EFFICIENT NLMS AND RLS ALGORITHMS FOR A CLASS OF NONLINEAR FILTERS USING PERIODIC INPUT SEQUENCES

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

V. John Mathews
University of Utah
Salt Lake City, UT, USA
mathews@ece.utah.edu

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

ABSTRACT
The paper discusses computationally efficient NLMS and RLS algorithms for a broad class of nonlinear filters using periodic input sequences. The class comprises all nonlinear filters whose output depends linearly on the filter coefficients. The algorithms presented in the paper are exact, suitable for identification and tracking of every nonlinear system in the class, and require a real-time computational effort of a single multiplication, an addition, and a subtraction per input sample. The transient and steady-state behavior of the algorithms are discussed and the effect of a model mismatch between the unknown system and the adaptive filter is also analyzed. The low computational complexity, good performance, and applicability of the algorithm to a large class of nonlinear systems make the approach of this paper a valuable alternative to the current techniques for nonlinear system identification.

Index Terms— Adaptive filters, adaptive signal processing, nonlinear filters.

1. INTRODUCTION
Identification and tracking of nonlinear systems is often a difficult task. Depending on the nonlinear system, the number of parameters to be estimated or updated can be much larger than that of a linear FIR filter with the same memory length, often resulting in exorbitant computational complexity. Even when the input signal is white Gaussian, the autocorrelation matrix of the input data may be non-diagonal and ill-conditioned [1]. In this paper we consider the class of nonlinear systems whose input-output relationships depend linearly on the system parameters [2]. This class of systems includes, in addition to linear filters, polynomial filters [1], the extended Volterra filters [3], the FLANN filters [4], and many other nonlinear structures [2]. Different approaches have been proposed in the literature to estimate these systems. One of the fundamental difficulties with nonlinear system identification is the complexity of the model structure and the correspondingly large computational complexity of the algorithms. For example, a generic pth order truncated Volterra system model with N sample memory has $O(N^p)$ coefficients. The lowest computational complexity for adaptive identification and tracking available today is $O(N^p)$ arithmetical operations per input signal sample. The goal of this paper is to present an adaptive system identification approach that reduces this computational burden substantially, and in a manner that is independent of the complexity of the system model.

Recently, the first author of this paper presented an algorithm for the identification and tracking of linear FIR systems using periodic input sequences [5], in which efficient NLMS and RLS algorithms that have a real-time computational effort of a single multiplication, an addition, and a subtraction per sample time were discussed. These algorithms do not evaluate the coefficients of the underlying system directly. Instead, they determine the coefficients of an equivalent representation, from which the impulse response can be easily computed. The paper also showed that the algorithms have convergence and tracking properties that can be better than or comparable to the NLMS algorithm for white noise input. In this paper, we show that the same algorithms can be applied to the identification and tracking of the class of nonlinear systems characterized by a linear dependence of the output on the coefficients of the system model. The resulting systems preserve the low computational complexity of their linear counterparts, and also share the good convergence and tracking properties exhibited by the adaptive linear filters of [5].

Several authors have described the use of periodic sequences for the identification of Volterra systems. For example, the properties of pseudorandom multilevel sequences, i.e., of periodic sequences, in the identification of Volterra and extended Volterra filters were studied in [3]. It was shown in [3] that a pseudorandom multilevel sequence of degree D is persistently exciting for an extended Volterra filter of order $P$ and memory length $N$ if and only if it has at least $P + 1$ distinct levels and $D \geq N$. An efficient algorithm for least-square parameter estimation was also proposed in [3]. The approach of this paper differs from that of [3] in two ways: 1) the derivations are applicable to a broader class of nonlinear filters, and 2) it deals with the NLMS and RLS algorithms for system identification and tracking.

The rest of the paper is organized as follows. In Section 2 the class of nonlinear filters whose outputs are linear in the coefficients is first reviewed and efficient NLMS and RLS algorithms using periodic sequences (PSEQ) are derived. The transient and steady-state behaviors of the algorithms are analyzed in Section 3. The effect of a model mismatch between the unknown system and the adaptive filter is discussed in Section 4. Experimental results are presented in Section 5. Concluding remarks are given in Section 6.

Throughout the paper, lowercase boldface letters are used to denote vectors, uppercase boldface letters are used to denote matrices, $E[\cdot]$ denotes mathematical expectation, $\| \cdot \|$ denotes the Euclidean norm, $\lfloor \cdot \rfloor$ is the largest integer smaller than or equal to the argument, and $I$ is an identity matrix of suitable dimensions.

2. EFFICIENT NLMS AND RLS ALGORITHMS
The class of nonlinear filters considered in this paper is characterized by a linear dependence of the filter output on the filter coefficients. The algorithms described in this section apply to non-recursive models with a finite memory of $N$ samples. The input-output relationship of these models can be expressed in vector form as:
\[ y(n) = h^T x_F(n), \]  
where \( h \) is a length-\( M \) coefficient vector and \( x_F(n) \) is an input data vector composed by \( M \) terms that are nonlinear functions of the last \( N \) input data samples \( x(n), \ldots, x(n - N + 1) \). This class is very broad and includes many commonly employed nonlinear filter structures [2]. In particular, in the experimental section, we will provide simulation results for a truncated second-order Volterra filters, where

\[
x_F(n) = [x(n), \ldots, x(n - N + 1), x^2(n), \ldots, x^2(n - N + 1), x(n)x(n - 1), \ldots, x(n)x(n - N + 2)x(n - N + 1), \ldots, x(n)x(n - N + 1)]^T
\]  
and the second-order FLANN filters, where

\[
x_F(n) = [x(n), \ldots, x(n - N + 1), \sin[x(n)], \ldots, \sin[x(n - N + 1)], \cos[x(n)], \ldots, \cos[x(n - N + 1)], \sin[2\pi x(n)], \ldots, \cos[2\pi x(n - N + 1)]]^T.
\]

Let us assume that the input sequence \( x(n) \) is periodic with period \( M \). Then, the input vector \( x_F(n) \) can take only one of \( M \) different values, say \( x_1, \ldots, x_M \). If the \( M \times M \) matrix

\[ X_F = [x_0, \ldots, x_{M-1}] \]

is invertible, we can define the \( M \times M \) matrix \( W \) such that

\[ WX_F = I \]  
Since \( W \) is invertible, we can find a vector \( e \) such that

\[ h = We. \]

Let \( w_0, \ldots, w_{M-1} \) represent the \( M \) columns of \( W \). These vectors are in general non-orthogonal (even when the input sequence is a perfect periodic sequence [5]) but they are linearly independent (for the invertibility of \( X_F \)) and can be used for expanding the input-output relationship of the nonlinear filter, i.e., the expression in (1) can be equivalently rewritten as

\[ y(n) = e^T W^T x_F(n) = \sum_{i=0}^{M-1} c_i w_i^T x_F(n), \]

where the vectors \( w_i \) are known once the periodic sequence is chosen and \( e = [e_1, \ldots, e_M]^T \) is a coefficient vector we wish to determine. The coefficient vector \( e \) characterizes the nonlinear filter in (1) as well as the coefficient vector \( h, \) and \( h \) can be easily estimated from the knowledge of \( e \) as in (6). Moreover, when the input sequence is periodic, at each time \( n \) we have \( x_F(n) = x_i \) with \( i = n \mod M \). Since

\[ w_i^T x_j = \begin{cases} 1 & \text{when } i = j, \\ 0 & \text{otherwise} \end{cases} \]

from (5), we have

\[ y(n) = c_i \text{ with } i = n \mod M. \]

This is the main property that was exploited in [5] to derive the efficient NLMS and RLS algorithms for periodic sequences.

In the NLMS algorithm, we want to find the coefficients \( c_i(n) \) that minimize the following mean-square cost function:

\[ J(n) = E \{[d(n) - y(n)]^2 \}, \]

with \( y(n) \) given in (7) and \( d(n) \) the desired signal. The coefficients are adapted with the gradient method,

\[ c_i(n + 1) = c_i(n) - \frac{\mu}{2} \frac{\partial J(n)}{\partial c_i(n)} \]

By approximating \( J(n) \) with \([d(n) - y(n)]^2\) and taking into account (9), it can be verified that

\[ \frac{\partial J(n)}{\partial c_i(n)} \approx \begin{cases} -2[d(n) - c_i(n)], & i = n \mod M \\ 0, & \text{otherwise} \end{cases} \]

Thus,

\[ c_i(n + 1) = \begin{cases} c_i(n) + \mu [d(n) - c_i(n)], & i = n \mod M \\ c_i(n), & \text{otherwise} \end{cases} \]

Similarly, it can be proved that the RLS algorithm that minimizes

\[ J(n) = \sum_{j=0}^{n} \lambda^{n-j} [d(j) - y(j)]^2, \]

with \( y(n) \) given in (7) and \( \lambda \) a forgetting factor, \( 0 < \lambda \leq 1 \), has adaptation equation

\[ c_i(n + 1) = \begin{cases} c_i(n) + \mu \{[h^T w_i] [d(n) - c_i(n)] \}, & i = n \mod M \\ c_i(n), & \text{otherwise} \end{cases} \]

For \( \lambda < 1 \), \( \mu(n) = \frac{1 - \lambda}{1 - \lambda^{M+1}} \) yields a perfect filter in the long run. The adaptation equations (13) and (15) are the exact NLMS and RLS algorithms, respectively, suitable for the identification and tracking of any nonlinear filter with input-output relationship (1). A significant advantage of these algorithms is their reduced computational complexity, as they only require a multiplication, an addition, and a subtraction during each iteration of the adaptive filter.

The adaptive filters (13) and (15) can identify the nonlinear filter in (1) only if the matrix \( X_F \) is invertible and has a reasonable condition number. By following a similar argument as in [3], it can be shown that, for extended Volterra filters of order \( P \), the matrix \( X_F \) is invertible only if the input sequence is composed by at least \( P + 1 \) symbols; less strict condition applies to truncated Volterra filters. We should point out that in all of our experimental results with random-phase multi-sine (RPMS) input sequences and Volterra, extended Volterra, or FLANN filters, the matrix \( X_F \) has been always invertible, even though it can become singular for some sequences (e.g., a periodic impulse sequence). According to [3], the condition number of \( X_F \) grows with the memory length and the order of the nonlinear filter (for Volterra filters, it grows exponentially with the order of the filter). Nevertheless, we have found experimentally that by properly choosing the periodic sequence (e.g., with RPMS sequences) the matrix \( X_F \) has acceptable condition numbers for low orders of nonlinearity and with reasonable values of the filter memory length.

### 3. Transient and Steady-State Analyses

In this section, we consider adaptation algorithms of the form (15), noting that \( \mu(n) = \mu \) results in the NLMS algorithm, and when \( \mu(n) \) varies with time as defined below (15), we get the RLS algorithm.

Assume that we wish to identify, using a periodic input sequence \( x(n) \), a nonlinear system with memory length \( N \) and input-output relationship

\[ d(n) = \tilde{h}^T x_F(n) + \nu(n) \]

\[ = \sum_{k=0}^{M-1} \tilde{c}_k w_i^T x_F(n) + \nu(n) \]

\[ = \tilde{c}_i + \nu(n), \quad \text{with } i = n \mod M, \]

1 An RPMS sequence is a periodic sequence with constant magnitude DFT and random phase.
Table 1. MSD(∞), MSDω(∞), MSE(∞) of the efficient LMS and RLS algorithms

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>μ(m) = μ</th>
<th>μ(m) = (1−λM)/2</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSD(∞)</td>
<td>(1−μM)/2</td>
<td>(1−μM)/2</td>
</tr>
<tr>
<td>MSDω(∞)</td>
<td>(1−μM)/2</td>
<td>(1−μM)/2</td>
</tr>
<tr>
<td>MSE(∞)</td>
<td>(1−μM)/2</td>
<td>(1−μM)/2</td>
</tr>
</tbody>
</table>

where h and ε = [ε0, ε1, ..., εM−1]T are related by (6), and ϵ(n) is a zero-mean stationary additive measurement noise, uncorrelated with x(n) and with power σε2. In order to simplify the algorithm analysis and avoid any transient effect on d(n), we assume that the unknown system has operated on the periodic input signal for at least N samples before the algorithm was activated.

Let us define the system errors τi(n) = c(n) − ci. By subtracting ci from both sides of (15) and taking into account (16), we obtain

τi(n+1) = \{ τi(n) + μ[(1−\lambda)]x(n−iucchini), i = n mod M otherwise. \}

(17)

The above equation fully characterizes the transient behavior of the adaptive algorithms.

From (17), it is easy to prove that the RLS algorithms in (15) and, for μ = 1, the NLMS algorithm in (13), are capable of identifying a noiseless non-linear system without the source signal. To see this for the RLS algorithm, we observe that μ[(1−\lambda)] = 1 for 0 ≤ n < M. It immediately follows, when ϵ(n) = 0, that

\[ τ_i(n+1) = 0 \] (18)

for i = n mod M and 0 ≤ n < M, such that τi(M) = 0 for all i.

We now want to study the mean-square error (MSE),

\[ \text{MSE}(n) = E[\|d(n) − y(n)\|^2] \] (19)

the mean-square deviation (MSD) of \( \epsilon(n) = [\epsilon_0(n), ..., \epsilon_{M−1}(n)]^T \),

\[ \text{MSD}_\epsilon(n) = E[\|\epsilon(n) − \epsilon\|^2] = E[\|\epsilon(n)\|^2] \] (20)

and the MSD of \( h(n) = \sum_{i=0}^{M−1} c_i(n)w_i \),

\[ \text{MSD}_\epsilon(n) = E[\|\epsilon(n) − \epsilon\|^2] = E[\|\epsilon(n)\|^2] \] (21)

By following the approach in [5] we can prove that the update expressions for MSE(n) and MSD(ω(n)) derived for linear system models apply also to the nonlinear filters (7). The derivations are not repeated here because of space limitations. A slightly different result is obtained for MSD(ω(n)), since, in contrast to the linear case, the Euclidean norms of the vectors \( w_i \) that generate the output signal y(n) are in general different from each other. Under the hypothesis that \( \tau_i(n) \) are zero mean and independent, it can be proved that

\[ \text{MSD}_\epsilon[(m+1)M+i] = [1 − \mu(m)]^2 \text{MSD}_\epsilon(mM+i) + \mu^2(m)\sigma^2 \sum_{i=0}^{M−1} \|w_i\|^2. \] (22)

From the update expressions for MSE, MSD(ω), and MSD(ω), we know that the transient and steady-state properties of the algorithms depend only on the step size μ(m) and the noise power σε2. In particular, MSD(ω) and MSE(ω) do not depend on the particular input periodic sequence. Only MSD(ω) depends on the choice of

the periodic sequence through \( \sum_{i=0}^{M−1} \|w_i\|^2 \). The algorithms are numerically stable and converge exponentially to the unknown system coefficients for \( 0 < \epsilon < \mu(m) < 2 − \epsilon < 2 \forall n \) and for some small positive constant ϵ.

The steady-state values of MSE, MSD(ω), and MSD(ω) computed from the corresponding update expressions are listed in Table 1 for μ(m) = μ and for μ(m) = (1−λM)/2.

4. MODEL MISMATCH

In this section we will analyze the behavior of the algorithm when there is a model mismatch between the adaptive filter and the unknown system to be identified. This analysis can be easily performed under the simplifying assumption that there is no measurement noise, i.e., ϵ(n) = 0. Under these conditions, the NLMS algorithm with μ = 1 and the RLS algorithm converge in just M samples and it can be proved that the filter they converge to is

\[ h(M) = \sum_{i=0}^{M−1} w_i d(i) = Wd, \] (23)

where \( d = [d(0), ..., d(M−1)]^T \). Proceed by considering different cases.

Case 1: The system is perfectly matched. In this case, we have d(i) = \( \hat{h}^T x_i \), d = \( Xf^T \hat{h} \), and h(M) = \( WXf^T \hat{h} = \hat{h} \). The algorithm is able to identify the unknown system.

Case 2: The system is overdetermined, i.e., the adaptive filter has more coefficients than needed.

\[ d(i) = [\hat{h}_0, ..., \hat{h}_L, 0, ..., 0]^T x_i, \] (24)

\[ h(M) = [\hat{h}_0, ..., \hat{h}_L, 0, ..., 0]^T. \] (25)

Thus, in the adaptive filter model, some coefficients are redundant and converge to zero, but the algorithm is still able to identify the unknown system.

Case 3: The system is underdetermined by memory, i.e., some terms are missing in the input data vector x(n) and the missing terms are delayed versions of those present in x(n). For simplicity, let us assume that

\[ d(n) = \hat{h}_1^T x_1(n) + \hat{h}_2^T x_2(n − 1), \] (26)

where some of the coefficients of \( \hat{h}_2 \) could be zero. We can generalize this model by including more delayed terms. Then,

\[ d(i) = \hat{h}_1^T x_1(n) + \hat{h}_2^T x_2(n − 1), \] (27)

\[ h(M) = Wx_f^T (n)\hat{h}_1 + Wx_f^T (n)U\hat{h}_2, \] (28)

where U a permutation matrix. Thus,

\[ h(M) = \hat{h}_1 + U\hat{h}_2 \] (29)

and the unknown system estimation is affected by an aliasing error that depends on \( \hat{h}_2 \).

Case 4: The system is underdetermined by order, i.e., in the input data vector x(n) some terms are missing and the missing terms are not delayed versions of those present in x(n). Then,

\[ d(n) = \hat{h}_1^T x_1(n) + \hat{h}_2^T x_2(n), \] (30)

with \( x_2(n) \) a nonlinear vector function of the input samples \( x(n), ..., x(n − N + 1) \), and

\[ d = X_f^T \hat{h}_1 + [x_1(0), ..., x_1(M−1)]^T \hat{h}_2 = X_f^T \hat{h}_1 + X_f^T V \hat{h}_2. \] (31)
with $V^T = [x_c(0), ..., x_c(M-1)]X^{-1}$. Thus,

$$h(M) = \hat{h}_1 + VH_2.$$  \hspace{1cm} (32)

In this case we also have an aliasing error that depends on $\hat{h}_2$. We refer to it as an aliasing error even though the matrix $V$ spreads the coefficients of $\hat{h}_2$ over all the coefficients of the adaptive filter.

It should be noted that, if we underestimate the memory length of a FLANN filter, the system is underdetermined by memory. On the other hand, if we underestimate the memory length of a truncated Volterra filter, the system is not only underdetermined by memory, but also by degree. In such situations, the system we want to identify depends on nonlinear terms that are not a delayed version of those present in the input vector to the adaptive filter, and thus they belong to Case 4.

The results presented in this section are not only interesting from a theoretical point of view, but are also useful because they determine which adaptive filter coefficients are affected by model underestimation.

**5. SIMULATION RESULTS**

In this section we provide some simulation results for the identification of a second-order truncated Volterra filter and a second-order FLANN filter, both of memory length 20 samples. The truncated Volterra filter has 440 coefficients and the FLANN filter has 100 coefficients in these simulations. The input signal is an RPMS sequence with unit power. The coefficients of the system to be identified are random generated using the additional constraint that the output power of the nonlinear part of the system is 10 dB below the output power of the linear part. The adaptive filter used for the identification was perfectly matched to the nonlinear system. For the chosen RPMS sequence, the condition number of $X$ is 41724 for the truncated Volterra filter, and 572.67 for the FLANN filter.

Figures 1 and 2 display the ensemble averages over 200 simulations of MSE($n$) [plot (a)], MSD$_c$($n$) [plot (b)], and MSD$_h$($n$) [plot (c)] for different output signal SNRs. Figure 1 shows the results for the second-order truncated Volterra filter adapted with the NLMS algorithm (13) with $\mu = 0.2$, while Figure 2 displays the corresponding results for the second-order FLANN filter adapted with the RLS algorithm (15) with $\lambda = 0.8$. In the figures, the dashed lines represent the theoretical values of $\text{MSE}(+\infty)$, $\text{MSE}_c(+\infty)$, or $\text{MSE}_h(+\infty)$ as appropriate, and as given in Table 1. In all experiments, the theoretical values of Table 1 match the results of the simulations accurately.

**6. CONCLUDING REMARKS**

We have discussed computationally efficient exact NLMS and RLS algorithms for a class of nonlinear filters that is characterized by a linear output dependence on the coefficients. The algorithms are suitable for identification and tracking of any nonlinear system in the class and require a computational effort of just a multiplication, an addition, and a subtraction per sample time. We have analyzed the transient and steady-state behavior of the algorithms and derived exact expressions for the steady-state values of the $\text{MSE}_c(n)$, $\text{MSE}_h(n)$, and $\text{MSE}(n)$. We also studied the behavior of the algorithms in case of a model mismatch. Preliminary simulation results demonstrate the validity of the algorithms as well as the accuracy of the derived theoretical results. Low computational complexity, good performance, and generality of the algorithms make the proposed approach a valid alternative to current techniques for identification of a large class of nonlinear systems.

**7. REFERENCES**


