy). From November 2018, he has been a Ph.D student at DII (Department of Information Engineering) at the same university. His current research interests are in the area of digital signal processing, including nonlinear audio system and audio processing. Dr. Terenzi is a member of the Acoustic Engineering Society (AES).

Stefania Cecchi (M'06) was born in Amandola, Italy, in 1979. She received the Laurea degree (with honors) in electronic engineering from the University of Ancona (now University Politecnica delle Marche, Italy) in 2004 and the Ph.D. degree in electronic engineering from the University Politecnica delle Marche (Ancona, Italy) in 2007. She was a Post Doc Researcher at DII (Department of Information Engineering) at the same university above, from February 2008 to October 2015, and Assistant Professor from November 2015 to October 2018. She is an Associate Professor at the same department since November 2018. She is the author or coauthor of several international papers. Her current research interests are in the area of digital signal processing, including adaptive DSP algorithms and circuits, speech, and audio processing. Dr. Cecchi is a member of the Acoustic Engineering Society (AES), Institute of Electrical and Electronics Engineers (IEEE), Italian Acoustical Association (AIA).