NONLINEAR CHANNEL IDENTIFICATION AND PERFORMANCE ANALYSIS WITH PSK INPUTS

G. Tong Zhou[1] *Georgios B. Giannakis*[2]

[1]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250, USA
[2]Department of Electrical Engineering, University of Virginia, Charlottesville, VA 22903-2442, USA

ABSTRACT

Many real life systems are nonlinear. We focus on the so-called Volterra models and the input-output aspect of the identification problem. To this date, closed form solutions for the Volterra kernels exist only for Gaussian inputs and they are usually in rather complicated forms. In this paper, we adopt a class of non-Gaussian inputs, namely the PSK (phase shift keying) sequences, which are common in digital communications systems. Such signals allow for simple *closed form* solutions of the Volterra kernels of *any* order. Kernels are estimated *separately* thus preventing error propagation. Signal independent zero-mean additive noise can also be tolerated even when its color and distribution are unknown. Because closed form solutions for the kernels are available, explicit variance expressions for their estimates can also be derived. Simulation results verify our theoretical findings.

1. INTRODUCTION

Nonlinear system identification is a subject of considerable importance and interest due to the prevalence of nonlinearities in real life problems, such as satellite communication links [1], magnetic recording channels [2], physiological systems [8], and plasma physics [6] (see e.g., [9] for a review). Nonlinearities can be modeled implicitly (such as with differential equations) or explicitly, and our interest lies in the latter. A broad class of nonlinear systems have polynomial functional representations, for which the Volterra and the Wiener representations are two examples [11].

An Lth-order discrete-time Volterra nonlinear system relates the output $x(t)$ and the input $w(t)$ via multidimensional convolutions, which are straightforward generalizations of the 1-D convolution:

$$\begin{aligned} x(t) &= h_0 + \sum_{l=1}^{L} \sum_{u_1,u_2,\ldots,u_l} h_l(u_1,u_2,\ldots,u_l) \\ &\times w(t-u_1)w(t-u_2)\ldots w(t-u_l) + v(t). \quad (1) \end{aligned}$$

In (1), h_0 is the dc term, $v(t)$ is the additive noise assumed to be zero-mean and independent of $w(t)$, and the discrete time index is $t = 0, 1, \ldots, T-1$. The l-dimensional functional $h_l(u_1,\ldots,u_l)$ is called the lth-order Volterra kernel and unless it is a delta function for all l's, the nonlinear process in (1) has memory. We also allow for complex input and complex kernels – a setup encountered with narrowband communication signals passing through bandpass Volterra channels (see [1, pp. 58-61]). In this case, $x(t)$ in (1) denotes the baseband representation of the channel response.

When the input $w(t)$ is zero-mean Gaussian, Koukoulas and Kalouptsidis have recently shown that the Volterra kernels can be found in closed form from the cross cumulants between the input and the output [7]. Their solutions are non-trivial and due to their sequential nature, error in one kernel estimate propagates to the next. However, two special cases are free of such shortcomings: (i) If the nonlinear system is lth-order homogeneous (i.e., the sum over l is absent in (1)), then we obtain Brillinger's formula [3]. (ii) If there are only two Volterra kernels h_{l_1} and h_{l_2} with $l_1 + l_2 =$odd. A special case of this is when $L = 2$, $l_1 = 1$, $l_2 = 2$, and Tick's method [12] can be derived. Other than (i) and (ii), it is in general difficult to obtain non-recursive closed-form expressions for the Volterra kernels.

In the Wiener theory of nonlinear systems, the polynomial nonlinearity is expressed in a different form, where the functionals are made orthogonal to each other through a Gram-Schmidt orthogonalization procedure, with respect to zero-mean white Gaussian input $w(t)$. The orthogonality property facilitates closed form solutions and separates kernel estimation through cross correlations [11]. However, the model becomes increasingly complex with larger L and the theory does not apply to non-Gaussian inputs.

In Volterra nonlinear system identification with stochastic inputs, a lot of attention has been given to Gaussian inputs because their higher-order moments are tractable. In this work, a class of complex valued non-Gaussian inputs, namely the PSK (phase shift keying) sequences, are shown to have many vanishing moments and hence can be used as an effective probing signal as well. Moreover, it turns out that with these sequences, the Volterra kernels are made orthogonal to each other, they are estimated separately, and the resulting solutions have very simple forms. Although similar observations have been made in [5], the kernel and variance expressions presented here are novel.

2. CHARACTERISTICS OF THE INPUT

Our input signal of choice has the following characteristics:

> $w(t)$ is i.i.d., drawn from a discrete alphabet set $\{r\ e^{j(2\pi k/M+\theta)}\}_{k=0}^{M-1}$, $M > 1$, with probability $1/M$ each. The constant modulus r and the angle θ are deterministic constants.

Regarding this class of inputs, we have the following key result:

Lemma 1 *The $(m+n)$th-order moment of w, where m copies of w and n copies of w^* are used, is*

$$E[w^m (w^*)^n] = r^{m+n}\ e^{j(m-n)\theta}\ \delta((m-n) \bmod M), \quad (2)$$

where $\quad \delta((m-n) \bmod M)$

$$= \begin{cases} 1, & \textit{if } (m-n) \textit{ is an integer multiple of } M, \\ 0, & \textit{otherwise.} \end{cases} \quad (3)$$

Proof:

$$E[w^m(w^*)^n] = r^{m+n}\, e^{j(m-n)\theta}\, \frac{1}{M}\sum_{k=0}^{M-1} e^{j2\pi k(m-n)/M} = r^{m+n}\, e^{j(m-n)\theta}\, \delta((m-n) \bmod M).$$

Note that our input is "rich enough" in frequencies (due to its whiteness), and with $M > L$ it is also "rich enough" in distinct values; hence, it is persistently exciting (see e.g., [5], [10]) for identification of Lth-order Volterra models.

3. VOLTERRA KERNEL ESTIMATION

Closed-form solutions of arbitrary Volterra kernels are derived here using the cross correlation between the input and the output. First, we shall recast the standard model (1) in a slightly different but more convenient form.

3.1. Symmetrized Volterra expansion

We assume w.l.o.g. that the lth-order Volterra kernel is symmetric in its arguments; e.g., $h_2(u_1,u_2) = h_2(u_2,u_1)$, because we can symmetrize the kernel if otherwise [11, p. 80-81]. Let us denote the number of distinguishable permutations of the l-tuple $(u_1,\ldots,u_l)$ by $P_l(u_1,\ldots,u_l)$. If $u_1 \neq u_2 \neq \ldots \neq u_l$, then $P_l = l!$. But when all u_i's are identical, we have $P_l = 1$. Therefore, P_l depends on how different the u_i's are. Let us group the u_i's into $(n+1)$ subsets where the u_i's are identical within each subset but are different between subsets. Suppose that the total number of u_i's in each subset is

$$m_1,\ m_2,\ \ldots,\ m_n,\ \ l-(m_1+m_2+\ldots+m_n).$$

It follows then that the number of symmetries in $h_l(u_1,\ldots,u_l)$ is simply the multinomial coefficient [5],

$$P_l(u_1,\ldots,u_l) = \binom{l}{m_1,m_2,\ldots,m_n}.$$

Next, let us use $\mathcal{R}_l$ to denote the non-redundant region of the $h_l(u_1,\ldots,u_l)$ kernel; i.e., $\mathcal{R}_l = \{(u_1,\ldots,u_l)|\ 0 \le u_1 \le u_2 \le \ldots \le u_l\}$. Many identical terms in (1) can be grouped together and hence the Volterra model can be expressed in a slightly different form,

$$x(t) = h_0 + \sum_{l=1}^{L}\ \sum_{(u_1,\ldots,u_l)\in\mathcal{R}_l} P_l(u_1,\ldots,u_l)h_l(u_1,\ldots,u_l) \times\ w(t-u_1)w(t-u_2)\ldots w(t-u_l) + v(t). \quad (4)$$

3.2. Closed form solutions

Using Lemma 1, we will establish the following fact:

$$E[w(t-u_1)\ldots w(t-u_l)] = 0, \qquad \forall\, 1 \le l \le L < M. \quad (5)$$

Since $1 \le l < M$ is not divisible by M, we have $E[w^l] = 0$ according to (2). This is true for (5) when all the u_i's are identical. If some of the u_i's are different, then due to independence, the lth-order moment in (5) splits and can be expressed as a product of lower order moments. All of these lower order moments are zero because their orders are still in $[1, M)$. Therefore, (5) always holds as long as $M > L$.

Because of (5) and the fact that $v(t)$ is zero-mean, we infer from (4) that $E[x(t)] = h_0$ for $M > L$ (if $v(t)$ has unknown nonzero mean, h_0 is not identifiable); i.e., the dc component can be obtained as the mean of the output data provided that $E[v(t)]$ is known. Next, let us consider the cross correlation (moment) between $x(t)$ and $1 \le k \le L$ conjugated and lagged copies of $w(t)$,

$$m_{xw\ldots w}(\tau_1,\ldots,\tau_k) \triangleq E[x(t)w^*(t-\tau_1)\ldots w^*(t-\tau_k)]. \quad (6)$$

Since the above moment is symmetric w.r.t. the τ's; e.g., $m_{xww}(\tau_1,\tau_2) = m_{xww}(\tau_2,\tau_1)$, we can assume w.l.o.g. that $\tau_1 \le \tau_2 \le \ldots \le \tau_k$.

Similar to (5), we find

$$h_0\ E[w^*(t-\tau_1)\ldots w^*(t-\tau_k)] = 0, \quad 1 \le k \le L < M, \quad (7)$$

and

$$E[v(t)w^*(t-\tau_1)\ldots w^*(t-\tau_k)] = E[v(t)]\ E[w^*(t-\tau_1)\ldots w^*(t-\tau_k)] = 0. \quad (8)$$

Therefore, the cross correlation becomes (c.f. (6)-(8)),

$$m_{xw\ldots w}(\tau_1,\ldots,\tau_k) = \sum_{l=1}^{L}\ \sum_{(u_1,\ldots,u_l)\in\mathcal{R}_l} P_l(u_1,\ldots,u_l)h_l(u_1,\ldots,u_l) \times\ E[w(t-u_1)\ldots w(t-u_l)w^*(t-\tau_1)\ldots w^*(t-\tau_k)]. \quad (9)$$

From $1 \le l$ and $k \le L$, we infer that $1-L \le (l-k) \le L-1$ is divisible by $M > L$ only when $l = k$. Hence, only the $l = k$ term survives in (9) and we can drop $\sum_{l=1}^{L}$ to obtain

$$m_{xw\ldots w}(\tau_1,\ldots,\tau_k) = \sum_{(u_1,\ldots,u_k)\in\mathcal{R}_k} P_k(u_1,\ldots,u_k)h_k(u_1,\ldots,u_k) \times\ E[w(t-u_1)\ldots w(t-u_k)w^*(t-\tau_1)\ldots w^*(t-\tau_k)]. \quad (10)$$

Equation (10) explains how we isolate the Volterra kernels of different orders by using PSK inputs and the appropriate cross correlation.

Although there are k unconjugated and k conjugated w's in the moment on the r.h.s. of (10), it will be non-zero only when we can equate every u_m with one τ_n; otherwise, unmatched w^* and w's factor out (due to independence) and the resulting lower order moments are zero according to Lemma 1. However, since we have assumed $u_1 \le u_2 \le \ldots \le u_k$ and $\tau_1 \le \tau_2 \le \ldots \le \tau_k$, this happens only when $u_i = \tau_i, \forall\, i = 1,\ldots,k$, in which case $E[w(t-u_1)\ldots w(t-u_k)$ $w^*(t-\tau_1)\ldots w^*(t-\tau_k)]$ equals $E\prod_{i=1}^{k}|w(t-u_i)|^2 = r^{2k}$. Therefore we can drop the sum over $\mathcal{R}_k$ in (10) and obtain a surprisingly simple formula linking the $(k+1)$st-order cross correlation and the kth-order Volterra kernel,

$$m_{xw\ldots w}(\tau_1,\ldots,\tau_k) = \left[r^{2k}\ P_k(u_1,\ldots,u_k)\right]\ h_k(u_1,\ldots,u_k),$$

$$\text{or,} \qquad h_k(u_1,\ldots,u_k) = \frac{m_{xw\ldots w}(\tau_1,\ldots,\tau_k)}{r^{2k}\ P_k(u_1,\ldots,u_k)}. \quad (11)$$

Consistent sample estimate of the cross correlation in (6) can be obtained from a single record of input/output data as usual:

$$\widehat{m}_{xw\ldots w}(\tau_1,\ldots,\tau_k) = \frac{1}{T}\sum_{t=0}^{T-1} x(t)w^*(t-\tau_1)\ldots w^*(t-\tau_k). \quad (12)$$

Remarks: (i) Our solution (11) is in simple closed form and only requires $M > L$. For parsimony, there should be an upper bound on the highest nonlinearity order for the Volterra model to be practical, and M can be taken above that upper bound. (ii) We do not impose *any* assumption on the kernel itself such as its memory or zero locations. (iii) Our algorithm tolerates additive noise of any color and distribution as long as it is independent of the input. (iv) The solution is linear and can be used as an initial estimate for other optimum nonlinear schemes such as the maximum likelihood method. (v) The kernels are estimated separately thus preventing error propagation. (vi) Eq. (11) depends on the constant modulus r but is independent of other constellation parameters such as the size M and the angle θ. However, the condition $M > L$ must be met.

The symmetry number $P_k(u_1, \ldots, u_k)$ can be computed beforehand. Specific results for the first- through fourth-order Volterra kernels are given below. All formulas assume that $M > L$.

$$h_1(\tau) = (1/r^2)\, m_{xw}(\tau), \tag{13}$$

$$h_2(\tau_1, \tau_2) = \begin{cases} (1/2r^4)\, m_{xww}(\tau_1, \tau_2), & \tau_1 \neq \tau_2, \\ (1/r^4)\, m_{xww}(\tau_1, \tau_2), & \tau_1 = \tau_2, \end{cases} \tag{14}$$

$$h_3(\tau_1, \tau_2, \tau_3) = \begin{cases} (1/6r^6)\, m_{xwww}(\tau_1, \tau_2, \tau_3), & \tau_1 \neq \tau_2 \neq \tau_3, \\ (1/3r^6)\, m_{xwww}(\tau_1, \tau_2, \tau_3), & (\tau_1 = \tau_2) \neq \tau_3, \\ (1/r^6)\, m_{xwww}(\tau_1, \tau_2, \tau_3), & \tau_1 = \tau_2 = \tau_3, \end{cases} \tag{15}$$

$$h_4(\tau_1, \tau_2, \tau_3, \tau_4) = \begin{cases} (1/24r^8) m_{xwwww}(\tau_1, \tau_2, \tau_3, \tau_4), \tau_1 \neq \tau_2 \neq \tau_3 \neq \tau_4, \\ (1/12r^8) m_{xwwww}(\tau_1, \tau_2, \tau_3, \tau_4), (\tau_1 = \tau_2) \neq \tau_3 \neq \tau_4, \\ (1/6r^8) m_{xwwww}(\tau_1, \tau_2, \tau_3, \tau_4), (\tau_1 = \tau_2) \neq (\tau_3 = \tau_4), \\ (1/4r^8) m_{xwwww}(\tau_1, \tau_2, \tau_3, \tau_4), \tau_1 \neq (\tau_2 = \tau_3 = \tau_4), \\ (1/r^8) m_{xwwww}(\tau_1, \tau_2, \tau_3, \tau_4), \tau_1 = \tau_2 = \tau_3 = \tau_4. \end{cases} \tag{16}$$

We emphasize that as long as $M > L$ is satisfied, these formulas are valid irrespective of the other nonlinear kernels present in the system. For example, if a QPSK ($M = 4$) input is used, then whether the system is linear ($L = 1$), linear-quadratic ($L = 2$), or linear-quadratic-cubic ($L = 3$), the first-order kernel is always estimated as in (13).

4. PERFORMANCE ANALYSIS

Sample kernel estimates are obtained by substituting (12) into (11). For example, the first-order kernel estimate is

$$\widehat{h}_1(\tau) \triangleq \frac{1}{Tr^2} \sum_{t=0}^{T-1} x(t) w^*(t-\tau), \qquad M \geq L. \tag{17}$$

Because we can express $\widehat{h}_k$ explicitly in terms of the true kernels and the input, it is possible to derive its variance expression. However, the task turns out to be quite tedious and so far we have derived the variance expression of (17) for linear-quadratic systems only. Additional variance expressions will be reported in [13].

Substituting (1) with $L = 2$ and $h_0 = 0$ into (17), we obtain

$$\widehat{h}_1(\tau) = \frac{1}{Tr^2} \sum_{t=0}^{T-1} \sum_u h_1(u) w^*(t-u) w(t-\tau) \tag{18}$$

$$+ \frac{1}{Tr^2} \sum_{t=0}^{T-1} \sum_{u_1, u_2} h_2(u_1, u_2) w^*(t-u_1) w^*(t-u_2) w(t-\tau) \tag{19}$$

$$+ \frac{1}{Tr^2} \sum_{t=0}^{T-1} v^*(t) w(t-\tau). \tag{20}$$

Let us denote (18) as $A(\tau)$, (19) as $B(\tau)$, and (20) as $C(\tau)$. The variance of $\widehat{h}_1(\tau)$ is then

$$\begin{aligned} \mathrm{var}\{\widehat{h}_1(\tau)\} \;=\; & \mathrm{cum}\{A(\tau), A^*(\tau)\} \\ & + \mathrm{cum}\{B(\tau), B^*(\tau)\} + \mathrm{cum}\{C(\tau), C^*(\tau)\} \\ & + \mathrm{cum}\{A(\tau), B^*(\tau)\} + \mathrm{cum}\{A^*(\tau), B(\tau)\} \\ & + \mathrm{cum}\{A(\tau), C^*(\tau)\} + \mathrm{cum}\{A^*(\tau), C(\tau)\} \\ & + \mathrm{cum}\{B(\tau), C^*(\tau)\} + \mathrm{cum}\{B^*(\tau), C(\tau)\}. \end{aligned}$$

By the multi-linearity property of cumulants [4], we find

$$\mathrm{cum}\{A(\tau), A^*(\tau)\} = \frac{1}{T^2 r^4} \sum_{t_1=0}^{T-1} \sum_{t_2=0}^{T-1} \sum_u \sum_v h_1(u) h_1(v)$$
$$\times\, \mathrm{cum}\{w^*(t_1-u) w(t_1-\tau), w(t_2-v) w^*(t_2-\tau)\}.$$

Next, we use the Leonov-Shiryaev formula [4] to express

$$\begin{aligned} & \mathrm{cum}\{w^*(t_1-u) w(t_1-\tau), w(t_2-v) w^*(t_2-\tau)\} \\ = \;& \mathrm{cum}\{w^*(t_1-u), w(t_1-\tau), w(t_2-v), w^*(t_2-\tau)\} \\ & + \mathrm{cum}\{w^*(t_1-u), w(t_2-v)\}\; \mathrm{cum}\{w(t_1-\tau), w^*(t_2-\tau)\} \\ = \;& -r^4\, \delta(u = v = \tau, t_2 = t_1) + r^4\, \delta(t_1 = t_2, u = v). \end{aligned}$$

Therefore, $\mathrm{cum}\{A(\tau), A^*(\tau)\} = T^{-1} \sum_{u \neq \tau} h_1^2(u)$. Other cumulant terms in $\mathrm{var}\{\widehat{h}_1(\tau)\}$ can be found similarly. The final variance of (17) is given by

$$\begin{aligned} \mathrm{var}\{\widehat{h}_1(\tau)\} = \; & \frac{\sigma_v^2}{Tr^2} + \frac{1}{T} \sum_{u \neq \tau} h_1^2(u) + \frac{2r^2}{T} \sum_{u_1, u_2} h_2^2(u_1, u_2) \\ & - \frac{r^2}{T} \sum_u h_2^2(u, u) + \frac{4r^2}{T} \left[\sum_u h_2(u, \tau) \right]^2 \\ & - \frac{4r^2}{T} \sum_u h_2^2(u, \tau) - \frac{4r^2}{T}\, h_2(\tau, \tau) \sum_{u \neq \tau} h_2(u, \tau), \end{aligned} \tag{21}$$

which is seen to depend on the data length T, the lag τ, the variance of the additive noise σ_v^2, the constellation parameter r, and all the true kernels. The estimator is consistent.

5. SIMULATIONS

We present here three numerical examples to illustrate the results developed in this paper.

Example 1: Suppose that there is an unknown linear-quadratic-cubic ($L = 3$) nonlinearity and we use QPSK symbols $\{-1-j,\ -1+j,\ 1-j,\ 1+j\}$ as the input which corresponds to $M = 4$, $r = \sqrt{2}$, and $\theta = \pi/4$. Table I shows the true kernels (h_2 and h_3 in their non-redundant regions), the mean and the standard deviation (std) of the sample estimates computed from (13)-(15) using $T = 4,096$ input/output data and 100 independent realizations. The parameter estimates are seen to be fairly accurate for the given data length, considering the fact that there are 13 unknowns and higher-order moments are used.

Example 2: We have here a 4th-order homogeneous nonlinear system. Additive noise $v(t)$ is present which is zero-mean white Gaussian with variance 0.5. A PSK sequence with $M = 5, r = 2$, and $\theta = 0$ is used as the input. Table II shows the non-redundant 4th-order kernels, the mean and the std of the sample estimates computed from (16) using $T = 4,096$ input/output data and 100 independent realizations. The estimates are seen to be fairly accurate despite of the high nonlinearity order and the presence of additive noise.

Example 3: The input here is QPSK and the system has linear and quadratic nonlinearities. The 1st- and 2nd-order Volterra kernels are identical to those of Example 1. Additive noise is also present which is zero-mean colored Gaussian with variance $\sigma_v^2 = 0.5$. We used $T = 8,192$ data and 200 independent realizations to verify (21). Table III shows the theoretical variance of $\widehat{h}_1(\tau)$ multiplied by T and the corresponding sample variance multiplied by T. Close agreement between the empirical and the theoretical values is observed.

TABLE I: RESULTS FOR EXAMPLE 1

τ	0	1	2	3	4
true $h_1(\tau)$	1.0000	0.5000	-0.8000	1.6000	0.4000
mean of $\widehat{h}_1(\tau)$	0.9992	0.5095	-0.8030	1.5759	0.3743
std of $\widehat{h}_1(\tau)$	0.1723	0.0970	0.1383	0.1299	0.1414

(τ_1, τ_2)	(0,0)	(0,1)	(1,1)
true $h_2(\tau_1, \tau_2)$	1.0000	0.6000	-0.3000
mean of $\widehat{h}_2(\tau_1, \tau_2)$	1.0049	0.6041	-0.2916
std of $\widehat{h}_2(\tau_1, \tau_2)$	0.1096	0.0891	0.0964

(τ_1, τ_2, τ_3)	(0,0,0)	(0,0,1)	(0,1,1)	(1,1,1)	(0,1,2)
true $h_3(\tau_1, \tau_2, \tau_3)$	1.0000	1.2000	0.8000	-0.5000	0.6000
mean of $\widehat{h}_3(\tau_1, \tau_2, \tau_3)$	1.0094	1.1983	0.7993	-0.4979	0.5993
std of $\widehat{h}_3(\tau_1, \tau_2, \tau_3)$	0.0741	0.0201	0.0265	0.0917	0.0093

TABLE II. RESULTS FOR EXAMPLE 2

$(\tau_1, \tau_2, \tau_3, \tau_4)$	(0,0,0,0)	(0,0,0,1)	(0,0,1,1)	(0,0,1,2)	(0,1,2,3)
true $h_4(\tau_1, \tau_2, \tau_3, \tau_4)$	1.0000	1.2000	0.8000	-0.5000	0.6000
mean of $\widehat{h}_4(\tau_1, \tau_2, \tau_3, \tau_4)$	0.9782	1.1909	0.8041	-0.5005	0.5997
std of $\widehat{h}_4(\tau_1, \tau_2, \tau_3, \tau_4)$	0.1802	0.0484	0.0281	0.0142	0.0040

TABLE III. RESULTS FOR EXAMPLE 3

τ	0	1	2	3	4
theoretical $T\text{var}\{\widehat{h}_1(\tau)\}$	13.7200	8.2300	9.2800	7.3600	9.7600
sample $T\text{var}\{\widehat{h}_1(\tau)\}$	13.7854	8.5657	10.1553	6.5336	9.7562

6. CONCLUSIONS AND DISCUSSION

We have developed a novel algorithm for input-output nonlinear system identification. The Volterra model is adopted and PSK sequences are used as input. These sequences are discrete, fairly easy to generate, and are common for communications applications. Our solution is in closed form, non-iterative, simple to implement, and works for Volterra kernels of arbitrary order. We have illustrated our algorithms with simulations. Currently, we are working on adaptive versions of the closed forms and variance expressions of the kernel estimates of arbitrary order, to be reported in [13]. Since AM-PM sequences can be regarded as PSK symbols on concentric circles, the results presented here may be extended to the larger AM-PM class of inputs – this constitutes another interesting future research direction.

REFERENCES

[1] S. Benedetto, E. Biglieri and V. Castellani, *Digital Transmission Theory,* Prentice-Hall, New Jersey, 1987.

[2] E. Biglieri, E. Chiaberto, G.P. Maccone, and E. Viterbo, "Compensation of nonlinearities in high-density magnetic recording channels," *IEEE Trans. on Magnetics,* pp. 5079-5086, Nov. 1994.

[3] D.R. Brillinger, "The identification of polynomial systems by means of higher-order spectra," *J. Sound and Vibration,* vol. 12, pp. 301-313, 1970.

[4] D. R. Brillinger, *Time Series: Data Analysis and Theory,* Holden-day Inc., San Francisco, 1981.

[5] W.A. Gardner and T.L. Archer, "Exploitation of cyclostationarity for identifying the Volterra kernels of nonlinear systems," *IEEE Trans. on Information Theory,* vol. 39, no. 2, pp. 535-542, March 1993.

[6] Y.C. Kim and E.J. Powers, "Digital bispectral analysis and its applications to nonlinear wave interactions," *IEEE Trans. on Plasma Science,* vol. 7, pp. 120-131, 1979.

[7] P. Koukoulas and N. Kalouptsidis, "Nonlinear system identification using Gaussian inputs," *IEEE Trans. on Signal Processing,* vol. 43, pp. 1831-1841, August 1995.

[8] P.Z. Marmarelis and V.Z. Marmarelis, *Analysis of Physiological Systems - The White Noise Approach,* New York: Plenum, 1978.

[9] V.J. Mathews, "Adaptive polynomial filters," *IEEE Signal Processing Magazine,* vol. 8, pp. 10-26, July 1991.

[10] R.D. Nowak and B.D. Van Veen, "Random and pseudorandom inputs for Volterra filter identification," *IEEE Trans. on Signal Processing,* vol. 42, pp. 2124-2135, August 1994.

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

[12] L.J. Tick, "The estimation of transfer functions of quadratic systems," *Technometrics,* vol. 3, pp. 562-567, Nov. 1961.

[13] G.T. Zhou and G.B. Giannakis, "Volterra nonlinear system identification with PSK inputs: Algorithms and performance analysis," *IEEE Trans. on Signal Processing,* in preparation.