LEAST-SQUARE APPROXIMATION OF SECOND-ORDER NONLINEAR SYSTEMS USING QUASI-PERFECT PERIODIC SEQUENCES

Giovanni L. Sicuranza
DIA - University of Trieste - Italy
sicuranza@univ.trieste.it

Alberto Carini
DiSBeF - University of Urbino - Italy
alberto.carini@uniurb.it

ABSTRACT
We consider the identification of nonlinear filters using periodic sequences. Perfect periodic sequences have already been proposed for this purpose. A periodic sequence is called perfect for a nonlinear filter if it causes the basis functions to be orthogonal and the autocorrelation matrix to be diagonal. In this paper, we introduce for the same purpose the quasi-perfect periodic sequences. We define a periodic sequence as quasi-perfect for a nonlinear filter if the resulting auto-correlation matrix is highly sparse. The sequence is obtained by means of a simple combinatorial rule and is formed by samples having few discrete levels. These characteristics allow an efficient implementation of the least-squares method for the approximation of certain linear-in-the-parameters nonlinear filters. A real-world experiment shows the good performance obtained.

Index Terms— Least-squares approximation, second-order nonlinear systems, quasi-perfect periodic sequences, sparse auto-correlation matrix.

1. INTRODUCTION
Least-squares (LS) algorithms belong to the class of estimation methods that can be solved exactly using measured values of input signal realizations. In principle, they are used for modeling unknown linear systems, by often applying stochastic input signals. The optimal least-square solution that minimizes the mean-square error (MSE) between the output of the unknown system and that of the modeling filter is obtained by multiplying the inverse auto-correlation matrix of the vector collecting the input samples by the cross-correlation vector between the unknown system output and the input samples. This procedure can be easily extended to the identification of nonlinear systems if the modeling filter is represented as a linear-in-the-parameters (LIP) nonlinear filter [1]. In fact, a LIP filter is characterized, as linear filters, by the linearity relationship of its output with respect to the filter coefficients. As a consequence, the optimal least-square solution is formally the same as for linear filters. The only difference is that, in place of the input samples, it is necessary to use the corresponding values of the basis functions defining the LIP nonlinear filter. The class of these filters is broad and includes, among others, Volterra filters [2], bilinear filters [2], functional link artificial neural networks (FLANN) [3], and piecewise linear regressors [4]. The Volterra filter, exploiting polynomial basis functions formed by products of input samples, is perhaps the most famous LIP model. In fact, in contrast to other members of this class, Volterra filters are universal approximators for causal, time-invariant, finite-memory, continuous, nonlinear systems, i.e., they can arbitrarily well approximate these filters according to the Stone-Weierstrass theorem [5]. However, their basis functions are not mutually orthogonal. To overcome this difficulty, two new LIP nonlinear filters have been recently introduced that replicate the construction rule of Volterra filters but use different basis functions. They are the even mirror Fourier nonlinear (EMFN) filter [6], [7], [8], based on even mirror symmetric trigonometric basis functions, and the Legendre filter [9], [10], based on Legendre polynomials. Both filters are universal approximators, in the sense specified above. Moreover, the basis functions of EMFN and Legendre filters are, respectively, mutually orthogonal for white uniform input signals in the interval [-1, +1]. As a consequence, EMFN and Legendre models can be simply estimated by means of the so-called cross-correlation method [11]. Indeed, in this situation, the auto-correlation matrix tends to become diagonal, as far as the number of the input samples increases, making it simple to compute its inverse. The cross-correlations between each basis function and the output of the unknown system need still to be calculated and eventually the result of the cross-correlation method coincides with the optimal LS solution. The drawback of the method, when stochastic inputs are used, is the huge number of input samples necessary to obtain a diagonal structure for the auto-correlation matrix. To overcome this difficulty, it is possible to resort to deterministic signals such as perfect periodic sequences (PPSs), able to guarantee the orthogonality of the basis functions on a finite interval for white uniform input signals in the interval [-1, +1]. A periodic sequence is perfect for a given filter if all the cross-
correlations between two different basis functions, estimated over a period of \( N \) samples, are equal to zero. PPSs have been exploited for the identification of linear filters \([12–14]\) and the derivation of efficient least-mean-squares adaptive algorithms \([15–17]\). More recently, it has been shown that PPSs can be developed even for EMFN filters \([18, 19]\) and Legendre filters \([10, 20]\). The PPSs are computed by deriving an undetermined system of nonlinear equations involving the EMFN and Legendre basis functions, and then solving it by means of the Newton-Raphson method. Since the number of equations exponentially increases with the order \( P \) and geometrically with the memory \( N \) of the nonlinear filter, only kernels of order \( P = 1, 2 \) and \( 3 \) are usually considered. It has been also shown in \([10, 19]\) that the number of nonlinear equations can be reduced exploiting symmetry and oddness conditions. As a consequence, the Newton-Raphson method becomes feasible, but at the expense of an increased length \( L \) of the PPS. PPSs for EMFN and Legendre filters of order \( P = 2 \) and \( P = 3 \) are available at the web site http://www.units.it/ipl/res_PSeqs.htm. In consideration of the diagonal structure of the auto-correlation matrix, the LS algorithm can be efficiently computed.

The aim of this paper is to define alternative, not perfect, sequences that, anyway, permit us to identify nonlinear systems using LS algorithms with reduced computational costs. To this purpose, it is necessary to relax some of the characteristics of the PPS, such as the diagonality of the auto-correlation matrix, and add some new constraints, such as the representation of the sequence samples with a small number of levels. The resulting sequence described in this paper is called a quasi-perfect periodic sequence (QPPS). It will be shown that, in contrast to PPSs, the QPPSs are the same for EMFN, Legendre and even Volterra filters. For sake of simplicity, the discussion is limited here to filters formed with only first- and second-order kernels.

The paper is organized as follows. In Section 2, basic notions on EMFN and Legendre filters are summarized. In Section 3, the generation rule of a QPPS for second-order EMFN, Legendre and Volterra filters is presented. Further details on QPPSs and experimental results are presented in Section 4. Conclusions follow in Section 5.

The following notation is used throughout the paper. Intervals are represented with square brackets, \( R_1 \) is the unit interval \([-1, 1]\), \( < x(n) >_L \) indicates time average over \( L \) successive samples of \( x(n) \), the operator \( \lceil \cdot \rceil \) indicates the first integer greater than or equal to the real argument.

### 2. BASIC NOTIONS ON EMFN AND LEGENDRE FILTERS

It has been shown in \([6, 7, 9, 10]\) that EMFN and Legendre filters are universal approximators for the input-output relationship of discrete-time, time-invariant, finite-memory, causal, continuous, nonlinear systems given by

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

where \( f \) is a real continuous function and \( x(n) \) belongs to \( R_1 \). The basis functions of EMFN filters are even mirror symmetric trigonometric functions, while those of Legendre filters are products of Legendre polynomials. Table 1 and Table 2 report the basis functions of order 1 and 2 of an EMFN and a Legendre filter, respectively, with memory of \( N \) samples. The Legendre polynomials \( \log_{k+1}(\xi) \) of Table 2 are obtained from the recursive relation

\[
\log_{k+1}(\xi) = \frac{2k+1}{k+1} \log_k(\xi) - \frac{k}{k+1} \log_{k-1}(\xi),
\]

where \( \log_0(\xi) = 1 \), \( \log_1(\xi) = \xi \), and \( k \) is the order of the basis function. For sake of simplicity, in the following derivations the basis function of order 0, which corresponds to an offset, is not taken into account. It is worth noting that the basis functions of a second-order Volterra filter are those given in Table 2, but replacing \( \log_2[x(n-i)] \) with \( x^2(n-i), i = 0, \ldots, N-1 \). By construction, EMFN and Legendre filters have the same complexity, i.e., number of coefficients, of a Volterra filter with the same order and memory. Therefore, all these filters, obtained in our case as a linear combination of the kernels of order 1 and 2, have a number of basis functions and a number of coefficients equal to

\[
N_T = \binom{N+2}{N} - 1.
\]

The input-output relationship of an EMFN, Legendre or Volterra filter, in general of a LiP nonlinear filter, with...
memory of $N$ samples, can be written as a linear combination of their respective basis functions $f_j(n)$

$$\hat{y}(n) = \sum_{j=1}^{N_T} h_j f_j(n).$$

(4)

Using a vector notation, (4) becomes

$$\hat{y}(n) = \mathbf{h} \cdot \mathbf{f}^T(n),$$

(5)

where

$$\mathbf{h} = [h_1, h_2, ..., h_{N_T}],$$

and

$$\mathbf{f}(n) = [f_1(n), f_2(n), ..., f_{N_T}(n)].$$

Then, the filter coefficients of the optimal LS solution are obtained, as for all the LIP nonlinear filters, by the equation

$$\mathbf{h}_o = \mathbf{R}_{ff}^{-1} \mathbf{y}_f,$$

(6)

since the output is linear with respect to the filter coefficients

In (6), $\mathbf{R}_{ff}^{-1}$ is the inverse of the $N_T \times N_T$ auto-correlation matrix of the basis functions of the LIP nonlinear filter and $\mathbf{P}$ is the $N_T \times 1$ cross-correlation vector between the output $y(n)$ of the unknown system and the basis functions. The elements of $\mathbf{R}_{ff}$ are computed as time averages on $L$ consecutive samples of the input $x(n)$ as follows,

$$r_{jk} = \langle f_j(n) f_k(n) \rangle_L, \quad j, k = 1, ..., N_T. $$

(7)

The elements of the vector $\mathbf{P}$ are similarly computed as time averages, on $L$ consecutive samples of the input $x(n)$, of the products of the output $y(n)$ and the basis functions $f_j(n)$,

$$p_j = \langle y(n) f_j(n) \rangle_L, \quad j = 1, ..., N_T. $$

(8)

3. QPPS FOR SECOND-ORDER FILTERS

Our goal is to define periodic sequences easy to generate and able to minimize the computational cost of the LS algorithm in (6). We consider the following constraints:

i) the derivation of the sequence is as simple as possible;

ii) each sample in the sequence is represented using only few levels, so that the computation of the basis functions can be done using small look-up tables;

iii) the resulting auto-correlation matrix, in general no more diagonal as for PPSs, is as sparse as possible, so that the computation and inversion of the auto-correlation matrix are not computationally intensive.

A sequence satisfying these constraints is called here a quasi-perfect periodic sequence (QPPS). To satisfy the first constraint, our choice is to resort to a combinatorial rule to generate the sequence. As far as the point ii) is concerned, the choice of a three values ($-1, 0, +1$) sequence led, in our experiments, to solutions affected by some cases by singularity or bad conditioning of $\mathbf{R}_{ff}$. These effects are greatly alleviated by using five values ($-1, -0.5, 0, +0.5, +1$) sequences. Therefore, in the following we will refer to this choice. Finally, to satisfy the third constraint, we proceed as it follows.

Let us consider an integer $M \geq \lceil (N + 1)/2 \rceil$. The QPPS is defined with the following rule:

i) List all the $2^M$ sub-sequences of length $M$ formed with $+1, -1$.

ii) Append to each sub-sequence at point i) a zero sub-sequence of length $N - M$.

iii) Take the list in i) and multiply each element by $-0.5$, obtaining a list of $2^M$ sub-sequences of length $M$ formed by $+0.5, -0.5$.

iv) Append to each sub-sequence at point iii) a zero sub-sequence of length $N - M$.

v) Define one period of the QPPS by concatenating all the resulting sub-sequences at points ii) and iv).

The length of the resulting QPPS is $L = 2 \cdot N \cdot 2^M$ samples. With this sequence, the auto-correlation matrix for second-order EMFN, Legendre and Volterra filters of memory $N$ is highly sparse. Indeed, with reference to second-order EMFN, the following cross-products between basis functions have zero average over a period:

$$\sin[\frac{\pi}{2} x(n-i)] \sin[\frac{\pi}{2} x(n-j)]$$

for any $i \neq j$ and $i, j = 0, ..., N-1$;

$$\cos[\pi x(n-i)] \cos[\pi x(n-j)]$$

for any $i, j = 0, ..., N-1$;

$$\sin[\frac{\pi}{2} x(n-i)] \sin[\frac{\pi}{2} x(n-l)] \sin[\frac{\pi}{2} x(n-m)]$$

for any $i, j, l \neq m$.

Then, the filter coefficients of the optimal LS solution are obtained, as for all the LIP nonlinear filters, by the equation

$$\mathbf{h}_o = \mathbf{R}_{ff}^{-1} \mathbf{P}_y.$$
of the remaining basis functions of order 2. This structure allows the determination of \( R_{ff} \) and its inverse with reduced computations. Indeed, the calculation of the elements of \( R_{ff} \) summing to zero are avoided and the computation of \( R_{ff}^{-1} \) requires the inversion of only an \( N \times N \) matrix and two diagonal matrices. Moreover, the basis functions can be computed by means of look-up-tables with only five entries, and thus further alleviating the computational cost.

Table 3 gives some of the characteristics of the QPPSs for a second-order EMFN, Legendre or Volterra filter, i.e., the memory \( N \), the total number of coefficients \( N_T \), the sequence length \( L \) and the density, in percent, of the auto-correlation matrix \( R_{ff} \). The last column reports, for comparison, the density, in percent, of the auto-correlation matrix for a PPS.

<table>
<thead>
<tr>
<th>( N )</th>
<th>( N_T )</th>
<th>( L )</th>
<th>Density</th>
<th>( 100/N_T )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>9</td>
<td>24</td>
<td>18.52</td>
<td>11.11</td>
</tr>
<tr>
<td>4</td>
<td>14</td>
<td>64</td>
<td>13.27</td>
<td>7.14</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>80</td>
<td>10.00</td>
<td>5.00</td>
</tr>
<tr>
<td>6</td>
<td>27</td>
<td>192</td>
<td>7.82</td>
<td>3.70</td>
</tr>
<tr>
<td>7</td>
<td>35</td>
<td>224</td>
<td>6.29</td>
<td>2.86</td>
</tr>
<tr>
<td>8</td>
<td>44</td>
<td>512</td>
<td>5.17</td>
<td>2.27</td>
</tr>
<tr>
<td>9</td>
<td>54</td>
<td>576</td>
<td>4.32</td>
<td>1.85</td>
</tr>
<tr>
<td>10</td>
<td>65</td>
<td>1280</td>
<td>3.67</td>
<td>1.54</td>
</tr>
<tr>
<td>11</td>
<td>77</td>
<td>1408</td>
<td>3.15</td>
<td>1.30</td>
</tr>
<tr>
<td>12</td>
<td>90</td>
<td>3072</td>
<td>2.74</td>
<td>1.11</td>
</tr>
<tr>
<td>13</td>
<td>104</td>
<td>3328</td>
<td>2.40</td>
<td>0.96</td>
</tr>
<tr>
<td>14</td>
<td>119</td>
<td>7168</td>
<td>2.13</td>
<td>0.84</td>
</tr>
<tr>
<td>15</td>
<td>135</td>
<td>7680</td>
<td>1.89</td>
<td>0.74</td>
</tr>
<tr>
<td>16</td>
<td>152</td>
<td>16384</td>
<td>1.70</td>
<td>0.66</td>
</tr>
<tr>
<td>17</td>
<td>170</td>
<td>17408</td>
<td>1.53</td>
<td>0.59</td>
</tr>
<tr>
<td>18</td>
<td>189</td>
<td>36864</td>
<td>1.39</td>
<td>0.53</td>
</tr>
<tr>
<td>19</td>
<td>209</td>
<td>38912</td>
<td>1.26</td>
<td>0.48</td>
</tr>
<tr>
<td>20</td>
<td>230</td>
<td>81920</td>
<td>1.15</td>
<td>0.43</td>
</tr>
</tbody>
</table>

4. EXPERIMENTAL RESULTS

4.1. Experiment 1.

In this experiment, we analyze the accuracy of the LS algorithm in the identification of an unknown system, described by a second-order EMFN structure, when a QPPS is used as input signal. The figure of merit is the mean-square deviation (MSD) between the coefficients of the unknown system and those of the modeling EMFN filter. The results shown in Figure 1 are averages over 100 independent tests, where the coefficients of the unknown system are chosen randomly in \( R_1 \). Both the unknown system and the modeling filter have memory \( N = 10 \). The dashed line indicates the MSD for the coefficients identified using an infinite precision PPS, the small circles indicate the MSD for the PPS quantized on finite numbers of bits, and the small squares indicate the MSD for the five levels QPPS. It can be noticed that the quantized PPS quickly degrades the performance when the number of bits decreases, whereas the five levels QPPS obviously gives the same performance on the whole range of quantization bits.

4.2. Experiment 2.

We consider the identification of a real nonlinear device, i.e., a saturated audiophile vacuum tube preamplifier Behringer Tube Ultragain MIC100. A QPPS has been fed to the preamplifier and the corresponding output has been recorded with a notebook at 8 kHz sampling frequency. The SNR is more than 50 dB. The system has a memory length lower than 20 samples, thus a QPPS for EMFN, Legendre, and Volterra filters with memory \( N = 20 \) and period \( L = 81920 \) samples has been used for the identification. By acting on the gain control of the preamplifier, different distortion levels can be set. At the different settings, the preamplifier has been identified with i) a linear filter of 20 samples memory, ii) a second order Volterra filter, iii) a second-order Legendre filter, and iv) a second-order EMFN filter all with memory of 20 samples and 230 coefficients.

Table 4 provides for increasing distortion levels: the second (Dist2) and third (Dist3) harmonic distortions measured on a tone at 1 kHz and at the maximum used volume, and the ratio in dB between the power of the residual identification error and the power of the preamplifier output signal for the Volterra, Legendre (Leg.), EMFN, and linear (Lin.) filters, respectively. We see from Table 4 that, apart from the smallest distortion levels, the linear filter is unable to efficiently model the nonlinearities introduced by the amplifier and gives always worst results than the Volterra and Legendre filters. For small nonlinearities, especially when the second order harmonic distortions prevails, the Volterra and Legendre filters provide the best model of the amplifier and give equivalent representations since they are both polynomial models that include a linear and quadratic term. In contrast, for small distortions, the EMFN filter provides poor results since it lacks a linear term. For larger distortion levels, when the third order harmonic distortion becomes significant, the Volterra and Legendre filters are unable to cope with the third order non-
Table 4. Identification results for the preamplifier Behringer MIC100

<table>
<thead>
<tr>
<th>Dist2</th>
<th>Dist3</th>
<th>Volterra</th>
<th>Leg.</th>
<th>EMFN</th>
<th>Lin.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.25%</td>
<td>0.19%</td>
<td>-23.6</td>
<td>-23.6</td>
<td>-16.6</td>
<td>-23.1</td>
</tr>
<tr>
<td>2.90%</td>
<td>0.64%</td>
<td>-22.8</td>
<td>-23.0</td>
<td>-16.8</td>
<td>-21.3</td>
</tr>
<tr>
<td>3.52%</td>
<td>0.84%</td>
<td>-22.5</td>
<td>-22.7</td>
<td>-17.0</td>
<td>-20.4</td>
</tr>
<tr>
<td>4.46%</td>
<td>1.28%</td>
<td>-21.6</td>
<td>-21.9</td>
<td>-17.3</td>
<td>-18.7</td>
</tr>
<tr>
<td>6.00%</td>
<td>2.46%</td>
<td>-19.4</td>
<td>-19.7</td>
<td>-18.0</td>
<td>-16.2</td>
</tr>
<tr>
<td>6.86%</td>
<td>6.06%</td>
<td>-16.0</td>
<td>-16.2</td>
<td>-18.5</td>
<td>-14.3</td>
</tr>
<tr>
<td>3.11%</td>
<td>11.32%</td>
<td>-14.0</td>
<td>-14.1</td>
<td>-18.0</td>
<td>-13.2</td>
</tr>
<tr>
<td>3.41%</td>
<td>14.87%</td>
<td>-13.0</td>
<td>-13.0</td>
<td>-17.4</td>
<td>-12.8</td>
</tr>
<tr>
<td>11.27%</td>
<td>17.22%</td>
<td>-11.7</td>
<td>-11.7</td>
<td>-15.6</td>
<td>-11.6</td>
</tr>
</tbody>
</table>

linearity and give results comparable with those of the linear filter. In these conditions the EMFN filter provides the best results thanks to the ability of its basis functions to generate also the higher order harmonics.

5. CONCLUSIONS

In this paper, the so-called QPPSs have been first introduced and then applied to the identification of EMFN, Legendre and Volterra filters. The properties of QPPSs allow their fast generation and the efficient implementation of the LS algorithm. For sake of simplicity, in this paper the discussion has been limited to filters formed with first- and second-order kernels, deferring the derivations for third-order filters to a paper presently in preparation. Indeed, complete or simplified third-order filters can often guarantee sufficiently good approximation performance. On the other hand, simplified structures allow the use of QPPSs of reduced length.

REFERENCES