Blind Identification of Bilinear Systems

Nicholas Kalouptsidis, *Senior Member, IEEE*, Panos Koukoulas, *Member, IEEE*, and V. John Mathews, *Fellow, IEEE*

***Abstract*—This paper is concerned with the blind identification of a class of bilinear systems excited by non-Gaussian higher order white noise. The matrix of coefficients of mixed input-output terms of the bilinear system model is assumed to be triangular in this work. Under the additional assumption that the system output is corrupted by Gaussian measurement noise, we derive an exact parameter estimation procedure based on the output cumulants of orders up to four. Results of the simulation experiments presented in the paper demonstrate the validity and usefulness of our approach.**

***Index Terms*—Bilinear systems, blind identification, high-order statistics, nonlinear system identification.**

I. Introduction

IDENTIFICATION of nonlinear systems is of primary importance in today's applications since many signals of interest are generated by nonlinear sources or are processed by nonlinear systems. There are several situations in which the inherent nonlinearities and distortions cannot be tolerated at a given level of performance, and hence, nonlinear processing techniques need to be employed. Such important examples include nonlinear echo cancellation, predistortion of nonlinear channels, equalization of communication channels where distortion is produced due to operation of amplifiers near to saturation region, linearization of loudspeaker nonlinearities, enhancement of noisy images, edge extraction, distortions in magnetic recording systems, motion of moored ships in ocean waves, control of industrial processes, physiological models, nuclear fission, and others [1], [7], [12], [25], [29], [32].

Conventional identification is concerned with the determination of an unknown system on the basis of input-output information in an uncertain environment. A given excitation drives the unknown system and the resulting response is measured. On the other hand, blind identification is concerned with the determination of an unknown system on the basis of output information only. In this latter case, information about the input that generates the measured output is limited. For instance, it may be *a priori* known or assumed that the input is white noise.

The need for tractable computational methods requires that the class of nonlinear models is properly restricted. Polynomial systems form a popular class of nonlinear systems [32]. Under relatively mild conditions, such systems are known to possess the universal approximation capability [4], [13]. This class of systems is defined by input-output relationships of the form

$$y(n) = f(u(n), u(n-1), \ldots \quad u(n-N), y(n-1), \ldots, y(n-M)) \tag{1}$$

where $u(n)$ and $y(n)$ represent the input and output signals, respectively, and $f(\cdots)$ is a polynomial in $N+M+1$ variables. Polynomial systems can be broadly classified into recursive and nonrecursive systems. Nonrecursive polynomial systems are obtained from (1) if f depends only on $u(n-i)$, $i=0,1,\ldots,N$. In this case, (1) takes the form of a truncated Volterra series expansion [36]:

$$\begin{aligned} y(n) = h_0 &+ \sum_{k_1=0}^{N_1} h_1(k_1)u(n-k_1) \\ &+ \sum_{k_1=0}^{N_2}\sum_{k_2=0}^{N_2} h_2(k_1,k_2)u(n-k_1)u(n-k_2) \\ &+ \cdots + \sum_{k_1=0}^{N_P}\sum_{k_2=0}^{N_P}\cdots\sum_{k_P=0}^{N_P} h_P(k_1,k_2,\ldots,k_P) \\ &\times u(n-k_1)u(n-k_2)\ldots u(n-k_P) \end{aligned} \tag{2}$$

where $h_m(k_1,\ldots,k_m)$ represents the mth-order Volterra kernel of the system, and $\max\{N_1,\ldots,N_P\}$ represents the memory of the system.

Conventional identification of a truncated Volterra series aims at estimating the Volterra kernels from either knowledge of the relevant statistics of the input and output signals or measurement of the input and output signal. The mean squared error (MSE) formulation and the least squares error (LSE) formulation enable the computation of the Volterra kernels via a linear system of equations. Algorithms for the estimation of the parameters of Volterra models based on input-output data have been extensively studied in the past [2], [10], [13], [14], [15], [17], [27], [31]–[33], [39], [40]. Most of these methods view the resulting linear regression as a multichannel setup. The Volterra parameters are then obtained by linear multichannel parameter estimation algorithms in batch or in adaptive form. Cumulants and polyspectra are employed in [19]–[22] to estimate symmetric Volterra kernels. These works derive closed-form solutions for the estimates when the input is a stationary, Gaussian, zero mean stochastic process or a linear process. In a similar manner, the identification of Volterra systems of second and third order for general stochastic inputs is treated in [14], [20], [21], and [33]. In general, for blind identification, the output statistics depend nonlinearly on the kernels even when the system is linear. Probably because of the complexity associated with such problems,

Manuscript received January 14, 2002; revised September 12, 2002. The associate editor coordinating the review of this paper and approving it for publication was Prof. Derong Li.

N. Kalouptsidis is with the Department of Informatics and Telecommunications, Division of Communications and Signal Processing, University of Athens, Athens, Greece (e-mail: kalou@di.uoa.gr).

P. Koukoulas is with the Electronics Division, Department of Radars and Automations, Hellenic Civil Aviation Authority, Ministry of Transport and Communications, Athens, Greece (e-mail: koukoula@di.uoa.gr).

V. J. Mathews is with the Department of Electrical and Computer Engineering, University of Utah, Salt Lake City, Utah 84112 USA (e-mail: mathews@ece.utah.edu).

Digital Object Identifier 10.1109/TSP.2002.806971

little is known about the blind identification of general Volterra systems [9], [24], [35].

Nonrecursive polynomial systems such as the truncated Volterra series expansion encounters serious limitations in practical applications due to the large number of coefficients that need to be estimated. Recursive polynomial models, just like linear IIR filters, can accurately represent many nonlinear systems with greater efficiency than truncated Volterra series representation. A special class of recursive nonlinear models is the class of bilinear systems. The input-output relationship of a bilinear system is given by

$$\begin{aligned} y(n) = &\sum_{i=1}^{K_a} a(i)y(n-i) \\ &+ \sum_{i=0}^{K_b} b(i)u(n-i) \\ &+ \sum_{i=1}^{K_{cy}}\sum_{j=0}^{K_{cu}} c(i,j)y(n-i)u(n-j) \end{aligned} \tag{3}$$

where $a(i)$, $b(i)$, and $c(i,j)$ represent the system coefficients, and the set $\{K_a, K_b, K_{cy}, K_{cu}\}$ corresponds to the order of the system. Several practical systems have been modeled by bilinear systems [2], [32]. The input-output bilinear representation of (3) is not equivalent to the original state-space bilinear model

$$\begin{aligned} x_{k+1} &= Ax_k + Nu_k \otimes x_k + Bu_k + w_k \\ y_k &= Cx_k + Du_k + v_k \end{aligned} \tag{4}$$

where v and w denote the measurement and process noise. If (3) is transformed into state-space, it involves polynomial nonlinearities between state variables. Conventional identification of bilinear state-space models has been studied in [5] and [6] using subspace identification methods. The more general class of state-affine systems, which provide finite dimensional realization of Volterra systems with separable kernels, has been treated in [8] using cumulants.

Conventional identification methods for input-output bilinear systems fall into equation error and output error methods. Equation error algorithms are straightforward to develop, and the mean square estimation error surface has a unique minimum. However, this unique minimum is, in general, biased. Output error algorithms are capable of estimating the coefficients without bias. Such enhanced performance is, however, determined by error surfaces that are nonlinear functions of the coefficient values. Consequently, they may contain local minima, and the estimation algorithms may not necessarily converge to the global minimum of their error surfaces. The parameter estimation for both types of methods can be carried out by the LMS algorithm, the extended least squares algorithms, or their variants [11], [32]. A different approach using cross-cumulant information is pursued in [23] and [41]. This approach divides the identification problem into successive solutions of triangular linear systems of equations by considering appropriate slices of the cross-cumulant sequences for each subproblem. Blind identification of bilinear systems has attracted limited attention so far [30], [37]. In these works, closed-form expressions that relate measurable statistics of the output signal to the unknown parameters are derived for a very restricted class of nonlinear system models and for Gaussian inputs. Consequently, the most common approach to estimating the parameters of the model is to resort to some form of numerical search algorithm that operates in an iterative manner [38].

In this paper, we consider the problem of blind identification of an input-output bilinear system where the matrix of coefficients of mixed terms is lower triangular. A new algorithm for the identification of bilinear system parameters is presented. The algorithm employs five stages and utilizes output cumulants up to order 4. The derivations are based on the application of the Leonov–Shiryaev theorem [28] to the output cumulants.

The rest of the paper is organized as follows. Section II contains a formal statement of the blind identification problem. The structure of the solution is described in Section III. The algorithm is described in Section IV. The details of the derivation of the identification structure are provided in Appendices A–C. A simulation example that verifies the accuracy of the derivations and demonstrates the quality of the estimates is given in Section V. Finally, Section VI contains our concluding remarks.

II. Problem Statement

We consider bilinear systems of the form

$$\begin{aligned} y(n) = &\sum_{i=1}^{K_a} a(i)y(n-i) \\ &+ \sum_{i=0}^{K_b} b(i)u(n-i) \\ &+ \sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}} c(i,j)y(n-i)u(n-j) + v(n) \end{aligned} \tag{5}$$

where $y(n)$ is the output of the system, $u(n)$ the input, and $v(n)$ the measurement noise. The input signal $u(n)$ cannot be accessed for measurement. The first term in (5) is characterized by the parameter vector $\mathbf{a}$ of size K_a

$$\mathbf{a} = [a(1), \ldots, a(K_a)]. \tag{6}$$

Extending standard terminology, we will refer to this term as the *linear AR part* of the bilinear model. Similarly, we will call the second term in (5), which is produced by the parameter vector $\mathbf{b}$ of size $K_b + 1$

$$\mathbf{b} = [b(0), b(1), \ldots, b(K_b)] \tag{7}$$

the *linear MA part* of the bilinear system. Finally, the third term is called *mixed part* and is accountable for the nonlinear behavior of the system. The parameters of the mixed part are described by a lower triangular matrix $\mathbf{C}$ with entries $c(i,j)$ and size $K_{cy} \times K_{cu}$.

The objective of this paper is to estimate the system parameters $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{C}$ using output information only. To make the analysis tractable, we make the following assumptions.

1) The measurement noise $v(n)$ is a zero mean Gaussian white process and is independent of the input signal $u(n)$.

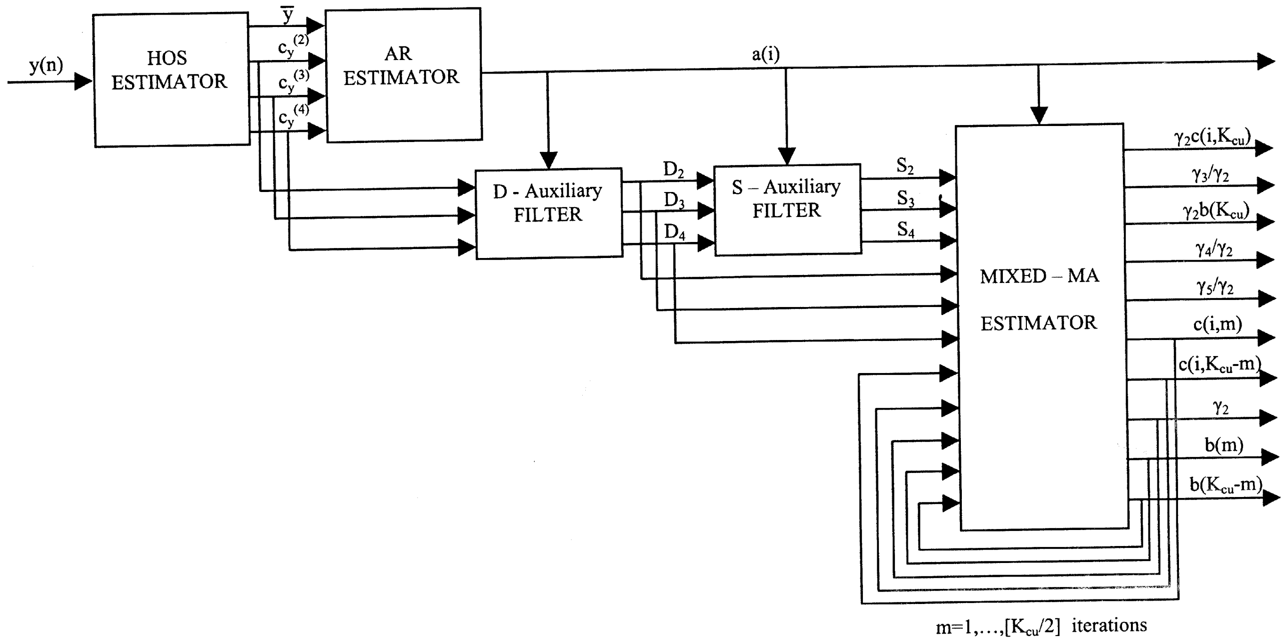

Fig. 1. Block diagram representation of the proposed algorithm.

2) The input signal is a non-Gaussian white process with zero mean value. This means that the cumulants of the input exist and are given by

$$\text{cum}[u(n), u(n-l_1), \ldots, u(n-l_{k-1})] = \gamma_k \delta(l_1, \ldots, l_{k-1}) \quad (8)$$

where $\delta(l_1, \ldots, l_{k-1})$ is the $(k-1)$-dimensional unit sample signal, and γ_k denotes the signal intensity of order k. For technical reasons that will become clear during the derivations, we assume that $\gamma_4 \neq 3\gamma_3^2/\gamma_2$.

3) The parameter vectors $\mathbf{a}$ and $\mathbf{b}$, as well as the coefficient matrix $\mathbf{C}$, are such that $y(n)$ is a stationary process. Sufficient conditions for the stationarity of bilinear processes are derived in [3], [18], and [26].

4) In order to find a unique solution and overcome the inherent scaling ambiguity of blind identification, we assume that $b(0) = 1$. We further assume that $c(K_{cu}, K_{cu}) \neq 0$ and that $K_{cu} \geq K_b$. Finally, we also assume that the system orders K_a, K_b, K_{cu}, and K_{cy} are known.

Based on the above assumptions, a closed-form solution is developed for the estimation of the parameters $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{C}$, using cumulants of $y(n)$ up to order 4. The main components of the method are presented next.

III. Organization of the Blind Estimation Algorithm

A block diagram representation of the blind estimation algorithm is provided in Fig. 1. It is formed by the cascade of several components. The first component is the "HOS estimator." It is fed with $y(n)$, which is the measurable output of the system we seek to identify, and estimates cumulants up to order 4. Efficient procedures for estimating cumulants both in terms of statistical and computational performance in the time as well as in the frequency domain have been extensively covered in the literature and will not be repeated here [34]. The coefficients of the linear AR part can be directly estimated from the cumulants. This function is performed by the box termed "AR estimator" in Fig. 1. The calculation of the coefficients of the linear MA part, the mixed part, and the statistics of the input requires knowledge of the AR coefficients and combinations of the cumulants derived using the estimated AR coefficient values. The combinations are generated by two "auxiliary" filters denoted by D and S in Fig. 1.

The outputs of the two auxiliary filters and of the AR estimator are fed into the "Mixed-MA" estimator to evaluate remaining parameters. The mixed-MA estimator contains an initialization module and the main module. The initialization module computes the K_{cu}th (last) column of $\mathbf{C}$ and $b(K_{cu})$, which is the last entry of the MA part, both scaled by the input variance. This module also calculates additional relationships between input cumulant intensities that are needed by the main module. The main module recursively computes the remaining columns of the mixed part together with the MA part. The functionality of every component is detailed in the next section.

IV. Algorithm Description

Our method utilizes suitably chosen slices of the output cumulants to estimate the system parameters. We will derive several relationships between these output cumulants and the unknown parameters using a list of properties presented in Appendix A. We first define the following input-output cross-cumulant sequences

$$g_1(m_1) = \text{cum}[y(n), u(n-m_1)] \quad (9)$$

$$g_2(m_1, m_2) = \text{cum}[y(n), y(n-m_1), u(n-m_2)] \quad (10)$$

$$g_3(m_1, m_2, m_3) = \text{cum}[y(n), y(n-m_1), y(n-m_2), u(n-m_3)] \quad (11)$$

$$g_4(m_1,m_2,m_3,m_4)=\text{cum}[y(n),y(n-m_1),y(n-m_2)\\ y(n-m_3),u(n-m_4)]. \quad (12)$$

The following relationships between the output cumulants and the system parameters are derived in Appendix B.

$$\begin{aligned} c_y^{(2)}(l_1)&=\text{cum}[y(n),y(n-l_1)]=\sum_{i=1}^{K_a}a(i)c_y^{(2)}(l_1-i)\\ &+\sum_{i=0}^{K_b}b(i)g_1(i-l_1)+\bar{y}\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}}c(i,j)g_1(j-l_1)\\ &+\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}}c(i,j)g_2(i-l_1,j-l_1),\quad l_1>0. \end{aligned} \quad (13)$$

$$\begin{aligned} c_y^{(3)}(l_1,l_2)&=\text{cum}[y(n),y(n-l_1),y(n-l_2)]\\ &=\sum_{i=1}^{K_a}a(i)c_y^{(3)}(l_1-i,l_2-i)\\ &+\sum_{i=0}^{K_b}b(i)g_2(l_2-l_1,i-l_1)\\ &+\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}}c(i,j)g_3(i-l_1,l_2-l_1,j-l_1)\\ &+\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}}c(i,j)c_y^{(2)}(l_2-i)g_1(j-l_1)\\ &+\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}}c(i,j)c_y^{(2)}(l_1-i)g_1(j-l_2)\\ &+\bar{y}\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}}c(i,j)g_2(l_2-l_1,j-l_1)\\ &l_1>0,\quad l_2>0. \end{aligned} \quad (14)$$

$$\begin{aligned} c_y^{(4)}(l_1,l_2,l_3)&=\text{cum}[y(n),y(n-l_1),y(n-l_2),y(n-l_3)]\\ &=\sum_{i=1}^{K_a}a(i)c_y^{(4)}(l_1-i,l_2-i,l_3-i)\\ &+\sum_{i=0}^{K_b}b(i)g_3(l_2-l_1,l_3-l_1,i-l_1)\\ &+\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}}c(i,j)g_4(i-l_1,l_2-l_1,l_3-l_1,j-l_1)\\ &+\bar{y}\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}}c(i,j)g_3(l_2-l_1,l_3-l_1,j-l_1)\\ &+\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}}c(i,j)c_y^{(2)}(l_1-i)g_2(l_3-l_2,j-l_2)\\ &+\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}}c(i,j)c_y^{(2)}(l_2-i)g_2(l_3-l_1,j-l_1)\\ &+\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}}c(i,j)c_y^{(2)}(l_3-i)g_2(l_2-l_1,j-l_1)\\ &+\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}}c(i,j)c_y^{(3)}(l_1-i,l_2-i)g_1(j-l_3)\\ &+\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}}c(i,j)c_y^{(3)}(l_1-i,l_3-i)g_1(j-l_2)\\ &+\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}}c(i,j)c_y^{(3)}(l_2-i,l_3-i)g_1(j-l_1)\\ &l_1>0,\quad l_2>0,\quad l_3>0. \end{aligned} \quad (15)$$

The above equations are considerably simplified if the lags are properly restricted. To this end the following proposition is useful. The proof is given in Appendix B.

Proposition 1: The following relations hold:

$$g_1(m_1)=0,\quad m_1<0 \quad (16)$$

$$g_2(m_1,m_2)=0,\quad m_1\geq m_2,\ m_2<0 \quad (17)$$

$$g_3(m_1,m_2,m_3)=0,\quad m_1\geq m_3,\ m_2\geq m_3\\ m_3<0 \text{ and} \quad (18)$$

$$g_4(m_1,m_2,m_3,m_4)=0,\quad m_1\geq m_4,\ m_2\geq m_4,\ m_3>m_4\\ m_4<0 \quad (19)$$

Next, we describe in detail each component of Fig. 1.

A. AR Estimator

The AR parameters $a(i)$ are determined with the aid of the following proposition.

Proposition 2: Let $L_1, L_2, L_3 > K_{cu}$. Then

$$c_y^{(2)}(L_1)=\sum_{i=1}^{K_a}a(i)c_y^{(2)}(L_1-i) \quad (20)$$

$$c_y^{(3)}(L_1,L_2)=\sum_{i=1}^{K_a}a(i)c_y^{(3)}(L_1-i,L_2-i), \quad \text{and} \quad (21)$$

$$c_y^{(4)}(L_1,L_2,L_3)=\sum_{i=1}^{K_a}a(i)c_y^{(4)}(L_1-i,L_2-i,L_3-i). \quad (22)$$

The proof is a direct application of Proposition 1 and (13)–(15). Proposition 2 states that the higher order output cumulant sequences behave in a manner that is identical to the covariance function of an autoregressive signal for sufficiently large values of the lag l. This property enables the computation of the $a(i)$ parameters via one of the above relations and a linear system Toeplitz solver such as a variant of the Levinson algorithm [13]. The simplest implementation for the AR parameter estimator relies on second-order statistics and (20). Collecting R successive output autocovariance lags in the range $K_{cu} < L_1 \leq K_{cu}+R$, we obtain an overdetermined system of linear equations in the unknown parameters.

$$\begin{pmatrix} c_y^{(2)}(K_{cu}+1)\\ c_y^{(2)}(K_{cu}+2)\\ \vdots\\ c_y^{(2)}(K_{cu}+R)\end{pmatrix}=$$

$$\begin{pmatrix} c_y^{(2)}(K_{cu}) & c_y^{(2)}(K_{cu}-1) & \dots & c_y^{(2)}(K_{cu}+1-K_a) \\ c_y^{(2)}(K_{cu}+1) & c_y^{(2)}(K_{cu}) & \dots & c_y^{(2)}(K_{cu}+2-K_a) \\ \vdots & \vdots & \ddots & \vdots \\ c_y^{(2)}(K_{cu}+R-1) & c_y^{(2)}(K_{cu}+R-2) & \dots & c_y^{(2)}(K_{cu}+R-K_a) \end{pmatrix} \begin{pmatrix} a(1) \\ a(2) \\ \vdots \\ a(K_a) \end{pmatrix} \tag{23}$$

from which we can easily solve for $\{a(1), a(2), \cdots a(K_a)\}$. In general, we obtain more accurate estimation performance by choosing $R > K_a$.

B. Mixed-MA Estimator

The estimation of the rest of the unknown parameters explicitly utilizes the estimated values of the AR coefficients. Two sets of auxiliary variables are first calculated using the auxiliary filters marked D and S in Fig. 1. The first set of variables are defined as

$$D_2(l_1) = -\sum_{i=0}^{K_a} a(i) c_y^{(2)}(l_1 - i). \tag{24}$$

Likewise we define

$$D_3(l_1, l_2) = -\sum_{i=0}^{K_a} a(i) c_y^{(3)}(l_1 - i, l_2 - i) \tag{25}$$

$$D_4(l_1, l_2, l_3) = -\sum_{i=0}^{K_a} a(i) c_y^{(4)}(l_1 - i, l_2 - i, l_3 - i) \tag{26}$$

where $a(0) = -1$ for all three definitions. In a similar manner, we define three new sets of auxiliary variables as linear combinations of D_2, D_3 and D_4 as follows:

$$S_2(m) = -\sum_{l=m}^{K_{cu}} a(l-m) D_2(l) \qquad m = K_{cu}, K_{cu}-1, \dots \tag{27}$$

$$S_3(m, K_{cu}) = -\sum_{l=m}^{K_{cu}} a(l-m) D_3(l, K_{cu}) \qquad m = K_{cu}, K_{cu}-1, \dots \tag{28}$$

$$S_4(m, K_{cu}, L) = -\sum_{l=m}^{K_{cu}} a(l-m) D_4(l, K_{cu}, L) \qquad L > K_{cu}, \quad m = K_{cu}, K_{cu}-1 \dots \tag{29}$$

We are now ready to estimate the remaining parameters. The mixed-MA estimator contains an initialization module and a main module. These modules are described separately next.

1) Initialization Module: The initialization module estimates $\gamma_2 c(\cdot, K_{cu}), \gamma_3/\gamma_2, \gamma_2 b(K_{cu}), \gamma_4/\gamma_2$, and γ_5/γ_2.

a) Computation of $\gamma_2 c(\cdot, K_{cu})$: The following system of linear equations is derived in Appendix C:

$$\begin{pmatrix} D_3(K_{cu}, K_{cu}+1) \\ D_3(K_{cu}, K_{cu}+2) \\ \vdots \\ D_3(K_{cu}, K_{cu}+R) \end{pmatrix} = \begin{pmatrix} c_y^{(2)}(1) & c_y^{(2)}(0) & \dots & c_y^{(2)}(K_{cu}+1-K_{cy}) \\ c_y^{(2)}(2) & c_y^{(2)}(1) & \dots & c_y^{(2)}(K_{cu}+2-K_{cy}) \\ \vdots & \vdots & \ddots & \vdots \\ c_y^{(2)}(R) & c_y^{(2)}(R-1) & \dots & c_y^{(2)}(K_{cu}+R-K_{cy}) \end{pmatrix} \begin{pmatrix} \gamma_2 c(K_{cu}, K_{cu}) \\ \gamma_2 c(K_{cu}+1, K_{cu}) \\ \vdots \\ \gamma_2 c(K_{cy}, K_{cu}) \end{pmatrix}. \tag{30}$$

Given measurements of the cumulant values in the above range, we can create an overdetermined set of linear equations in the unknown parameters $\gamma_2 c(K_{cu}, K_{cu}), \cdots, \gamma_2 c(K_{cy}, K_{cu})$ by choosing $R \geq K_{cy} - K_{cu} + 1$ equations. Recall that $c(i,j) = 0$ if $i < j$, and thus, only the parameters explicitly involved in (30) are nonzero and need to be estimated. Solving for $\gamma_2 c(\cdot, K_{cu})$ is straightforward.

b) Estimation of γ_3/γ_2: The following relationship is established in Appendix C:

$$\frac{\gamma_3}{\gamma_2} = \left[D_4(K_{cu}, K_{cu}, L) - 2\gamma_2 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) \times c_y^{(3)}(K_{cu}-i, L-i)\right] / D_3(K_{cu}, L) \tag{31}$$

where L is such that $D_3(K_{cu}, L) \neq 0$. The right-hand side of (31) involves quantities available from previous steps.

c) Calculation of $\gamma_2 b(K_{cu}), \gamma_4/\gamma_2$ and γ_5/γ_2: The rest of the calculations in the initialization module are performed in a similar way. The following equations are derived in Appendix C and can be used directly to estimate the three remaining quantities.

$$\gamma_2 b(K_{cu}) = D_2(K_{cu}) - \bar{y}\gamma_2 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) - \gamma_3 c(K_{cu}, K_{cu}) \tag{32}$$

$$\frac{\gamma_4}{\gamma_2} = \frac{D_3(K_{cu}, K_{cu}) - \gamma_3 b(K_{cu}) - \bar{y}\gamma_3 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) - 2\gamma_2 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) c_y^{(2)}(K_{cu}-i)}{\gamma_2 c(K_{cu}, K_{cu})} \tag{33}$$

and

$$\begin{aligned} \frac{\gamma_5}{\gamma_2} = \frac{1}{\gamma_2 c(K_{cu}, K_{cu})} \Big(& D_4(K_{cu}, K_{cu}, K_{cu}) \\ & - \gamma_4 b(K_{cu}) - \bar{y}\gamma_4 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) \\ & - 3\gamma_2 \sum_{i=K_{cu}}^{K_{cu}} c(i, K_{cu}) c_y^{(3)}(K_{cu}-i, K_{cu}-i) \\ & - 3\gamma_3 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu})\, c_y^{(2)}(K_{cu}-i) \Big). \end{aligned} \tag{34}$$

2) Main Module: The main module involves $[(K_{cu})/(2)]$ recurrent steps, where $[(.)]$ denotes the largest integer smaller than or equal to (.). During the mth step, we estimate the mth and $(K_{cu} - m)$th columns of $\mathbf{C}$, $b(m)$ and $b(K_{cu} - m)$. The input variance γ_2 is also estimated in the main module. At the end of $[K_{cu}/2]$ recursions, all the unknown parameters are estimated.

For each $m = 1, 2, \ldots, [(K_{cu})/(2)]$, the computations in this module are organized into three stages. In what follows, we outline the steps involved in each of the stages. All the derivations are given in Appendix C.

Stage 1: The first stage utilizes a linear system of equations to determine the following set of parameters during the mth step:

1) the mth column of $\mathbf{C}$ of length $K_{cy} - m + 1$, denoted by $\mathbf{c}_m = (c(m,m), \ldots, c(K_{cy}, m))^T$;
2) the first m entries of the $(K_{cu} - m)$th column of $\mathbf{C}$ scaled by the input variance. We denote these parameters as

$$\gamma_2 \mathbf{c}^1_{K_{cu}-m} = (\gamma_2 c(K_{cu}-m, K_{cu}-m), \ldots \gamma_2 c(K_{cu}-1, K_{cu}-m))^T. \quad (35)$$

3) a linear combination of the remaining terms of the $K_{cu} - m$ column $\mathbf{c}^2_{K_{cu}-m}$ with the last column of C, $\mathbf{c}_{K_{cu}}$ of length $K_{cy} - K_{cu} + 1$ and given by

$$\begin{aligned}\mathbf{d}_m &= \gamma_2 \mathbf{c}^2_{K_{cu}-m} + b(m)\gamma_2 \mathbf{c}_{K_{cu}} \\ &= (\gamma_2 c(K_{cu}, K_{cu}-m) + b(m)\gamma_2 c(K_{cu}, K_{cu}), \ldots \\ &\quad \gamma_2 c(K_{cy}, K_{cu}-m) + b(m)\gamma_2 c(K_{cy}, K_{cu}))^T. \quad (36)\end{aligned}$$

4) The scalar quantity

$$r_m = b(m)\frac{\gamma_3}{\gamma_2} + \gamma_2 c(m,m). \quad (37)$$

The computation of the parameters in items 1–4 is performed by solving an overdetermined system of linear equations of the form

$$(\mathbf{F} \quad \mathbf{G} \quad \mathbf{P} \quad \mathbf{Q}) \begin{pmatrix} \mathbf{c}_m \\ \gamma_2 \mathbf{c}^1_{K_{cu}-m} \\ \mathbf{d}_m \\ r_m \end{pmatrix} = \mathbf{E} \quad (38)$$

where the vector of the unknown parameters contains $2K_{cy} - K_{cu} + 3$ elements, and the right-hand-side vector $\mathbf{E}$ has length $R \geq 2K_{cy} - K_{cu} + 3$. The s-th element of this vector is

$$\begin{aligned}e_s &= S_4(K_{cu} - m, K_{cu}, K_{cu} + s) \\ &+ \sum_{i=K_{cu}-m}^{K_{cu}} a(i - K_{cu} + m)\gamma_2 \\ &\times \sum_{n=K_{cu}}^{K_{cy}} c(n, K_{cu}) c_y^{(3)}(i - n, K_{cu} + s - n) \\ &- \gamma_2 \sum_{j=K_{cu}-m+1}^{K_{cu}-1} \left[b(j) + \bar{y} \sum_{i=j}^{K_{cy}} c(i,j) \right] \\ &\times \sum_{n=j-K_{cu}+m}^{K_{cy}} c(n, j - K_{cu} + m) \\ &\times c_y^{(3)}(m - n, s + m - n) - \gamma_2 \sum_{j=K_{cu}-m+1}^{K_{cu}-1} \\ &\times \sum_{i=j}^{K_{cy}} c(i,j) \sum_{n=j-K_{cu}+m}^{K_{cy}} c(n, j - K_{cu} + m) \\ &\times \Big[c_y^{(4)}(i - K_{cu} + m - n, m - n, s + m - n) + c_y^{(2)}(K_{cu} - i) \\ &\times c_y^{(2)}(s + m - n) + c_y^{(2)}(K_{cu} + s - i) c_y^{(2)}(m - n) \\ &+ \bar{y} c_y^{(3)}(K_{cu} - i, K_{cu} + s - i) \Big] \\ &- \gamma_2 \sum_{j=K_{cu}-m+1}^{K_{cu}-1} \sum_{i=j}^{K_{cy}} c(i,j) c_y^{(3)}(K_{cu} - i, K_{cu} + s - i) \\ &\times \left[b(j - K_{cu} + m) + c(j - K_{cu} + m, j - K_{cu} + m)\frac{\gamma_3}{\gamma_2} \right] \\ &- \gamma_3 \sum_{j=K_{cu}-m+1}^{K_{cu}-1} c(j,j) \sum_{n=j-K_{cu}+m}^{K_{cy}} \\ &\times c(n, j - K_{cu} + m) c_y^{(3)}(m - n, s + m - n). \quad (39)\end{aligned}$$

The matrix $\mathbf{F}$ has dimensions $R \times K_{cy} - m + 1$. The (s_1, s_2) element of $\mathbf{F}$, with $1 \leq s_1 \leq R$ and $1 \leq s_2 \leq K_{cy} - m + 1$, is given by the expression

$$\begin{aligned}f_{s_1,s_2} &= D_2(K_{cu}) c_y^{(3)}(1 - s_2, s_1 - s_2 + 1) \\ &+ \left(D_3(K_{cu}, K_{cu}) - \gamma_2 \sum_{n=K_{cu}}^{K_{cy}} c(n, K_{cu}) c_y^{(2)}(K_{cu} - n) \right) \\ &\quad \times c_y^{(2)}(s_1 - s_2 + 1) \\ &+ \gamma_2 \sum_{n=K_{cu}}^{K_{cy}} c(n, K_{cu}) \\ &\times c_y^{(4)}(n - K_{cu} + 1 - s_2, 1 - s_2, s_1 - s_2 + 1) \\ &+ \gamma_3 \sum_{n=K_{cu}}^{K_{cy}} c(n, K_{cu}) c_y^{(3)}(n - K_{cu} + 1 - s_2, s_1 - s_2 + 1) \\ &+ \gamma_4 \sum_{n=K_{cu}}^{K_{cy}} c(n, K_{cu}) c_y^{(2)}(K_{cu} + s_1 - n)\delta(s_2 - 1) \\ &+ \gamma_2 \sum_{n=K_{cu}}^{K_{cy}} c(n, K_{cu}) c_y^{(2)}(K_{cu} + s_1 - n) c_y^{(2)}(1 - s_2) \\ &+ \bar{y}\gamma_3 \sum_{n=K_{cu}}^{K_{cy}} c(n, K_{cu}) c_y^{(2)}(K_{cu} + s_1 - n) \\ &+ \gamma_3 \sum_{n=K_{cu}}^{K_{cy}} c(n, K_{cu}) \\ &\times c_y^{(3)}(K_{cu} - n, K_{cu} + s_1 - n)\delta(s_2 - 1) \\ &+ \bar{y}\gamma_2 \sum_{n=K_{cu}}^{K_{cy}} c(n, K_{cu}) c_y^{(3)}(K_{cu} - n, K_{cu} + s_1 - n). \quad (40)\end{aligned}$$

The matrix $\mathbf{G}$ has size $R \times m$, and its (s_1, s_2)th entry is

$$g_{s_1,s_2} = c_y^{(3)}(m + 1 - s_2, s_1 - s_2 + m + 1). \quad (41)$$

The matrix $\mathbf{P}$ has $R \times K_{cy} - K_{cu} + 1$ entries given by

$$p_{s_1,s_2} = c_y^{(3)}(1 - s_2, s_1 - s_2 + 1). \tag{42}$$

Finally, the vector $\mathbf{Q}$ has R entries given by

$$q_{s_1} = \gamma_2 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) c_y^{(2)}(K_{cu} + s_1 - i). \tag{43}$$

Stage 2: The second stage determines $b(m)$ and γ_2. Recall from (37) that

$$r(m) = b(m)\frac{\gamma_3}{\gamma_2} + \gamma_2 c(m,m). \tag{44}$$

Suppose $m = m^*$ is the first integer for which $c(m^*, m^*) \neq 0$. Then, for every $m < m^*$, (44) gives $b(m) = r(m)\gamma_2/\gamma_3$. Thus, all $b(m)$ for $m < m^*$ are determined. Next, suppose $m = m^*$. It is shown in Appendix C that $b(m)$ and γ_2 also satisfy a linear equation of the form

$$\begin{aligned}
&S_3(K_{cu} - m, K_{cu}) - \gamma_2 \sum_{j=K_{cu}-m+1}^{K_{cu}} \left(b(j) + \bar{y} \sum_{i=j}^{K_{cy}} c(i,j) \right) \\
&\times \sum_{n=j-K_{cu}+m}^{K_{cy}} c(n, j - K_{cu} + m)\, c_y^{(2)}(m - n) - \gamma_2 \\
&\times \left(b(K) + \bar{y} \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) \right) \left[\frac{\gamma_4}{\gamma_2} c(m,m) + \frac{\gamma_3}{\gamma_2} \bar{y} \sum_{n=m}^{K_{cy}} c(n,m) \right] \\
&- \gamma_2 \sum_{j=K_{cu}-m+1}^{K_{cu}-1} \sum_{i=j}^{K_{cy}} c(i,j) c_y^{(2)}(K_{cu} - i) \\
&\times \Bigg[b(j - K_{cu} + m) + \frac{\gamma_3}{\gamma_2} c(j - K_{cu} + m, j - K_{cu} + m) \\
&\quad + \bar{y} \sum_{n=j-K_{cu}+m}^{K_{cy}} c(n, j - K_{cu} + m) \Bigg] \\
&- \sum_{i=K_{cu}}^{K_{cy}} (\gamma_2 c(i, K_{cu} - m) + b(m)\gamma_2 c(i, K_{cu}))\, c_y^{(2)}(K_{cu} - i) \\
&- c(m,m)\gamma_3 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) c_y^{(2)}(K_{cu} - i) \\
&- \bar{y} \sum_{i=m}^{K_{cy}} c(i,m)\gamma_2 \sum_{n=k_{cu}}^{k_{cy}} c(n, K_{cu}) c_y^{(2)}(K_{cu} - n) \\
&- \gamma_2 \sum_{i=K_{cu}-m}^{K_{cu}-1} c(i, K_{cu} - m) c_y^{(2)}(K_{cu} - i) \\
&- \bar{y}\gamma_4 c(K_{cu}, K_{cu}) \sum_{n=m}^{K_{cy}} c(n,m) \\
&- \gamma_2 \sum_{j=K_{cu}-m+1}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i,j) \sum_{n=j-K_{cu}+m}^{K_{cy}} \\
&\times c(n, j - K_{cu} + m) c_y^{(3)}(i - K_{cu} + m - n, m - n) \\
&- \gamma_3 \sum_{j=K_{cu}-m+1}^{K_{cu}} c(j,j) \sum_{n=j-K_{cu}+m}^{K_{cy}} c(n, j - K_{cu} + m) \\
&\times c_y^{(2)}(m - n) - \gamma_5 c(K_{cu}, K_{cu}) c(m,m) \\
&- \gamma_3 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) \sum_{n=m}^{K_{cy}} c(n,m) c_y^{(2)}(i - K_{cu} + m - n) \\
&+ \sum_{i=K_{cu}-m}^{K} a(i - K_{cu} + m)\gamma_2 \sum_{n=K_{cu}}^{K_{cy}} c(n, K_{cu}) c_y^{(2)}(i - n) \\
&= b(m) \left[\gamma_4 c(K_{cu}, K_{cu}) + \gamma_3 \left(b(K_{cu}) + \bar{y} \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) \right) \right] \\
&\quad + \gamma_2 \Bigg[3\gamma_3 c(K_{cu}, K_{cu}) c(m,m) + c(m,m)\gamma_2 \\
&\quad \times \left(b(K_{cu}) + \bar{y} \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) \right) \Bigg].
\end{aligned} \tag{45}$$

We note that all quantities in the above equation except $b(m)$ and γ_2 are either measurable statistics of the output signal or parameters that have been estimated in previous steps. Consequently, we can solve the system of two linear equations (44) and (45) to estimate $b(m)$ and γ_2. The determinant of the matrix associated with the two equations is

$$c(m,m)c(K_{cu}, K_{cu}) \left(\gamma_4 - \frac{3\gamma_3^2}{\gamma_2} \right). \tag{46}$$

Since $c(m,m) \neq 0$ for $m = m^*$, the assumptions stated in Section II ensure that the determinant does not vanish. Therefore, $b(m^*)$ and γ_2 are uniquely determined. The remaining parameters $b(m)$ for $m > m^*$ are readily computed from (44).

Stage 3: This stage completes the estimation of the $(K_{cu} - m)$th column of $\mathbf{C}$. It also estimates $b(K_{cu} - m)$. Recall that we estimated the product of γ_2 and the first m entries of the $(K_{cn} - m)$ th column of $\mathbf{C}$ in step 2 of Stage 1. Since we computed γ_2 in Stage 2 of the recursion, it is now straightforward to estimate the first m elements of the $(K_{cu} - m)$th column of $\mathbf{C}$. Similarly, we note that we estimated

$$\mathbf{d}_m = \gamma_2 \mathbf{c}^2_{K_{cu}-m} + b(m)\gamma_2 \mathbf{c}_{K_{cu}} \tag{47}$$

in step 3 of Stage 1. Since all variables except $\mathbf{c}^2_{K_{cu}-m}$ have been estimated at this time, we can solve for the entries of the $(K_{cu} - m)$th column that were not computed earlier from the above equation. The only other parameter that is estimated in the mth recursion is $b(K_{cu} - m)$. This parameter is estimated from the auxiliary sequence S_2 and the expression

$$\begin{aligned}
\gamma_2 b(K_{cu} - m) = {} & S_2(K_{cu} - m) - \bar{y}\gamma_2 \sum_{i=K_{cu}-m}^{K_{cy}} c(i, K_{cu} - m) \\
& - \gamma_3 c(K_{cu} - m, K_{cu} - m) - \sum_{j=K_{cu}-m+1}^{K_{cu}}
\end{aligned}$$

$$\times c(j,j)\left[\gamma_3 b(j-K_{cu}+m)+\left(\gamma_4+\gamma_2^2\right)\right.$$
$$\times c(j-K_{cu}+m, j-K_{cu}+m)+\bar{y}\gamma_3$$
$$\left.\times \sum_{n=j-K_{cu}+m}^{K_{cy}} c(n, j-K_{cu}+m)\right]$$
$$-\sum_{j=K_{cu}-m+1}^{K_{cu}}\left(b(j)+\bar{y}\sum_{i=j}^{K_{cy}} c(i,j)\right)$$
$$\times\left[\gamma_2 b(j-K_{cu}+m)\right.$$
$$+\gamma_3 c(j-K_{cu}+m, j-K_{cu}+m)+\bar{y}\gamma_2$$
$$\left.\times \sum_{n=j-K_{cu}+m}^{K_{cy}} c(n, j-K_{cu}+m)\right]$$
$$-\gamma_2 \sum_{j=K_{cu}-m+1}^{K_{cu}} \sum_{i=j+1}^{K_{cy}} c(i,j)$$
$$\times \sum_{n=j-K_{cu}+m}^{K_{cy}} c(n, j-K_{cu}+m)$$
$$\times c_y^{(2)}(i-K_{cu}+m-n). \quad (48)$$

This completes the set of calculations necessary to perform the blind estimation of the bilinear system parameters.

V. Simulation Results

In this section, we present the results of a simulation experiment illustrating the performance of the algorithm. The method is applied to a bilinear system of the form

$$y(n)=\sum_{i=1}^{2} a(i)y(n-i)+\sum_{i=0}^{1} b(i)u(n-i)+\sum_{j=1}^{3}\sum_{i=j}^{4} c(i,j)y(n-i)u(n-j)$$

where $K_a = 2$, $K_b = 1$, $K_{cu} = 3$, and $K_{cy} = 4$, with $\mathbf{a} = [-0.1 \quad 0.02]$, $\mathbf{b} = [1 \quad -0.4]$, and

$$\mathbf{C}=\begin{pmatrix} 0.1 & 0 & 0 \\ 0 & -0.05 & 0 \\ 0 & 0 & 0.3 \\ -0.1 & 0.05 & 0.1 \end{pmatrix}.$$

The input sequence $u(n)$ is a pseudorandom binary sequence (PRBS) generated by a linear feedback shift register. The characteristic polynomial of the register is a primitive polynomial. To reduce the realization dependency, the parameter estimates were averaged over 100 Monte Carlo runs. For each experiment, a new PRBS input is generated of length $2^{14}-1$. The mean and the variance of the estimated parameters against the true ones are shown in Table I.

TABLE I
TRUE AND ESTIMATED PARAMETERS FOR A PRBS INPUT SEQUENCE OF 16383 SAMPLES (100 MONTE CARLO RUNS)

Parameters	True Value	Mean	Variance
$a(1)$	-0.1	-0.0976	$4.58\ 10^{-7}$
$a(2)$	0.02	0.0225	$4.41\ 10^{-7}$
$b(1)$	-0.4	-0.3801	$6.57\ 10^{-4}$
$c(1,1)$	0.1	0.0999	$1.67\ 10^{-4}$
$c(2,1)$	0	-0.0032	$2.14\ 10^{-5}$
$c(3,1)$	0	-0.0015	$8.88\ 10^{-7}$
$c(4,1)$	-0.1	-0.1063	$4.22\ 10^{-5}$
$c(2,2)$	-0.05	-0.0267	$6.92\ 10^{-5}$
$c(3,2)$	0	0.0121	$4.61\ 10^{-5}$
$c(4,2)$	0.05	0.0594	$3.97\ 10^{-5}$
$c(3,3)$	0.3	0.2741	$4.08\ 10^{-4}$
$c(4,3)$	0.1	0.0911	$4.47\ 10^{-5}$
γ_2	3.24	3.4946	$4.78\ 10^{-2}$

VI. Concluding Remarks

This paper dealt with the blind identification of bilinear systems with measurements corrupted by Gaussian noise. The excitation is non-Gaussian white noise. The parameters are determined via a sequence of linear systems involving cumulant slices of orders less than four. Simulations validating the proposed method were supplied. One issue that should be pointed out with regard to this work is the need for good experiment design conditions. This, in our case, translates to inputs with good white characteristics in higher order cumulants. Recent work in this direction utilizing dual BCH sequences, Gold sequences, and sequences generated by modulo 2 addition of maximal length sequences of relatively prime periods is reported in [16].

Appendix A
Basic Properties of Input Output Cumulants

In this appendix, some basic properties of input–output cumulants are derived. They are heavily used in the derivation of the blind estimation algorithm.

Property 1: Let $z_1(n-1), z_2(n-1), \ldots, z_k(n-1)$ be functions of $u(n-i)$ and $y(n-i)$ for $i \geq 1$. Recall that $u(n)$ and $y(n)$ are the input and output signals, respectively, of the bilinear system and that $u(n)$ is a higher order white sequence with zero mean value. Then

$$\text{cum}[u(n), z_1(n-1), z_2(n-1), \ldots, z_k(n-1)] = 0. \quad (49)$$

Proof: The conclusion follows immediately from the fact that $u(n)$ is white.

Property 2: Let $z_1(n-1), z_2(n-1), \dots, z_k(n-1)$ be as defined in Property 1. Then

$$\text{cum}[u(n), u(n), \dots, u(n), z_1(n-1) \; z_2(n-1), \dots, z_k(n-1)] = 0. \tag{50}$$

Proof: If a random variable X is independent of the random variables $Y_1, Y_2, \dots, Y_k$, then [34]

$$\text{cum}[X, X, \dots, X, Y_1, Y_2, \dots, Y_k] = 0. \tag{51}$$

Since $u(n)$ is independent of $z_1(n-1), \dots, z_k(n-1)$, the result follows.

Property 3: For the same set of definitions of the signals as in Property 1

$$\text{cum}[y(n), u(n), z_1(n-1), z_2(n-1), \dots, z_k(n-1)] = 0. \tag{52}$$

Proof: Let

$$z_{k+1}(n-1) = \sum_{i=1}^{K_a} a(i)y(n-i) + \sum_{i=1}^{K_b} b(i)u(n-i) + \sum_{j=1}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i,j)y(n-i)u(n-j) \tag{53}$$

so that the bilinear expression becomes

$$y(n) = u(n) + z_{k+1}(n-1) + v(n). \tag{54}$$

Substituting (54) for $y(n)$ in the cumulant expression gives

$$\begin{aligned} &\text{cum}[y(n), u(n), z_1(n-1), z_2(n-1), \dots, z_k(n-1)] = \\ &\quad \text{cum}[u(n), u(n), z_1(n-1), z_2(n-1), \dots, z_k(n-1)] \\ &\quad +\text{cum}[u(n), z_1(n-1), z_2(n-1), \dots, z_k(n-1), z_{k+1}(n-1)] \\ &\quad +\text{cum}[v(n), u(n), z_1(n-1), z_2(n-1), \dots, z_k(n-1)]. \end{aligned} \tag{55}$$

All terms on the right-hand side are zero due to Properties 1 and 2 and independence assumptions of $v(n)$ and $u(n)$. This completes the proof. Multilinearity of cumulants [34] leads to the following straightforward generalization of the above result:

$$\text{cum}[y(n), \dots, y(n), u(n), \dots, u(n), z_1(n-1) \; z_2(n-1), \dots, z_k(n-1)] = 0. \tag{56}$$

Property 4: Let $y(n)$ and $u(n)$ be the output and input of the bilinear system in (5). Then

$$\text{cum}[y(n), \dots, y(n), u(n), \dots, u(n)] = \gamma_{N+P} \tag{57}$$

where $y(n)$ appears N times, and $u(n)$ appears P times in the above expression.

Proof: Suppose first that $N = P = 1$, and let $y(n) = u(n) + z_{k+1}(n-1) + v(n)$, as in (54). Substituting for $y(n)$ and employing Property 1, we get

$$\begin{aligned} \text{cum}[y(n), u(n)] &= \text{cum}[u(n), u(n)] + \text{cum}[z_{k+1}(n-1), u(n)] \\ &\quad + \text{cum}[v(n), u(n)] = \gamma_2. \end{aligned} \tag{58}$$

Multilinearity of cumulants proves the generalization given as Property 4. Properties 1–4 in combination with the Leonov-Shiryaev theorem [28] form the main tools for the computation of the output statistics.

APPENDIX B
DERIVATION OF OUTPUT CUMULANT EXPRESSIONS

We establish (13)–(15) and Proposition 1 in this Appendix. Using (5) and the multilinearity of cumulants, we obtain

$$\begin{aligned} c_y^{(2)}(l_1) &= \text{cum}[y(n), y(n-l_1)] \\ &= \sum_{i=1}^{K_a} a(i)\text{cum}[y(n-i), y(n-l_1)] \\ &\quad + \sum_{i=0}^{K_b} b(i)\text{cum}[u(n-i), y(n-l_1)] \\ &\quad + \sum_{j=1}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i,j)\text{cum}[y(n-i)u(n-j), y(n-l_1)] \\ &\quad + \text{cum}[v(n), y(n-l_1)]. \end{aligned} \tag{59}$$

The last term on the right-hand-side of (59) is zero since $v(n)$ and $y(n-l_1)$ are independent variables. Application of the Leonov-Shiryaev theorem [28] and the zero mean assumption of the input signal to the third term on the right-hand side of (59) give

$$\begin{aligned} \text{cum}[y(n-i)u(n-j), y(n-l_1)] &= \text{cum}[y(n-l_1) \\ & y(n-i), u(n-j)] + \bar{y}\text{cum}[y(n-l_1), u(n-j)]. \end{aligned} \tag{60}$$

Substituting the above result into (59) and making use of the cross-cumulant definitions in Section IV result in (13). A similar approach is used to derive (14). Thus

$$\begin{aligned} c_y^{(3)}(l_1, l_2) &= \text{cum}[y(n), y(n-l_1), y(n-l_2)] \\ &= \sum_{i=1}^{K_a} a(i)\text{cum}[y(n-i), y(n-l_1), y(n-l_2)] \\ &\quad + \sum_{i=0}^{K_b} b(i)\text{cum}[u(n-i), y(n-l_1), y(n-l_2)] \\ &\quad + \sum_{j=1}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i,j)\text{cum}[y(n-i) \\ &\quad \times u(n-j), y(n-l_1), y(n-l_2)] \\ &\quad + \text{cum}[v(n), y(n-l_1), y(n-l_2)]. \end{aligned} \tag{61}$$

Applying the Leonov-Shiryaev theorem to the third term on the right-hand side gives

$$\begin{aligned}\text{cum}&[y(n-i)u(n-j), y(n-l_1), y(n-l_2)]\\ &=\text{cum}[y(n-l_1), y(n-i), y(n-l_2), u(n-j)]\\ &\quad+\text{cum}[y(n-i), y(n-l_2)]\text{cum}[y(n-l_1), u(n-j)]\\ &\quad+\text{cum}[y(n-i), y(n-l_1)]\text{cum}[y(n-l_2), u(n-j)]\\ &\quad+\bar{y}\text{cum}[y(n-l_1), y(n-l_2), u(n-j)].\end{aligned} \tag{62}$$

Substituting the above result and the cross-cumulant definitions in (61), along with the use of the independence property of cumulants, we obtain (14). Equation (15) is derived in a similar manner.

Next, we turn to Proposition 1. First, we present a lemma describing the recursive structure of the cross-cumulant sequences. The proof of the proposition is a direct consequence of the lemma.

Lemma 1: The following recursions hold.

$$g_1(m_1)=\begin{cases}\sum_{i=1}^{K_a} a(i)g_1(m_1-i)\\ \quad\times+\gamma_2\sum_{i=0}^{K_b} b(i)\delta(m_1-i)\\ \quad\times+\gamma_3\sum_{j=1}^{K_{cu}} c(j,j)\delta(m_1-j)\\ +\bar{y}\gamma_2\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}} c(i,j)\delta(m_1-j);\\ \quad m_1>0\\ \gamma_2;\quad m_1=0\\ 0;\quad m_1<0\end{cases} \tag{63}$$

$$g_2(m_1,m_2)=\begin{cases}\sum_{i=1}^{K_a} a(i)g_2(m_1-i,m_2-i)\\ \quad\times+\gamma_3\sum_{i=0}^{K_b} b(i)\delta(m_1-i)\delta(m_2-i)\\ +\gamma_4\sum_{j=1}^{K_{cu}} c(j,j)\delta(m_1-j)\delta(m_2-j)\\ \quad+\gamma_2\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}} c(i,j)c_y^{(2)}(m_1-i)\delta(m_2-j)\\ +\gamma_2^2 c(m_2,m_2)\delta(m_1-m_2)\\ \quad+\bar{y}\gamma_3\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}} c(i,j)\delta(m_1-j)\delta(m_2-j)\\ \quad m_1\geq m_2\geq 0\\ 0;\quad m_1>0, m_2=0\\ \gamma_3;\quad m_1=m_2=0\\ 0;\quad m_1\geq m_2<0\end{cases} \tag{64}$$

$$g_3(m_1,m_2,m_3)=\begin{cases}\sum_{i=1}^{K_a} a(i)g_3(m_1-i,m_2-i,m_3-i)\\ +\sum_{i=m_3}^{K_{cy}} c(i,m_3)[\gamma_2 c_y^{(3)}\\ \quad\times(m_1-i,m_2-i)+\gamma_3 c_y^{(2)}\\ \quad\times(m_2-i)\delta(m_1-m_3)]\\ +\gamma_4\sum_{i=0}^{K_b} b(i)\delta(m_1-i)\delta(m_2-i)\\ \quad\times\delta(m_3-i)\\ +\gamma_5\sum_{j=1}^{K_{cu}} c(j,j)\delta(m_1-j)\delta(m_2-j)\\ \quad\times\delta(m_3-j)\\ +\gamma_3\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}} c(i,j)c_y^{(2)}(m_1-i)\\ \quad\times\delta(m_2-j)\delta(m_3-j)\\ +3\gamma_3\gamma_2\sum_{j=1}^{K_{cu}} c(j,j)\delta(m_1-j)\\ \quad\times\delta(m_2-j)\delta(m_3-j)\\ +\bar{y}\gamma_4\sum_{i=m_3}^{K_{cy}} c(i,m_3)\delta(m_1-m_2)\\ \quad\times\delta(m_3-m_2);\quad m_1\geq m_3,\\ \quad m_2\geq m_3,\quad m_3\geq 0\\ 0;\quad m_1>0,\quad m_2>0,\quad m_3=0\\ \gamma_4;\quad m_1=m_2=m_3=0\\ 0;\quad m_1\geq m_3,\ m_2\geq m_3,\ m_3<0\end{cases} \tag{65}$$

and

$$g_4(m_1,m_2,m_3,m_4)=\begin{cases}\sum_{i=1}^{K_a} a(i)g_4(m_1-i,m_2-i,m_3-i,m_4-i)\\ +\sum_{i=m_4}^{K_{cy}} c(i,m_4)[\gamma_2 c_y^{(4)}\\ \quad\times(m_1-i,m_2-i,m_3-i)\\ +\gamma_3 c_y^{(3)}(m_2-i,m_3-i)\\ \quad\times\delta(m_1-m_4)\\ +\gamma_3 c_y^{(3)}(m_1-i,m_3-i)\\ \quad\times\delta(m_2-m_4)\\ +\gamma_4 c_y^{(2)}(m_3-i)\delta(m_1-m_4)\\ \quad\times\delta(m_2-m_4)];\quad m_1\geq m_4,\\ \quad m_2\geq m_4,\quad m_3>m_4,\\ \quad m_4\geq 0\\ 0;\quad m_1>0,\quad m_2>0,\quad m_3>0,\\ \quad m_4=0\\ \gamma_5;\quad m_1=m_2=m_3=m_4=0\\ 0;\quad m_1\geq m_4,\quad m_2\geq m_4,\\ \quad m_3>m_4,\quad m_4<0.\end{cases} \tag{66}$$

Proof: We start by proving (63). Property 1 ensures that $g_1(m_1) = 0$ for $m_1 < 0$. The initialization stage $g_1(0) = \gamma_2$ follows directly from Property 4. Substituting the bilinear equation (5) into the definition of g_1 gives

$$\begin{aligned}g_1(m_1)&=\text{cum}[y(n), u(n-m_1)]\\ &=\sum_{i=1}^{K_a} a(i)\text{cum}[y(n-i), u(n-m_1)]\end{aligned}$$

$$+\sum_{i=0}^{K_b} b(i)\text{cum}[u(n-i), u(n-m_1)] + \sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}} c(i,j)\text{cum}[y(n-i)u(n-j), u(n-m_1)] + \text{cum}[v(n), u(n-m_1)]\,; \qquad m_1 \geq 0. \tag{67}$$

Since $v(n)$ and $u(n-m_1)$ are independent, the last term on the right-hand side of the above expression is zero. Using the Leonov–Shiryaev theorem on the third term of the right-hand side yields

$$\text{cum}[y(n-i)u(n-j), u(n-m_1)] = \text{cum}[y(n-i)\, u(n-j), u(n-m_1)] + \bar{y}\text{cum}[u(n-j), u(n-m_1)]. \tag{68}$$

Among the terms of the form $\text{cum}[y(n-i), u(n-j), u(n-m_1)]$, with $1 \leq j \leq K_{cu}$ and $j \leq i \leq K_{cy}$, only the term $\text{cum}[y(n-m_1), u(n-m_1), u(n-m_1)]$ is nonzero due to Property 1. When $i = j = m_1$, Property 4 implies that $\text{cum}[y(n-m_1), u(n-m_1), u(n-m_1)] = \gamma_3$. Applying these results along with the fact that $u(n)$ is white leads to (63).

Next, we prove (64). Properties 1 and 3 ensure that $g_2(m_1, m_2) = 0$ for $m_1 \geq m_2 < 0$. The initialization $g_2(0,0) = \gamma_3$ follows directly from Property 4. Substituting the bilinear equation (5) into the definition of g_2 gives

$$\begin{aligned} g_2(m_1, m_2) =& \text{cum}[y(n), y(n-m_1), u(n-m_2)] \\ =& \sum_{i=1}^{K_a} a(i)\text{cum}[y(n-i), y(n-m_1), u(n-m_2)] \\ &+ \sum_{i=0}^{K_b} b(i)\text{cum}[u(n-i), y(n-m_1), u(n-m_2)] \\ &+ \sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}} c(i,j)\text{cum}[y(n-i)u(n-j), y(n-m_1), u(n-m_2)] \\ &+ \text{cum}[v(n), y(n-m_1), u(n-m_2)] \\ & m_1 \geq m_2 \geq 0. \end{aligned} \tag{69}$$

The last term on the right-hand side of the above equation is zero. Let us focus now on the second term in the right-hand side of the same equation. For any $m_2 \geq 0$, $m_1 \geq m_2$ and $i < m_2$, $\text{cum}[u(n-i), y(n-m_1), u(n-m_2)] = 0$ because of Property 2. If $i > m_2$, the above cumulant is again zero due to Property 1. Hence, the second term becomes

$$\gamma_3 \sum_{i=0}^{K_b} b(i)\delta(m_1 - i)\delta(m_2 - i) \tag{70}$$

and is nonzero only when $0 \leq m_2 \leq K_b$ and $m_1 = m_2 = i$. Using the Leonov–Shiryaev theorem, the third term can be expanded as

$$\begin{aligned} &\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}} c(i,j)\text{cum}[y(n-i), u(n-j), y(n-m_1), u(n-m_2)] \\ &+ \sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}} c(i,j) \times \text{cum}[y(n-i), y(n-m_1)]\text{cum}[u(n-j), u(n-m_2)] \\ &+ \sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}} c(i,j) \times \text{cum}[y(n-i), u(n-m_2)]\text{cum}[u(n-j), y(n-m_1)] \\ &+ \bar{y}\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}} c(i,j)\text{cum}[u(n-j), y(n-m_1), u(n-m_2)]. \end{aligned} \tag{71}$$

Arguing as before, we find that $\text{cum}[y(n-i), u(n-j), y(n-m_1), u(n-m_2)] = 0$ for $j \neq m_2$ and any i. Moreover, for $j = m_2$, the only nonzero term occurs when $i = j = m_1 = m_2$. Hence, the first term becomes

$$\gamma_4 \sum_{j=1}^{K_{cu}} c(j,j)\delta(m_1 - j)\delta(m_2 - j)\,. \tag{72}$$

In a similar manner, we can show that the second term is

$$\gamma_2 \sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}} c(i,j)c_y^{(2)}(m_1 - i)\delta(m_2 - j) \tag{73}$$

and that the third term is

$$\sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}} c(i,j)g_1(m_2 - i)g_1(j - m_1). \tag{74}$$

Recall that for the calculations of interest here, $i \geq j$, and $m_1 \geq m_2$. Since $g_1(m_1) = 0$ for $m_1 < 0$, the only nonzero term results when $i = j = m_1 = m_2$. Hence, the previous term becomes

$$\gamma_2^2 c(m_2, m_2)\delta(m_1 - m_2). \tag{75}$$

Reasoning as before, the last term of (71) becomes

$$\bar{y}\gamma_3 \sum_{j=1}^{K_{cu}}\sum_{i=j}^{K_{cy}} c(i,j)\delta(m_1 - j)\delta(m_2 - j). \tag{76}$$

Substituting the above results in (69) and making use of the cross-cumulant definitions, we obtain (64). Similar arguments are employed to establish (65) and (66).

APPENDIX C
MIXED-MA ESTIMATOR

Initialization Module

We begin the derivations by considering the relationship for $c_y^{(3)}(l_1, l_2)$ for $l_1 = K_{cu}$ and $l_2 = L > K_{cu}$. Using the simplifications possible through Lemma 1, we get

$$c_y^{(3)}(K_{cu}, L) = \sum_{i=1}^{K_a} a(i)c_y^{(3)}(K_{cu} - i, L - i) + \gamma_2 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu})c_y^{(2)}(L - i). \tag{77}$$

We observe from the definition of $D_3(l_1, l_2)$ in (25) that

$$D_3(K_{cu}, L) = c_y^{(3)}(K_{cu}, L) - \sum_{i=1}^{K_a} a(i) c_y^{(3)}(K_{cu} - i, L - i). \quad (78)$$

Substituting this result in (77) gives

$$D_3(K_{cu}, L) = \gamma_2 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) c_y^{(2)}(L - i). \quad (79)$$

Successive evaluation of (78) for $R \geq K_{cy} - K_{cu} + 1$ values of L leads to (30).

To derive (31), we evaluate (15) for $l_1 = K_{cu}, l_2 = K_{cu}$, and $l_3 = L > K_{cu}$. Applying Lemma 1 to the various terms in (15) for these choices of the parameters, we get

$$\begin{aligned} c_y^{(4)}(K_{cu}, K_{cu}, L) = &\sum_{i=1}^{K_a} a(i) c_y^{(4)}(K_{cu} - i, K_{cu} - i, L - i) \\ &+ 2\gamma_2 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) c_y^{(3)}(K_{cu} - i, L - i) \\ &+ \gamma_3 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) c_y^{(2)}(L - i). \end{aligned} \quad (80)$$

Substituting the definitions for $D_4(K_{cu}, K_{cu}, L)$ from (26) and the expression for $D_3(K_{cu}, L)$ from (78) in (80) results in

$$\begin{aligned} D_4(K_{cu}, K_{cu}, L) = &2\gamma_2 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) c_y^{(3)}(K_{cu} - i, L - i) \\ &+ \frac{\gamma_3}{\gamma_2} D_3(K_{cu}, L). \end{aligned} \quad (81)$$

The expression for γ_3/γ_2 in (31) follows from this result.

Next, we consider (13) for $l_1 = K_{cu}$. Again, using Lemma 1, we get

$$\begin{aligned} c_y^{(2)}(K_{cu}) = &\sum_{i=1}^{K_a} a(i) c_y^{(2)}(K_{cu} - i) + \gamma_2 b(K_{cu}) \\ &+ \bar{y}\gamma_2 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) + \gamma_3 c(K_{cu}, K_{cu}). \end{aligned} \quad (82)$$

Substituting for $D_2(K_{cu})$ from (24) in the above equation and rearranging the terms gives

$$\gamma_2 b(K_{cu}) = D_2(K_{cu}) - \bar{y}\gamma_2 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) - \gamma_3 c(K_{cu}, K_{cu}) \quad (83)$$

which is identical to (32).

In a similar manner, (33) follows from (14) for the case when $l_1 = l_2 = K_{cu}$. Then

$$\begin{aligned} D_3(K_{cu}, K_{cu}) = &\gamma_3 b(K_{cu}) + \bar{y}\gamma_3 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) \\ &+ \gamma_4 c(K_{cu}, K_{cu}) \\ &+ 2\gamma_2 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) c_y^{(2)}(K_{cu} - i). \end{aligned} \quad (84)$$

Finally, (34) follows from (15) with $l_1 = l_2 = l_3 = K_{cu}$. Then

$$\begin{aligned} D_4(K_{cu}, K_{cu}, K_{cu}) = &\gamma_4 b(K_{cu}) \\ &+ \bar{y}\gamma_4 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) + \gamma_5 c(K_{cu}, K_{cu}) \\ &+ 3\gamma_2 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) c_y^{(3)}(K_{cu} - i, K_{cu} - i) \\ &+ 3\gamma_3 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) c_y^{(2)}(K_{cu} - i). \end{aligned} \quad (85)$$

This completes the derivation of the initialization module.

Main Module

We first substitute (24) in (13) and apply Lemma 1 to get

$$\begin{aligned} D_2(l) = &\sum_{j=l}^{K_{cu}} g_1(j - l) \left(b(j) + \bar{y} \sum_{i=j}^{K_{cy}} c(i, j) \right) \\ &+ \sum_{j=l}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i, j) g_2(i - l, j - l), \quad l > 0. \end{aligned} \quad (86)$$

We then multiply both sides of (86) by $a(l - m)$ and add the result over l in the range $m \leq l \leq K_{cu}$. This operation gives

$$\begin{aligned} S_2(m) &= -\sum_{l=m}^{K_{cu}} a(l - m) D_2(l) \\ &= -\sum_{l=m}^{K_{cu}} a(l - m) \sum_{j=l}^{K_{cu}} g_1(j - l) \left(b(j) + \bar{y} \sum_{i=j}^{K_{cy}} c(i, j) \right) \\ &\quad - \sum_{l=m}^{K_{cu}} a(l - m) \sum_{j=l}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i, j) g_2(i - l, j - l) \\ &\quad 0 < m \leq K_{cu}. \end{aligned} \quad (87)$$

If we change the order of summations in the first term of the right-hand side, we have

$$\begin{aligned} &-\sum_{l=m}^{K_{cu}} a(l - m) \sum_{j=l}^{K_{cu}} g_1(j - l) \left(b(j) + \bar{y} \sum_{i=j}^{K_{cy}} c(i, j) \right) \\ &= \sum_{j=m}^{K_{cu}} \left(b(j) + \bar{y} \sum_{i=j}^{K_{cy}} c(i, j) \right) \left(-\sum_{l=m}^{j} a(l - m) g_1(j - l) \right) \\ &= \sum_{j=m}^{K_{cu}} \left(b(j) + \bar{y} \sum_{i=j}^{K_{cy}} c(i, j) \right) \left(-\sum_{n=0}^{j-m} a(n) g_1(j - m - n) \right). \end{aligned} \quad (88)$$

Substituting (63) for g_1 on the right-hand side of the above equation transforms this term to

$$\sum_{j=m}^{K_{cu}} \left(b(j) + \bar{y} \sum_{i=j}^{K_{cy}} c(i,j) \right) \times \left[\gamma_2 b(j-m) + \gamma_3 c(j-m,j-m) + \bar{y}\gamma_2 \sum_{l=j-m}^{K_{cy}} c(l,j-m) \right]. \tag{89}$$

Likewise, by changing the order of summations in the second term of the right-hand side of (87), we find that

$$\begin{aligned} &-\sum_{l=m}^{K_{cu}} a(l-m) \sum_{j=l}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i,j) g_2(i-l,j-l) \\ &= \sum_{j=m}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i,j) \left(-\sum_{l=m}^{j} a(l-m) g_2(i-l,j-l) \right) \\ &= \sum_{j=m}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i,j) \left(-\sum_{n=0}^{j-m} a(n) g_2(i-m-n,j-m-n) \right). \end{aligned} \tag{90}$$

Applying (64) to the above equation, its right-hand side becomes

$$\begin{aligned} &\sum_{j=m}^{K_{cu}} c(j,j) \left(\gamma_3 b(j-m) + \left(\gamma_4 + \gamma_2^2 \right) c(j-m,j-m) + \right. \\ &\qquad \left. + \bar{y}\gamma_3 \sum_{l=j-m}^{K_{cy}} c(l,j-m) \right) \\ &\quad + \gamma_2 \sum_{j=m}^{K_{cu}} \sum_{i=j+1}^{K_{cy}} c(i,j) \sum_{l=j-m}^{K_{cy}} c(l,j-m) c_y^{(2)}(i-m-l). \end{aligned} \tag{91}$$

If we substitute (89) and (91) into (87), we get

$$\begin{aligned} S_2(m) =& \sum_{j=m}^{K_{cu}} \left(b(j) + \bar{y} \sum_{i=j}^{K_{cy}} c(i,j) \right) \left[\gamma_2 b(j-m) \right. \\ & \left. + \gamma_3 c(j-m,j-m) + \bar{y}\gamma_2 \sum_{l=j-m}^{K_{cy}} c(l,j-m) \right] \\ & + \sum_{j=m}^{K_{cu}} c(j,j) \left[\gamma_3 b(j-m) \right. \\ & \left. + \left(\gamma_4 + \gamma_2^2 \right) c(j-m,j-m) + \bar{y}\gamma_3 \sum_{l=j-m}^{K_{cy}} c(l,j-m) \right] \\ & + \gamma_2 \sum_{j=m}^{K_{cu}} \sum_{i=j+1}^{K_{cy}} c(i,j) \sum_{l=j-m}^{K_{cy}} c(l,j-m) c_y^{(2)}(i-m-l) \\ & 0 < m \le K_{cu}. \end{aligned} \tag{92}$$

Using a similar approach, we next derive expressions for $S_3(m, K_{cu})$ and $S_4(m, K_{cu}, L)$. We only outline the derivation of $S_3(m, K_{cu})$ below. Application of Lemma 1 to (14) for $l_2 = K_{cu}$ gives

$$\begin{aligned} c_y^{(3)}(l, K_{cu}) =& \sum_{i=1}^{K_a} a(i) c_y^{(3)}(l-i, K_{cu}-i) \\ & + \sum_{i=l}^{K_b} b(i) g_2(K_{cu}-l, i-l) \\ & + \sum_{j=l}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i,j) g_3(i-l, K_{cu}-l, j-l) \\ & + \sum_{j=l}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i,j) c_y^{(2)}(K_{cu}-i) g_1(j-l) \\ & + \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) c_y^{(2)}(l-i) g_1(0) \\ & + \bar{y} \sum_{j=l}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i,j) g_2(K_{cu}-l, j-l),\ l > 0. \end{aligned} \tag{93}$$

Substituting for $D(l, K_{cu})$ from (25) into the above equation results in

$$\begin{aligned} D_3(l, K_{cu}) =& \sum_{j=l}^{K_{cu}} g_2(K_{cu}-l, j-l) \left(b(j) + \bar{y} \sum_{i=j}^{K_{cy}} c(i,j) \right) \\ & + \sum_{j=l}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i,j) g_3(i-l, K_{cu}-l, j-l) \\ & + \sum_{j=l}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i,j) c_y^{(2)}(K_{cu}-i) g_1(j-l) \\ & + \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu}) c_y^{(2)}(l-i) g_1(0), \quad l > 0. \end{aligned} \tag{94}$$

Multiplying both sides of (94) with $a(l-m)$ and adding the results over l in the range $m \le l \le K_{cu}$ gives

$$\begin{aligned} S_3(m, K_{cu}) =& \\ & -\sum_{l=m}^{K} a(l-m) D_3(l, K_{cu}) = \\ & -\sum_{l=m}^{K_{cu}} a(l-m) \sum_{j=l}^{K_{cu}} g_2(K_{cu}-l, j-l) \left(b(j) + \bar{y} \sum_{i=j}^{K_{cy}} c(i,j) \right) \\ & -\sum_{l=m}^{K_{cu}} a(l-m) \sum_{j=l}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i,j) g_3(i-l, K_{cu}-l, j-l) \\ & -\sum_{l=m}^{K_{cu}} a(l-m) \sum_{j=l}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i,j) c_y^{(2)}(K_{cu}-i) g_1(j-l) \end{aligned}$$

$$-\sum_{l=m}^{K_{cu}} a(l-m) \sum_{i=K_{cu}}^{K_{cy}} c(i,K_{cu})c_y^{(2)}(l-i)g_1(0)$$
$$0 < m \le K_{cu}. \quad (95)$$

As was done before, we change the order of the summations in the first term of the right-hand side for (95), and we obtain

$$\sum_{j=m}^{K_{cu}} \left(b(j) + \bar{y} \sum_{i=j}^{K_{cy}} c(i,j) \right) \times \left(-\sum_{n=0}^{j-m} a(n) g_2(K_{cu}-m-n, j-m-n) \right). \quad (96)$$

Applying (64) to the above expression results in

$$\gamma_2 \sum_{j=m}^{K_{cu}} \left(b(j) + \bar{y} \sum_{i=j}^{K_{cy}} c(i,j) \right) \sum_{l=j-m}^{K_{cy}} c(l,j-m) c_y^{(2)}(K_{cu}-m-l)$$
$$+ \left(b(K_{cu}) + \bar{y} \sum_{i=K_{cu}}^{K_{cy}} c(i,K_{cu}) \right) \times$$
$$\left[\gamma_3 b(K_{cu}-m) + \left(\gamma_4+\gamma_2^2\right) c(K_{cu}-m, K_{cu}-m) + \bar{y}\gamma_3 \sum_{l=K_{cu}-m}^{K_{cy}} c(l, K_{cu}-m) \right]. \quad (97)$$

In a similar manner, we can show that the second term of the right-hand side of (95) is equal to

$$\gamma_2 \sum_{j=m}^{K_{cu}} \sum_{i=j}^{K_{cy}} \sum_{l=j-m}^{K_{cy}} c(i,j)c(l,j-m)c_y^{(3)}(i-m-l, K_{cu}-m-l)$$
$$+\gamma_3 \sum_{j=m}^{K_{cu}} \sum_{l=j-m}^{K_{cy}} c(j,j)c(l,j-m)c_y^{(2)}(K_{cu}-m-l)$$
$$+\gamma_3 \sum_{i=K_{cu}}^{K_{cy}} \sum_{l=K_{cu}-m}^{K_{cy}} c(i,K_{cu})c(l,K_{cu}-m)c_y^{(2)}(i-m-l)$$
$$+c(K_{cu},K_{cu}) \times$$
$$\left[\gamma_4 b(K_{cu}-m) + (\gamma_5+3\gamma_3\gamma_2) c(K_{cu}-m, K_{cu}-m) + \bar{y}\gamma_4 \sum_{l=K_{cu}-m}^{K_{cy}} c(l,K_{cu}-m) \right]. \quad (98)$$

Similarly, the third term of the right-hand-side of (95) becomes

$$\sum_{j=m}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i,j) c_y^{(2)}(K_{cu}-i) \times \left[\gamma_2 b(j-m) + \gamma_3 c(j-m,j-m) + \bar{y}\gamma_2 \sum_{l=j-m}^{K_{cy}} c(l,j-m) \right]. \quad (99)$$

Substitution of (97)–(99) into (95) gives

$$S_3(m,K_{cu}) =$$
$$\gamma_2 \sum_{j=m}^{K_{cu}} \left(b(j) + \bar{y} \sum_{i=j}^{K_{cy}} c(i,j) \right) \times \sum_{l=j-m}^{K_{cy}} c(l,j-m) c_y^{(2)}(K_{cu}-m-l)$$
$$+ \left(b(K_{cu}) + \bar{y} \sum_{i=K_{cu}}^{K_{cy}} c(i,K_{cu}) \right)$$
$$\times \left[\gamma_3 b(K_{cu}-m) + \left(\gamma_4+\gamma_2^2\right) c(K_{cu}-m,K_{cu}-m) + \bar{y}\gamma_3 \sum_{l=K_{cu}-m}^{K_{cy}} c(l,K_{cu}-m) \right] + \sum_{j=m}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i,j) c_y^{(2)}(K_{cu}-i)$$
$$\times \left[\gamma_2 b(j-m) + \gamma_3 c(j-m,j-m) + \bar{y}\gamma_2 \sum_{l=j-m}^{K_{cy}} c(l,j-m) \right]$$
$$+\gamma_2 \sum_{j=m}^{K_{cu}} \sum_{i=j}^{K_{cy}} \sum_{l=j-m}^{K_{cy}} c(i,j)c(l,j-m)c_y^{(3)}(i-m-l, K_{cu}-m-l)$$
$$+\gamma_3 \sum_{j=m}^{K_{cu}} \sum_{l=j-m}^{K_{cy}} c(j,j)c(l,j-m)c_y^{(2)}(K_{cu}-m-l)$$
$$+\gamma_3 \sum_{i=K_{cu}}^{K_{cy}} \sum_{l=K_{cu}-m}^{K_{cy}} c(i,K_{cu})c(l,K_{cu}-m)c_y^{(2)}(i-m-l)$$
$$+c(K_{cu},K_{cu})$$
$$\times \left[\gamma_4 b(K_{cu}-m) + (\gamma_5+3\gamma_3\gamma_2) c(K_{cu}-m,K_{cu}-m) + \bar{y}\gamma_4 \sum_{l=K_{cu}-m}^{K_{cy}} c(l,K_{cu}-m) \right]$$
$$-\sum_{l=m}^{K} a(l-m)\gamma_2 \sum_{i=K}^{K_{cy}} c(i,K_{cu}) c_y^{(2)}(l-i), \quad 0 < m \le K_{cu}. \quad (100)$$

A similar analysis leads to the following expression for $S_4(m, K_{cu}, L)$:

$$S_4(m,K_{cu},L) =$$
$$\gamma_2 \sum_{j=m+1}^{K_{cu}} \left(b(j) + \bar{y} \sum_{i=j}^{K_{cy}} c(i,j) \right) \times \sum_{l=j-m}^{K_{cy}} c(l,j-m) c_y^{(3)}(K_{cu}-m-l, L-m-l)$$
$$+\gamma_3 \left(b(K_{cu}) + \bar{y} \sum_{i=K_{cu}}^{K_{cy}} c(i,K_{cu}) \right)$$

$$
\begin{aligned}
&\times \sum_{l=K_{cu}-m}^{K_{cy}} c(l, K_{cu}-m) c_y^{(2)}(L-m-l) \\
&+\gamma_2 \sum_{j=m+1}^{K_{cu}} \sum_{i=j}^{K_{cy}} \sum_{l=j-m}^{K_{cy}} c(i,j)c(l,j-m) \\
&c_y^{(4)}(i-m-l, K_{cu}-m-l, L-m-l) \\
&+\gamma_3 \sum_{j=m+1}^{K_{cu}} \sum_{l=j-m}^{K_{cy}} c(j,j)c(l,j-m)c_y^{(3)}(K_{cu}-m-l, L-m-l) \\
&+\gamma_3 \sum_{i=K_{cu}}^{K_{cy}} \sum_{l=K_{cu}-m}^{K_{cy}} c(i, K_{cu})c(l, K_{cu}-m) \\
&c_y^{(3)}(i-m-l, L-m-l) \\
&+\gamma_4 c(K_{cu}, K_{cu}) \sum_{l=K_{cu}-m}^{K_{cy}} c(l, K_{cu}-m) c_y^{(2)}(L-m-l) \\
&+\gamma_2 \sum_{j=m+1}^{K_{cu}} \sum_{i=j}^{K_{cy}} \sum_{l=j-m}^{K_{cy}} c(i,j)c(l,j-m) \\
&c_y^{(2)}(K_{cu}-i)c_y^{(2)}(L-m-l) \\
&+\gamma_2 \sum_{j=m}^{K_{cu}} \sum_{i=j}^{K_{cy}} \sum_{l=j-m}^{K_{cy}} c(i,j)c(l,j-m) \\
&c_y^{(2)}(L-i)c_y^{(2)}(K_{cu}-m-l) \\
&+\sum_{i=K_{cu}}^{K_{cu}} c(i, K_{cu})c_y^{(2)}(L-i) \Bigg[\gamma_3 b(K_{cu}-m) \\
&+(\gamma_4+\gamma_2^2)\, c(K_{cu}-m, K_{cu}-m) \\
&+\bar{y}\gamma_3 \sum_{l=K_{cu}-m}^{K_{cy}} c(l, K_{cu}-m) \Bigg] \\
&+\sum_{j=m}^{K_{cu}} \sum_{i=j}^{K_{cy}} c(i,j)c_y^{(3)}(K_{cu}-i, L-i) \\
&\times \left[\gamma_2 b(j-m) + \gamma_3 c(j-m, j-m) + \bar{y}\gamma_2 \sum_{l=j-m}^{K_{cy}} c(l, j-m) \right] \\
&-\sum_{l=m}^{K} a(l-m)\gamma_2 \sum_{i=K_{cu}}^{K_{cy}} c(i, K_{cu})c_y^{(3)}(l-i, L-i) \\
&0 < m \leq K_{cu}.
\end{aligned}
\tag{101}
$$

The substitution $m \longleftarrow K_{cu} - m$ in (101), (100), and (92) leads to (38), (45), and (48).

REFERENCES

[1] S. Benedetto and E. Biglieri, "Nonlinear equalization of digital satellite channels," *IEEE J. Select. Areas Commun.*, vol. SAC-1, pp. 57–62, Jan. 1983.

[2] S. A. Billings, "Identification of nonlinear systems—a survey," *Proc. Inst. Elect. Eng. D*, vol. 127, no. 6, pp. 272–285, Nov. 1980.

[3] T. Bose and M. Q. Chen, "BIBO stability of the discrete bilinear systems," *Digital Signal Process. Rev. J.*, vol. 5, no. 3, 1995.

[4] S. Boyd, "Volterra Series: Engineering Fundamentals," Ph.D. dissertation, Univ. California, Berkeley, 1985.

[5] W. Favoreel, B. De Moor, and P. Van Overschee, "Subspace identification of bilinear systems subject to white inputs," *IEEE Trans. Automat. Contr.*, vol. 44, pp. 1157–1165, June 1999.

[6] W. Favoreel, "Subspace methods for identification and control of linear and bilinear systems," Ph.D. dissertation, Dept. Elect. Eng., Katholieke Univ. Leuven, Leuven, Belgium, 1999.

[7] W. A. Frank, "An efficient approximation to the quadratic Volterra filter and its application in real-time loudspeaker linearization," *Signal Process.*, vol. 45, no. 1, pp. 97–114, July 1995.

[8] G. Gatt and N. Kalouptsidis, "Identification of discrete-time state affine state space models using cumulants," *Automatica*, vol. 38, pp. 1663–1681, Oct. 2002.

[9] G. B. Giannakis and E. Serpedin, "Linear multichannel blind equalizers of nonlinear FIR Volterra channels," *IEEE Trans. Signal Processing*, vol. 45, pp. 67–81, Jan. 1997.

[10] G. O. Glentis, P. Koukoulas, and N. Kalouptsidis, "Efficient algorithms for Volterra system identification," *IEEE Trans. Signal Processing*, vol. 47, pp. 3042–3057, Nov. 1999.

[11] S. Han, J. Kim, and K. Sung, "Extended generalized total least squares method for the identification of bilinear systems," *IEEE Trans. Signal Processing*, vol. 44, pp. 1015–1018, Apr. 1996.

[12] R. Hermann, "Volterra modeling of digital magnetic saturation recording channels," *IEEE Trans. Magn.*, vol. 26, no. 5, pp. 2125–2127, Sept. 1990.

[13] N. Kalouptsidis, *Signal Processing Systems: Theory and Design*. New York: Wiley, 1997.

[14] K. I. Kim and E. J. Powers, "A digital method of modeling quadratically nonlinear systems with a general random input," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1758–1769, Nov. 1988.

[15] T. Koh and E. J. Powers, "Second-order Volterra filtering and its application to nonlinear system identification," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-33, pp. 1445–1455, Dec. 1985.

[16] N. Kolokotronis, G. Gatt, and N. Kalouptsidis, "On the generation of sequences simulating higher order white noise for a system identification," *IEEE Trans. Signal Processing*, submitted for publication.

[17] M. J. Korenberg and L. D. Paarmann, "Orthogonal approaches to time-series analysis and system identification," *IEEE Signal Processing Mag.*, vol. 8, pp. 29–43, July 1991.

[18] S. Kotsios and N. Kalouptsidis, "BIBO stability criteria for a certain class of discrete nonlinear systems," *Int. J. Contr.*, vol. 58, no. 3, pp. 707–730, Sept. 1993.

[19] P. Koukoulas and N. Kalouptsidis, "Nonlinear system identification using Gaussian inputs," *IEEE Trans. Signal Processing*, vol. 43, pp. 1831–1841, Aug. 1995.

[20] ——, "Third order Volterra system identification," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process.*, Munich, Germany, Apr. 1997, pp. 2405–2408.

[21] ——, "Second-order Volterra system identification," *IEEE Trans. Signal Processing*, vol. 48, pp. 3574–3577, Dec. 2000.

[22] P. Koukoulas, V. Tsoulkas, G. Gatt, and N. Kalouptsidis, "Identification of discrete-time Volterra series using higher order statistics," in *Preprints first Nonlinear Contr. Network Pedagogical School*, Athens, Greece, Sept. 1999, pp. 149–201.

[23] P. Koukoulas, V. Tsoulkas, and N. Kalouptsidis, "A cumulant based algorithm for the identification of input output quadratic systems," *Automatica*, vol. 38, pp. 391–407, Mar. 2002.

[24] P. Koukoulas and N. Kalouptsidis, "Blind identification of second order Hammerstein series," *Signal Process.*, vol. 83, pp. 213–234, Jan. 2003.

[25] G. Lazzarin, S. Pupolin, and A. Sarti, "Nonlinearity compensation in digital radio systems," *IEEE Trans. Commun.*, vol. 42, pp. 988–999, Feb./Apr. 1994.

[26] J. Lee and V. J. Mathews, "A stability condition for certain bilinear systems," *IEEE Trans. Signal Processing*, vol. 42, pp. 1871–1873, July 1994.

[27] ——, "A fast recursive least squares adaptive second-order Volterra filter and its performance analysis," *IEEE Trans. Signal Processing*, vol. 41, pp. 1087–1102, Mar. 1993.

[28] V. P. Leonov and A. N. Shiryaev, "On a method of calculation of semi-invariants," *Theory Prob. Appl.*, vol. 4, pp. 319–329, 1959.

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

[30] V. J. Mathews and T. K. Moon, "Parameter estimation for a bilinear time series model," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Toronto, ON, Canada, May 1991, pp. 3513–3516.

[31] V. J. Mathews, "Adaptive polynomial filters," *IEEE Signal Processing Mag.*, vol. 8, pp. 10–26, July 1991.

[32] V. J. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*. New York: Wiley, 2000.

[33] S. W. Nam and E. J. Powers, "Application of higher order spectral analysis to cubically nonlinear system identification," *IEEE Trans. Signal Processing*, vol. 42, pp. 1746–1765, July 1994.
[34] C. L. Nikias and A. P. Petropulu, *Higher-Order Spectral Analysis: A Nonlinear Signal Processing Framework*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
[35] S. Prakriya and D. Hatzinakos, "Blind identification of LTI-ZMNL-LTI nonlinear channel models," *IEEE Trans. Signal Processing*, vol. 43, pp. 3007–3013, Dec. 1995.
[36] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*. New York: Wiley, 1980.
[37] S. A. Sesay and T. Subba Rao, "Difference equations for higher order moments and cumulants for the bilinear time series model BP(p,0,p,1)," *J. Time Series Anal.*, vol. 12, no. 2, pp. 159–177, Feb. 1991.
[38] T. Subba Rao, "On the theory of bilinear time series models," *J. Roy. Soc. Ser. B*, vol. 43, pp. 244–255, 1981.
[39] M. A. Syed and V. J. Mathews, "Lattice algorithms for recursive least squares adaptive second order Volterra filtering," *IEEE Trans. Circuits Syst. II*, vol. 41, pp. 202–214, Mar. 1994.
[40] ——, "QR-decomposition based algorithms for adaptive Volterra filtering," *IEEE Trans. Circuits Syst. I*, vol. 40, pp. 372–382, June 1993.
[41] V. Tsoulkas, P. Koukoulas, and N. Kalouptsidis, "Identification of input output bilinear systems using cumulants," *IEEE Trans. Signal Processing*, vol. 49, pp. 2753–2761, Nov. 2001.

Nicholas Kalouptsidis (SM'85) was born in Athens, Greece, on September 13, 1951. He received the B.Sc. degree in mathematics (with highest honors) from the University of Athens in 1973 and the M.S. and Ph.D. degrees in systems science and mathematics from Washington University, St. Louis, MO, in 1975 and 1976, respectively.

He has held visiting positions at Washington University; Politecnico di Torino, Torino, Italy; Northeastern University, Boston, MA; and CNET Lannion, Paris, France. He has been an Associate Professor and Professor with the Department of Physics, University of Athens. In the Fall of 1998, he was a Clyde Chair Professor with the School of Engineering, University of Utah, Salt Lake City. He is currently a Professor with the Department of Informatics and Telecommunications, University of Athens. He is the author of the textbook *Signal Processing Systems: Theory and Design* (New York: Wiley 1997) and coeditor, with S. Theodoridis, of the book *Adaptive System Identification and Signal Processing Algorithms* (Englewood Cliffs, NJ: Prentice-Hall, 1993). His research interests are in system theory and signal processing.

Panos Koukoulas (M'02) was born in Athens, Greece, on November 17, 1965. He received the B.Sc. degree in physics in 1989, the M.S. degree in electronics and communications in 1993, and the Ph.D. degree in informatics in 1997, all from the University of Athens.

From 1997 to 2000, he was with the Access Network & Wireless Communications Department, Intracom S.A., Athens, where he worked on aspects of digital modem design. He is currently with the Radar and Automation Department of the Hellenic Civil Aviation Authority, Athens. His research interests include nonlinear system identification, higher order statistics, and digital signal processing.

V. John Mathews (F'02) received the B. E. (Hons.) degree in electronics and communication engineering from the University of Madras, Madras, India in 1980 and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Iowa, Iowa City, in 1981 and 1984, respectively.

At the University of Iowa, he was a Teaching/Research Fellow from 1980 to 1984 and a Visiting Assistant Professor with the Department of Electrical and Computer Engineering from 1984 to 1985. He joined at the University of Utah, Salt Lake City, in 1985, where he is now Professor and Chairman of the Department of Electrical and Computer Engineering. His research interests are in adaptive filtering, nonlinear filtering, image compression, and application of signal processing techniques in communication systems and biomedical engineering. He is the author of the book *Polynomial Signal Processing* (New York: Wiley), which he co-authored with Prof. G. L. Sicuranza, University of Trieste, Trieste, Italy.

Prof. Mathews has served as a member of the Signal Processing Theory and Methods Technical Committee, the Education Committee, and the Conference Board of the IEEE Signal Processing Society. He is currently the Vice President— for Finance of the IEEE Signal Processing Society. He is a past associate editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING LETTERS. He was the General Chairman of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) in 2001.