BLIND LINEARIZATION, AND IDENTIFICATION OF NONLINEAR SYSTEMS – A LEAST SQUARES, P^{TH} ORDER INVERSE APPROACH

Gil M. Raz and Barry D. Van Veen

Department of Electrical and Computer Engineering
University of Wisconsin – Madison
1415 Engineering Drive, Madison, WI 53706
raz@caddis.ece.wisc.edu

ABSTRACT

A deterministic approach to blind nonlinear channel equalization and identification is presented. This approach applies to nonlinear channels that can be approximately linearized by finite memory, finite order Volterra filters. Both the Volterra equalizers and the linearized channels are identified. This method also applies to blind identification of linear IIR channels. General conditions for existence and uniqueness are discussed and numerical examples are given.

1. INTRODUCTION

Identification and equalization of nonlinear systems are problems of practical interest in many problems where the systems cannot be accurately modeled as linear. Blind linear identification and equalization problems have been studied by many researchers; however, relatively little work has been done in the field of blind nonlinear system identification and equalization. Blind approaches have been proposed for restricted classes of nonlinearities such as linear-zero memory-linear systems [2] or strict Volterra models of nonlinearity [1]. Blind finite alphabet methods [5] suffer from ambiguity in the solution. The work in [1] describes a very elegant approach for finding linear FIR equalizers for nonlinear channels. However, the results in [1] only apply to nonlinear channels that can be exactly described by Volterra filters.

In this paper we consider blind equalization of a very general class of nonlinear systems, those that can be approximately linearized by finite order, finite memory Volterra filters. We generalize the approach in [7] to show that under rather general conditions it is possible to blindly determine P^{th} order linearizing equalizers and the linearized channels represented by the cascade of the nonlinear channel and the P^{th} order equalizer. As a special case we apply this approach to blindly identify linear IIR channels.

This paper is organized as follows. Section 2 introduces notation and develops the problem statement. The equations needed for linearization and identification are derived in Section 3. In Section 4 rudimentary conditions for identifiability and linearizability are given. Examples illustrating the effectiveness of our approach are given in Section 5.

This work was supported in part by Army Research Office under grant DAAL04-93-G-0208 and the Wisconsin Alumni Research Foundation

2. PRELIMINARIES

2.1. Notation

Bold face symbols denote vectors, $(\cdot)^T$ denotes the transposition operator, and $*$ represents convolution. $T_L[\mathbf{x}]$, describes a Toeplitz matrix with $L+1$ columns constructed from the vector $\mathbf{x}$ as follows.

$$T_L[\mathbf{x}] = \begin{bmatrix} x(0) & 0 & 0 & \cdots & 0 \\ x(1) & x(0) & 0 & \cdots & 0 \\ \vdots & & & \ddots & \\ x(L) & & \cdots & & x(0) \\ \vdots & & & & \vdots \\ x(N) & & \cdots & & x(N-L) \end{bmatrix} \quad (1)$$

The notation $L_{\mathbf{h}}$ denotes the memory of the FIR filter represented by the impulse response vector $\mathbf{h}$. Hence, we may write the convolution $y(l) = (x*h)(l)$ in vector form as $\mathbf{y}=\mathbf{x}*\mathbf{h}=T_{L_{\mathbf{h}}}[\mathbf{x}]\mathbf{h}$. Finally, the notation $\mathbf{h}(z)$ denotes the polynomial $\sum_{l=0}^{L_{\mathbf{h}}} h(l)z^{-l}$.

2.2. Diagonal Coordinate System

The Volterra filter [4] is used in this paper to implement an equalizer to a nonlinear system. In order to facilitate the use of linear algebraic methods, we express the Volterra filter output as a parallel combination of linear filters applied to nonlinear combinations of the input. We call this a diagonal coordinate system (DCS) representation since the linear filter coefficients are obtained from the diagonals of the sampled hypercube defined by the Volterra kernels.

Let $x(l)$ be the input to a Volterra filter G of order P and memory L. The output sequence $y(l)$ is commonly written as

$$y(l) = \sum_{p=1}^{P}\sum_{l_1=0}^{L}\cdots\sum_{l_p=0}^{L} G_p(l_1,\ldots,l_p)x(l-l_1)\cdots x(l-l_p) \quad (2)$$

without loss of generality the kernels $G_p(l_1,\ldots,l_p)$ are assumed to be symmetric under permutation of the variables

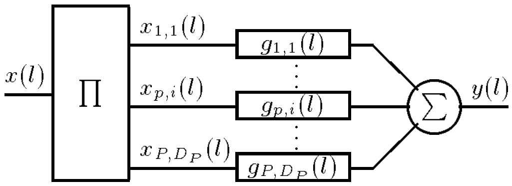

Figure 1: Volterra filter in Diagonal Coordinate System

$l_1, \ldots, l_p$ hence (2) is rewritten as

$$y(l) = \sum_{p=1}^{P} \sum_{0 \le l_1 \le \cdots \le l_p \le L} G_p^{\mathrm{symm}}(l_1, \ldots, l_p) x(l - l_1) \cdots x(l - l_p) \tag{3}$$

There are $D_p = D(p, L) = (p + L - 1)!/((p-1)!L!)$ possible combinations of the parameters $l_1, \ldots, l_p$ for a given polynomial order p [3]. We introduce the change of coordinates (DCS) $l_1 = l$ and $l_{\alpha+1} = l + r_\alpha$ for $\alpha = 1, \ldots, p-1$ and the notation

$$\begin{aligned} x_{p,i}(l) &= x(l) \prod_{\alpha=1}^{p-1} x(l - r_\alpha^{(i)}) \\ g_{p,i}(l) &= G_p^{\mathrm{symm}}(l, l + r_1^{(i)}, \ldots, l + r_{p-1}^{(i)}) \end{aligned} \tag{4}$$

where the superscript i denotes a given combination of the r_α and $i = 1, \ldots, D_p$. Using this notation we write the Volterra filter output (3) in DCS format

$$y(l) = \sum_{p=1}^{P} \sum_{i=1}^{D_p} (x_{p,i} * g_{p,i})(l) \tag{5}$$

Thus the Volterra filter output is given by a sum of linear filters acting on the various p^{th} order products of the input with itself as illustrated in Figure 1.

Note that we may rewrite (5) in vector form as

$$\begin{aligned} \mathbf{y} &= G[\mathbf{x}] \\ &= \left[T_{L_{1,1}}[\mathbf{x}_{1,1}] \cdots T_{L_{P,D_P}}[\mathbf{x}_{P,D_P}]\right] \mathbf{g} \\ &= \mathcal{T}[\mathbf{x}]\, \mathbf{g} \end{aligned} \tag{6}$$

where $\mathbf{g}$ is the vector of concatenated impulse responses $\mathbf{g} = \left[\mathbf{g}_{1,1}{}^T \cdots \mathbf{g}_{P,D_P}{}^T\right]^T$ and $L_{p,i}$ is the length of $\mathbf{g}_{p,i}$.

2.3. Problem Statement

Assume that oversampling or a sensor array is used to generate M outputs representing M nonlinear channels $H^{(m)}[\cdot]$, $m = 1, \ldots, M$ applied to an input $s(l)$. Further, suppose the following is true:

Assumption 2.1 *For all channels represented by the operators $H^{(m)}[\cdot]$, $m = 1, \ldots, M$ there are integers P, L such that there exist P^{th} order Volterra filters $G^{(m)}[\cdot]$ of maximum memory L whose cascade connection with $H^{(m)}[\cdot]$ create linear systems $Q^{(m)}[\cdot]$. That is, $Q^{(m)}[\cdot] = G^{(m)}[H^{(m)}[\cdot]]$ has negligible nonlinear behavior.*

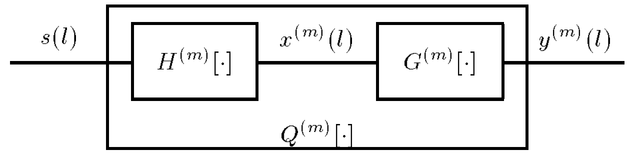

Figure 2: Channel m with linearizing equalizer.

The problem is as follows. Given only the observations of the channel outputs, determine the Volterra filter equalizers and the linear channels that result from the cascade connection of the channels with their respective equalizers. Further, recover the input signals $s(l)$.

3. IDENTIFICATION EQUATIONS

To simplify notation we assume identical filter lengths for all channels, e.g., $L_{\mathbf{q}^{(m)}} = L_{\mathbf{q}}$ and $L_{p,i}^{(m)} = L_{p,i}$. The output of the overall channel $Q^{(m)}[\cdot]$ depicted in Figure 2 is expressed as

$$\mathbf{y}^{(m)} = \mathbf{s} * \mathbf{q}^{(m)} = G^{(m)}[\mathbf{x}^{(m)}] \tag{7}$$

where $G^{(m)}[\mathbf{x}^{(m)}] = \mathcal{T}\left[\mathbf{x}^{(m)}\right] \mathbf{g}^{(m)}$ is defined analogously to (6).

We now proceed with a derivation resembling that given in [7]. Consider any two different channels m and n. Equation (7) implies

$$\mathbf{y}^{(m)} * \mathbf{q}^{(n)} = \mathbf{s} * \mathbf{q}^{(m)} * \mathbf{q}^{(n)} = \mathbf{s} * \mathbf{q}^{(n)} * \mathbf{q}^{(m)} = \mathbf{y}^{(n)} * \mathbf{q}^{(m)} \tag{8}$$

Note that if the linearizing filters $\mathbf{g}_{p,i}$ were known, then the results from [7] could be directly applied to identify the linearized channels represented by $\mathbf{q}^{(m)}$. In the problem at hand, we need to find not only the vectors $\mathbf{q}^{(m)}$ but also the vectors $\mathbf{g}^{(m)}$.

Write (7) in matrix form as follows

$$\begin{aligned} \mathbf{y}^{(m)} &= T_{L_{\mathbf{q}}}[\mathbf{s}]\mathbf{q}^{(m)} \\ &= \mathcal{T}[\mathbf{x}^{(m)}]\mathbf{g}^{(m)} \\ &= \sum_{i,p=1,1}^{P,D_p} \mathbf{x}_{p,i}^{(m)} * \mathbf{g}_{p,i}^{(m)} \end{aligned} \tag{9}$$

By using (9) in conjunction with (8) we find identification equations that define the vectors $\mathbf{g}^{(m)}$. Denote $\mathbf{q}_{p,i}^{(m,n)} = \mathbf{g}_{p,i}^{(m)} * \mathbf{q}^{(n)}$ and expand the left hand side of (8) to obtain

$$\begin{aligned} \mathbf{y}^{(m)} * \mathbf{q}^{(n)} &= \left(\sum_{i,p=1,1}^{P,D_p} \mathbf{x}_{p,i}^{(m)} * \mathbf{g}_{p,i}^{(m)} \right) * \mathbf{q}^{(n)} \\ &= \sum_{i,p=1,1}^{P,D_p} \mathbf{x}_{p,i}^{(m)} * \mathbf{q}_{p,i}^{(m,n)} \\ &= \mathcal{T}\left[\mathbf{x}^{(m)}\right] \mathbf{q}^{(m,n)} \end{aligned} \tag{10}$$

where $\mathbf{q}^{(m,n)} = \left[(\mathbf{q}_{1,1}^{(m,n)})^T \cdots (\mathbf{q}_{P,D_P}^{(m,n)})^T\right]^T$. We can now rewrite (8) as

$$\mathcal{T}[\mathbf{x}^{(m)}]\mathbf{q}^{(m,n)} = \mathcal{T}[\mathbf{x}^{(n)}]\mathbf{q}^{(n,m)} \tag{11}$$

Let $\mathcal{X}^{(m,n)} = [\mathcal{T}[\mathbf{x}^{(m)}] \vdots -\mathcal{T}[\mathbf{x}^{(n)}]]$. Equation (11) implies that the vector $[(\mathbf{q}^{(m,n)})^T(\mathbf{q}^{(n,m)})^T]^T$ is in the null space of $\mathcal{X}^{(m,n)}$.

If the null space of $\mathcal{X}^{(m,n)}$ is rank 1, then we can determine all the vectors $\mathbf{q}^{(m,n)}$ up to a constant multiplicative factor. As we shall see, a necessary condition for a rank 1 null space is that the polynomials $\mathbf{q}_{p,i}^{(m,n)}(z)$ have no common zeros over m, n. Assuming for the moment that this is indeed the case, we can write the polynomial $\mathbf{q}_{p,i}^{(m,n)}(z)$ as $\mathbf{q}_{p,i}^{(m,n)}(z) = \mathbf{g}_{p,i}^{(m)}(z)\mathbf{q}^{(n)}(z)$. In order to perform this factorization of $\mathbf{q}^{(m,n)}(z)$ we take $\mathbf{g}_{p,i}^{(m)}(z)$ as the common factor of the polynomials $\mathbf{q}_{p,i}^{(m,1)}(z)$ through $\mathbf{q}_{p,i}^{(m,M)}(z)$. Similarly we find $\mathbf{q}^{(n)}(z)$ by finding the common factor of $\mathbf{q}_{1,1}^{(1,n)}(z)$ through $\mathbf{q}_{P,D_P}^{(M,n)}(z)$.

The procedure developed here also applies to identification of linear IIR channels. The special case of a Volterra filter whit $P=1$ and $D_P=1$ is a linear FIR filter with impulse response vector $\mathbf{g}_{1,1}$. Since Figure 2 implies that $\mathbf{h}^{(m)}(z)\mathbf{g}_{1,1}^{(m)}(z) = \mathbf{q}^{(m)}(z)$ where $\mathbf{q}^{(m)}(z)$ is also FIR, we conclude that

$$\mathbf{h}^{(m)}(z) = \frac{\mathbf{q}^{(m)}(z)}{\mathbf{g}_{1,1}^{(m)}(z)} \tag{12}$$

That is, the zeros of $\mathbf{g}_{1,1}^{(m)}(z)$ are the poles of the channel, while the zeros of $\mathbf{q}^{(m)}(z)$ are the zeros of the channel.

4. IDENTIFIABILITY CONDITIONS

Due to space constraints we present only rudimentary identifiability results. More extensive results will be presented in a future publication. The results given here are an extension of those in [7].

Theorem 4.1 *The multichannel system $H^{(m)}[\cdot]$ for $M \geq 2$ can be uniquely linearized and identified by solving the equations*

$$\mathcal{X}^{(m,n)}[(\mathbf{q}^{(m,n)})^T(\mathbf{q}^{(n,m)})^T]^T = 0 \quad (1 \leq m \leq n \leq M) \tag{13}$$

iff 1)assumption 2.1 holds, 2) the matrices $\mathcal{X}^{(m,n)}$ are rank $\left(\sum_{p=1}^{P}\sum_{i=1}^{D_p}(L_{p,i}+1)\right) - 1$ (i.e., the $\mathcal{X}^{(m,n)}$ have rank 1 null space) and, 3) There is a subset of $\mathbf{g}_{p,i}^{(m)}(z)$ of size greater or equal to 2 that is coprime (or have no common roots) with respect to p, i, m.

Proof: By construction, $[(\mathbf{q}^{(m,n)})^T(\mathbf{q}^{(n,m)})^T]^T$ is in the null space of $\mathcal{X}^{(m,n)}$. The rank condition assures uniqueness up to a common multiplicative constant. Hence, the only problem left is that of uniquely factoring $\mathbf{q}_{p,i}^{(m,n)}(z)$ in the form of the product $\mathbf{q}_{p,i}^{(m,n)}(z) = \mathbf{g}_{p,i}^{(m)}(z)\mathbf{q}^{(n)}(z)$. This is guaranteed by the third assumption. The common roots of $\mathbf{q}_{p,i}^{(m,n)}(z)$ corresponding to the m, p, and i of the subset in assumption 3 are exactly the roots of $\mathbf{q}^{(n)}(z)$.

Note that the ability to uniquely factorize $\mathbf{q}_{p,i}^{(m,n)}(z)$ stems from both the redundancy introduced by having multiple channels (m) and the redundancy introduced by the DCS representation of the equalizer (p, i). The latter represents a new form of redundancy in the context of blind equalization. Clearly, in the special case of blind linear IIR channel identification the redundancy is introduced solely by the multiple channel construction.

Theorem 4.1 gives the general conditions for unique channel and equalizer identifiability. More explicit theorems, in the spirit of [7], are needed to provide further insight into the requirements on the channels and the input signal. One such theorem is stated next without proof. The proof is similar in spirit to that for the equivalent theorem in [7].

Theorem 4.2 *The channels cannot be uniquely identified if, given p, i, there is a common root shared by $\mathbf{q}_{p,i}^{(m,n)}$ for $1 \leq m, n \leq M$.*

This implies that a necessary condition for unique identifiability is that there be no common root for $\mathbf{q}^{(m)}$ over $1 \leq m \leq M$. Moreover, theorem 4.2 in conjunction with the third condition of theorem 4.1 implies no common roots may exist for $\mathbf{g}_{p,i}^{(m)}$ over $1 \leq m \leq M$.

Note that while the conditions for identifiability are stated in terms of the $\mathbf{g}_{p,i}^{(m)}$ and $\mathbf{q}^{(m)}$, the actually represent conditions on the nonlinear channels $H^{(m)}[\cdot]$. This follows from the fact that $\mathbf{g}_{p,i}^{(m)}$ and $\mathbf{q}^{(m)}$ define the inverse of the channel operators $H^{(m)}[\cdot]$.

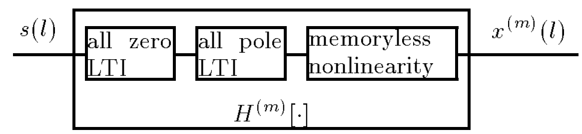

Figure 3: Block diagram of unknown channels in example.

5. NUMERICAL RESULTS

5.1. Example 1

To illustrate the proposed algorithm we solve a simple example of nonlinear channel linearization and identification. Throughout this example we use zero and pole information to describe linear filters since the linearization and identification are accurate up to a common multiplicative factor.

Assume there are 3 channels of the form depicted in Figure 3. The zeros, poles and nonlinearity for each channel are specified in Table 1. The input sequence used here is from the alphabet $[-5,-3,-1,1,3,5]$ with equiprobable symbols.

It is obvious by inspection that there exists a Volterra filter of polynomial order 3 and memory 2 that serves as

m	zeros	poles	nonlinearity
1	0.5	$-0.25 \pm 0.37i$	Inverse of
2	0.4	$-0.25 \pm 0.3i$	operator
3	0.35	$-0.24 \pm 0.25i$	$(\cdot) + (\cdot)^3$

Table 1: Channel parameters in example.

m	n	p	i	zeros		
1	2	1	1	0.4	$-0.25 - 0.37i$	$-0.25 + 0.37i$
1	2	3	1	0.4	$-0.25 - 0.37i$	$-0.25 + 0.37i$
2	1	1	1	0.5	$-0.25 - 0.3i$	$-0.25 + 0.3i$
2	1	3	1	0.5	$-0.25 - 0.3i$	$-0.25 + 0.3i$
1	3	1	1	0.35	$-0.25 - 0.37i$	$-0.25 + 0.37i$
1	3	3	1	0.35	$-0.25 - 0.37i$	$-0.25 + 0.37i$
3	1	1	1	0.5	$-0.24 - 0.25i$	$-0.24 + 0.25i$
3	1	3	1	0.5	$-0.24 - 0.25i$	$-0.24 + 0.25i$
2	3	1	1	0.35	$-0.25 - 0.3i$	$-0.25 + 0.3i$
2	3	3	1	0.35	$-0.25 - 0.3i$	$-0.25 + 0.3i$
3	2	1	1	0.4	$-0.24 - 0.25i$	$-0.24 + 0.25i$
3	2	3	1	0.4	$-0.24 - 0.25i$	$-0.24 + 0.25i$

Table 2: Zeros of $\mathbf{q}_{p,i}^{(m,n)}(z)$.

an exact linearizer. Specifically, the Volterra filter is constructed by the cascade connection of the memoryless nonlinearity $f(u) = u + u^3$ followed by a linear filter of memory 2 whose zeros are equal to the poles in the channel. Indeed, when implementing the above algorithm for this example, ($L_{p,i} = 3$ for all p, i,) we get, in this noiseless case, the zeros of the filters $\mathbf{q}_{p,i}^{(m,n)}$ corresponding to the expected zeros and poles of the "unknown" channel, as summarized in Table 2. An input of less than 20 symbols is more than sufficient to identify the channels. By appropriately grouping the zeros in Table 2 we determine the DCS filters $\mathbf{g}_{p,i}^{(m)}$ of the linearizer and the linear FIR filters $\mathbf{q}^{(n)}$ of the linearized channels. For example, the only zeros common to rows corresponding to $m = 2$, $p = 3$, and $i = 1$ are $-0.25 \pm 0.3i$. These are the zeros of $\mathbf{g}_{3,1}^{(2)}(z)$. We find that in this case $\mathbf{g}_{1,1}^{(2)}(z) = \mathbf{g}_{3,1}^{(2)}(z)$ and indeed

$$x_{1,1}^{(2)}(z)\mathbf{g}_{1,1}^{(2)}(z) + x_{3,1}^{(2)}(z)\mathbf{g}_{3,1}^{(2)}(z) = s(z)\mathbf{q}^{(2)}(z) \qquad (14)$$

where $\mathbf{q}^{(2)}(z)$ is the first order polynomial with zero 0.4. We notice that $\mathbf{q}^{(2)}(z)$ corresponds to the unique common zero in all the rows where $n = 2$, as expected.

5.2. Example 2

To illustrate the robustness of the algorithm we add noise and consider nonlinear channels which cannot be exactly equalized by a finite order Volterra filter. The channel poles and zeros are the same as in Example 1 and the memoryless nonlinearity is $w(u) = u - 0.1u^3$. By inspection it is clear that the inverse operator has infinite polynomial order. However, we will approximately equalize the channels

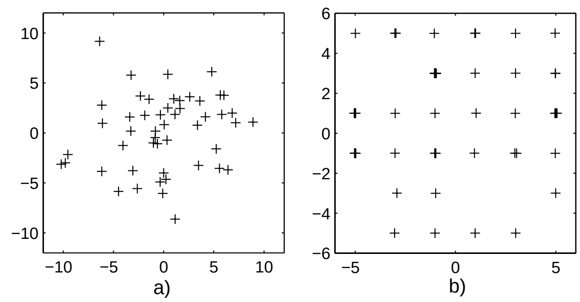

Figure 4: Constellations for Example 2: a) Before equalization. b) After equalization.

using 3^{rd} order Volterra filters as in the previous example. The channel outputs are corrupted by independent white Gaussian noises, that are independent of the input. The signal to noise ratio is 50 dB. In this example the input symbols are complex and equiprobable with alphabet $[-5, -3, -1, 1, 3, 5] \times i[-5, -3, -1, 1, 3, 5]$.

In order to factorize $\mathbf{q}_{p,i}^{(m,n)}(z)$ we first round the estimated roots to the second decimal place, then search for the common roots.

The results of equalization using only 50 input symbols is shown in Figure 4. If we project each equalized symbol to the closest symbol in the alphabet, then the equalized symbols exactly match the input symbols.

6. REFERENCES

[1] G. B. Giannakis, E. Serpedin, "Linear Multichannel Blind Equalizers of Nonlinear FIR Volterra Channels," *IEEE Trans. Sig. Proc.,* vol. 45, no. 1, pp. 67-81, Jan 1997.

[2] S. Prakriya, D. Hatzinakos, "Blind identification of LTI-ZMNL-LTI nonlinear channel models," *IEEE Trans. Signal Proc.,* vol. 43, pp. 3007-3013, Dec 1995.

[3] G. M. Raz, B. D. Van Veen, "Baseband Volterra filters for implementing carrier based nonlinearities," *IEEE Trans. Signal Proc.,* to appear, Jan 1998.

[4] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems.* New York: Wiley, 1980.

[5] M. K. Tsatsanis, H. A. Cirpan, "Blind Identification of Nonlinear Channels Excited by Discrete Alphabet Inputs," *SSAP Proceedings of the 1996 8'th IEEE Signal Processing Workshop on Statistical Signal and Array Processing,* pp. 176-179 Jun 1996.

[6] B. D. Van Veen, R. D. Nowak, G. M. Raz, "Diagonal Coordinate System Approximations for Volterra Filters," *Signal Processing,* to appear.

[7] G. Xu, H. Liu, L. Tong, T. Kailath, "A Least-Squares Approach to Blind Channel Identification," *IEEE Trans. Sig. Proc.,* vol. 43, no. 12, Dec 1995.