

PERFECT PERIODIC SEQUENCES FOR LEGENDRE NONLINEAR FILTERS

Alberto Carini*

Stefania Cecchi[†]

Laura Romoli[†]

Giovanni L. Sicuranza[‡]

* DiSBeF - University of Urbino - Italy

[†] DII - Università Politecnica delle Marche - Italy

[‡] DIA - University of Trieste - Italy

ABSTRACT

The paper shows that perfect periodic sequences can be developed and used for the identification of Legendre nonlinear filters, a sub-class of linear-in-the-parameters nonlinear filters recently introduced in the literature. A periodic sequence is perfect for the identification of a nonlinear filter if all cross-correlations between two different basis functions, estimated over a period, are zero. Using perfect periodic sequences as input signals, the unknown nonlinear system and its most relevant basis functions can be identified with the cross-correlation method. The effectiveness and efficiency of this approach is illustrated with experimental results involving a real nonlinear system.

Index Terms— Nonlinear system identification, linear-in-the-parameters nonlinear filters, Legendre nonlinear filters, perfect periodic sequences, cross-correlation method

1. INTRODUCTION

Legendre nonlinear (LN) filters have been recently introduced in the literature [1]. They belong to the class of linear-in-the-parameters (LIP) nonlinear filters, which is characterized by the property that the filter output depends linearly on the filter coefficients. The class comprises many popular finite-memory and infinite-memory nonlinear filters, including the well known truncated Volterra filters [2]. LN filters share many of the characteristics of Volterra filters. Indeed, they are polynomial filters whose basis functions are product of Legendre polynomials of the input signal samples. They include a linear term formed by the first order basis functions. The LN filter basis functions form an algebra that satisfies all the requirements of the Stone-Weierstrass approximation theorem [3]. Consequently, these filters can arbitrarily well approximate any causal, time invariant, finite-memory, continuous, nonlinear system, as well as the Volterra filters. Moreover, LN filters present further interesting properties originated by the orthogonality of Legendre polynomials. Specifically, the basis functions are mutually orthogonal for white uniform input signals in $[-1, +1]$. This property is particularly appealing since it allows the derivation of gradient de-

scent algorithms with fast convergence speed and of efficient identification algorithms. Since the LN filter basis functions are orthogonal, the expansion of a nonlinear function using their basis functions is a generalized Fourier series [4]. Other LIP nonlinear filters present similar orthogonality properties. In particular they can be found in even mirror Fourier nonlinear (EMFN) filters. These filters derive from the truncation of a multidimensional Fourier series expansion of an even mirror periodic repetition of the nonlinear function we want to approximate. They are based on trigonometric basis functions that satisfies the requirements of the Stone-Weierstrass theorem and thus, they can also arbitrarily well approximate any causal, time invariant, finite-memory, continuous, nonlinear system. In terms of modelling performance, it has been shown that EMFN filters can often model strong nonlinearities better than Volterra filters [5]. However, in contrast to Volterra and LN filters, EMFN filters do not have a linear term among the basis functions. Thus, for weak or medium nonlinearities Volterra and LN filters should be preferred [1].

In [6], it was shown that perfect periodic sequences (PPSs) can be developed for the efficient identification of EMFN filters. PPSs have been extensively studied and proposed as inputs for linear system identification [7] and in this context they have found application in signal processing [8], information theory [9], communications [10], and acoustics [11]. A periodic sequence is called perfect for a certain modeling filter if all cross-correlations between two of its basis functions, estimated over a period, are zero. By applying a PPS as input signal, it is possible to model an unknown system exploiting the cross-correlation method, i.e., computing the cross-correlation between the basis functions and the system output. The most relevant basis functions, i.e., those that guarantee the most compact representation of the nonlinear system according to some information criterion, can also be easily estimated. In [6], the existence of PPSs for EMFN filters was conjectured from the orthogonality property of the basis functions for white uniform input signals in $[-1, +1]$. LN filters satisfy the same property and in this paper we show that PPSs can be developed also for their efficient identification. The resulting approach based on PPSs and cross-correlation is one of the most efficient identification methods for polynomial systems.

This work was supported in part by DiSBeF Research Grant.

The paper is organized as follows. LN filters are reviewed in Section 2. PPSs for LN filters are introduced in Section 3. Nonlinear systems identification using PPS is discussed in Section 4. Experimental results about the construction of PPSs for LN filters and about the identification of a real nonlinear system are presented in Section 5. Concluding remarks are given in Section 6.

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

2. LEGENDRE NONLINEAR FILTERS

Differently from [1], in this paper we introduce the LN filters starting from a normalized set of Legendre polynomials. This set is obtained by applying the following recursive relation

$$L_{i+1}(\xi) = \frac{2i+1}{i+1} \sqrt{\frac{2i+3}{2i+1}} x L_i(\xi) - \frac{i}{i+1} \sqrt{\frac{2i+3}{2i-1}} L_{i-1}(\xi) \quad (1)$$

with $L_0(\xi) = 1$ and $L_1(\xi) = \sqrt{3}\xi$. With this choice, $L_i(\xi) = \sqrt{2i+1} \text{Leg}_i(\xi)$, with $\text{Leg}_i(\xi)$ the i -th Legendre polynomial. The set of polynomials $L_i(x)$ is orthogonal in \mathbb{R}_1 , since

$$\int_{-1}^{+1} L_i(x) L_j(x) dx = 2\delta_{ij}, \quad (2)$$

and can be used to arbitrarily well approximate any continuous function $f(\xi)$ defined in \mathbb{R}_1 .

Let us assume we want to approximate the input-output relationship of a discrete-time, time-invariant, finite-memory, causal, continuous, nonlinear system given by

$$y(n) = f[x(n), x(n-1), \dots, x(n-N+1)], \quad (3)$$

where f is a real N -dimensional continuous function and $x(n)$ belongs to \mathbb{R}_1 . According to [1], a set of basis functions that satisfy all the requirements of Stone-Weierstrass theorem and can arbitrarily well approximate (3) is formed by writing the polynomials $L_i(\xi)$ for $\xi = x(n)$, $\xi = x(n-1)$, \dots , $\xi = x(n-N+1)$, and by multiplying in any possible manner the polynomials of different variable, taking care of avoiding repetitions. The order of each N -dimensional basis function is then defined as the sum of the orders of the constituent Legendre polynomials. These LN basis functions are orthogonal and have unit power (i.e., are orthonormal) for white uniform inputs in \mathbb{R}_1 . Thus, we can easily find an unbiased estimate for the coefficients of the LN filter approximating (3) using the cross-correlation method [12]. In fact, the coefficient g_l of the basis function $f_l(n)$ is given by

$$g_l = E[f_l(n)y(n)], \quad (4)$$

where the expectation can be estimated using time averages.

The LN basis functions of order 1, 2, 3 are shown in Table 1, while the basis function of order 0 is equal to 1. An LN filter of order K and memory N is the linear combination of

Table 1. Basis functions of LN nonlinear filters

Order 1:

$$L_1[x(n)], \dots, L_1[x(n-N+1)]$$

Order 2:

$$\begin{aligned} &L_2[x(n)], \dots, L_2[x(n-N+1)] \\ &L_1[x(n)]L_1[x(n-1)], \dots, \\ &\quad L_1[x(n-N+2)]L_1[x(n-N+1)] \\ &L_1[x(n)]L_1[x(n-2)], \dots, \\ &\quad L_1[x(n-N+3)]L_1[x(n-N+1)], \\ &\quad \vdots \\ &L_1[x(n)]L_1[x(n-N+1)]. \end{aligned}$$

Order 3:

$$\begin{aligned} &L_3[x(n)], \dots, L_3[x(n-N+1)] \\ &L_2[x(n)]L_1[x(n-1)], \dots, \\ &\quad L_2[x(n-N+2)]L_1[x(n-N+1)], \\ &\quad \vdots \\ &L_2[x(n)]L_1[x(n-N+1)], \\ &L_1[x(n)]L_2[x(n-1)], \dots, \\ &\quad L_1[x(n-N+2)]L_2[x(n-N+1)], \\ &\quad \vdots \\ &L_1[x(n)]L_2[x(n-N+1)], \\ &L_1[x(n)]L_1[x(n-1)]L_1[x(n-2)], \dots \\ &\quad \vdots \\ &L_1[x(n)]L_1[x(n-N+2)]L_1[x(n-N+1)]. \end{aligned}$$

$$\mathcal{N}_T(K, N) = \binom{N+K}{N} \quad (5)$$

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

To obtain a reasonable estimate for the coefficients applying the cross-correlation technique to stochastic inputs a huge number (millions) of samples is needed [12, page 77]. To overcome this problem, in the next section we introduce PPSs for LN filters, i.e., periodic sequences that guarantee the orthogonality of the basis functions on a finite time interval. Using these sequences, it is possible to obtain an exact estimate of the coefficients of the LN filter applying again (4), with the expectations replaced by time averages on one or a few periods of the PPS.

3. PPS FOR LN FILTERS

The development of PPSs for LN filters follows the same approach of [6].

Let us consider a sequence x_0, x_1, \dots, x_{L-1} of period L . Such a sequence is perfect for an LN filter of order K and

memory N if all cross-correlations between two different basis functions, estimated over a period, are zero, i.e., if

$$\langle f_l(n) \cdot f_m(n) \rangle_L = 0, \quad (6)$$

for all $f_l(n) \in S_{f,n}(K, N)$, $f_m(n) \in S_f(K, N)$ with $f_l(n) \neq f_m(n)$. Together with the conditions in (6) it is convenient to impose

$$\langle f_l(n) \cdot f_l(n) \rangle_L = 1, \quad (7)$$

for all $f_l(n) \in S_{f,n}(K, N)$ and $f_l(n) \neq 1$. The system of nonlinear equations defined in (6) and (7) is equivalent to the following simpler system

$$\langle f_l(n) \rangle_L = 0, \quad (8)$$

for all $f_l(n) \in S_{f,n}(2K, N)$. In fact, the product of two basis functions of order k and h , respectively, can be expanded in the sum of basis functions of maximum order $k+h$ [13]. Each basis function in (8) appears in the expansion of at least one of the products in (6). Moreover, imposing (8), both (6) and (7) are satisfied. Indeed, if we expand the products in (6), we find a linear combination of basis functions different from $f_1(n) = 1$, while if we expand the terms in (7) we find $f_1(n)$ plus a linear combination of other basis functions.

The system in (8) has $Q = N_T(2K - 1, N)$ equations in the L variables x_0, x_1, \dots, x_{L-1} . For sufficiently large L , it is an under-determined system of nonlinear equations that may have infinite solutions. In our experiments, it has always been possible to find a solution for it. For this purpose, any algorithm for solving nonlinear equation systems can be used. We found particularly effective the Newton-Raphson method, implemented as described in [14, ch. 9.7], with the only modification of reflecting the variables x_0, x_1, \dots, x_{L-1} in \mathbb{R}_1 when they exceeded the range. In our simulation, the iterations started from a random distribution of x_0, x_1, \dots, x_{L-1} in \mathbb{R}_1 and the Jacobian matrix was computed analytically. Employing numerical methods, only an approximate solution can be obtained. Nevertheless, the cross-correlations between basis functions can be made as small as desired, selecting an appropriate precision in the stop-condition of the Newton-Raphson method. The number of iterations necessary for the Newton-Raphson to converge depends on the selected precision and on the ratio L/Q .

The number of equations Q in (8) increases exponentially with the order K and geometrically with the memory N . Also for low orders and memory lengths Q can be unacceptably large. We can reduce the number of equations and variables imposing a specific structure to the periodic sequence. For example, the following conditions allow to almost halve the number of equations and variables:

- 1) Symmetry: if the PPS is formed with the terms a_1, a_2, \dots, a_M and reversed ones a_M, a_{M-1}, \dots, a_1 , then for any couple of symmetric basis functions, only one has to be considered.
- 2) Oddness: if the PPS is formed with the terms a_1, a_2, \dots, a_M and the negated ones $-a_1, -a_2, \dots, -a_M$, then all odd basis

functions have a priori zero average.

- 3) Oddness-1: if the PPS is formed with the terms a_1, a_2, \dots, a_M and those obtained by alternatively negating one every two terms, $a_1, -a_2, a_3, -a_4, \dots, -a_M$, then all Odd-1 functions have a priori zero average.

We define Odd-1 all those basis functions that change their sign by alternatively negating one every two sample, e.g., $L_1[x(n)]L_1[x(n-1)]$.

Similarly it is possible to impose a priori a zero average of Odd-2, Odd-4, \dots , functions by alternatively negating two every four samples, four every eight samples, and so on. Two or more conditions can also be considered together. The reduction in the number of equations is paid with a longer period of the resulting PPS but it is often determinant to solve the system in (8). Indeed, the Newton-Raphson algorithm has memory and processing time requirements that grow with Q^3 . The method is effective only for not too large orders K and memory lengths N . Another strategy to tackle the computational complexity of the system in (8) for large orders and memory lengths is to resort to simplified models, as done for Volterra filters in [15].

4. IDENTIFICATION USING PPS

In this Section we describe how PPSs can be used to identify a time-invariant, finite-memory, causal, continuous, nonlinear system. Let us assume that the input-output relationship of the nonlinear system is expressed as a linear combination of LN basis functions up to order K and memory of N samples,

$$y(n) = \sum_l g_l f_l(n). \quad (9)$$

Using a PPS input, the coefficients g_l can be estimated by computing the cross-correlation between the output of the system and each basis function over a period mL , where $m \in \mathbb{N}^+$ and L is the PPS period,

$$\hat{g}_l = \langle f_l(n)y(n) \rangle_{mL}. \quad (10)$$

When the input-output relationship in (3) is a linear combination of LN basis functions with memory N and maximum order greater than K , using a PPS for LN filters of order K and memory N the identification is affected by an error. It is possible to prove that this error affects mainly the coefficients of the higher-order basis functions, while, in general, it has only a marginal influence on the coefficients of the lower-order basis functions. The identification is also affected when the input-output relationship of the system to be identified is a linear combination of LN basis functions with order K but memory greater than N . In this case, the error affects mainly the coefficients of basis functions associated with the most recent samples $x(n), x(n-1), \dots$, while, in general, the coefficients of basis functions associated with less recent samples $x(n-N+1), x(n-N+2), \dots$ are only marginally affected. The proofs will be included in a journal paper in preparation.

Since the basis functions are orthogonal on a PPS period, they can be easily ranked according to the mean square error

Table 2. Results of Newton-Raphson method

Seq.	Q	M	L	Iter.	Max XC
1	3003	12012	12012	55	8.1 E-16
2	2232	8928	17856	42	9.1 E-16
3	1567	6268	12536	36	6.2 E-16
4	1116	4465	17860	57	3.8 E-16
5	593	2373	18980	48	5.2 E-16

(MSE) they produce, which for the l -th basis function is

$$\delta\text{MSE}_l = \langle f_l(n)y(n) \rangle_{mL}^2. \quad (11)$$

To obtain a compact representation for the nonlinear system, (10) and (11) can be combined with the minimization of an information criterion. Common criteria, exploited in the experiments of the next Section, are the Akaike's information criterion (AIC) [16], the Final Prediction Error (FPE) [16], the Khundrin's law of iterated logarithm criterion (LILC) [17], and the Bayesian information criterion (BIC) [18].

5. EXPERIMENTAL RESULTS

The experimental results aim to illustrate the generation of PPSs and their ability to identify real nonlinear systems.

In the first experiment, we show how the Newton-Raphson method can solve the system in (8). In particular, we develop PPSs for an order 3, memory 10 LN filter (i.e., with 286 coefficients), considering as stop condition the maximum absolute value of the averages of the basis functions over a period to be less than 10^{-15} . Table 2 summarizes the results obtained applying the Newton-Raphson method to the full system in (8) (Seq. 1) and to the reduced systems obtained exploiting the sequence oddness (Seq. 2), symmetry (Seq. 3), oddness and oddness-1 (Seq. 4), oddness, oddness-1, and symmetry (Seq. 5). Table 2 provides the number of equations Q of the system in (8), the number of independent variables M , the period L of the sequence, the number of iterations (Iter.) necessary for the Newton-Raphson method to converge, and the maximum cross-correlation (Max XC) between the basis functions of the resulting sequence. We can notice that by considering $M \simeq 4Q$, the Newton-Raphson method converges within $48 \div 57$ iterations. In our simulations, we have found that $M \simeq 3Q \div 4Q$ is generally sufficient to obtain convergence within a reasonable number of iterations (i.e., in less than 100 iterations). It should be noted that the larger is the ratio M/Q , the smaller is the number of iterations needed to the Newton-Raphson method to converge.

In the second experiment, we consider the identification of an audiophile vacuum tube preamplifier, Behringer Tube Ultrgain Mic 100. With the selected settings, the preamplifier introduces, on a sinusoidal input at 1 kHz, a second and third order harmonic distortion of 8.7% and 4.2%, respectively. The harmonic distortion is defined as the ratio, in percent, between the magnitude of each harmonic and that of the fundamental frequency. At 8 kHz sampling frequency, the

Table 3. Results of identification of Behringer MIC100.

Filter	Information Criterion	Selected Bases	MSE
LN	AIC(4)	433	$7.49E - 4$
	FPE	681	$7.47E - 4$
	LILC	356	$7.49E - 4$
	BIC	182	$7.52E - 4$
Volterra	AIC(4)	434	$7.49E - 4$
	FPE	682	$7.47E - 4$
	LILC	357	$7.49E - 4$
	BIC	183	$7.52E - 4$
EMFN	AIC(4)	430	$8.85E - 4$
	FPE	654	$8.83E - 4$
	LILC	351	$8.86E - 4$
	BIC	164	$8.89E - 4$

system has a memory length lower than 20 samples. Thus, a PPS for a LN of order 3, memory 20, exploiting oddness, oddness-1, and symmetry, and with period of $L = 357956$ samples has been fed to the preamplifier input and the corresponding output has been recorded with a notebook. Table 3 shows the number of terms selected by the AIC (with parameter 4), FPE, LILC, and BIC information criteria, and the corresponding MSE for (i) the LN filter, estimated on a period L with the cross-correlation method of (10), (ii) a Volterra filter and (iii) an EMFN filter, estimated with the method of [19] on the same data. For a linear filter of memory 20 the MSE is $1.63E - 3$. LN and Volterra filters provide almost identical results. The basis functions of the LN filter are a linear combination of those of Volterra filter, and viceversa. Thus, both filters provide the same MSE, and there is only a little difference in the number of basis functions selected by the information criteria. In contrast, the EMFN filter gives slightly worse results because it lacks a linear term and the preamplifier in this case introduces a mild nonlinearity. The main advantage of the proposed method is the remarkable computational efficiency. The method of [19] has been chosen for comparison purposes since it is one of the most computationally efficient identification methods for LIP nonlinear systems available in the literature. Nevertheless, the computational cost of the method of [19] is of order TBS^2 operations, i.e., multiplications and additions, with T the number of samples used for the identification, B the number of candidate basis functions, and S the number of selected basis functions. In contrast, the cross-correlation method has a computational cost of only TB operations. In our experiment, while the execution of the cross-correlation required a processing time of few minutes, the method of [19] requested hours of simulation.

Figure 1 shows the order and the diagonal number of the first 400 selected basis functions for LN and EMFN filters (those for the Volterra filter are almost identical to the LN filter ones). By definition the "diagonal number" of a basis function is the maximum time difference between the samples involved in its expression ($L[x(n)]L[x(n-3)]$) has diagonal

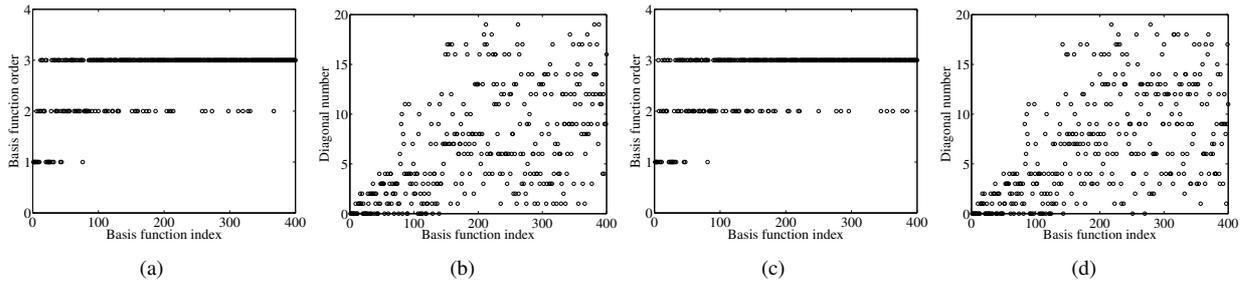


Fig. 1. Order and diagonal number of the first 400 selected basis functions for (a)-(b) the LN filter, (c)-(d) the EMFN filter.

number 3). We can see that low diagonal numbers are selected in the first terms. Thus, if a very compact representation is desired, the system could be modelled with a simplified LN, EMFN or Volterra filter with maximum diagonal number 5.

6. CONCLUSIONS

It has been shown that periodic sequences that guarantee perfect orthogonality of LN filter basis functions on a finite period can be developed. Using the cross-correlation approach, with PPSs as input signals, is one of the most efficient identification methods for nonlinear systems. A compact representation can also be identified by ranking the basis functions according to the MSE reduction they produce. Experimental results involving identification of a real nonlinear system testify the effectiveness of the approach.

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

REFERENCES

- [1] A. Carini, S. Cecchi, M. Gasparini, and G. L. Sicuranza, "Introducing Legendre nonlinear filters," in *Proc. of ICASSP 2014*, Florence, Italy, May 2014.
- [2] V. J. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*, Wiley, New York, 2000.
- [3] W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill, New York, 1976.
- [4] R. L. Herman, *An Introduction to Fourier and complex analysis with applications to the spectral analysis of signals*, 5th edition, Apr. 2012.
- [5] A. Carini and G. L. Sicuranza, "Fourier nonlinear filters," *Signal Processing*, vol. 94, pp. 183–194, 2014.
- [6] A. Carini and G. L. Sicuranza, "Perfect periodic sequences for identification of even mirror fourier nonlinear filters," in *Proc. of ICASSP 2014*, Florence, Italy, May 2014.
- [7] 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, Sept. 1983.
- [8] C. Antweiler, "Multi-channel system identification with perfect sequences," in *Advances in Digital Speech Transmission*, U. Heute R. Martin and C. Antweiler, Eds. SPIE, 2008, pp. 171–198, John Wiley & Sons.
- [9] I. Mercer, "Merit factor of Chu sequences and best merit factor of polyphase sequences," *IEEE Trans. on Information Theory*, vol. 59, no. 9, pp. 6083–6086, 2013.
- [10] M. Soltanian and P. Stoica, "Design of perfect phase-quantized sequences with low peak-to-average-power ratio," in *Prof. of EUSIPCO 2012*, 2012, pp. 2576–2580.
- [11] C. Antweiler and G. Enzner, "Perfect sequence lms for rapid acquisition of continuous-azimuth head related impulse responses," in *Proc. of WASPAA '09*, 2009, pp. 281–284.
- [12] V. Marmarelis, *Nonlinear Dynamic Modeling of Physiological Systems*, John Wiley & Sons, Hoboken, NJ, USA, 2004.
- [13] E. A. Hylleraas, "Linearization of products of Jacobi polynomials," *Mathematica Scandinavica*, vol. 10, pp. 189–200, 1962.
- [14] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in C : the art of scientific computing*, Cambridge University Press, New York, NY, USA, 1995.
- [15] A. Fermo, A. Carini, and G. L. Sicuranza, "Low-complexity nonlinear adaptive filters for acoustic echo cancellation in GSM handset receivers," *European Trans. on Telecomm.*, vol. 14, no. 2, pp. 161–169, 2003.
- [16] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Signal Processing*, vol. 19, no. 6, pp. 716–723, 1974.
- [17] E. J. Hannan and B. G. Quinn, "The determination of the order of an autoregression," *J. R. Statist. Soc. B*, vol. 41, no. 2, pp. 190–195, 1979.
- [18] G. Schwartz, "Estimating the dimension of a model," *Ann. Statist.*, vol. 6, pp. 416–464, 1978.
- [19] K. Li, J.-X. Peng, and G. W. Irwin, "A fast nonlinear model identification method," *Automatic Control, IEEE Transactions on*, vol. 50, no. 8, pp. 1211–1216, 2005.