# Introducing Stochastic Functional Link Polynomial Filters

Simone Orcioni[1], Alberto Carini[2], Stefania Cecchi[1] and Massimo Conti[1]
[1]*Department of Information Engineering, Università Politecnica delle Marche* Ancona, Italy
[2]*Department of Engineering and Architecture, University of Trieste* Trieste, Italy
s.orcioni@staff.univpm.it

***Abstract*—The class of Functional Link Polynomials (FLiP) filters is very broad and includes many popular nonlinear filters, as the well-known Volterra and the Wiener nonlinear filters. They are linear in the parameters and can approximate arbitrarily well any discrete-time, time invariant, finite memory, continuous nonlinear system. This work extends the approximation capability of FLiP filters to systems that together with the previous properties have random parameters. This is achieved by extending the FLiP representation to random coefficients and applying the Kosambi–Karhunen–Loève theorem to the coefficients.**

## I. Introduction

Functional link polynomial (FLiP) filters [1], [2] form a very broad class of linear in the parameters (LIP) filters that includes many of the most popular nonlinear filters, like Volterra filters [3], [4] or Wiener nonlinear (WN) filters [4], [5]. It includes also even mirror Fourier nonlinear (EMFN) filters [6], Legendre nonlinear (LN) filters [7], Chebyshev nonlinear (CN) filters [8] and others [2].

FLiP filters consist of linear combinations of products of time-delayed univariate functions, following the constructive rule of Volterra filter. In this way, the resulting basis functions form an algebra that satisfies all conditions of the Stone-Weierstrass theorem [2], allowing FLiP filters to be universal approximators, i.e., they can arbitrarily well approximate any discrete-time, time invariant, finite memory, continuous nonlinear system. Furthermore, the LIP property allows the use of a projection method for the identification, i.e., the output can be projected on each basis function, provided that these functions form an orthogonal set.

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 [1]. In this case, the projection can be performed by estimating the expected value of the product of the filter output and basis functions. If these expected values are computed with time averages, the cross-correlation method results.

As an alternative to the classic cross-correlation methods that use stochastic inputs [9]–[13], a novel family of deterministic periodic sequences, the orthogonal periodic sequences (OPSs) has been recently proposed for identification by means of cross-correlation [14], [15]. They can identify any orthogonal or non-orthogonal FLiP filter, including Volterra filters. Moreover, 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.

The objective of this work is to extend the approximation capabilities of FLiP filters from discrete-time, time invariant, finite memory, continuous nonlinear systems to stochastic ones, where the filter coefficients are stochastic variables whose distribution has to be determined. Nonlinear stochastic filters have been considered only recently in the literature with very few contributions [16], [17]. In these works, Volterra series has been extended to model stochastic systems, by expanding the kernels with random Kautz functions. Monte Carlo simulations and the least-squares method were used to identify the coefficients of the stochastic Kautz functions [16], [17]. In our work, the more general class of FLiP filters is extended to stochastic systems by using linear stochastic process modelling, like Kosambi–Karhunen–Loève transform (KLT). In discrete time domain the KLT matches the principal component analysis (PCA), that performs a transformation of the initial features into an equal number of uncorrelated vectors. This transformation of basis could also be seen as a linear representation of a discrete and finite stochastic process, corresponding to the discrete time case of the KLT transform. PCA [18], [19] is used in a wide field of applications, from classical face recognition and object recognition [20]–[22] to, more recently, ECG beat classification [23], speaker identification [24], and blind image quality assessment [25].

The rest of the paper is organized as follows. The deterministic FLiP filters are reviewed in Section II. The stochastic FLiP filters are introduced in Section III. Section IV presents some experimental results, involving the identification of a cascade system composed of a real nonlinear system (a vacuum tube preamplifier), and a simulated stochastic one. Section V provides the concluding remarks. The following notation is used throughout the paper: $\mathbb{E}\{\cdot\}$ indicates expected value, bold letters are used for vectors and matrices, and Greek letters are used for random quantities.

## II. Deterministic FLiP filter

FLiP filters are universal approximators: they can arbitrarily well approximate any discrete-time, time-invariant, finite memory, continuous nonlinear system,

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

where $f$ is a continuous $M$-dimensional function from $\mathbb{R}_1^M$ to $\mathbb{R}$, $x$ is the system input, with $x(n) \in [-1, +1]$, and $M$ is the length of its memory.

FLiP filters have been proved to be universal approximators [2]. 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, $g_{2i+1}[\xi]$ 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 $2i$.

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 variables 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 [2].

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 input samples involved in basis function product. A FLiP filter of order $P$, memory $M$, diagonal number $D$ (with $D < M$) is the linear combination of all FLiP basis functions, with order, memory, and diagonal number up to $P$, $M$, and $D$, respectively.

For sake of clarity, the FLiP basis functions up to order 3, diagonal number $D$, and memory $M$ are given in Table I. The FLiP filter has $N$ coefficients with

$$N = \binom{D+P+1}{D+1} + \binom{D+P}{D+1}(M-1-D) \tag{3}$$

where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

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 $\sigma_x^2$,

$$\{1, \xi, \xi^2 - \sigma_x^2, \xi^3 - 3\sigma_x^2\xi, \xi^4 - 6\sigma_x^2\xi^2 + 3\sigma_x^4, \dots\} \tag{4}$$

and others, as discussed in [2].

### A. Matrix representation of FLiP filters

The input-output relationship of any FLiP filter of order $K$, with memory of $N$ samples, can be written as

$$y(n) = \sum_{i=0}^{N-1} h_i \ (f_i x)(n) \tag{5}$$

TABLE I
BASIS FUNCTIONS $(f_k x)(n)$ OF FLIP FILTER

Order 0:
$$g_0[x(n)] = 1.$$

Order 1:
$$g_1[x(n)], \dots, g_1[x(n-M+1)].$$

Order 2:
$$\begin{gathered} g_2[x(n)], \dots, g_2[x(n-M+1)], \\ g_1[x(n)]g_1[x(n-1)], \dots, g_1[x(n-M+2)]g_1[x(n-M+1)], \\ g_1[x(n)]g_1[x(n-2)], \dots, g_1[x(n-M+3)]g_1[x(n-M+1)], \\ \vdots \\ g_1[x(n)]g_1[x(n-D)], \dots, g_1[x(n-M+D+1)]g_1[x(n-M+1)], \end{gathered}$$

Order 3:
$$\begin{gathered} g_3[x(n)], \dots, g_3[x(n-M+1)], \\ g_2[x(n)]g_1[x(n-1)], \dots, g_2[x(n-M+2)]g_1[x(n-M+1)], \\ \vdots \\ g_2[x(n)]g_1[x(n-D)], \dots, g_2[x(n-M+D+1)]g_1[x(n-M+1)], \\ g_1[x(n)]g_2[x(n-1)], \dots, g_1[x(n-M+2)]g_2[x(n-M+1)], \\ \vdots \\ g_1[x(n)]g_2[x(n-D)], \dots, g_1[x(n-M+D+1)]g_2[x(n-M+1)], \\ g_1[x(n)]g_1[x(n-1)]g_1[x(n-2)], \dots \\ g_1[x(n-N+3)]g_1[x(n-M+2)]g_1[x(n-M+1)], \\ \vdots \\ g_1[x(n)]g_1[x(n-D+1)]g_1[x(n-D)], \dots \\ g_1[x(n-M+D+1)]g_1[x(n-M+2)]g_1[x(n-M+1)], \end{gathered}$$

where $(f_i x)$, $i = 0, ..., N-1$, are basis functions taken from Table I and derived from multiplication of $g_i(x(n))$ functions. Using vector notation, (5) becomes [2]

$$y(n) = \mathbf{h}^{\mathrm{T}} \mathbf{f}_n \tag{6}$$

$$\mathbf{h} = [h_0, h_1, ..., h_{N-1}]^{\mathrm{T}} \tag{7}$$

$$\mathbf{f}_n = [(f_0 x)(n), (f_1 x)(n), ..., (f_{N-1} x)(n)]^{\mathrm{T}} \ .$$

## III. STOCHASTIC FLIP FILTER

Let us suppose to deal with a discrete-time stochastic system, with finite memory. It can be expressed as

$$\upsilon(\boldsymbol{\gamma}, n) = \varphi(\boldsymbol{\gamma}, x(n), x(n-1), \dots, x(n-M+1)) \tag{8}$$

where $\varphi$ is a nonlinear function, and $\boldsymbol{\gamma}$ is a vector of an unknown number of normal random variables (r.v.).

The stochastic equivalent of (6) as a stochastic FLiP filter can be written as

$$\upsilon(\boldsymbol{\gamma}, n) = \boldsymbol{\eta}(\boldsymbol{\gamma})^{\mathrm{T}} \boldsymbol{f}_n \tag{9}$$

where the stochastic vector $\boldsymbol{\eta}(\boldsymbol{\gamma})$ has the vector $\mathbf{h}$ in (6) as its realizations.

### A. Representation of stochastic coefficients with Kosambi–Karhunen–Loève Transform

If the stochastic process $\boldsymbol{\eta}(\boldsymbol{\gamma})$ is linear in the stochastic parameters $\boldsymbol{\gamma}$, it can be represented with the Kosambi–Karhunen–Loève Transform (KLT). In the following the dependence of $\boldsymbol{\eta}$ from $\boldsymbol{\gamma}$ will be omitted for simplicity. Since the stochastic process $\boldsymbol{\eta} = \boldsymbol{\eta}(\boldsymbol{\gamma}) = [\eta_1, \dots, \eta_N]^{\mathrm{T}}$, is discrete and finite, the KLT series has a finite number of terms and it coincides with PCA

$$\boldsymbol{\eta} = \mathbb{E}\{\boldsymbol{\eta}\} + \sum_{i=1}^{N} \kappa_i \boldsymbol{w}_i \tag{10}$$

where $\kappa_i$ and $\boldsymbol{w}_i$ can be obtained as it follows.

Data for the PCA representation can be obtained by identifying many realizations of the stochastic system. The PCA transforms the set of initial features so obtained into an equal number of uncorrelated vectors. These new vectors are obtained by means of projections of the original features on a new orthogonal basis that represents the directions of maximum variance of the data. This basis can be obtained as the eigenvectors of the covariance matrix of the data.

Let us be

$$\boldsymbol{C}_{\boldsymbol{\eta\eta}} = \mathbb{E}\{(\boldsymbol{\eta} - \boldsymbol{h}_e)(\boldsymbol{\eta} - \boldsymbol{h}_e)\} \tag{11}$$

the covariance matrix of the random vector $\boldsymbol{\eta}$, and $\boldsymbol{h}_e = \mathbb{E}\{\boldsymbol{\eta}\}$ its expected value. Being $\boldsymbol{C}_{\boldsymbol{\eta\eta}}$ a positive definite symmetric matrix it possesses a spectral representation

$$\boldsymbol{C}_{\boldsymbol{\eta\eta}} = \boldsymbol{W}\boldsymbol{L}\boldsymbol{W}^{\mathrm{T}} \tag{12}$$

where $\boldsymbol{L}$ is a diagonal matrix whose entries, $l_i$, are the eigenvalues of $\boldsymbol{C}_{\boldsymbol{\eta\eta}}$; whilst the unitary matrix $\boldsymbol{W}$ contains the eigenvectors

$$\boldsymbol{W} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_N] \ . \tag{13}$$

If the joint probability density function (pdf) of the source variables $\boldsymbol{\eta}$ is normal, as assumed in the following, this transformation decorrelates them, so that the pdf of the transformed variables is also normal with a diagonal covariance matrix, and the eigenvalues $l_i$ are the variances of $\kappa_i$. If the pdf of the original variables is not normal, the pdf of the transformed variable can anyway be represented as a Gaussian mixture model and its parameters estimated for examples with the expectation maximization algorithm [26], [27].

We can represent $\boldsymbol{\eta}$ as a linear combination of latent variables

$$\boldsymbol{\eta} - \boldsymbol{h}_e = \sum_{i=1}^{N} \boldsymbol{w}_i^{\mathrm{T}} \boldsymbol{\eta} \ \boldsymbol{w}_i = \sum_{i=1}^{N} \kappa_i \boldsymbol{w}_i = \sum_{i=1}^{N} \sqrt{l_i} \nu_i \boldsymbol{w}_i \tag{14}$$

where the random variables $\kappa_i$ are the projection of $\boldsymbol{\eta}$ on the eigenvectors, the so called principal directions, and can be written as the product of their standard deviation $\sqrt{l_i}$ with a unitary normal r.v. $\nu_i$.

The number of latent variables used to express the stochastic process can be reduced by using only the first $K$ eigenvectors, ordered according to their eigenvalues, resulting in

$$\boldsymbol{W}_K = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K] \tag{15}$$

$$\boldsymbol{\eta} = \boldsymbol{h}_e + \sum_{i=1}^{K} \sqrt{l_i} \nu_i \boldsymbol{w}_i \tag{16}$$

$$\frac{\sum_{i=1}^{K} l_i}{\sum_{i=1}^{M} l_i} \geq c \tag{17}$$

where $c$ is the fraction of the output variance of stochastic process we want to take into account.

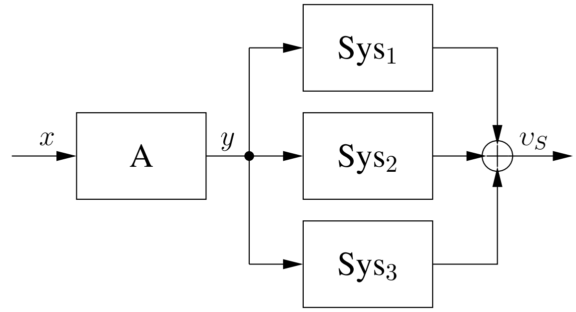


Fig. 1. The stochastic system.

TABLE II
PARAMETERS OF STOCHASTIC SYSTEMS $\mathrm{SYS}_i$.

| $\mathbb{E}\{\alpha_i\}$ | $\sigma_{\alpha_i}$ | $t_i$ |
|---|---|---|
| $-2$ | 3 | 0 |
| 3 | 1 | $2T_\mathrm{S}$ |
| 0 | 2 | $6T_\mathrm{S}$ |

Eventually, using the complexity reduction (9) becomes

$$\upsilon_K(\boldsymbol{\gamma}, n) = \boldsymbol{h}_e^{\mathrm{T}} \boldsymbol{f}_n + \sum_{i=1}^{K} \sqrt{l_i} \nu_i \boldsymbol{w}_i^{\mathrm{T}} \boldsymbol{f}_n. \tag{18}$$

## IV. Experimental results

In this section, we want to test the ability of an estimated stochastic FLiP model to produce the same output of a given stochastic system, related to a real vacuum tube preamplifier for audio applications. To this end, since we will consider a normal stochastic process that is completely defined by the second order moment, the goodness of approximation will be tested by comparing the covariance matrices.

We consider the continuous time stochastic system of Fig. 1 consisting of a non-linear deterministic system $A$ followed by three linear stochastic systems, $\mathrm{Sys}_i$ with $i = 1, \ldots, 3$. The output $y$ of the system $A$ is the input to three different systems, assumed linear with respect to the input and the random parameters for the sake of simplicity.

The overall output is the sum of the three stochastic terms

$$\upsilon_S(t) = \sum_{i=1}^{3} \alpha_i y(t - t_i) \tag{19}$$

where $t \in \mathbb{R}$, $\alpha_i$ and $t_i$ are the parameters of each system $\mathrm{Sys}_i$. $\alpha_i$ are normal r.v. with mean values and standard deviations reported in Tab. II and $t_i$ are deterministic delays also reported in Tab. II and taken for simplicity multiple of the sampling period $T_s$. This system is representative of many real systems since it represents a mixing of different delay paths with random gains.

The system $A$ is a real vacuum tube preamplifier, the Behringer MIC100. To show the connection between the deterministic system model and the stochastic system model, the MIC100 was identified in discrete time domain with a deterministic Volterra filter with $P = 2$, $M = 25$, and $D = 2$,

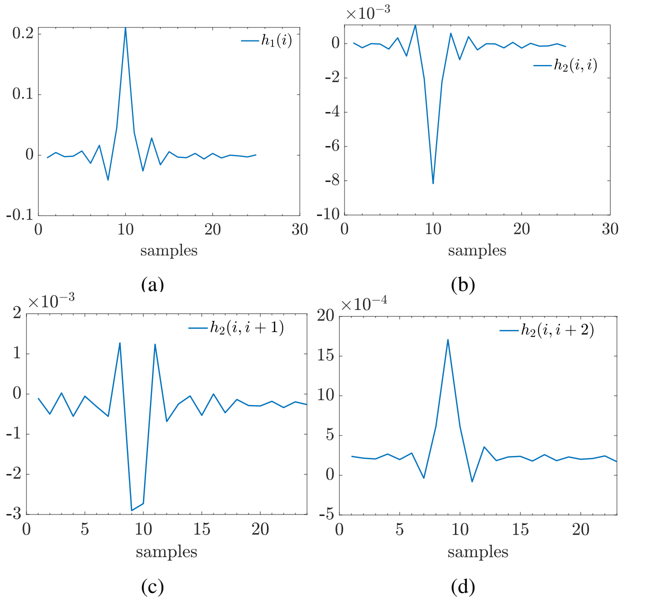


Fig. 2. Volterra kernels of the Behringer MIC100 amplifier; (a) linear kernel (b) main diagonal of second order kernel (c), (d) other diagonals of second order kernel.

$$y_P(n) = h_0 + \sum_{i=0}^{M-1} h_1(i)x(n-i) + \sum_{d=0}^{D} \sum_{i=1}^{M-1-d} h_2(i,i+d)x(n-i)x(n-i+d). \quad (20)$$

The identification was performed with a 16 kHz sampling frequency, using an OPS with a period of 65 536 samples [15]. The deterministic kernels of the amplifier are shown in Fig. 2. Defining the Normalized Mean Square Error, between the output of a system $\boldsymbol{y}_S$ and its model $\boldsymbol{y}_M$, as

$$\text{NMSE} = \frac{(\boldsymbol{y}_S - \boldsymbol{y}_M)^{\mathrm{T}}(\boldsymbol{y}_S - \boldsymbol{y}_M)}{\boldsymbol{y}_S^{\mathrm{T}}\boldsymbol{y}_S}, \quad (21)$$

the NMSE of the Volterra model of the deterministic Behringer MIC100 was equal to $0.72\%$.

The amplifier output corresponding to this OPS was then used as input to the systems $\text{Sys}_i$ to obtain $1000$ realizations of the overall output $\upsilon_S$ with a Monte Carlo method. These realizations are then used to obtain the stochastic Volterra model (18), using the procedure described in the previous section. The order, memory, and diagonal number of the stochastic model were chosen in accordance with the MIC100 model, adopting $P = 2$, $M = 25$, and $D = 2$. Each realization $\boldsymbol{h}$ of the stochastic kernel $\boldsymbol{\eta}(\boldsymbol{\gamma})$ is organized as $\boldsymbol{h} = [h_0, .., h_1(i), .., h_2(i,i), .., h_2(i,i+1), .., h_2(i,i+2), ..]^{\mathrm{T}}$. Following the section III-A these realizations were used to build the stochastic Volterra model in (18), whose kernels $\boldsymbol{h_e}$ and $\boldsymbol{w}_i$, for $i = 1, 2, 3$ are shown in Fig. 3. Only 3 covariance eigenvalues were different from zero, as could be expected from the 3 independent r.v. present in the system. The vector of standard deviations of the normal variables associated with each eigenvector is $\sigma_\nu = [0.6772, 0.4557, 0.2234]^{\mathrm{T}}$.

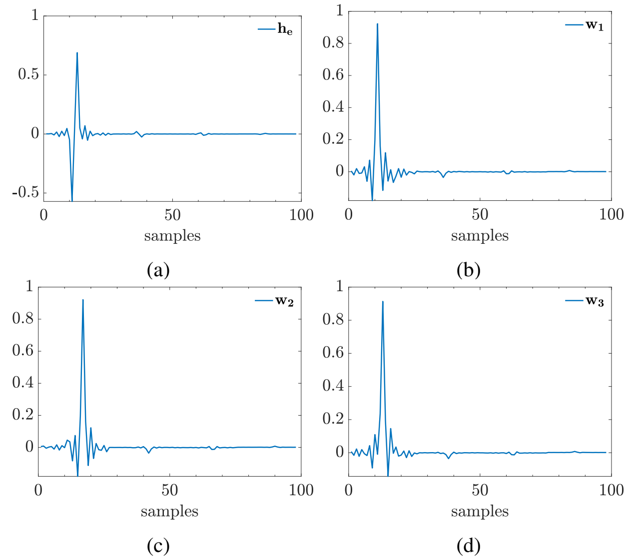


Fig. 3. Vectors composing the stochastic Volterra model of system in Fig. 1; (a) expected value $\boldsymbol{h_e}$ (b) $\boldsymbol{w}_1$, (c) $\boldsymbol{w}_2$, (d) $\boldsymbol{w}_3$.

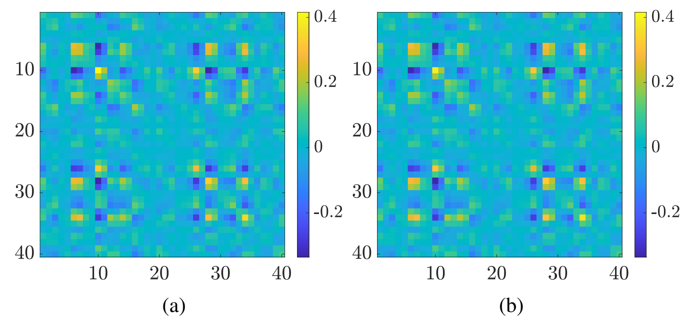


Fig. 4. Covariance matrix of a short segment of $\upsilon_S$ (a) obtained from the system Fig. 1, (b) obtained from the stochastic Volterra model.

To test the stochastic model, a white uniform noise in the range $[-1, 1]$ of $200\,000$ samples has been applied to the system in Fig. 1 and to the stochastic Volterra model.

The covariance matrices of a short segment of the outputs so obtained are calculated and represented in Fig. 4. The two covariance matrices are in excellent agreement and the NMSE between the two is equal to $0.92\%$, thus showing the convergence in distribution between the stochastic Volterra model and the system in Fig. 1.

## V. Conclusion

In this work, the class of FLiP filter is extended to model systems with stochastic parameters by using the KLT transform applied to the FLiP filter coefficients. In the experimental results, a system consisting of a real part (an amplifier) and a simulated part (a mixing of delay paths with random gain) has been identified with a stochastic Volterra filter. The covariance matrices of the outputs of the system and of the model show an excellent agreement, proving the convergence in distribution between the model and the system. In future works, a technique to model stochastic processes that non-linearly depend on random parameters will be introduced.

## 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," in *Adaptive Learning Methods for Nonlinear System Modeling*, D. Comminello and J. C. Príncipe, Eds. Elsevier, 2018, ch. 2, pp. 15–46.

[3] V. Volterra, *Theory of Functionals and of Integrals and Integro-Differential Equations*. London and Glasgow: Blackie and Son Limited, 1930.

[4] V. J. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*. New York: John Wiley and Sons, 2000.

[5] M. Schetzen, "Nonlinear system modeling based on the Wiener theory," *Proceedings of the IEEE*, vol. 69, no. 12, pp. 1557–1573, 1981.

[6] A. Carini and G. L. Sicuranza, "Fourier nonlinear filters," *Signal Processing*, vol. 94, pp. 183–194, Jan. 2014.

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

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

[9] 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.

[10] P. Koukoulas and N. Kalouptsidis, "Nonlinear system identification using Gaussian inputs," *IEEE Transactions on Signal Processing*, vol. 43, no. 43, pp. 1831–1841, 1995.

[11] M. Pirani, S. Orcioni, and C. Turchetti, "Diagonal kernel point estimation of nth-order discrete Volterra-Wiener systems," *EURASIP J. Adv. Signal Process.*, vol. 2004, no. 12, pp. 1807–1816, Sep. 2004.

[12] 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.

[13] 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.

[14] 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 Int. Conf. Acoust. Speech Signal Process.* IEEE, May 2019, pp. 5486–5490.

[15] A. Carini, S. Orcioni, A. Terenzi, and S. Cecchi, "Orthogonal periodic sequences for the identification of functional link polynomial filters," *IEEE Trans. Signal Process.*, no. c, pp. 1–1, 2020.

[16] L. G. Villani, S. da Silva, and A. Cunha, "Damage detection in uncertain nonlinear systems based on stochastic Volterra series," *Mech. Syst. Signal Process.*, vol. 125, pp. 288–310, 2019.

[17] L. G. G. Villani, S. da Silva, and A. Cunha, "An optimizationless stochastic Volterra series approach for nonlinear model identification," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 44, no. 6, p. 260, jun 2022.

[18] H. Hotelling, "Analysis of a complex of statistical variables into principal components." *Journal of Educational Psychology*, vol. 24, no. 6, pp. 417–441, sep 1933.

[19] I. T. Jolliffe, *Principal Component Analysis*, ser. Springer Series in Statistics. New York: Springer-Verlag, 2002.

[20] M. Kirby and L. Sirovich, "Application of the Karhunen-Loève procedure for the characterization of human faces," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 103–108, Jan. 1990.

[21] B. Moghaddam and A. Pentland, "Probabilistic visual learning for object representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 696–710, Jul. 1997.

[22] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[23] G. Biagetti, P. Crippa, A. Curzi, S. Orcioni, and C. Turchetti, "A multi-class ECG beat classifier based on the truncated KLT representation," in *Proceedings of the 2014 UKSim-AMSS 8th European Modelling Symposium (EMS 2014)*. IEEE Computer Society, 2014, pp. 93–98.

[24] ——, "Speaker identification with short sequences of speech frames," in *Proceedings of the 4th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2015)*, vol. 2. SCITEPRESS (Science and Technology Publications,Lda.), 2015, pp. 178–185.

[25] C. Yang, X. Zhang, P. An, L. Shen, and C.-C. J. Kuo, "Blind image quality assessment based on multi-scale KLT," *IEEE Transactions on Multimedia*, vol. 23, pp. 1557–1566, 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9113755/

[26] M. Figueiredo and A. Jain, "Unsupervised learning of finite mixture models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 3, pp. 381–396, mar 2002.

[27] G. Biagetti, P. Crippa, L. Falaschetti, S. Orcioni, and C. Turchetti, "An investigation on the accuracy of truncated DKLT representation for speaker identification with short sequences of speech frames," *IEEE Transactions on Cybernetics*, vol. 47, no. 12, pp. 4235–4249, 2017.